NAVAL POSTGRADUATE SCHOOL
Monterey, California

DISSERTATION

ANALYSIS OF NON-GAUSSIAN PROCESSES USING THE WIENER MODEL OF DISCRETE NONLINEAR SYSTEMS

by

Atalla Ibrahim Hashad

December 1994

Dissertation Advisor: Charles W. Therrien

Approved for public release, distribution is unlimited.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	December 1994	Doctoral Dissertation

4. TITLE AND SUBTITLE
ANALYSIS OF NON-GAUSSIAN PROCESSES USING THE WIENER MODEL OF DISCRETE NONLINEAR SYSTEMS

5. FUNDING NUMBERS

6. AUTHOR(S)
Atalla Ibrahim Hashad

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Naval Postgraduate School
Monterey, CA 93943-5000

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)

10. SPONSORING / MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES
The views expressed in this thesis are those of the author and do not reflect the official policy or position of the Department of Defense or the United States Government.

12a. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for public release; distribution unlimited.

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

Fundamental results developed by Wiener in the 1950's are combined with new work in the area of higher-order statistics to develop and explore a general model for nonlinear stochastic processes. The Wiener model is developed for discrete nonlinear systems and its orthogonality properties are analyzed to characterize its output statistics. An efficient structured procedure for computing the k^{th}-order statistics of the model output is formulated in both the time and frequency domains. Explicit formulas that exploit the structure of the Wiener model are given for computing the cumulants and polyspectra. A necessary condition for a discrete random process to be representable by the Wiener model is discussed. A computationally efficient procedure is given for matching the model output cumulants to estimated cumulants for a given process by minimizing the squared magnitude of the error. Examples of applying this procedure to given sets of data are presented.

14. SUBJECT TERMS	15. NUMBER OF PAGES
Wiener model, nonlinear systems, nonlinear time series, nonlinear models, discrete Laguerre functions	200
	16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	UL

Approved for public release; distribution is unlimited

ANALYSIS OF NON-GAUSSIAN PROCESSES USING THE WIENER MODEL OF DISCRETE NONLINEAR SYSTEMS

Atalla Ibrahim Hashad
Colonel, Egyptian Army
B.S.E.E., Military Technical College, Egypt, 1976
M.S., Military Technical College, Egypt, 1990

DOCTOR OF PHILOSOPHY IN ELECTRICAL ENGINEERING

from the

NAVAL POSTGRADUATE SCHOOL
December 1994

Author: _____
Atalla Ibrahim Hashad

Approved by:

Roberto Cristi Assoc. Professor of Electrical & Computer Engineering	Peter A. W. Lewis Professor of Operations Research & Computer Engineering
Herschel H. Loomis, Jr. Professor of Electrical & Computer Engineering	Murali Tummala Assoc. Professor of Electrical & Computer Engineering
Amr Zaky Asst. Professor of Computer Science	Charles W. Therrien Professor of Electrical & Computer Engineering Dissertation Supervisor

Approved by: _____
Michael A. Morgan, Chairman, Department of
Electrical and Computer Engineering

Approved by: _____
Richard S. Elster, Dean of Instruction

ABSTRACT

Fundamental results developed by Wiener in the 1950's are combined with new work in the area of higher-order statistics to develop and explore a general model for nonlinear stochastic processes. The Wiener model is developed for discrete nonlinear systems and its orthogonality properties are analyzed to characterize its output statistics. An efficient structured procedure for computing the k^{th}-order statistics of the model output is formulated in both the time and frequency domains. Explicit formulas that exploit the structure of the Wiener model are given for computing the cumulants and polyspectra. A necessary condition for a discrete random process to be representable by the Wiener model is discussed. A computationally efficient procedure is given for matching the model output cumulants to estimated cumulants for a given process by minimizing the squared magnitude of the error. Examples of applying this procedure to given sets of data are presented.

TABLE OF CONTENTS

LIST OF FIGURES

I. INTRODUCTION

In the field of signals and systems, as in other scientific areas, we often have measured data that we wish to explain. This data is considered to be explained when we are able to develop an appropriate mathematical model which can produce data "statistically similar" to the data that has been originally measured. Another criterion for "goodness of modeling" could be prediction. When the detailed variation of the data is sufficiently complicated, or when we want to to model an entire *class* of data with different detailed characteristics, a stochastic model is often appropriate. The work in this dissertation develops and explores a general stochastic model for data which combines fundamental results developed by Wiener in the 1950's with new work in the area of higher order statistics. The latter work was brought about by a desire on the part of researchers to move beyond linear Gaussian models and has been bolstered by new developments in high speed computation and digital signal processing.

Although a significant amount of work has already been done, there are still many areas where researchers continue to work on building models to characterize stochastic processes. It is appropriate to cite some of this work (as well as some of the limitations) here.

The stationary linear Gaussian model has dominated work on time series analysis for a large number of years. In this model a time series (or discrete time random process) is represented as the output of a linear time-invariant system driven by zero-mean white Gaussian noise. The two basic assumptions of this model, namely the Gaussianity of the input process and the linearity of the system, has enabled researchers to develop a large number of procedures to identify the system param-

eters such that the model output dependably represents the data. When the linear system is represented by a difference equation of finite order, the model is called a linear Gaussian ARMA (AutoRegressive Moving Average) model of the stochastic processes.

Among engineers, some of the most recognizable contributions to linear Gaussian modeling are due to Wiener [1]. While Wiener concentrated primarily on continuous time processes because of the need to process signals with analog electronic devices, others, such as Wold [2] and Kolmogorov [3] had similar results for discrete time processes. Wiener showed how the second-order statistics of the modeled process can be used to identify the linear system to within a complex scale factor. He showed conditions for which the complex power spectral function of the process can be factored to determine the transfer function of the linear system. In this case, if the modeled process is assumed Gaussian, the model is called the innovations representation of the linear Gaussian process.

The class of linear Gaussian ARMA models has considerable limitations in representing stochastic processes. Among these limitations are the following [4, 5]:

1. Linear difference equations do not allow stable periodic solutions independent of the initial state.

2. The models are not suited for modeling data exhibiting sudden bursts of very large amplitude at irregular time epochs.

3. The models are not suitable for data exhibiting a strong asymmetry (about the mean).

4. The models may not be best for data exhibiting strong cyclicity. Nonlinear approximation for the regression functions may be more appropriate in these

cases, while the regression functions for linear Gaussian models are inherently linear.

5. Since linear Gaussian models are inherently time-reversible, these models are not suitable if the data exhibits time irreversibility.

It can therefore be seen that the linear Gaussian model is incapable of representing a significantly wide class of processes. Since only *Gaussian* processes are completely characterized by their second-order statistics, the higher-order statistics of non-Gaussian processes provide information that can be used to build an adequately dependable model. Recently the problem of modeling non-Gaussian processes has been closely associated with the higher-order statistical analysis of random processes. If the modeled process is far from Gaussian a number of approaches have been followed in attempting to build a reliable model. Among these approaches are "local linear", Multi Variate Adaptive Regression Splines (MARS) and neural nets. In one general approach the model consists of a white non-Gaussian process driving a linear system, while in another approach the model is comprised of a white Gaussian or non-Gaussian process driving a nonlinear system. In addition, there are a whole host a methods that focus on non-Gaussian and or nonlinear processes of one form or another but are specific in the types of processes that they seek to represent. It is safe to say that none of the "general" and none of the specific methods overcomes all of the limitations enumerated above. We can however review some of these methods briefly.

A considerable amount of work has been done in building linear non-Gaussian models for stochastic processes, mainly due to the availability of both analytical and computational tools for higher-order statistics and the relative simplicity of results that arise from the linear model. In such models the process is assumed to be representable as the output of a linear system driven by a k^{th}-order stationary white

non-Gaussian process with known statistics. For such a process the mean is constant and the other moments up to k^{th}-order exist and are functions of only the lag differences. Cumulants and polyspectra of the modeled process are fairly easy to compute and can be used to identify the transfer function of the linear system to within a gain and phase ambiguity. The methods that have been proposed to identify the linear system in this type of models include both parametric and non-parametric methods. In the parametric methods the system is characterized by its ARMA parameters and the transfer function is expressed in terms of these parameters. Then matching a set of the model output statistics to those of the modeled processs enables one to construct a set of equations solvable for the system parameters. In [6, 7] Tugnait uses both the second- and fourth-order statistics to recover the poles and zeros of the linear system transfer functions. Giannakis and Mendel use the output cumulants to determine the AR and the MA orders of the system [8]. Fonollosa and Vidal use a linear combination of the output cumulants to identify the linear system in such models [9]. Several others have also made significant contributions.

Non-parametric methods for identifying the transfer function of the linear system are also used. Lii and Rosenblatt [10] provide a method to estimate the amplitude of the transfer function from the power spectral density function of the process and then show how to estimate the phase from a k^{th}-order polyspectrum. Nikias [11] shows how to estimate both maximum-phase and minimum-phase transfer functions by applying a convolution between the third-order cumulant of the modeled process and the same order complex cepstrum.

Although linear models for non-Gaussian processes have a significant analytical support, they are incapable of representing a large class of non-Gaussian processes. In fact, recent analyses have shown that at least from a theoretical point of view, the set of linear non-Gaussian processes is vanishingly small [12]. From a practical point

of view however it is advisable before modeling a process to test for its linearity. Tong [4] presents linearity tests of stochastic processes in both time and frequency domains. Tekalp [13, 12] establishes the necessary and sufficient conditions for a non-Gaussian process to be representable by a linear non-Gaussian model.

A class of linear and nonlinear stochastic processes can be specified by the dynamics of a system the output of which represents the modeled process. A significant amount of work has been done to specify such dynamic systems. For nonlinear processes the system dynamic representation (e.g., the difference equation) has a kind of nonlinearity according to which different classes of the stochastic outputs can be discriminated. In his book [4], Tong presents some examples of threshold models based on piecewise linearity. He also presents different types of autoregressive models with nonlinear autoregression and other types of specific models. Recent work by Stevens [14], Lewis and Stevens [15] and Lewis , Ray and Stevens [16] has applied Friedman's MARS methodology to give a practical way of applying threshold models.

Another broad class of nonlinear random processes can be represented as the output of a nonlinear system driven by a white random process. The input-output relations of the nonlinear system are described in the Taylor series-like representation with memory known as the Volterra series [17, 18]. The model is completely known when the kernels of this representation are identified. Powers *et al.* present applications of using a finite-order Volterra representation to identify nonlinear systems [19, 20]. In these applications both the input and the output sequences are accessible; therefore different order cross-moments between the input and the output can be estimated and used to obtain the kernels of the representation. In this application the Volterra representation must be finite. Other similar work can be found in [21] and references therein.

The Volterra series representation is limited in its application because a rather

severe condition is imposed on the magnitude of the input process to guarantee convergence of the series. Further the individual kernels cannot be identified separately unless the representation is assumed finite. To overcome these limitations Wiener proposed a model of nonlinear systems which is associated with an orthogonal functional representation of random processes. He called this the G-functional representation because the functionals are orthogonal when the input process is white Gaussian [22, 17]. In the literature the structure and analysis of this model is based upon second-order statistics of the model output. The model kernels are identified via different order cross-correlations between the input and the output. So far, no one has proposed a method for determining the model parameters from output measurements alone as is the case when we try to fit a stochastic model to measured data. However, the use of higher-order statistics of the output suggests the possibility of a new approach where this could be done.

In this dissertation we introduce the Wiener model for nonlinear systems as a general model for a class of nonlinear processes that are necessarily neither periodic nor bandlimited. We also develop the complete details of the representation in discrete time. Wiener's original development of nonlinear systems was for continuous time and there are a few significant differences. Following this we provide a complete statistical analysis of the model. In particular, we develop expressions for the higher-order statistics of the model output in the time and frequency domains as functions of the model parameters. Finally we use these developments in an algorithm to identify the model parameters such that the model output statistics match those of a measured data sequence.

The remainder of the thesis is organized as follows. In Chapter II we summarize the properties of higher-order statistics of discrete random processes that are useful to our application. New methods for finding cumulants and polyspectra of

linear and nonlinear *functionals* of Gaussian random processes are also formulated in this chapter. In Chapter III we describe the Wiener model of discrete nonlinear systems. The orthogonality of the discrete G-functionals is used to specify the model structure and an analysis of this structure is performed. In Chapter IV we formulate the model output cumulants and polyspectra and develop an efficient procedure to compute them. This comprises the main body of the work. The Wiener model has unique structure that simplifies the computation and which has not heretofore been exploited. In Chapter V we consider the problem of modeling a discrete random process using the Wiener model. We present the test of linearity and provide a definition of the class of discrete random processes that can be represented using the nonlinear model. We then give two examples of nonlinear processes resulting from the discrete Wiener model that match estimated cumulants of some measured data. The nonlinear equations resulting from this modeling problem are solved using the Extended Kalman filter technique. Finally, in Chapter VI, we present conclusions and recommendations for future continuation of this research.

II. HIGHER-ORDER MOMENTS, CUMULANTS AND POLYSPECTRA

Higher-order statistics and polyspectra can be applied in a wide variety of problems in signal processing and system theory. By "higher-order" we mean the statistics of orders more than the second. Among higher-order statistics applications are the problems where either non-Gaussianity, nonminimum phase, colored noise, or nonlinearity are of substantial interest [23]. In the theory of nonlinear systems, a system output can be modeled by defining the input-output relation using a Taylor series-like representation. Thus, computing an output statistic of any order involves the computation of higher order statistics of the input. Moreover, the output of a nonlinear system is in general non-Gaussian even if the input is Gaussian.

A. HIGHER-ORDER MOMENTS AND CUMULANTS

This section begins with a general set of definitions for moments and cumulants and then proceeds to discuss properties of special interest in this thesis.

1. Definitions of Moments and Cumulants

The k-dimensional vector $\mathbf{x} = [x_1, x_2, \cdots, x_k]^T$ is used here to denote a collection of real-valued random variables x_i for which the joint moments up to a sufficient order exist. Let $\mathbf{u} = [u_1, u_2, \cdots, u_k]^T$ denote a set of real variables (i.e., $\mathbf{u} \in \mathcal{R}^k$); then the moment generating function is defined by

$$\phi_{\mathbf{x}}^{m}(\mathbf{u}) = \mathcal{E}\left\{e^{j\mathbf{u}^T\mathbf{x}}\right\} \tag{2.1}$$

where $\mathcal{E}$ denotes statistical expectation. The k^{th} order moment

$$\mathrm{M}_{\mathbf{x}}^{(k)} = \mathcal{E}\{x_1 x_2 \cdots x_k\} \tag{2.2}$$

is the coefficient of the term $j^k u_1 u_2 \cdots u_k$ in the Taylor series expansion of $\phi_{\mathbf{x}}^m(\mathbf{u})$. In general the moment of order $m = \nu_1 + \nu_2 + \cdots + \nu_k$, namely

$$\mathrm{M}_{\mathbf{x}}^{(\nu_1+\nu_2+\cdots+\nu_k)} = \mathcal{E}\{x_1^{\nu_1} x_2^{\nu_2} \cdots x_k^{\nu_k}\} \tag{2.3}$$

is the coefficient of the term

$$\frac{j^{\nu_1+\nu_2+\cdots+\nu_k}}{\nu_1!\nu_2!\cdots\nu_k!} u_1^{\nu_1} u_2^{\nu_2} \cdots u_k^{\nu_k}$$

in the same expansion.

The cumulant generating function is defined by

$$\phi_{\mathbf{x}}^c(\mathbf{u}) = \ln \phi_{\mathbf{x}}^m(\mathbf{u}) \tag{2.4}$$

and the k^{th} order cumulant denoted by

$$\mathrm{C}_{\mathbf{x}}^{(k)} = \mathrm{cum}\{x_1 x_2 \cdots x_k\} \tag{2.5}$$

is the coefficient of the term $j^k u_1 u_2 \cdots u_k$ in the Taylor series expansion of $\phi_{\mathbf{x}}^c(\mathbf{u})$. Likewise

$$\mathrm{C}_{\mathbf{x}}^{(\nu_1+\nu_2+\cdots+\nu_k)} = \mathrm{cum}\{x_1^{\nu_1} x_2^{\nu_2} \cdots x_k^{\nu_k}\} \tag{2.6}$$

is the coefficient of the term

$$\frac{j^{\nu_1+\nu_2+\cdots+\nu_k}}{\nu_1!\nu_2!\cdots\nu_k!} u_1^{\nu_1} u_2^{\nu_2} \cdots u_k^{\nu_k}$$

in the same expansion. The k^{th}-order cumulant can be expressed in terms of the moments up to the k^{th} order by the moment-cumulant relation

$$\mathrm{C}_{\mathbf{x}}^{(k)} = \sum_{k^{(1)}+k^{(2)}+\cdots+k^{(q)}=k} \frac{(-1)^{(q-1)}}{q} \frac{k!}{k^{(1)}!k^{(2)}!\cdots k^{(q)}!} \prod_{p=1}^{q} \mathrm{M}_{\mathbf{x}}^{(k^{(p)})} \tag{2.7}$$

where the summation is over all the possible values of q. Similarly the cumulant-moment relation is

$$\mathrm{M}_{\mathbf{x}}^{(k)} = \sum_{k^{(1)}+k^{(2)}+\cdots+k^{(q)}=k} \frac{1}{q!} \frac{k!}{k^{(1)}!k^{(2)}!\cdots k^{(q)}!} \prod_{p=1}^{q} \mathrm{C}_{\mathbf{x}}^{(k^{(p)})} \tag{2.8}$$

In particular, if the exponents are such that $\nu_1 = \nu_2 = \cdots = \nu_k = 1$ then from (2.7) we have

$$\mathrm{C}_{\mathbf{x}}^{(k)} = \sum_{q} (-1)^{(q-1)}(q-1)!\, \mathcal{E}\left\{\prod_{i\in\kappa_1} x_i\right\} \cdots \mathcal{E}\left\{\prod_{i\in\kappa_q} x_i\right\} \tag{2.9}$$

where κ_1, κ_2, $\cdots$ and κ_q are partitions of the set of integers $\{1, 2, \cdots, k\}$ and q is the number of partitions. Further, from (2.8) it follows that

$$\mathcal{E}\{x_1, x_2, \cdots, x_k\} = \sum_{q} D_{\kappa_1} D_{\kappa_2} \cdots D_{\kappa_q} \tag{2.10}$$

where

$$D_{\kappa_p} = \mathrm{cum}(x_{j_1}, x_{j_2}, \cdots, x_{j_m}) \qquad p = 1, 2, 3, \cdots, q \tag{2.11}$$

where the j_i are the elements in the partition κ_p and the summation in (2.10) is over all the possible partitions. Tables (2.1-2.3) show the moment-cumulant relations for orders 2, 3 and 4. The moment is the sum of all cumulant terms in the same column. Likewise the cumulant is the sum of all moment terms in its same column. The value of q and the corresponding partitions are also shown.

partitions			moment	cumulant
q	κ_1	κ_2	$\mathrm{M}_{\mathbf{x}}^{(2)}$	$\mathrm{C}_{\mathbf{x}}^{(2)}$
1	1,2		$\mathrm{C}^{(2)}(x_1x_2)$	$\mathcal{E}\{x_1x_2\}$
2	1	2	$\mathrm{C}^{(1)}(x_1)\mathrm{C}^{(1)}(x_2)$	$-\mathcal{E}\{x_1\}\mathcal{E}\{x_2\}$

TABLE 2.1: Cumulants and moments for order $k = 2$.

Let us denote a discrete random process defined on the integers $-\infty < n < \infty$ as $\{x(n)\}$ or simply as $x(n)$ when there is no possibility of confusion. Unless otherwise

partitions				moment	cumulant
q	κ_1	κ_2	κ_3	$M_x^{(3)}$	$C_x^{(3)}$
1	1,2,3			$C^{(3)}(x_1x_2x_3)$	$\mathcal{E}\{x_1x_2x_3\}$
2	1	2,3		$C^{(1)}(x_1)C^{(2)}(x_2x_3)$	$-\mathcal{E}\{x_1\}\mathcal{E}\{x_2x_3\}$
2	2	1,3		$C^{(1)}(x_2)C^{(2)}(x_1x_3)$	$-\mathcal{E}\{x_2\}\mathcal{E}\{x_1x_3\}$
2	3	1,2		$C^{(1)}(x_3)C^{(2)}(x_1x_2)$	$-\mathcal{E}\{x_3\}\mathcal{E}\{x_1x_2\}$
3	1	2	3	$C^{(1)}(x_1)C^{(1)}(x_2)C^{(1)}(x_3)$	$\mathcal{E}\{x_1\}\mathcal{E}\{x_2\}\mathcal{E}\{x_3\}$

TABLE 2.2: Cumulants and moments for order $k = 3$.

partitions					moment	cumulant
q	κ_1	κ_2	κ_3	κ_4	$M_x^{(4)}$	$C_x^{(4)}$
1	1,2,3,4				$C^{(4)}(x_1x_2x_3x_4)$	$\mathcal{E}\{x_1x_2x_3x_4\}$
2	1,2	3,4			$C^{(2)}(x_1x_2)C^{(2)}(x_3x_4)$	$-\mathcal{E}\{x_1x_2\}\mathcal{E}\{x_3x_4\}$
2	1,3	2,4			$C^{(2)}(x_1x_3)C^{(2)}(x_2x_4)$	$-\mathcal{E}\{x_1x_3\}\mathcal{E}\{x_2x_4\}$
2	1,4	2,3			$C^{(2)}(x_1x_4)C^{(2)}(x_2x_3)$	$-\mathcal{E}\{x_1x_4\}\mathcal{E}\{x_2x_3\}$
2	1	2,3,4			$C^{(1)}(x_1)C^{(3)}(x_2x_3x_4)$	$-\mathcal{E}\{x_1\}\mathcal{E}\{x_2x_3x_4\}$
2	2	1,3,4			$C^{(1)}(x_2)C^{(3)}(x_1x_3x_4)$	$-\mathcal{E}\{x_2\}\mathcal{E}\{x_1x_3x_4\}$
2	3	1,2,4			$C^{(1)}(x_3)C^{(3)}(x_1x_2x_4)$	$-\mathcal{E}\{x_3\}\mathcal{E}\{x_1x_2x_4\}$
2	4	1,2,3			$C^{(1)}(x_4)C^{(3)}(x_1x_2x_3)$	$-\mathcal{E}\{x_4\}\mathcal{E}\{x_1x_2x_3\}$
3	1,2	3	4		$C^{(2)}(x_1x_2)C^{(1)}(x_3)C^{(1)}(x_4)$	$-\mathcal{E}\{x_1x_2\}\mathcal{E}\{x_3\}\mathcal{E}\{x_4\}$
3	1,3	2	4		$C^{(2)}(x_1x_3)C^{(1)}(x_2)C^{(1)}(x_4)$	$2\mathcal{E}\{x_1x_3\}\mathcal{E}\{x_3\}\mathcal{E}\{x_4\}$
3	1,4	2	3		$C^{(2)}(x_1x_4)C^{(1)}(x_2)C^{(1)}(x_3)$	$2\mathcal{E}\{x_1x_4\}\mathcal{E}\{x_2\}\mathcal{E}\{x_4\}$
3	2,3	1	4		$C^{(2)}(x_2x_3)C^{(1)}(x_1)C^{(1)}(x_3)$	$2\mathcal{E}\{x_2x_3\}\mathcal{E}\{x_1\}\mathcal{E}\{x_4\}$
3	2,4	1	3		$C^{(2)}(x_2x_4)C^{(1)}(x_1)C^{(1)}(x_3)$	$2\mathcal{E}\{x_2x_4\}\mathcal{E}\{x_1\}\mathcal{E}\{x_3\}$
3	3,4	1	2		$C^{(2)}(x_3x_4)C^{(1)}(x_1)C^{(1)}(x_2)$	$2\mathcal{E}\{x_3x_4\}\mathcal{E}\{x_1\}\mathcal{E}\{x_2\}$
4	1	2	3	4	$C^{(1)}(x_1)C^{(1)}(x_2)C^{(1)}(x_3)C^{(1)}(x_4)$	$-6\mathcal{E}\{x_1\}\mathcal{E}\{x_2\}\mathcal{E}\{x_3\}\mathcal{E}\{x_4\}$

TABLE 2.3: Cumulants and moments for order $k = 4$.

specified, the random process will be assumed to take on real values, i.e., $x(n) \in \mathcal{R}$. A random process $\{x(n)\}$ is defined to be *m^{th}-order stationary* if for each $k \leq m$ the cumulant of order k for samples of the random process $x(n_0), x(n_1), \cdots, x(n_{k-1})$ is a function of only the lag differences $l_1 = n_1 - n_0, l_2 = n_2 - n_0, \cdots, l_{k-1} = n_{k-1} - n_0$. For an m^{th}-order stationary random process, its k^{th}-order moment and *cumulant functions* are denoted by

$$\mu_{\mathbf{x}}^{(k)}(l_1, l_2, \cdots, l_{k-1}) = \mathcal{E}\{(x(n), x(n+l_1), x(n+l_2), \cdots, x(n+l_{k-1}))\} \qquad k \leq m \tag{2.12}$$

and

$$C_{\mathbf{x}}^{(k)}(l_1, l_2, \cdots, l_{k-1}) = \text{cum}(x(n), x(n+l_1), x(n+l_2), \cdots, x(n+l_{k-1})) \qquad k \leq m \tag{2.13}$$

For brevity, we will also use the notation $\mu_{\mathbf{x}}^{(k)}(\boldsymbol{l})$ and $C_{\mathbf{x}}^{(k)}(\boldsymbol{l})$ where $\boldsymbol{l}$ is the vector argument

$$\boldsymbol{l} = [l_1, l_2, \cdots, l_{k-1}]^T$$

The size of the vector $\boldsymbol{l}$ is implicit with the order of the moment or the cumulant.

When there is no ambiguity it is common to refer to the moment and cumulant functions as simply "moments" or "cumulants" of the process. For a zero-mean random process the second, third and fourth-order cumulants are given by

$$
\begin{aligned}
C_{\mathbf{x}}^{(2)}(l) &= \mathcal{E}\{x(n)x(n+l)\} = \mu_{\mathbf{x}}^{(2)}(l) && (a)\\
C_{\mathbf{x}}^{(3)}(l_1,l_2) &= \mathcal{E}\{x(n)x(n+l_1)x(n+l_2)\} = \mu_{\mathbf{x}}^{(3)}(l_1,l_2) && (b)\\
C_{\mathbf{x}}^{(4)}(l_1,l_2,l_3) &= \mathcal{E}\{x(n)x(n+l_1)x(n+l_2)x(n+l_3)\} - C_{\mathbf{x}}^{(2)}(l_1)C_{\mathbf{x}}^{(2)}(l_2-l_3)\\
&\quad -C_{\mathbf{x}}^{(2)}(l_2)C_{\mathbf{x}}^{(2)}(l_3-l_1) - C_{\mathbf{x}}^{(2)}(l_3)C_{\mathbf{x}}^{(2)}(l_1-l_2)\\
&= \mu_{\mathbf{x}}^{(4)}(l_1,l_2,l_3) - \mu_{\mathbf{x}}^{(2)}(l_1)\mu_{\mathbf{x}}^{(2)}(l_2-l_3) - \mu_{\mathbf{x}}^{(2)}(l_2)\mu_{\mathbf{x}}^{(2)}(l_3-l_1)\\
&\quad -\mu_{\mathbf{x}}^{(2)}(l_3)\mu_{\mathbf{x}}^{(2)}(l_1-l_2) && (c)
\end{aligned}
\tag{2.14}
$$

These formulas follow directly from the corresponding formulas in Tables 2.1-2.3 for random variables.

2. Properties of the Cumulant

Cumulants have properties that follow from the cumulant definition and are important for their application in digital signal processing and system theory. Among these properties are the following [24]:

1. If $x(n)$ is a Gaussian random process its cumulants of orders higher than second are zero. Since the moment generating function of jointly Gaussian random variables with mean vector $\boldsymbol{\mu}$ and covariance matrix $\mathbf{C}$ is given by

$$\phi_{\mathbf{x}}^{m}(\mathbf{u}) = e^{j\mathbf{u}^T\boldsymbol{\mu}-\frac{1}{2}\mathbf{u}^T\mathbf{C}\mathbf{u}} \tag{2.15}$$

the natural logarithm of this function (the cumulant generating function) can be considered as a power series with zero coefficients for terms corresponding to powers greater than 2. Thus the cumulants of all orders greater than second are zero. This is a property of *any* Gaussian process, white or colored.

2. For a zero-mean random process, the third and fourth-order cumulants can be defined as

$$C_{\mathbf{x}}^{(k)}(l_1, l_2, \cdots, l_{k-1}) = M_{\mathbf{x}}^{(k)}(l_1, l_2, \cdots, l_{k-1}) - \mathcal{E}\{g(n+l_1)g(n+l_2)\cdots g(n+l_{k-1})\} \quad \text{for } k = 3 \text{ or } 4 \tag{2.16}$$

where $g(n)$ is a Gaussian process with the same second-order statistics as $x(n)$ [23]. (This does not apply to cumulants of orders higher than fourth). Therefore, the cumulant can provide a measure of the process "difference from Gaussianity." In general, if $x(n)$ is a stationary random process with variance σ^2, the coefficient of Kurtosis is given by:

$$\begin{aligned} &\frac{\mathcal{E}\left\{(x(n) - \mathcal{E}\{x(n)\})^4\right\}}{\sigma^4} - 3 \\ &= \frac{\text{cum}(x(n), x(n), x(n), x(n)) + 3 \cdot \text{cum}(x(n), x(n)) \cdot \text{cum}(x(n), x(n))}{\sigma^4} - 3 \\ &= \frac{C_{\mathbf{x}}^{(4)}(0,0,0) + 3(C_{\mathbf{x}}^{(2)}(0))^2}{\sigma^4} - 3 \end{aligned} \tag{2.17}$$

which equals zero for Gaussian processes because in this case $C_{\mathbf{x}}^{(4)}(0,0,0) = 0$ and $C_{\mathbf{x}}^{(2)}(0) = \sigma^2$.

3. If a random process consisting of independent identically-distributed (i.i.d) random variables has a symmetric distribution (such as Laplace, Uniform, Gaussian and Bernoulli-Gaussian) its third-order cumulant is zero. This is *not* true for non-symmetrically distributed processes (such as exponential, Rayleigh and K-distributions). The coefficient of skewness is

$$\frac{\mathcal{E}\left\{(x(n) - \mathcal{E}\{x(n)\})^3\right\}}{\sigma^3} = \frac{\text{cum}(x(n), x(n), x(n))}{\sigma^3} = \frac{C_{\mathbf{x}}^{(3)}(0,0)}{\sigma^3} \tag{2.18}$$

which is zero for symmetrically distributed processes.

4. If each of the random variables x_i, $i = 1, 2, \cdots, k$ is scaled by a constant a_i the cumulant becomes

$$\text{cum}(a_1 x_1, a_2 x_2, \cdots, a_k x_k) = \left(\prod_{i=1}^{k} a_i \right) \text{cum}(x_1, x_2, \cdots, x_k) \tag{2.19}$$

5. Cumulants are symmetric with respect to their arguments, i.e.,

$$\text{cum}(x_1, x_2, \cdots, x_k) = \text{cum}(x_{i_1}, x_{i_2}, \cdots, x_{i_k}) \tag{2.20}$$

where $(i_1, i_2, \cdots, i_k)$ is any permutation of the indices $\{1, 2, \cdots, k\}$. This implies for example that the third order cumulant of a stationary real random process has six regions of symmetry in the l_1, l_2 plane such that

$$\begin{aligned} &C_{\mathbf{x}}^{(3)}(l_1, l_2) = C_{\mathbf{x}}^{(3)}(l_2, l_1) = C_{\mathbf{x}}^{(3)}(-l_1, l_2 - l_1) = \\ &C_{\mathbf{x}}^{(3)}(l_2 - l_1, -l_1) = C_{\mathbf{x}}^{(3)}(-l_2, l_1 - l_2) = C_{\mathbf{x}}^{(3)}(l_1 - l_2, -l_2) \end{aligned} \tag{2.21}$$

The principal region for which $l_1 \geq l_2 \geq 0$ is called the *non-redundant* region of support of $C_{\mathbf{x}}^{(3)}(l_1, l_2)$. If the value of the cumulant at any point (l_1, l_2) in this region is known, the values of the cumulant at corresponding points in the five other regions of support defined by (2.21) are also known (see Fig.2.1). The number of regions of symmetry increases rapidly with the order of the cumulant. The fourth order cumulant of a real random process has 24 regions of symmetry in the space of l_1, l_2, l_3 .

6. Cumulants are additive with respect to their arguments, i.e.,

$$\text{cum}(x_0 + x_1, x_2, \cdots, x_k) = \text{cum}(x_0, x_2, \cdots, x_k) + \text{cum}(x_1, x_2, \cdots, x_k) \tag{2.22}$$

Therefore the cumulants of sums of random variables are equal to the sums of their cumulants.

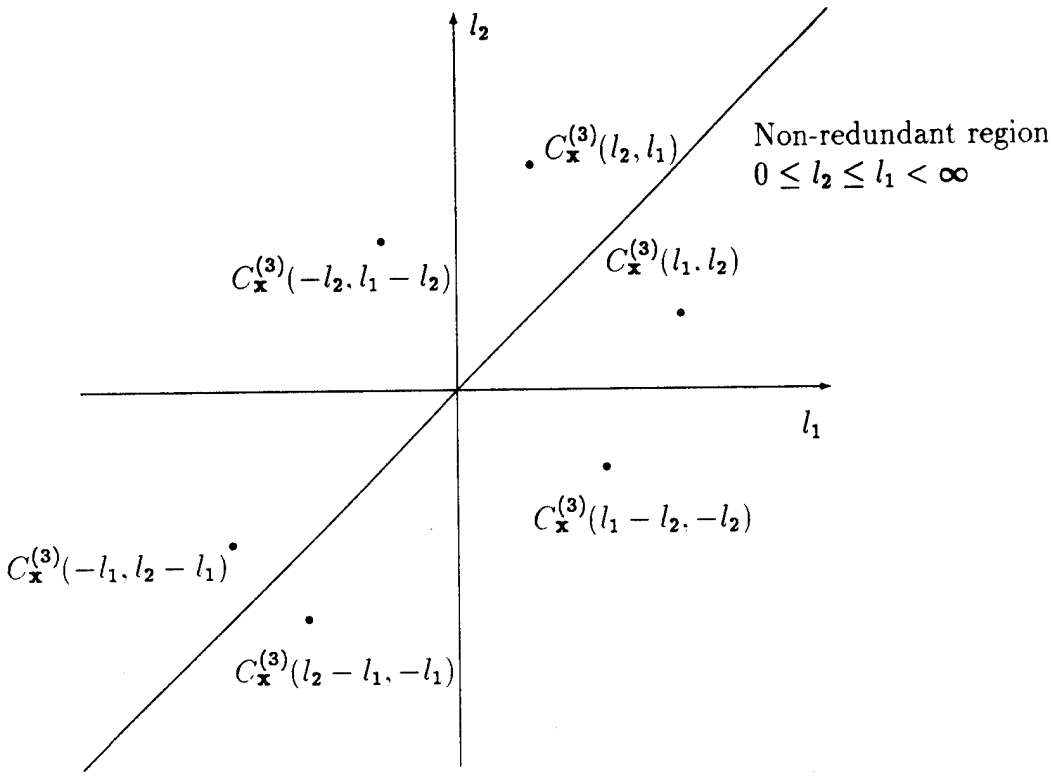

Figure 2.1: Regions of symmetry of $C_{\mathbf{x}}^{(3)}(l_1, l_2)$.

7. If constants are added to the values of the random variables the cumulant does not change its value, i.e.,

$$\text{cum}(a_1 + x_1, a_2 + x_2, \cdots, a_k + x_k) = \text{cum}(x_1, x_2, \cdots, x_k) \tag{2.23}$$

where $a_1, a_2, \cdots, a_k$ are constants.

8. If the sets of random variables $\{x_i\}$ and $\{y_i\}$ are statistically independent then

$$\text{cum}(x_1 + y_1, x_2 + y_2, \cdots, x_k + y_k) = \text{cum}(x_1, x_2, \cdots, x_k) + \text{cum}(y_1, y_2, \cdots, y_k) \tag{2.24}$$

9. If a subset of the random variables $\{x_i\}$ for $i = 1, 2, \cdots, k$ is independent of the rest, then

$$\text{cum}(x_1, x_2, \cdots, x_k) = 0 \tag{2.25}$$

This implies that for a sequence $w(n)$ of independent identically-distributed (i.i.d.) random variables, the cumulant of any order is the multi-dimensional *delta function*, i.e.,

$$C_{\mathbf{w}}^{(k)}(l_1, l_2, \cdots, l_{k-1}) = \gamma_w^{(k)} \delta(l_1)\delta(l_2)\cdots\delta(l_{k-1}); \qquad \text{for } w(n) \text{ i.i.d.} \tag{2.26}$$

where $\gamma_w^{(k)}$ is a real-valued constant and $\delta(l)$ is the unit sample function, i.e.,

$$\delta(l) = \begin{cases} 1 & \text{for } l = 0 \\ 0 & \text{for } l \neq 0 \end{cases} \tag{2.27}$$

Such a process is referred to as *higher-order white noise process.*

3. Advantages of the Application of Higher-Order Cumulants

Cumulants and their Fourier transforms (the polyspectra) have recently gained increasing interest in many diverse fields of application. However for decades,

due to the lack of analytical and computational tools for the higher-order statistics applications, only the statistics up to the second order (correlation and power spectrum) were used to any significant extent in signal and system analysis. Although the application of higher-order statistics involves an extensive amount of computation and requires a much larger set of data compared to the second-order approach, there exist applications for which the second-order methods cannot serve appropriately [23].

The second-order statistics are "blind" to the phase of a random process while the higher-order statistics reveal both amplitude and phase information. In the fields of linear system identification and process modeling, the solution of a problem is not unique if only second-order statistics are applied. The system that results from this approach can be either minimum-phase or one of many possible associated nonminimum-phase systems, because in any of these solutions the underlying second-order statistics are the same. When higher-order statistics are applied to the same problem the phase information that is revealed from the data can serve to determine which of these systems is the required solution.

Cumulants, on the other hand, are "blind" to Gaussian processes. Therefore, cumulant applications are insensitive to any additive independent measurement noise if it is Gaussian. Thus cumulant-based methods can boost the signal-to-noise ratio when signals are (at least partially) corrupted with Gaussian noise.

When the processes are non-Gaussian or nonlinear (i.e., generated from a nonlinear mechanism) higher-order statistics are essential in most cases. In real-world applications, processes which are truly non-Gaussian have traditionally been dealt with as though they were Gaussian due to the lack of appropriate mathematical tools. Similarly, nonlinear systems and models are frequently analyzed after being linearized in the vicinity of an operating point. In this case the solution is reliable only

within a range that ensures its stability (i.e., the region of convergence of the linear representation). The development of higher-order statistics tools, both analytical and computational, can and has provided solutions that are free from both Gaussianity and linearity limitations.

When higher-order statistics are applied to problems in signal and system analysis, cumulants, instead of moments, are used for many mathematical and practical reasons. Among these reasons are:

1. For higher-order (non-Gaussian) white noise cumulants of any order (not moments) are multi-dimensional delta functions (see (2.26)) and their Fourier transforms are multi-dimensional constant functions in the frequency domain. This has an enormous effect on the simplicity of the computation of system output statistics when the input is white noise. Moments do not share this property, in general.

2. The cumulant of a sum of independent random processes is the sum of their cumulants (see (2.24)). The same is not true for moments. This enables working with cumulants as *operators*.

Although the above applies to all higher-order cumulants, in many cases the third-order cumulant may not serve the purpose of the application and one may need to work with fourth-order cumulants. This is clear in the case of white symmetrically-distributed processes, for which the third-order cumulant is identically zero. The fourth-order cumulant is also used in cases of processes that have a relatively small third-order cumulant and a much larger fourth-order cumulant.

4. Moments of Jointly Gaussian Random Variables

Let $\mathbf{y} = [y_1, y_2, \cdots, y_k]^T$ be a vector of real-valued zero-mean jointly Gaussian random variables. Their joint probability density function is thus given by

$$f_{\mathbf{y}}(\mathbf{y}) = \frac{1}{(2\pi)^{\frac{k}{2}}|\mathbf{C}|^{\frac{k}{2}}} e^{-\frac{1}{2}\mathbf{y}^T\mathbf{C}^{-1}\mathbf{y}} \tag{2.28}$$

where $\mathbf{C}$ is the covariance matrix with entries

$$C_{ij} = C_{ji} = \mathcal{E}\{y_i y_j\} = \mathcal{E}\{y_j y_i\} \tag{2.29}$$

From the definition of the moment generating function (2.1) it is easy to show that

$$\Phi_y^m(\mathbf{u}) = \epsilon^{-\frac{1}{2}\mathbf{u}^T\mathbf{C}\mathbf{u}} \tag{2.30}$$

Expanding this in a power series yields

$$\begin{aligned}
\Phi_y^m(\mathbf{u}) &= \sum_{m=0}^{\infty} \frac{1}{m!}\left[-\frac{1}{2}\mathbf{u}^T\mathbf{C}\mathbf{u}\right]^m \\
&= \sum_{m=0}^{\infty} \frac{1}{m!}\left(-\frac{1}{2}\right)^m \left[\sum_{j_1=1}^{k}\sum_{j_2=1}^{k} C_{j_1 j_2} u_1 u_2\right]^m \\
&= \sum_{m=0}^{\infty} \frac{1}{m!}\left(-\frac{1}{2}\right)^m \left[\sum_{j_1=1}^{k}\sum_{j_2=1}^{k} \mathcal{E}\{y_{j_1} y_{j_2}\} u_1 u_2\right]^m \\
&= \sum_{m=0}^{\infty} \frac{1}{m!}\frac{j^{2m}}{2^m} \left[\sum_{j_1=1}^{k}\sum_{j_2=1}^{k} \mathcal{E}\{y_{j_1} y_{j_2}\} u_1 u_2\right]^m
\end{aligned} \tag{2.31}$$

To find the value of the moment $\mathcal{E}\{y_1 y_2 \cdots y_k\}$, the expansion (2.31) is compared to the Taylor series expansion of the moment generating function and the required value of the moment is found as the coefficient of the term $j^k u_1 u_2 \cdots u_k$. This comparison reveals the following properties [24]:

1. The sum of the exponents of the u_j's in (2.31) is always even and equals $2m$, for $m = 0, 1, 2, \cdots$. Therefore, in the Taylor series expansion, the coefficients are equal to zero if the sum of the exponents in $u_1^{\nu_1} u_2^{\nu_2} \cdots u_k^{\nu_k}$ is odd or if the

number of variables k in $u_1 u_2 \cdots u_k$ is odd. Thus *the moments of the product of jointly Gaussian random variables is zero if their number is odd.*

2. If the number of random variables is even and equals $2p$, where p is a positive integer, the required average of their product is the coefficient of $u_1 u_2 \cdots u_{2p}$ in the term

$$\frac{1}{p!2^p}\left[\sum_{j_1=1}^{k}\sum_{j_2=1}^{k}\mathcal{E}\{y_{j_1}y_{j_2}\}\,u_1 u_2\right]^p \tag{2.32}$$

The required coefficient thus takes the form

$$\mathcal{E}\{y_{j_1}y_{j_2}\cdots y_{2p}\} = \frac{1}{p!2^p}\sum\prod\mathcal{E}\{y_{j_1}y_{j_2}\} \tag{2.33}$$

where the sum of products is such that such that none of the subscripts appears more than once in the product of the p pairs of the form $\mathcal{E}\{y_{j_1}y_{j_2}\}$ and the summation is done over *all the possible pairing permutations.*

To simplify the expression (2.33) we first notice that $\mathcal{E}\{y_{j_1}y_{j_2}\} = \mathcal{E}\{y_{j_2}y_{j_1}\}$. This means that every 2^p of the terms under the summation are equal and can be replaced by just one term multiplied by the factor 2^p. Also, since the order of the terms in the product has no significance, there are $p!$ identical terms for the same pairing configuration. As a result of these considerations, the required expectation of product of Gaussian variables in its simplest form is

$$\mathcal{E}\{y_{j_1}y_{j_2}\cdots y_{2p}\} = \sum\prod\mathcal{E}\{y_{j_1}y_{j_2}\} \tag{2.34}$$

where the summation here is *over all the possible* <u>*distinct*</u> *pairing permutations.* The number of these distinct permutations equals

$$N_{\text{perm}} = \frac{(2p)!}{p!2^p} \tag{2.35}$$

As discussed in Chapter III, the output of a "well-behaved" nonlinear system can be described by a series-like representation with or without memory. This series

can be considered as a multinomial of functionals of the input. When a moment (cumulant) of the output is computed, the average of the product of these functionals raised to some powers is encountered. As shown in Chapter IV, the output cumulant computation involves the expectation of the product of Gaussian variables raised to some powers in the form

$$\mu_{\mathbf{y}}^{(\nu_1+\nu_2+\cdots+\nu_k)} = \mathcal{E}\{y_1^{\nu_1} y_2^{\nu_2} \cdots y_k^{\nu_k}\} \tag{2.36}$$

The value of this moment is zero if the sum of the exponents $\sum_{i=1}^{k} \nu_i$ is odd. As explained in Appendix A, when the sum of the exponents in (A.1) is even the value of the moment can be computed by first constructing the $k \times k$ correlation matrix

$$\mathbf{C} = \begin{bmatrix} C(1,1) & C(1,2) & C(1,3) & \cdots & C(1,k) \\ C(2,1) & C(2,2) & C(2,3) & \cdots & C(2,k) \\ C(3,1) & C(3,2) & C(3,3) & \cdots & C(3,k) \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ C(k,1) & C(k,2) & C(k,3) & \cdots & C(k,k) \end{bmatrix} \tag{2.37}$$

where

$$C(i,j) = \mathcal{E}\{y_i y_j\} \tag{2.38}$$

and the same size matrix $\mathbf{M}$ of *non-negative integer* entries such that

$$\mathbf{M} = \begin{bmatrix} 2i_1 & i_2 & i_3 & \cdots & i_k \\ i_{k+1} & 2i_{k+2} & i_{k+3} & \cdots & i_{2k} \\ i_{2k+1} & i_{2k+2} & 2i_{2k+3} & \cdots & i_{3k} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ i_{(k-1)k+1} & i_{(k-1)k+2} & i_{(k-1)k+3} & \cdots & 2i_{k^2} \end{bmatrix} \tag{2.39}$$

The matrix will be called the *multiplicity matrix*. Note that for clarity it is prefered to use k^2 linearly indexed variables i_j to represent the entries of $\mathbf{M}$ rather than variables with dual (column,row) index. We also force the diagonal entries to assume even values for reasons that are explained in Appendix A. The value of the moment is then given by

$$\mu_{\mathbf{y}}^{(\nu_1+\nu_2+\cdots+\nu_k)} =$$

$$\nu_1!\nu_2!\cdots\nu_k!\sum\frac{1}{2^{\frac{\mathrm{tr}(M)}{2}}}\prod_{j_1=1}^{k}\frac{(C(j_1,j_1))^{\frac{M(j_1,j_1)}{2}}}{(\frac{M(j_1,j_1)}{2})!}\prod_{\substack{j_1=1\\ j_2\neq j_1}}^{k}\frac{(C(j_1,j_2))^{M(j_1,j_2)}}{(M(j_1,j_2))!} \quad (2.40)$$

where the summation is over all the possible configurations of the matrix **M** such that the sum of its entries along the rows and the columns satisfy the condition

$$\sum_{j=1}^{k}M(j,i)+\sum_{j=1}^{k}M(i,j)-M(i,i)=\nu_i \quad (2.41)$$

This means that the sum of the matrix entries in the i^{th} row and column equals the exponent of y_i in (A.1). The first product corresponds to the product of the diagonal entries of the correlation matrix **C** (the autocorrelation) while the second corresponds to that of the off-diagonal ones (the cross-correlation).

B. POLYSPECTRA OF RANDOM PROCESS

Moment and cumulant functions provide information about the different order correlations among the components of random processes. They demonstrate the dependence of these correlations on the values of the time lag differences among these components. In many applications information about the frequency content of the process, the power distribution over the frequency components, and the coupling between the different frequency components are of significant interest. This information can be obtained from the frequency functions known as the polyspectra. In this section we introduce the definitions of the polyspectra, their properties, and the spectral representation of a random process that leads to the procedure of computing the polyspectra.

1. Definitions and Properties of Polyspectra

The multi-dimensional z-transform of the k^{th}-order cumulant of a random process $x(n)$ is defined as the k^{th}-order *complex cross-polyspectrum*

$$S_{\mathbf{x}}^{(k)}(z_1, z_2, \cdots, z_{k-1}) = \sum_{l_1=-\infty}^{\infty} \cdots \sum_{l_{k-1}=-\infty}^{\infty} C_{\mathbf{x}}^{(k)}(l_1, \cdots, l_{k-1}) z_1^{-l_1} z_2^{-l_2} \cdots z_{k-1}^{-l_{k-1}} \quad (2.42)$$

where $z_1, z_2, \cdots$ and z_{k-1} are complex-valued variables in the region of convergence of $S_{\mathbf{x}}^{(k)}$. The Fourier transform of the cumulant will be called the k^{th}-order *cross-polyspectrum* and is obtained by letting $z_i = e^{j\omega_i}$, namely [1]

$$S_{\mathbf{x}}^{(k)}(\omega_1, \omega_2, \cdots, \omega_{k-1}) = \sum_{l_1=-\infty}^{\infty} \cdots \sum_{l_{k-1}=-\infty}^{\infty} C_{\mathbf{x}}^{(k)}(l_1, \cdots, l_{k-1}) e^{-j(l_1\omega_1 + \cdots + l_{k-1}\omega_{k-1})} \quad (2.43)$$

A sufficient condition for the existence of this quantity is

$$\sum_{l_1=-\infty}^{\infty} \cdots \sum_{l_{k-1}=-\infty}^{\infty} |C_{\mathbf{x}}^{(k)}(l_1, \cdots, l_{k-1})| < \infty \quad (2.44)$$

The polyspectra are defined for all real values of the frequency variables ω_i but are periodic in each variable with a period 2π. For the third-order case this quantity is called the *bispectrum* and for the fourth-order case it is called the *trispectrum*. As in the case of moments and cumulants we shall sometimes use the notation $S_{\mathbf{x}}^{(k)}(\boldsymbol{z})$ to denote $S_{\mathbf{x}}^{(k)}(z_1, z_2, \cdots, z_{k-1})$ and $S_{\mathbf{x}}^{(k)}(\boldsymbol{\omega})$ to denote $S_{\mathbf{x}}^{(k)}(\omega_1, \omega_2, \cdots, \omega_{k-1})$ where

$$\boldsymbol{z} = [z_1, z_2, \cdots, z_{k-1}]^T$$

and

$$\boldsymbol{\omega} = [\omega_1, \omega_2, \cdots, \omega_{k-1}]^T$$

The properties of cumulants discussed in Section A.2. lead to corresponding properties of the polyspectra, namely:

[1] The notation $S_{\mathbf{x}}^{(k)}(\omega_1, \omega_2, \cdots, \omega_{k-1})$ rather than $S_{\mathbf{x}}^{(k)}(e^{j\omega_1}, e^{j\omega_2}, \cdots, e^{j\omega_{k-1}})$ is slightly abusive but will be used for notational simplicity. The meaning will be clear from the context.

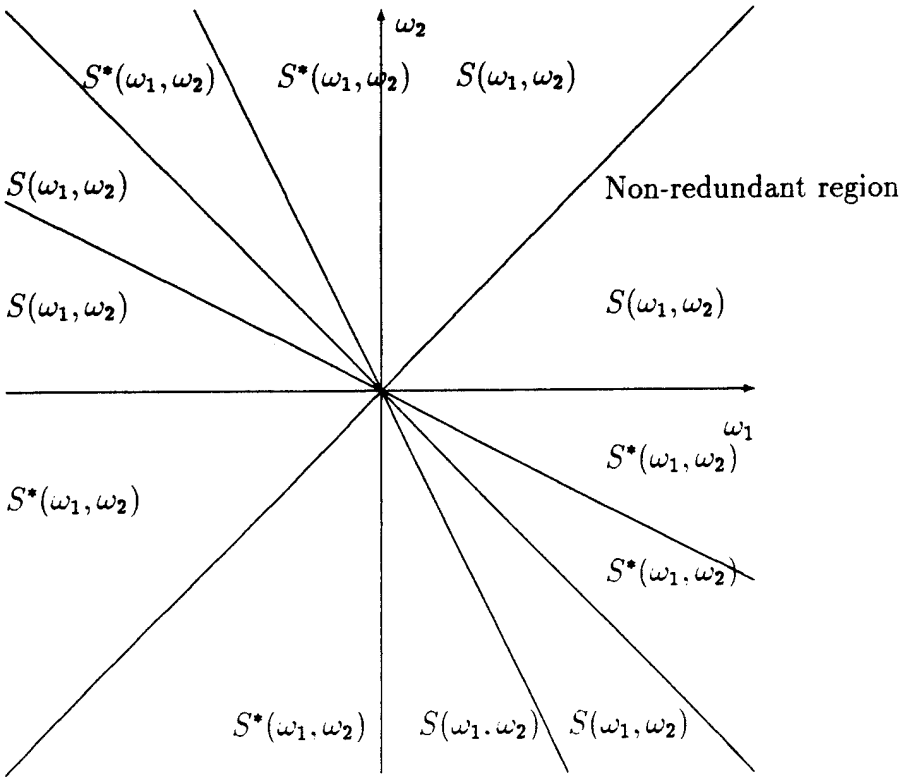

Figure 2.2: Regions of symmetry of the Bispectrum of a real random process.

1. The bispectrum, trispectrum and all higher-order polyspectra are zero for Gaussian processes.

2. The bispectrum is zero for symmetrically-distributed processes which are i.i.d.

3. When the random variables are scaled, their cross-polyspectrum is scaled by a factor that equals the product of the individual scaling factors.

4. Polyspectra have symmetry properties similar to those of cumulants. For example, the bispectrum of a real random process has the twelve regions of symmetry shown in Fig. 2.2.

5. Polyspectra are additive with respect to their arguments

$$S^{(k)}_{(x_0+x_1,x_2,\cdots,x_k)}(\omega) = S^{(k)}_{(x_0,x_2,\cdots,x_k)}(\omega) + S^{(k)}_{(x_1,x_2,\cdots,x_k)}(\omega) \tag{2.45}$$

6. Polyspectra are insensitive to the addition of constants to the random variables.

7. The polyspectra of the sum of independent random variables equals the sum of their individual polyspectra.

8. If a subset of a collection of the random variables is independent of the rest, their joint polyspectra are zero.

9. For processes consisting of i.i.d. samples, the polyspectra are constant, i.e., independent of frequency.

Similar properties also apply to the complex cross-polyspectrum $S^{(k)}_{\mathbf{x}}(\boldsymbol{z})$

2. Spectral Representation of a Stationary Random Process

The second-order statistics of a random process are fully described by the covariance function and its Fourier transform, the power spectral density function. These two quantities, are related by the Fourier transform pair

$$\begin{aligned} C^{(2)}_{\mathbf{x}}(l) &= \frac{1}{2\pi}\int_{-\pi}^{\pi} S^{(2)}_{\mathbf{x}}(\omega)\epsilon^{j\omega l}d\omega \\ S^{(2)}_{\mathbf{x}}(\omega) &= \sum_{l=-\infty}^{\infty} C^{(2)}_{\mathbf{x}}(l)\epsilon^{-j\omega l} \end{aligned} \tag{2.46}$$

The power spectral distribution function $\Psi_{\mathbf{x}}(w)$ is a nondecreasing function of ω that equals zero at $\omega = -\pi$ and is obtained by integrating the power spectral density function

$$\Psi_{\mathbf{x}}(\omega) = \int_{-\pi}^{\omega} S^{(2)}_{\mathbf{x}}(\omega)d\omega \tag{2.47}$$

The covariance function can then be described as the Stieltjes integral

$$C_{\mathbf{x}}^{(2)}(l) = \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{\omega l} d\Psi_{\mathbf{x}}(\omega) \tag{2.48}$$

since

$$d\Psi_{\mathbf{x}}(\omega) = S_{\mathbf{x}}^{(2)}(\omega)d\omega \tag{2.49}$$

The representation is embodied in the "Wiener-Khintchine" theorem [18]. The inversion of (2.48) gives

$$\Psi_{\mathbf{x}}(\omega) = \sigma^2(\omega+\pi) + \sum_{-\infty}^{-1} \frac{e^{-j\omega l}}{-jl} C_{\mathbf{x}}^{(2)}(l) + \sum_{1}^{\infty} \frac{e^{-j\omega l}}{-jl} C_{\mathbf{x}}^{(2)}(l) \tag{2.50}$$

Similar to this representation there exists a spectral representation of the zero-mean random process $x(n)$. The development of this representation starts with defining the complex-valued random process $Z_{\mathbf{x}}(\omega)$ with orthogonal increments such that

$$\begin{aligned} &\mathcal{E}\{dZ_{\mathbf{x}}(\omega)\} = 0 && (a) \\ &\mathcal{E}\{dZ_{\mathbf{x}}(\omega_1)dZ_{\mathbf{x}}^*(\omega_2)\} = \delta_{\omega_1\omega_2} d\Psi_{\mathbf{x}}(\omega_1) && (b) \end{aligned} \tag{2.51}$$

where

$$\delta_{\omega_1\omega_2} = \begin{cases} 1 & \text{for } \omega_1 = \omega_2 \\ 0 & \text{otherwise} \end{cases} \tag{2.52}$$

Then the zero-mean random process can be represented as

$$x(n) = \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{j\omega n} dZ_{\mathbf{x}}(\omega) \tag{2.53}$$

The interpretation of (2.53) is that the process $x(n)$ is represented as a sum of complex exponentials in the continuous range $[-\pi, \pi)$ with random amplitude $|dZ_{\mathbf{x}}(\omega)|$ and random phase $\angle dZ_{\mathbf{x}}(\omega)$. In this case, $d\Psi_{\mathbf{x}}(\omega)$ represents the mean-square amplitude of the component of $x(n)$ at frequency ω. The value of $\Psi_{\mathbf{x}}(\pi)$ is the average total process power and $S_{\mathbf{x}}(j\omega)d\omega$ is the contribution to the total process power from

the components in the frequency range between ω and $\omega + d\omega$. The spectral representation of the random process is therefore related to the spectral representation of its covariance function (the power spectral density function).

By inverting (2.43), the k^{th}-order cumulant is given by

$$C_{\mathbf{x}}^{(k)}(l_1, \cdots, l_{k-1}) = \frac{1}{(2\pi)^{k-1}} \int_{-\pi}^{\pi} \cdots \int_{-\pi}^{\pi} S_{\mathbf{x}}^{(k)}(\omega_1, \omega_2, \cdots, \omega_{k-1}) e^{j(l_1\omega_1 + \cdots + l_{k-1}\omega_{k-1})} d\omega_1 d\omega_2 \cdots d\omega_{k-1} \tag{2.54}$$

As an example, the third-order cumulant for the zero-mean process is given by (2.14.b); so substituting (2.53) we have

$$\begin{aligned} C_{\mathbf{x}}^{(3)}(l_1, l_2) &= \mathcal{E}\{x(n)x(n+l_1)x(n+l_2)\} \\ &= \mathcal{E}\left\{ \frac{1}{(2\pi)^3} \int_{-\pi}^{\pi} e^{j\omega_1 n} dZ_{\mathbf{x}}(\omega_1) \int_{-\pi}^{\pi} e^{j\omega_2(n+l_1)} dZ_{\mathbf{x}}(\omega_2) \int_{-\pi}^{\pi} e^{j\omega_3(n+l_2)} dZ_{\mathbf{x}}(\omega_3) \right\} \\ &= \frac{1}{(2\pi)^3} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} e^{jn(\omega_1+\omega_2+\omega_3)} e^{j(\omega_2 l_1 + \omega_3 l_2)} \mathcal{E}\{dZ_{\mathbf{x}}(\omega_1) dZ_{\mathbf{x}}(\omega_2) dZ_{\mathbf{x}}(\omega_3)\} \end{aligned} \tag{2.55}$$

Since the process is assumed stationary, the cumulant value must be independent of n. Since this can only be true if $\omega_1 + \omega_2 + \omega_3 = 0$, it follows that

$$C_{\mathbf{x}}^{(3)}(l_1, l_2) = \frac{1}{(2\pi)^2} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} e^{j(\omega_2 l_1 + \omega_3 l_2)} \mathcal{E}\{dZ_{\mathbf{x}}(\omega_2) dZ_{\mathbf{x}}(\omega_3) dZ_{\mathbf{x}}(-\omega_2 - \omega_3)\} \tag{2.56}$$

Comparing this result with (2.54) we find that

$$S_{\mathbf{x}}^{(3)}(\omega_1, \omega_2) d\omega_1 d\omega_2 = \mathcal{E}\{dZ_{\mathbf{x}}(\omega_1) dZ_{\mathbf{x}}(\omega_2) dZ_{\mathbf{x}}(-\omega_1 - \omega_2)\} \tag{2.57}$$

By a similar analysis it can be shown that the cross-bispectrum satisfies

$$S_{xyw}^{(3)}(\omega_1, \omega_2) d\omega_1 d\omega_2 = \mathcal{E}\{dZ_{\mathbf{x}}(\omega_1) dZ_y(\omega_2) dZ_w(-\omega_1 - \omega_2)\} \tag{2.58}$$

For the fourth-order cumulant we can write

$$
\begin{aligned}
C_{\mathbf{x}}^{(4)}(l_1,l_2,l_3) \quad &= M_{\mathbf{x}}^{(4)}(l_1,l_2,l_3) - M_{\mathbf{x}}^{(2)}(l_1)M_{\mathbf{x}}^{(2)}(l_2-l_3) \\
&\quad -M_{\mathbf{x}}^{(2)}(l_2)M_{\mathbf{x}}^{(2)}(l_3-l_1) - M_{\mathbf{x}}^{(2)}(l_3)M_{\mathbf{x}}^{(2)}(l_1-l_2) \\
&= \frac{1}{(2\pi)^4}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi} e^{jn(\omega_1+\omega_2+\omega_3+\omega_4)} e^{j(\omega_1 l_1+\omega_2 l_2+\omega_3 l_3)} \\
&\qquad \times \mathcal{E}\left\{dZ_{\mathbf{x}}(\omega_1)dZ_{\mathbf{x}}(\omega_2)dZ_{\mathbf{x}}(\omega_3)dZ_{\mathbf{x}}(\omega_4)\right\} \\
&\quad -\frac{1}{(2\pi)^2}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi} e^{jn(\omega_1+\omega_4)} e^{j(\omega_1 l_1)} \mathcal{E}\left\{dZ_{\mathbf{x}}(\omega_1)dZ_{\mathbf{x}}(\omega_4)\right\} \\
&\quad \times \frac{1}{(2\pi)^2}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi} e^{jn(\omega_2+\omega_3)} e^{j\omega_2(l_2-l_3)} \mathcal{E}\left\{dZ_{\mathbf{x}}(\omega_2)dZ_{\mathbf{x}}(\omega_3)\right\} \\
&\quad -\frac{1}{(2\pi)^2}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi} e^{jn(\omega_2+\omega_4)} e^{j(\omega_2 l_2)} \mathcal{E}\left\{dZ_{\mathbf{x}}(\omega_2)dZ_{\mathbf{x}}(\omega_4))\right\} \\
&\quad \times \frac{1}{(2\pi)^2}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi} e^{jn(\omega_1+\omega_3)} e^{j\omega_3(l_3-l_1)} \mathcal{E}\left\{dZ_{\mathbf{x}}(\omega_1)dZ_{\mathbf{x}}(\omega_3)\right\} \\
&\quad -\frac{1}{(2\pi)^2}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi} e^{jn(\omega_3+\omega_4)} e^{j(\omega_3 l_3)} \mathcal{E}\left\{dZ_{\mathbf{x}}(\omega_3)dZ_{\mathbf{x}}(\omega_4))\right\} \\
&\quad \times \frac{1}{(2\pi)^2}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi} e^{jn(\omega_1+\omega_2)} e^{j\omega_1(l_1-l_2)} \mathcal{E}\left\{dZ_{\mathbf{x}}(\omega_1)dZ_{\mathbf{x}}(\omega_2)\right\} \qquad (2.59)
\end{aligned}
$$

In this case also to satisfy the stationarity condition (2.59) becomes

$$
\begin{aligned}
C_{\mathbf{x}}^{(4)}(l_1,l_2,l_3) = \quad & \frac{1}{(2\pi)^3}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi} e^{j(\omega_1 l_1+\omega_2 l_2+\omega_3 l_3)} \\
& \mathcal{E}\left\{dZ_{\mathbf{x}}(\omega_1)dZ_{\mathbf{x}}(\omega_2)dZ_{\mathbf{x}}(\omega_3)dZ_{\mathbf{x}}(-\omega_1-\omega_2-\omega_3)\right\} \\
- \quad & \frac{1}{(2\pi)^3}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi} e^{j(\omega_1 l_1+\omega_2 l_2+\omega_3 l_3)} \\
& \mathcal{E}\left\{dZ_{\mathbf{x}}(\omega_1)dZ_{\mathbf{x}}(-\omega_1)\right\}\mathcal{E}\left\{dZ_{\mathbf{x}}(\omega_2)dZ_{\mathbf{x}}(-\omega_2)\right\}\delta(\omega_2+\omega_3) \\
- \quad & \frac{1}{(2\pi)^3}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi} e^{j(\omega_1 l_1+\omega_2 l_2+\omega_3 l_3)} \\
& \mathcal{E}\left\{dZ_{\mathbf{x}}(\omega_2)dZ_{\mathbf{x}}(-\omega_2)\right\}\mathcal{E}\left\{dZ_{\mathbf{x}}(\omega_3)dZ_{\mathbf{x}}(-\omega_3)\right\}\delta(\omega_1+\omega_3) \\
- \quad & \frac{1}{(2\pi)^3}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi} e^{j(\omega_1 l_1+\omega_2 l_2+\omega_3 l_3)} \\
& \mathcal{E}\left\{dZ_{\mathbf{x}}(\omega_3)dZ_{\mathbf{x}}(-\omega_3)\right\}\mathcal{E}\left\{dZ_{\mathbf{x}}(\omega_1)dZ_{\mathbf{x}}(-\omega_1)\right\}\delta(\omega_1+\omega_2)
\end{aligned}
\qquad (2.60)
$$

Comparing this result with the definition of the trispectrum by letting $k = 4$ in (2.54)

yields the relation

$$S_{\mathbf{x}}^{(4)}(\omega_1,\omega_2,\omega_3)d\omega_1 d\omega_2 d\omega_3 = \mathcal{E}\left\{dZ_{\mathbf{x}}(\omega_1)dZ_{\mathbf{x}}(\omega_2)dZ_{\mathbf{x}}(\omega_3)dZ_{\mathbf{x}}(-\omega_1-\omega_2-\omega_3)\right\}$$
$$-\mathcal{E}\left\{dZ_{\mathbf{x}}(\omega_1)dZ_{\mathbf{x}}(-\omega_1)\right\}\mathcal{E}\left\{dZ_{\mathbf{x}}(\omega_2)dZ_{\mathbf{x}}(-\omega_2)\right\}\delta(\omega_2+\omega_3)$$
$$-\mathcal{E}\left\{dZ_{\mathbf{x}}(\omega_2)dZ_{\mathbf{x}}(-\omega_2)\right\}\mathcal{E}\left\{dZ_{\mathbf{x}}(\omega_3)dZ_{\mathbf{x}}(-\omega_3)\right\}\delta(\omega_1+\omega_3)$$
$$-\mathcal{E}\left\{dZ_{\mathbf{x}}(\omega_3)dZ_{\mathbf{x}}(-\omega_3)\right\}\mathcal{E}\left\{dZ_{\mathbf{x}}(\omega_1)dZ_{\mathbf{x}}(-\omega_1)\right\}\delta(\omega_1+\omega_2) \qquad (2.61)$$

Similarly we can show that the cross-trispectrum of the processes x, y, w and v is given by

$$S_{xywv}^{(4)}(\omega_1,\omega_2,\omega_3)d\omega_1 d\omega_2 d\omega_3 =$$
$$\mathcal{E}\left\{dZ_{\mathbf{x}}(\omega_1)dZ_y(\omega_2)dZ_w(\omega_3)dZ_v(-\omega_1-\omega_2-\omega_3)\right\}$$
$$-\mathcal{E}\left\{dZ_{\mathbf{x}}(\omega_1)dZ_y(-\omega_1)\right\}\delta(\omega_1+\omega_2)\mathcal{E}\left\{dZ_w(\omega_3)dZ_v(-\omega_3)\right\}$$
$$-\mathcal{E}\left\{dZ_{\mathbf{x}}(\omega_1)dZ_w(-\omega_1)\right\}\delta(\omega_1+\omega_3)\mathcal{E}\left\{dZ_y(\omega_2)dZ_v(-\omega_2)\right\}$$
$$-\mathcal{E}\left\{dZ_{\mathbf{x}}(\omega_1)dZ_v(-\omega_1)\right\}\mathcal{E}\left\{dZ_y(\omega_2)dZ_w(-\omega_2)\right\}\delta(\omega_2+\omega_3) \qquad (2.62)$$

Now let us specialize these relations to the case of a zero-mean Gaussian process $g(n)$. Since for this process the cumulants of orders higher than the second are identically zero, the left hand sides of (2.57) and (2.61) are identically zero. Thus for this process

$$\mathcal{E}\left\{dZ_g(\omega_1)dZ_g(\omega_2)dZ_g(-\omega_1-\omega_2)\right\} = 0 \qquad (2.63)$$

for all ω_1 and ω_2, and

$$\mathcal{E}\left\{dZ_g(\omega_1)dZ_g(\omega_2)dZ_g(\omega_3)dZ_g(-\omega_1-\omega_2-\omega_3)\right\} =$$
$$\mathcal{E}\left\{dZ_g(\omega_1)dZ_g(-\omega_1)\right\}\delta(\omega_1+\omega_2)\mathcal{E}\left\{dZ_g(\omega_3)dZ_g(-\omega_3)\right\}$$
$$+\mathcal{E}\left\{dZ_g(\omega_1)dZ_g(-\omega_1)\right\}\delta(\omega_1+\omega_3)\mathcal{E}\left\{dZ_g(\omega_2)dZ_g(-\omega_2)\right\}$$
$$+\mathcal{E}\left\{dZ_g(\omega_1)dZ_g(-\omega_1)\right\}\mathcal{E}\left\{dZ_g(\omega_2)dZ_g(-\omega_2)\right\} \qquad (2.64)$$

This means that for the Gaussian process the value of

$$\mathcal{E}\{dZ_g(\omega_1)dZ_g(\omega_2)dZ_g(\omega_3)dZ_g(\omega_4)\}$$

is nonzero on the manifold defined by $\omega_1 + \omega_2 + \omega_3 + \omega_4 = 0$. Moreover on this manifold it is nonzero only on the three proper submanifolds defined by

- $(\omega_1 + \omega_2 = 0) \cap (\omega_3 + \omega_4 = 0)$
- $(\omega_1 + \omega_3 = 0) \cap (\omega_2 + \omega_4 = 0)$
- $(\omega_1 + \omega_4 = 0) \cap (\omega_2 + \omega_3 = 0)$

The value is given by

$$\begin{aligned}\mathcal{E}\{dZ_g(\omega_1)dZ_g(\omega_2)dZ_g(\omega_3)dZ_g(\omega_4)\} = \\ S_g(\omega_1)S_g(\omega_3)\delta(\omega_1+\omega_2)\delta(\omega_3+\omega_4)d\omega_1 d\omega_2 d\omega_3 d\omega_4 \\ +S_g(\omega_1)S_g(\omega_2)\delta(\omega_1+\omega_3)\delta(\omega_2+\omega_4)d\omega_1 d\omega_2 d\omega_3 d\omega_4 \\ +S_g(\omega_1)S_g(\omega_2)\delta(\omega_1+\omega_4)\delta(\omega_2+\omega_3)d\omega_1 d\omega_2 d\omega_3 d\omega_4 \end{aligned} \tag{2.65}$$

which has a three-dimensional region of support defined by ω_1, ω_2 and ω_3; we use four frequency arguments for clarity however and the condition

$$\omega_1 + \omega_2 + \omega_3 + \omega_4 = 0$$

to make the notation correct.

Now assume we have a Gaussian process that is also white. Denote this process by $w(n)$. In this case the right hand side of (2.65) becomes

$$\begin{aligned}\mathcal{E}\{dZ_w(\omega_1)dZ_w(\omega_2)dZ_w(\omega_3)dZ_w(\omega_4)\} = \\ \sigma_o^4[\delta^c(\omega_1+\omega_2)\delta^c(\omega_3+\omega_4) + \delta^c(\omega_1+\omega_3)\delta^c(\omega_2+\omega_4) + \delta^c(\omega_1+\omega_4)\delta^c(\omega_2+\omega_3)] \\ d\omega_1 d\omega_2 d\omega_3 d\omega_4 \end{aligned} \tag{2.66}$$

where $\delta^c(\omega)$ is the continuous time unit impulse function, not to be confused with the dirac delta $\delta_{\omega_1\omega_2}$ defined in (2.52) or the (discrete time) unit sample function $\delta(l)$ defined in (2.27). In general since the process $w(n)$ is zero-mean white Gaussian, all the cumulants *except the second-order cumulant (the covariance) and the associated power spectral density function* are equal to zero. Therefore when we apply this characteristic to any order polyspectra expressions similar to (2.57) and (2.61), equations (2.63) and (2.65) can be generalized as follows:

$$\mathcal{E}\{dZ_w(\omega_1)dZ_w(\omega_2)\cdots dZ_w(\omega_k)\} = \begin{cases} 0 & \text{for odd } k \\ \sigma_o^k \delta_\omega^k d\omega & \text{for even } k \end{cases} \tag{2.67}$$

where we use the notation

$$\begin{aligned} d\omega &= d\omega_1 d\omega_1 \cdots d\omega_k \\ \delta_\omega^k &= \sum \prod_{i,j} \delta(\omega_i + \omega_j) \end{aligned} \tag{2.68}$$

and where the product is over the $\frac{k}{2}$ pairs of ω_i and ω_j and the summation is over all possible pairing permutations.

C. MOMENTS AND SPECTRA OF FUNCTIONALS OF A WHITE GAUSSIAN PROCESS

1. Moments and Spectra of Linear Functionals

For a zero-mean white Gaussian process $w(n)$, the power spectral density function $S_w^{(2)}(\omega)$ has the same value σ_o^2 for all frequencies; this is the average process power. Therefore the spectral representation of w is given by

$$w(n) = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{\omega n} dZ_w(\omega) \tag{2.69}$$

where

$$\begin{aligned} &\mathcal{E}\{dZ_w(\omega)\} = 0 \\ &\mathcal{E}\{dZ_w(\omega_1)dZ_w^*(\omega_2)\} = d\Psi_x(\omega_1)\delta_{\omega_1\omega_2} \end{aligned} \tag{2.70}$$

and

$$d\Psi_w(\omega) = S_w(\omega)d\omega = \sigma_o^2 d\omega \tag{2.71}$$

If this process drives a linear system with transfer function $H(\omega)$ then the output $y(n)$ is another zero-mean Gaussian process that also has a spectral representation in the continuous region $[-\pi, \pi)$ in the form of

$$y(n) = \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{j\omega n} dZ_y(\omega) \tag{2.72}$$

The components of the input random process that comprise its spectral representation are modified by the magnitude and phase of the transfer function to give the components of the output process with random magnitude $|H(j\omega)||dZ_w(\omega)|$ and random phase $\angle H(\omega) + \angle dZ_w(\omega)$. In general

$$dZ_y(\omega) = H(\omega)dZ_w(\omega) \tag{2.73}$$

Which satisfies conditions similar to (2.51) because

$$\begin{aligned} \mathcal{E}\{dZ_y(\omega)\} &= H(\omega)\mathcal{E}\{dZ_w(\omega)\} = 0 \\ \mathcal{E}\{dZ_y(\omega_1)dZ_y^*(\omega_2)\} &= H(\omega_1)H^*(\omega_2)\mathcal{E}\{dZ_w(\omega_1)dZ_w^*(\omega_2)\} \\ &= \delta_{\omega_1\omega_2}|H(\omega_1)|^2\sigma_o^2 d\omega_1 \end{aligned} \tag{2.74}$$

The output power spectral density function has the value

$$S_y^{(2)}(\omega) = \sigma_o^2|H(\omega)|^2 \tag{2.75}$$

If the system impulse response is denoted by $h(n)$ then the output autocorrelation function is given by

$$R_y(l) = \mathcal{E}\{y(n)y(n+l)\} = \sigma_o^2 r_h(l) \tag{2.76}$$

where $r_h(l)$ is the *system correlation sequence* given by

$$r_h(l) = \sum_{n=-\infty}^{\infty} h(n+l)h(n) \tag{2.77}$$

which is the Fourier inverse of $|H(\omega)|^2$.

If a set of linear systems with transfer functions $H_1(\omega), H_2(\omega), \cdots, H_k(\omega_k)$ and corresponding impulse responses $h_1(n), h_2(n), \cdots, h_k(n)$ are driven by the same zero-mean white Gaussian process $w(n)$ with variance σ_o^2, the outputs $y_1, y_2, \cdots, y_k$ are zero-mean Gaussian and have the spectral representations

$$y_j(n) = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{j\omega n} dZ_{y_j}(\omega) \quad ; \quad j = 1, 2, \cdots, k \tag{2.78}$$

for which

$$dZ_{y_j}(\omega) = H_j(j\omega) dZ_w(\omega) \tag{2.79}$$

In this case the cross-correlation function is then given by

$$R_{y_i y_j}(l) = \mathcal{E}\{y_i(n) y_j(n+l)\} = \sigma_o^2 r_{ij}(l) \tag{2.80}$$

where

$$r_{ij}(l) = \sum_{n=-\infty}^{\infty} h_i(n+l) h_j(n) \tag{2.81}$$

The quantity $r_{ij}(l)$ will be called the *system cross-correlation sequence.* The cross-spectral density function is then given by

$$\begin{aligned} S_{y_i y_j}(\omega) d\omega &= \mathcal{E}\{dZ_{y_i}(\omega) dZ^*_{y_j}(\omega)\} \\ &= \sigma_o^2 H_i(\omega) H_j^*(\omega) d\omega \\ &= \sigma_o^2 s_{ij}(\omega) d\omega \end{aligned} \tag{2.82}$$

where

$$s_{ij}(\omega) = s^*_{ji}(\omega) = H_i(\omega) H_j*(\omega) \tag{2.83}$$

We refer to the last quantity as the *system cross-spectral function.*

2. Moments for Nonlinear Functionals of a White Gaussian Process

A set of nonlinear processes $u_j(n)$ can be formed by multiplying the outputs of the linear systems raised to some integer powers ν_{ij} to give

$$u_j(n) = y_1^{\nu_{j1}}(n) y_2^{\nu_{j2}}(n) \cdots y_k^{\nu_{jk}}(n) \tag{2.84}$$

The u_j are functionals of the white Gaussian process $w(n)$ because each of the y_j is given by

$$y_j(n) = \sum_{i=-\infty}^{\infty} w_j(n-i) h_j(i) \tag{2.85}$$

The value of the m^{th}-order cross-moment function

$$\begin{aligned}
&\mu_u^{(m)}(\boldsymbol{l}) = \mathcal{E}\{u_1(n) u_2(n+l_1) \cdots u_m(n+l_{m-1})\} \\
&= \mathcal{E}\{y_1^{\nu_{11}}(n) y_2^{\nu_{12}}(n) \cdots y_k^{\nu_{1k}}(n) y_1^{\nu_{21}}(n+l_1) y_2^{\nu_{22}}(n+l_1) \cdots y_k^{\nu_{2k}}(n+l_1) \cdots \\
&\times y_1^{\nu_{m1}}(n+l_{m-1}) y_2^{\nu_{m2}}(n+l_{m-1}) \cdots y_k^{\nu_{mk}}(n+l_{m-1})\}
\end{aligned} \tag{2.86}$$

is zero if the sum of the exponents is odd. For an even sum of exponents the procedure explained in the Subsection A.4. is followed: the $mk \times mk$ matrix $\mathbf{C}_u^{(m)}$ is constructed as

$$\mathbf{C}_u^{(m)} = \begin{bmatrix}
\mathbf{R}(0) & \mathbf{R}(l_1) & \mathbf{R}(l_2) & \cdots & \mathbf{R}(l_{m-1}) \\
\mathbf{R}^T(l_1) & \mathbf{R}(0) & \mathbf{R}(l_2 - l_1) & \cdots & \mathbf{R}(l_{m-1} - l_1) \\
\mathbf{R}^T(l_2) & \mathbf{R}^T(l_2 - l_1) & \mathbf{R}(0) & \cdots & \mathbf{R}(l_{m-1} - l_2) \\
\vdots & \vdots & \vdots & \vdots & \vdots \\
\mathbf{R}^T(l_{m-1}) & \mathbf{R}^T(l_{m-1} - l_1) & \mathbf{R}^T(l_{m-1} - l_2) & \cdots & \mathbf{R}(0)
\end{bmatrix} \tag{2.87}$$

where the $k \times k$ matrix $\mathbf{R}(n)$ is given by

$$\mathbf{R}(n) = \begin{bmatrix}
r_{11}(n) & r_{12}(n) & r_{13}(n) & \cdots & r_{1k}(n) \\
r_{21}(n) & r_{22}(n) & r_{23}(n) & \cdots & r_{2k}(n) \\
r_{31}(n) & r_{32}(n) & r_{33}(n) & \cdots & r_{3k}(n) \\
\vdots & \vdots & \vdots & \vdots & \vdots \\
r_{k1}(n) & r_{k2}(n) & r_{k3}(n) & \cdots & r_{kk}(n)
\end{bmatrix} \tag{2.88}$$

and $r_{ij}(n)$ are the system cross-correlation sequences (see (2.81)). A multiplicity matrix $\mathbf{M}$ of size corresponding to $\mathbf{C}_u^{(m)}$ is defined as

$$\mathbf{M} = \begin{bmatrix} \mathbf{M}(0) & \mathbf{M}(l_1) & \mathbf{M}(l_2) & \cdots & \mathbf{M}(l_{m-1}) \\ \mathbf{M}^T(l_1) & \mathbf{M}(0) & \mathbf{M}(l_2 - l_1) & \cdots & \mathbf{M}(l_{m-1} - l_1) \\ \mathbf{M}^T(l_2) & \mathbf{M}^T(l_2 - l_1) & \mathbf{M}(0) & \cdots & \mathbf{R}(l_{m-1} - l_2) \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \mathbf{M}^T(l_{m-1}) & \mathbf{M}^T(l_{m-1} - l_1) & \mathbf{M}^T(l_{m-1} - l_2) & \cdots & \mathbf{M}(0) \end{bmatrix} \tag{2.89}$$

where each block $\mathbf{M}(l)$ represents the multiplicities of the $r_{ij}(n)$ in the block $\mathbf{R}(l)$. Although equation (A.15) is still applicable, we can benefit from the fact that $C(i,j) = C(j,i)$ to save computations in this case and in similar cases where the matrix $\mathbf{C}_u^{(m)}$ is symmetric. The multiplicity matrix $\mathbf{M}$ can be modified because the two entries $M(i,j)$ and $M(j,i)$ result in pairing $M(i,j) + M(j,i)$ elements from both y_i and y_j to give $C(i,j)^{(M(i,j)+M(i,j))/2}$. Therefore we let both entries of $\mathbf{M}$ be equal such that each of them equals the value that their sum would have if the notation in (A.15) is used. The diagonal elements are not changed because they still have the same meaning. Subsequently the off-diagonal elements need to be examined in the upper triangle only, and the condition on the sum of the elements along the columns and the rows becomes that *the sum of the elements of the* j^{th} *row equals the sum of the elements of the* j^{th} *column and each equals* ν_j.

Before proceeding, let us define some notation that will be useful in the remainder of this discussion. For a k-dimensional vector $\boldsymbol{\nu}$ with elements $\nu_1, \nu_2, \cdots$ and ν_k, we dfine the *additive reduction* or simply the *reduction* of $\boldsymbol{\nu}$ as the sum of its components. If we denote this by $\Sigma[\boldsymbol{\nu}]$ we have the definition

$$\Sigma[\boldsymbol{\nu}] = \sum_{i=1}^{k} \nu_k \tag{2.90}$$

For a $k \times l$ matrix $\mathbf{M}$, with elements $M(i,j)$ we define the reduction as a vector with elements equal to the sum of the rows of $\mathbf{M}$. Thus

$$\Sigma[\mathbf{M}] = \mathbf{m} \qquad \text{where} \qquad m_i = \sum_{j=1}^{l} M(i,j) \tag{2.91}$$

Now since each of the cross-correlation functions equals the corresponding system cross-correlation sequence multiplied by the input variance σ_o^2, and since the total number of the cross-correlation functions is always equal to half the total sum of exponents, then each term in the expression of the moment has the value $(\sigma_o^2)^{\frac{\Sigma[\boldsymbol{\nu}]}{2}}$ multiplied by the product of the system cross-correlation sequences r_j that result from each pairing permutation. Therefore the moment in this case is given by substituting in (A.15)

$$\begin{aligned}\mu_u^{(m)}(\boldsymbol{l}) &= \mathcal{E}\{u_1(n)u_2(n+l_1)\cdots u_m(n+l_{m-1})\} \\ &= (\sigma_o^2)^{\frac{\Sigma[\boldsymbol{\nu}]}{2}} \prod_{i=1}^{mk} \boldsymbol{\nu}(i)! \sum_{\Sigma[\mathbf{M}]=\boldsymbol{\nu}} \frac{1}{2^{\frac{tr(\mathbf{M})}{2}}} \prod_{j_1=1}^{mk} \frac{(C_v^{(m)}(j_1,j_1))^{\frac{M(j_1,j_1)}{2}}}{(\frac{M(j_1,j_1)}{2})!} \\ &\quad \times \prod_{j_2=j_1+1}^{mk} \frac{(C_v^{(m)}(j_1,j_2))^{M(j_1,j_2)}}{M(j_1,j_2)!} \end{aligned} \tag{2.92}$$

The condition $\Sigma[\mathbf{M}] = \boldsymbol{\nu}$ in the subscript of the summation means that the summation is done over all the possible values of the matrix $\mathbf{M}$ that satisfy the condition.

3. Spectra of Nonlinear Functionals of a White Gaussian Process

In the frequency domain the $(m-1)$-dimensional *moment cross-spectral function* is defined as

$$\begin{aligned} &S_u^{(m)}(\omega_1,\omega_2,\cdots,\omega_{m-1}) \\ &= \sum_{l_1=-\infty}^{\infty} \cdots \sum_{l_{m-1}=-\infty}^{\infty} \mu_u^{(m)}(l_1,l_2,\cdots,l_{m-1}) \epsilon^{-j(\omega_1 l_1+\omega_2 l_2+\cdots+\omega_{m-1}l_{m-1})} \end{aligned} \tag{2.93}$$

with inverse transform relation

$$\begin{aligned} &\mu_u^{(m)}(l_1,l_2,\cdots,l_{m-1}) \\ &= (\frac{1}{2\pi})^{m-1} \underbrace{\int_{-\pi}^{\pi} \cdots \int_{-\pi}^{\pi}}_{m-1} S_u^{(m)}(\omega_1,\omega_2,\cdots,\omega_{m-1}) \epsilon^{j(\omega_1 l_1+\omega_2 l_2+\cdots+\omega_{m-1}l_{m-1})} \\ &\qquad\qquad d\omega_1 d\omega_2 \cdots d\omega_{m-1} \end{aligned} \tag{2.94}$$

Equating the right hand side of this equation and (2.86) yields

$$
\begin{aligned}
&(\frac{1}{2\pi})^{m-1} \underbrace{\int_{-\pi}^{\pi} \cdots \int_{-\pi}^{\pi}}_{m-1} S_u^{(m)}(\omega_1, \omega_2, \cdots, \omega_{m-1}) e^{j(\omega_1 l_1 + \omega_2 l_2 + \cdots + \omega_{m-1} l_{m-1})} \\
&\qquad\qquad d\omega_1 d\omega_2 \cdots d\omega_{m-1} \\
&= \mathcal{E}\{u_1(n) u_2(n + l_1) \cdots u_m(n + l_{m-1})\} \\
&= \mathcal{E}\{y_1^{\nu_{11}}(n) y_2^{\nu_{12}}(n) \cdots y_k^{\nu_{1k}}(n) y_1^{\nu_{21}}(n + l_1) y_2^{\nu_{22}}(n + l_1) \cdots y_k^{\nu_{2k}}(n + l_1) \cdots \\
&\quad \times y_1^{\nu_{m1}}(n + l_{m-1}) y_2^{\nu_{m2}}(n + l_{m-1}) \cdots y_k^{\nu_{mk}}(n + l_{m-1})\}
\end{aligned} \tag{2.95}
$$

Now substituting each $y(n)$ with its spectral representation in (2.78) we obtain

$$
\begin{aligned}
y^{\nu}(n) &= \left[\frac{1}{2\pi} \int_{-\pi}^{\pi} e^{j\omega n} dZ_y(\omega)\right]^{\nu} \\
&= (\frac{1}{2\pi})^{\nu} \underbrace{\int_{-\pi}^{\pi} \cdots \int_{-\pi}^{\pi}}_{\nu} e^{jn(\omega_1 + \omega_1 + \cdots + \omega_{\nu})} dZ_y(\omega_1) dZ_y(\omega_2) \cdots dZ_y(\omega_{\nu}) \\
&= (\frac{1}{2\pi})^{\nu} \underbrace{\int_{-\pi}^{\pi} \cdots \int_{-\pi}^{\pi}}_{\nu} H(\omega_1) H(\omega_2) \cdots H(\omega_{\nu}) e^{jn(\omega_1 + \omega_1 + \cdots + \omega_{\nu})} \\
&\qquad\qquad dZ_w(\omega_1) dZ_w(\omega_2) \cdots dZ_w(\omega_{\nu})
\end{aligned} \tag{2.96}
$$

Let us define $\omega_{\iota_1 \iota_2 \iota_3}$ as new variables of integration that are necessary when we substitute (2.96) into (2.95). Then the right hand side of (2.95) takes the form

$$
\begin{aligned}
&\mathcal{E}\left\{(\frac{1}{2\pi})^{P} \underbrace{\int_{-\pi}^{\pi} \cdots \int_{-\pi}^{\pi}}_{P} \left(\prod_{\iota_1=1}^{m} \prod_{\iota_2=1}^{k} \prod_{\iota_3=1}^{\nu_{\iota_1 \iota_2}} H_{\iota_2}(\omega_{\iota_1 \iota_2 \iota_3}) dZ_w(\omega_{\iota_1 \iota_2 \iota_3})\right)\right. \\
&\left. \times e^{jn\left(\sum_{\iota_1=1}^{m} \sum_{\iota_2=1}^{k} \sum_{\iota_3=1}^{\nu_{\iota_1 \iota_2}} \omega_{\iota_1 \iota_2 \iota_3}\right)} e^{j\left(\sum_{\iota_1=2}^{m} \sum_{\iota_2=1}^{k} \sum_{\iota_3=1}^{\nu_{\iota_1 \iota_2}} l_{\iota_1 - 1} \omega_{\iota_1 \iota_2 \iota_3}\right)} \right\}
\end{aligned} \tag{2.97}
$$

where P equals the sum of all the exponents in (2.95), i.e.,

$$
P = \Sigma[\boldsymbol{\nu}] \tag{2.98}
$$

and must be even so that the expression is not equal to zero. The stationarity condition implies that

$$\sum_{\iota_1=1}^{m}\sum_{\iota_2=1}^{k}\sum_{\iota_3=1}^{\nu_{\iota_1\iota_2}}\omega_{\iota_1\iota_2\iota_3}=0 \tag{2.99}$$

Now interchanging the order of the integration and expectation in (2.97) produces

$$(\frac{1}{2\pi})^P\int_{-\pi}^{\pi}\cdots\int_{-\pi}^{\pi}\left(\prod_{\iota_1\iota_2\iota_3}H_{\iota_2}(\omega_{\iota_1\iota_2\iota_3})\right)e^{j\left(\sum_{\iota_1=2}^{m}\sum_{\iota_2}\sum_{\iota_3}l_{\iota_1-1}\omega_{\iota_1\iota_2\iota_3}\right)}\mathcal{E}\left\{\prod_{\iota_1\iota_2\iota_3}dZ_w(\omega_{\iota_1\iota_2\iota_3})\right\} \tag{2.100}$$

where the integration is taken over the manifold defined by (2.99) and the product notation used here is equivalent to the triple product expression in (2.97). By substituting the expectation in the integrand with its value given by (2.67), the expression becomes

$$(\sigma_o^2)^{\frac{P}{2}}(\frac{1}{2\pi})^P\underbrace{\int_{-\pi}^{\pi}\cdots\int_{-\pi}^{\pi}}_{P}\left(\prod_{\iota_1\iota_2\iota_3}H_{\iota_2}(j\omega_{\iota_1\iota_2\iota_3})\right)\epsilon^{j\left(\sum_{\iota_1=2}^{m}\sum_{\iota_2}\sum_{\iota_3}l_{\iota_1-1}\omega_{\iota_1\iota_2\iota_3}\right)}\delta_\omega^P\prod_{\iota_1\iota_2\iota_3}d\omega_{\iota_1\iota_2\iota_3} \tag{2.101}$$

Since δ_ω^P makes this expression equal to zero outside the manifolds defined by the $P!/((\frac{P}{2})!2^{\frac{P}{2}})$ permutations of the product of delta-dirac functions $\delta(\omega_i+\omega_j)$; the expression (2.101) is a sum of terms represented by $\frac{P}{2}$ multiple integrations. Each term corresponds to a certain submanifold of that defined by (2.99) and it is multiplied by the number of permutations that result in the same integrand. The product of the P transfer functions $H(\omega)$ is replaced by the product of $\frac{P}{2}$ system spectral density functions $S_{ij}(\omega)$ and the sum of lag-frequency products $l_\iota\omega_\iota$ in the complex exponential is reduced correspondingly because

$$\begin{aligned}(\frac{1}{2\pi})^2\int_{-\pi}^{\pi}\int_{-\pi}^{\pi}H_1(j\omega_1)H_2(j\omega_2)\delta(\omega_1+\omega_2)\epsilon^{j(l_1\omega_1+l_2\omega_2)}d\omega_1d\omega_2\\ =\frac{1}{2\pi}\int_{-\pi}^{\pi}s_{12}(\omega_1)\epsilon^{j\omega_1(l_1-l_2)}d\omega_1\end{aligned} \tag{2.102}$$

When this result is applied to (2.101) we can develop a procedure to calculate the value of $S_u^{(m)}(\omega_1,\omega_2,\cdots,\omega_{m-1})$ given the linear system transfer functions (or the sys-

tem cross-spectral density functions) and the input variance σ_o^2.

4. Second-Order Cross-Spectral Function of Functionals of a White Gaussian Process

Let us continue to consider the functionals $u_i[w(n)]$ of the zero-mean white Gaussian process $w(n)$ defined by (2.84) and (2.85). Note that in general these functionals do not have zero mean. The cross-correlation for any two of these functionals is given by

$$\begin{aligned}\mu_u^{(2)}(l) &= \mathcal{E}\{u_1(n)u_2(n+l)\} \\ &= \mathcal{E}\{y_1^{\nu_{11}}(n)y_2^{\nu_{12}}(n)\cdots y_k^{\nu_{1k}}(n)y_1^{\nu_{21}}(n+l)y_2^{\nu_{22}}(n+l)\cdots y_k^{\nu_{2k}}(n+l)\}\end{aligned} \quad (2.103)$$

and its frequency transform (the cross-power spectral density function) is defined as

$$S_u^{(2)}(\omega) = \sum_{l=-\infty}^{\infty} \mu_u^{(2)}(l)\epsilon^{-j\omega l} \quad (2.104)$$

If the sum of exponents in (2.103) has an even value P then the expectation is not identically zero and we can compute the value of the moment as described in (2.92). The right hand side of (2.95) is obtained as a sum of terms each corresponding to a possible permutation pairs of frequencies resulting in the expression (2.101). We start the procedure by constructing the $2k \times 2k$ matrix of system cross-spectral functions

$$\mathcal{S}_u^{(2)}(\omega) = \begin{bmatrix} \mathbf{S}^0(\omega) & \mathbf{S}^\omega(\omega) \\ (\mathbf{S}^\omega)^{*T}(\omega) & \mathbf{S}^0(\omega) \end{bmatrix} \quad (2.105)$$

where each term $\mathbf{S}^0(\omega)$ or $\mathbf{S}^\omega(\omega)$ is of the form

$$\mathbf{S}(\omega) = \begin{bmatrix} s_{11}(\omega) & s_{12}(\omega) & s_{13}(\omega) & \cdots & s_{1k}(\omega) \\ s_{21}(\omega) & s_{22}(\omega) & s_{23}(\omega) & \cdots & s_{2k}(\omega) \\ s_{31}(\omega) & s_{32}(\omega) & s_{33}(\omega) & \cdots & s_{3k}(\omega) \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ s_{k1}(\omega) & s_{k2}(\omega) & s_{k3}(\omega) & \cdots & s_{kk}(\omega) \end{bmatrix} \quad (2.106)$$

with the system cross-spectral function s_{ij} defined in (2.83). The significance of the superscript (0 or ω) is explained below. A $2k \times 2k$ multiplicity matrix $\mathbf{M}$ is also constructed in the same way as for the moment computation procedure.

Let us now assume that the i^{th} frequency pairing permutation in (2.101) results in a component $S_u^{(2)i}(\omega)$ of $\mathcal{S}_u^{(2)}(\omega)$ which is obtained by summing all the possible components. In this component the frequency pairing is over all transfer functions that comprise the functionals u_1 and u_2. Those transfer functions of u_1 do not have their frequency argument appearing in the lag-frequency product of the complex exponential. Therefore if the frequency arguments of two such transfer functions are paired together they do not affect the lag-frequency exponent and do not have their system cross-spectral function multiplied by a complex exponential of its frequency argument. This spectral function is an entry of the upper left diagonal block $\mathbf{S}^0$ of $\mathcal{S}_u^{(2)}$. If the frequency arguments of two transfer functions that come from u_2 are paired together then the sum of these two arguments is zero and they therefore disappear from the lag-frequency complex exponential. The resulting spectral function is an entry of the lower right diagonal block $\mathbf{S}^0$. On the other hand if a frequency argument of a transfer function of u_1 is paired with one of u_2 , this frequency argument appears in the lag-frequency product in the complex exponential. This spectral function is located in $\mathbf{S}^\omega$.

For a specific configuration of the frequency pairing let us assume that p_1 system cross-spectral functions lie in the upper triangles of the two diagonal blocks denoted $\mathbf{S}^0$, and p_2 are entries of $\mathbf{S}^\omega$ such that $p_1 + p_2 = \frac{P}{2}$. Then from (2.104) and (2.101)

$$S_u^{(2)i}(\omega) = \sum_{l=-\infty}^{\infty} e^{-j(\omega l)}(\sigma_o^2)^{\frac{P}{2}}(\frac{1}{2\pi})^{\frac{P}{2}} \underbrace{\int_{-\pi}^{\pi}\cdots\int_{-\pi}^{\pi}}_{\frac{P}{2}} s_1(\omega_1)s_2(\omega_2)\cdots s_{p_1}(\omega_{p_1})$$

$$\times s_{p_1+1}(\omega_{p_1+1})s_{p_1+2}(\omega_{p_1+2})\cdots s_{\frac{P}{2}}(\omega_{\frac{P}{2}})e^{jl(\omega_{p_1+1}+\omega_{p_1+2}+\omega_{\frac{P}{2}})}d\omega_1 d\omega_2\cdots d\omega_{\frac{P}{2}} \tag{2.107}$$

(Here we used a single digit in the subscript of the spectral functions for simplicity; they are actually indexed with two digits according to the indices of the correlated system outputs). Since each of the first p_1 spectral functions can be integrated separately, the result of this integration is a constant multiplying the rest of the integrals. The value of this constant is given by

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} s_{ij}(\omega)d\omega = r_{ij}(0) \tag{2.108}$$

where r_{ij} is the system cross-correlation sequence defined in (2.81). Therefore (2.107) reduces to

$$S_u^{(2)i}(\omega) = (\sigma_o^2)^{\frac{P}{2}} r_1(0)r_2(0)\cdots r_{p_1}(0)\sum_{l=-\infty}^{\infty}\epsilon^{-j\omega l}(\frac{1}{2\pi})^{p_2}\underbrace{\int_{-\pi}^{\pi}\cdots\int_{-\pi}^{\pi}}_{p_2}$$

$$s_{p_1+1}(\omega_{p_1+1})s_{p_1+2}(\omega_{p_1+2})\cdots s_{\frac{P}{2}}(\omega_{\frac{P}{2}})\epsilon^{jl(\omega_{p_1+1}+\omega_{p_1+2}+\cdots+\omega_{\frac{P}{2}})}d\omega_{p_1+1}\cdots d\omega_{\frac{P}{2}} \tag{2.109}$$

Now interchanging the order of the summation and integration we obtain

$$S_u^{(2)i}(\omega) = (\sigma_o^2)^{\frac{P}{2}} r_1(0)r_2(0)\cdots r_{p_1}(0)(\frac{1}{2\pi})^{p_2}\underbrace{\int_{-\pi}^{\pi}\cdots\int_{-\pi}^{\pi}}_{p_2}$$

$$s_{p_1+1}(\omega_{p_1+1})s_{p_1+2}(\omega_{p_1+2})\cdots s_{\frac{P}{2}}(\omega_{\frac{P}{2}})\sum_{l=-\infty}^{\infty}\epsilon^{-jl(\omega-\omega_{p_1+1}-\omega_{p_1+2}\cdots-\omega_{\frac{P}{2}})}d\omega_{p_1+1}\cdots d\omega_{\frac{P}{2}} \tag{2.110}$$

where we have used the fact that formally the Fourier transform of $\epsilon^{-j\omega l)}$ results in an impulse $2\pi\delta^c(\omega)$ in frequency. The summation in the integrand equals $2\pi\delta(\omega -$

$\omega_{p_1+1} - \omega_{p_1+2} \cdots - \omega_{\frac{p}{2}}$) which reduces the number of integrations to $p_2 - 1$ over the manifold defined by

$$\omega_{p_1+1} + \omega_{p_1+2} + \cdots + \omega_{\frac{p}{2}} = \omega \tag{2.111}$$

In other words we can conclude this development by the expression

$$S_u^{(2)i}(\omega) = (\sigma_o^2)^{\frac{p}{2}} S^{0(i)} \text{conv}^{(i)}(S^{\omega}(\omega)) \tag{2.112}$$

where $S^{0(i)}$ represents the product of the system cross-correlation sequences $r_{j_1 j_2}(0)$ that are located in the diagonal blocks and included in the i^{th} permutation. The value of $\text{conv}^{(i)}(S^{\omega}(\omega))$ is obtained by performing multiple convolution on the system cross-spectral functions located in the off-diagonal block $\mathbf{S}^{\omega}$ and included in the i^{th} permutation. Specifically, for p spectral functions $s_1, s_2, \cdots, s_p$ of $\mathbf{S}^{\omega}$ included in the permutation this operation is defined as

$$\begin{aligned}&\text{conv}^{(i)}(S^{\omega}(\omega)) = \\ &(\frac{1}{2\pi})^{p-1} \underbrace{\int_{-\pi}^{\pi} \cdots \int_{-\pi}^{\pi}}_{p-1} s_1(\omega - \theta_1) s_2(\theta_1 - \theta_2) \cdots s_{p-1}(\theta_{p-2} - \theta_{p-1}) s_p(\theta_{p-1}) \\ &\qquad\qquad d\theta_1 d\theta_2 \cdots d\theta_{p-1}\end{aligned} \tag{2.113}$$

Finally, we obtain an expression similar to the expression of the value of the moment

$$S_u^{(2)}(\omega) = \prod_{i=1}^{2k} \boldsymbol{\nu}(i)! \sum_{\Sigma[\mathbf{M}]=\boldsymbol{\nu}} \frac{S_u^{(2)i}(\omega)}{2^{\frac{\text{tr}(\mathbf{M})}{2}}} \prod_{j_1=1}^{2k} \frac{1}{(\frac{M(j_1,j_1)}{2})!} \prod_{j_2=j_1+1}^{2k} \frac{1}{M(j_1,j_2)!} \tag{2.114}$$

where $\boldsymbol{\nu}$ is the vector of exponents appearing in (2.103).

5. Higher-Order Cross-Spectral Density Functions of Functionals of a White Gaussian Process

By generalizing the procedure developed above, a procedure can be developed to find an expression for the value of third, fourth and other higher-order

cross-spectral density functions. The third order cross-spectral density function is defined as

$$S_u^{(3)}(\omega_1, \omega_2) = \sum_{l_1=-\infty}^{\infty} \sum_{l_2=-\infty}^{\infty} \mu_u^{(3)}(l_1, l_2) e^{-j(\omega_1 l_1 + \omega_2 l_2)} \tag{2.115}$$

We start the procedure by constructing the $3k \times 3k$ matrix of system cross-spectral functions

$$\mathcal{S}_u^{(3)} = \begin{bmatrix} \mathbf{S}^0 & \mathbf{S}^{\omega_1} & \mathbf{S}^{\omega_2} \\ (\mathbf{S}^{\omega_1})^{*T} & \mathbf{S}^0 & \mathbf{S}^{\omega_1 - \omega_2} \\ (\mathbf{S}^{\omega_2})^{*T} & (\mathbf{S}^{\omega_1 - \omega_2})^{*T} & \mathbf{S}^0 \end{bmatrix} \tag{2.116}$$

where each of the $k \times k$ matrices is defined as in (2.106). The associated $3k \times 3k$ multiplicity matrix $\mathbf{M}$ is also constructed. We assume that for a specific frequency argument permutation (specific configuration of the matrix $\mathbf{M}$) in the expression (2.101) some system cross-spectral functions are entries of the diagonal blocks denoted $\mathbf{S}^0$. The result of these functions is a constant value taken out of the integrations. The rest of the functions are divided into three groups. The first group consists of p_1 functions that are entries of the block $\mathbf{S}^{\omega_1}$. The second group consists of p_2 functions that are entries of the block $\mathbf{S}^{\omega_2}$, while the third group consists of p_3 functions in the block $\mathbf{S}^{\omega_1 - \omega_2}$. Each of the associated frequency arguments appear in the lag-frequency product of the complex exponential. If the sum of the exponents of the Gaussian functions y_i in the moment expression has an even value p, the corresponding component $S_u^{(3)i}$ takes the form

$$
\begin{aligned}
S_u^{(3)i}(\omega_1,\omega_2) &= (\sigma_o^2)^{\frac{p}{2}}\left(\prod_i S^{0(i)}\right)\sum_{l_1=-\infty}^{\infty}\sum_{l_2=-\infty}^{\infty} e^{-j(\omega_1 l_1+\omega_2 l_2)}(\frac{1}{2\pi})^{p_1+p_2+p_3} \\
&\underbrace{\int_{-\pi}^{\pi}\cdots\int_{-\pi}^{\pi}}_{p_1+p_2+p_3} s_{11}(\omega_{11})\cdots s_{1p_1}(\omega_{1p_1})s_{21}(\omega_{21})\cdots s_{2p_2}(\omega_{2p_2})s_{31}(\omega_{31})\cdots s_{3p_3}(\omega_{3p_3}) \\
&\times e^{jl_1(\omega_{11}+\cdots+\omega_{1p_1})}e^{jl_2(\omega_{21}+\cdots+\omega_{2p_2})}e^{j(l_1-l_2)(\omega_{31}+\cdots+\omega_{3p_3})}d\omega_{11}d\omega_{12}\cdots\cdots d\omega_{3p_3} \\
&= (\sigma_o^2)^{\frac{p}{2}}\left(\prod_i S^{0(i)}\right)(\frac{1}{2\pi})^{p_1+p_2+p_3} \\
&\times \underbrace{\int_{-\pi}^{\pi}\cdots\int_{-\pi}^{\pi}}_{p_1+p_2+p_3} s_{11}(\omega_{11})\cdots s_{1p_1}(\omega_{1p_1})s_{21}(\omega_{21})\cdots s_{2p_2}(\omega_{2p_2})s_{31}(\omega_{31})\cdots s_{3p_3}(\omega_{3p_3}) \\
&\times \sum_{l_1=-\infty}^{\infty} e^{-jl_1(\omega_1-\omega_{11}-\cdots-\omega_{1p_1}-\omega_{31}-\cdots-\omega_{3p_3})} \\
&\times \sum_{l_2=-\infty}^{\infty} e^{-jl_2(\omega_2-\omega_{21}-\cdots-\omega_{2p_2}+\omega_{31}+\cdots+\omega_{3p_3})}d\omega_{11}d\omega_{12}\cdots\cdots d\omega_{3p_3}
\end{aligned}
\tag{2.117}
$$

where ω_{ij} are the new variables of integration that are needed when we substitute (2.102) into (2.101). After substituting the two summations with the values $\delta^c(\Sigma[\boldsymbol{\omega}])$, this expression equals zero outside the manifold defined by the two relations

$$
\begin{aligned}
\omega_1 &= \omega_{11}+\cdots+\omega_{1p_1}+(\omega_{31}+\cdots+\omega_{3p_3}) \\
\omega_1 &= \omega_{21}+\cdots+\omega_{2p_2}-(\omega_{31}+\cdots+\omega_{3p_3})
\end{aligned}
\tag{2.118}
$$

If we now define the argument

$$
\theta = \omega_{31}+\cdots+\omega_{3p_3} \tag{2.119}
$$

we have

$$
\omega_{11}+\cdots+\omega_{1p_1} = \omega_1-\theta \tag{2.120}
$$

and

$$
\omega_{21}+\cdots+\omega_{2p_2} = \omega_2+\theta \tag{2.121}
$$

and the number of integrations is reduced to $p_1 + p_2 + p_3 - 2$. This can be repeated successively as follows. The first group of functions is integrated $p_1 - 1$ times over the manifold defined by (2.120) to give

$$S^{(i)}_{conv1}(\omega_1 - \theta) = \\ (\frac{1}{2\pi})^{p_1-1} \underbrace{\int_{-\pi}^{\pi} \cdots \int_{-\pi}^{\pi}}_{p_1-1} s_{11}(\omega_1 - \theta - \theta_1) s_{12}(\theta_1 - \theta_2) \cdots s_{1p_1}(\theta_{p_1-1}) d\theta_1 d\theta_2 \cdots d\theta_{p_1-1} \tag{2.122}$$

The second group is then integrated $p_2 - 1$ times over the manifold defined by (2.121) to give

$$S^{(i)}_{conv2}(\omega_2 + \theta) = \\ (\frac{1}{2\pi})^{p_2-1} \underbrace{\int_{-\pi}^{\pi} \cdots \int_{-\pi}^{\pi}}_{p_2-1} s_{21}(\omega_2 + \theta - \theta_1) s_{22}(\theta_1 - \theta_2) \cdots s_{2p_2}(\theta_{p_2-1}) d\theta_1 d\theta_2 \cdots d\theta_{p_2-1} \tag{2.123}$$

and the third group is integrated on the manifold defined by (2.119) to give

$$S^{(i)}_{conv3}(\theta) = \\ (\frac{1}{2\pi})^{p_3-1} \underbrace{\int_{-\pi}^{\pi} \cdots \int_{-\pi}^{\pi}}_{p_3-1} s_{31}(\theta - \theta_1) s_{32}(\theta_1 - \theta_2) \cdots s_{3p_3}(\theta_{p_3-1}) d\theta_1 d\theta_2 \cdots d\theta_{p_3-1} \tag{2.124}$$

The last integral gives us the value of the required component

$$S^{(3)i}_u(\omega_1, \omega_2) = (\sigma_o^2)^{\frac{p}{2}} \left(\prod S^{0(i)} \right) \\ \times (\frac{1}{2\pi}) \int_{-\pi}^{\pi} S^{(i)}_{conv1}(\omega_1 - \theta) S^{(i)}_{conv2}(\omega_2 + \theta) S^{(i)}_{conv3}(\theta) d\theta \tag{2.125}$$

and the total value of the third-order cross-spectral density function is therefore

$$S^{(3)}_u(\omega_1, \omega_2) = \prod_{i=1}^{3k} \boldsymbol{\nu}(i)! \sum_{\Sigma[\mathbf{M}]=\boldsymbol{\nu}} \frac{S^{(3)i}_u(\omega)}{2^{\frac{tr(\mathbf{M})}{2}}} \prod_{j_1=1}^{3k} \frac{1}{(\frac{M(j_1,j_1)}{2})!} \prod_{j_2=j_1+1}^{3k} \frac{1}{M(j_1, j_2)!} \tag{2.126}$$

Finally an analogous procedure can be explained to compute the fourth-order cross-spectral density function

$$S_u^{(4)}(\omega_1,\omega_2,\omega_3)=\sum_{l_1=-\infty}^{\infty}\sum_{l_2=-\infty}^{\infty}\sum_{l_3=-\infty}^{\infty}\mu_u^{(4)}(l)e^{-j(\omega_1 l_1+\omega_2 l_2+\omega_3 l_3)} \tag{2.127}$$

The $4k \times 4k$ matrix of system cross-spectral functions is defined as

$$\mathcal{S}_u^{(4)}=\begin{bmatrix} \mathbf{S}^0 & \mathbf{S}^{\omega_1} & \mathbf{S}^{\omega_2} & S^{\omega_3} \\ (\mathbf{S}^{\omega_1})^{*T} & \mathbf{S}^0 & \mathbf{S}^{\omega_1-\omega_2} & \mathbf{S}^{\omega_1-\omega_3} \\ (\mathbf{S}^{\omega_2})^{*T} & (\mathbf{S}^{\omega_1-\omega_2})^{*T} & \mathbf{S}^0 & \mathbf{S}^{\omega_2-\omega_3} \\ (\mathbf{S}^{\omega_3})^{*T} & (\mathbf{S}^{\omega_1-\omega_3})^{*T} & (\mathbf{S}^{\omega_2-\omega_3})^{*T} & \mathbf{S}^0 \end{bmatrix} \tag{2.128}$$

and the same size multiplicity matrix is constructed to compute all the distinct permutations of frequency pairing and the number of their occurances. In this case we get six groups of the product of the system cross-spectral functions located in the upper off-diagonal blocks. Let us assume that the number of spectral functions in the j^{th} group is p_j and their sum is P. The frequency arguments are denoted by $\omega_{j1},\omega_{j2},\cdots,\omega_{jp_j}$. In this case three intermediate frequency arguments $\theta_1,\theta_2,\theta_3$ corresponding to the fourth, fifth and sixth groups located in $\mathbf{S}^{\omega_1-\omega_2}$, $\mathbf{S}^{\omega_1-\omega_3}$ and $\mathbf{S}^{\omega_2-\omega_3}$ respectively are defined such that

$$\theta_1=\omega_{41}+\omega_{42}+\cdots+\omega_{4p_4} \tag{2.129}$$

$$\theta_2=\omega_{51}+\omega_{52}+\cdots+\omega_{5p_5} \tag{2.130}$$

$$\theta_3=\omega_{61}+\omega_{62}+\cdots+\omega_{6p_6} \tag{2.131}$$

By procedures similar to those described for the third order polyspectrum the component $S_u^{(4)i}(\omega_1,\omega_2,\omega_3)$ is obtained by performing the following multiple convolutions:

$$S^{(i)}_{conv1}(\omega_1-\theta_1-\theta_2)=$$
$$(\frac{1}{2\pi})^{p_1-1}\underbrace{\int_{-\pi}^{\pi}\cdots\int_{-\pi}^{\pi}}_{p_1-1}s_{11}(\omega_1-\theta_1-\theta_2-\vartheta_1)s_{12}(\vartheta_1-\vartheta_2)\cdots s_{1p_1}(\vartheta_{p_1-1})d\vartheta_1 d\vartheta_2\cdots d\vartheta_{p_1-1}$$
$$S^{(i)}_{conv2}(\omega_2+\theta_1-\theta_3)=$$
$$(\frac{1}{2\pi})^{p_2-1}\underbrace{\int_{-\pi}^{\pi}\cdots\int_{-\pi}^{\pi}}_{p_2-1}s_{21}(\omega_2+\theta_1-\theta_3-\vartheta_1)s_{22}(\vartheta_1-\vartheta_2)\cdots s_{2p_2}(\vartheta_{p_2-1})d\vartheta_1 d\vartheta_2\cdots d\vartheta_{p_2-1}$$
$$S^{(i)}_{conv3}(\omega_3+\theta_2+\theta_3)=$$
$$(\frac{1}{2\pi})^{p_3-1}\underbrace{\int_{-\pi}^{\pi}\cdots\int_{-\pi}^{\pi}}_{p_3-1}s_{31}(\omega_3+\theta_2+\theta_3-\vartheta_1)s_{32}(\vartheta_1-\vartheta_2)\cdots s_{3p_3}(\vartheta_{p_3-1})d\vartheta_1 d\vartheta_2\cdots d\vartheta_{p_3-1}$$
$$S^{(i)}_{conv4}(\theta_1)=$$
$$(\frac{1}{2\pi})^{p_4-1}\underbrace{\int_{-\pi}^{\pi}\cdots\int_{-\pi}^{\pi}}_{p_4-1}s_{41}(\theta_1-\vartheta_1)s_{42}(\vartheta_1-\vartheta_2)\cdots s_{4p_4}(\vartheta_{p_4-1})d\vartheta_1 d\vartheta_2\cdots d\vartheta_{p_4-1}$$
$$S^{(i)}_{conv5}(\theta_2)=$$
$$(\frac{1}{2\pi})^{p_5-1}\underbrace{\int_{-\pi}^{\pi}\cdots\int_{-\pi}^{\pi}}_{p_5-1}s_{51}(\theta_2-\vartheta_1)s_{52}(\vartheta_1-\upsilon_2)\cdots s_{5p_5}(\vartheta_{p_5-1})d\vartheta_1 d\vartheta_2\cdots d\vartheta_{p_5-1}$$
$$S^{(i)}_{conv6}(\theta_3)=$$
$$(\frac{1}{2\pi})^{p_6-1}\underbrace{\int_{-\pi}^{\pi}\cdots\int_{-\pi}^{\pi}}_{p_6-1}s_{61}(\theta_3-\vartheta_1)s_{62}(\vartheta_1-\vartheta_2)\cdots s_{6p_6}(\vartheta_{p_6-1})d\vartheta_1 d\vartheta_2\cdots d\vartheta_{p_6-1} \quad (2.132)$$

Finally the spectral function component $S^{(4)i}_u(\omega_1,\omega_2,\omega_3)$ is obtained by performing the triple integration

$$S^{(4)i}_u(\omega_1,\omega_2,\omega_3)=(\sigma^2_o)^{\frac{P}{2}}\left(\prod S^{0(i)}\right)\left(\frac{1}{2\pi}\right)^3\int_{-\pi}^{\pi}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi}S^{(i)}_{conv1}(\omega_1-\theta_1-\theta_2)$$
$$\times S^{(i)}_{conv2}(\omega_2+\theta_1-\theta_3)S^{(i)}_{conv3}(\omega_3+\theta_2+\theta_3)S^{(i)}_{conv4}(\theta_1)S^{(i)}_{conv5}(\theta_2)S^{(i)}_{conv6}(\theta_3)d\theta_1 d\theta_2 d\theta_3 \quad (2.133)$$

and the expression for the fourth-order cross-spectral function becomes

$$S_u^{(4)}(\omega_1,\omega_2,\omega_3) = \prod_{i=1}^{4k} \boldsymbol{\nu}(i)! \sum_{\boldsymbol{\Sigma}[\mathbf{M}]=\boldsymbol{\nu}} \frac{S_u^{(4)i}(\omega)}{2^{\frac{\mathrm{tr}(\mathbf{M})}{2}}} \prod_{j_1=1}^{4k} \frac{1}{(\frac{M(j_1,j_1)}{2})!} \prod_{j_2=j_1+1}^{4k} \frac{1}{M(j_1,j_2)!} \quad (2.134)$$

In the same way cross-spectral density functions of any order for the functionals of a white Gaussian process can be obtained through a procedure of successive convolutions.

III. THE WIENER MODEL FOR A DISCRETE NONLINEAR PROCESS

A discrete time random process can be modeled as the output of a discrete system driven by another discrete time random process whose statistics are known up to a sufficient order. This system is generally specified by a set of parameters that can be of either finite or infinite size. The model is successfully constructed when the system parameters are chosen such that the statistics of the system output match the statistics of the modeled process up to a required order. Here the underlying system, although nonlinear, is assumed to be stable and causal (non-anticipative). This assumption satisfies the realizability requirement of the model and (as shown later) generates a structured system architecture. In this chapter we present the Volterra series representation of discrete nonlinear systems and then develop the Wiener model of discrete nonlinear systems and their associated random processes. Although most of the published work on the Wiener model is formulated for continuous time, we adapt the theory in this dissertation to the discrete time case. Consequently, there are several extensions and new results that are developed in this chapter prior to the main results of this dissertation which are presented in the later chapters.

A. THE VOLTERRA SERIES REPRESENTATION OF NONLINEAR SYSTEMS

The output $\{x(n)\}$ of certain "well-behaved" nonlinear systems can be related to the input sequence $\{w(n)\}$ by a Taylor series-like representation developed by the mathematician Vito Volterra and first applied to the study of nonlinear systems by

Norbert Wiener [22]. The details of this series representation of nonlinear systems can be found in several places [18, 22, 17]. We give a brief outline of the theory here.

1. Higher-Order Volterra Kernels and Operators

The Volterra series representation for a causal nonlinear system with input $w(n)$ and output $x(n)$ can be written as

$$x(n) = h_0 + \sum_{k=0}^{\infty} h_1(k)w(n-k) + \sum_{k_1=0}^{\infty}\sum_{k_2=0}^{\infty} h_2(k_1,k_2)w(n-k_1)w(n-k_2) + \cdots + \sum_{k_1=0}^{\infty}\cdots\sum_{k_p=0}^{\infty} h_p(k_1,\ldots,k_p)w(n-k_1)\cdots w(n-k_p) + \cdots \quad (3.1)$$

where $h_p(k_1,\ldots,k_p)$ is called the p^{th}-order Volterra kernel which is identically zero for any of the $k_i < 0$. An equivalent way of representing the system is by a series of operators

$$x(n) = \mathcal{H}_0[w(n)] + \mathcal{H}_1[w(n)] + \mathcal{H}_2[w(n)] + \cdots + \mathcal{H}_p[w(n)] + \cdots \quad (3.2)$$

where

$$\mathcal{H}_p[w(n)] \stackrel{\text{def}}{=} \sum_{k_1=0}^{\infty}\cdots\sum_{k_p=0}^{\infty} h_p(k_1,\ldots,k_p)w(n-k_1)\cdots w(n-k_p) \quad (3.3)$$

is the p^{th}-order Volterra operator or *functional*. In this respect the following is worth noting:

- The Volterra series is a power series with memory. If the input is scaled by a factor a the new output is given by

  $$x(n) = \mathcal{H}_0[w(n)] + a\mathcal{H}_1[w(n)] + a^2\mathcal{H}_2[w(n)] + \cdots + a^p\mathcal{H}_p[w(n)] + \cdots \quad (3.4)$$

  Thus it is inherently nonlinear.

- There is no loss of generality if the Volterra kernels are considered to be symmetric with respect to their arguments because, if they are not, there exists a

procedure by which any asymmetric kernel can be made symmetric. In particular, for any asymmetric kernel $h_p^{\dagger}(k_1, \ldots, k_p)$ the symmetric one is given by

$$h_p(k_1, \ldots, k_p) = \frac{1}{p!} \sum h_p^{\dagger}(k_{i_1}, \ldots, k_{i_p}) \tag{3.5}$$

where the summation is taken over all the $p!$ possible arrangements of the k_i's . Any system constructed using asymmetric kernels can also be constructed with the symmetric one [17].

- The individual Volterra functionals are *homogeneous* since

$$\mathcal{H}_p[aw(n)] = a^p \mathcal{H}_p[w(n)] \tag{3.6}$$

That is, a change in the input magnitude results in a change of the output magnitude but not a change of the output waveform.

To discuss the Volterra system in the frequency domain let $X(\omega)$ and $W(\omega)$ be the Fourier transforms of $x(n)$ and $w(n)$ respectively.[1] Thus $X(\omega)$ and $x(n)$ are related by the Fourier transform pair

$$X(\omega) = \sum_{n=-\infty}^{\infty} x(n) e^{-jn\omega}$$

$$x(n) = \frac{1}{2\pi} \int_{-\pi}^{\pi} X(\omega) e^{jn\omega} d\omega$$

and $W(\omega)$ and $w(n)$ are related similarly. Further, let the generalized transfer functions be defined as the multidimensional Fourier transforms of the kernels

$$H_p(\omega_1, \ldots, \omega_p) = \sum_{k_1=0}^{\infty} \cdots \sum_{k_p=0}^{\infty} h_p(k_1, \ldots, k_p) e^{-j(k_1\omega_1 + \cdots + k_p\omega_p)} \tag{3.7}$$

[1] For the present discussion $w(n)$ and $x(n)$ are assumed to be deterministic and the corresponding Fourier transforms are assumed to exist.

which are symmetric functions of the frequency arguments ω_i if the kernels h_p are symmetric. The response of the p^{th} order Volterra operator

$$x_p(n) = \mathcal{H}_p[w(n)] \tag{3.8}$$

can be obtained by first defining the associated (artificial) functional

$$x_{(p)}(n_1, n_2, \cdots, n_p) \stackrel{\text{def}}{=} \sum_{k_1=0}^{\infty} \cdots \sum_{k_p=0}^{\infty} h_p(k_1, \cdots, k_p) w(n_1 - k_1) \cdots w(n_p - k_p) \tag{3.9}$$

and its p-dimensional Fourier transform given by

$$X_{(p)}(\omega_1, \omega_2, \cdots, \omega_p) = H_p(\omega_1, \omega_2, \cdots, \omega_p) W(\omega_1) W(\omega_2) \cdots W(\omega_p) \tag{3.10}$$

Then the response of the p^{th}-order Volterra operator is given in the frequency domain by [17, 18]

$$X_p(\omega) = \frac{1}{(2\pi)^{p-1}} \int_{-\pi}^{\pi} \cdots \int_{-\pi}^{\pi} X_{(p)}(\omega - \theta_1, \theta_1 - \theta_2, \cdots, \theta_{p-1}) d\theta_1 d\theta_2 \cdots d\theta_{p-1} \tag{3.11}$$

which is a $(p-1)$-dimensional convolution, generally not simple to evaluate. The system output in the frequency domain is then given as the sum of the terms

$$X(\omega) = X_0(\omega) + X_1(\omega) + \cdots + X_p(\omega) + \cdots \tag{3.12}$$

2. Limitations of the Volterra Representation of Nonlinear Systems

The practical application of Volterra series in nonlinear system theory has two main difficulties. One concerns the measurements of the Volterra kernels and the second concerns the convergence of the Volterra series.

The measurement of a general system's Volterra kernel is possible only if the contribution of each of the Volterra operators can be separated from the total system response. For the work done in this direction, the nonlinear system to be

identified is assumed to be representable by a finite order Volterra series and each kernel is obtained by isolating the higher order operators. There is no exact method to isolate an individual Volterra operator for a system that is not of finite order unless approximations are assumed possible.

The problem of convergence of the Volterra series representation of a nonlinear system is similar to that encountered in the Taylor series representation of a function. This is evident from equation (3.4). This representation is expected to converge only for a limited range of the magnitude of the system input. Moreover, nonlinear memoryless systems that include saturating elements cannot be characterized by a Volterra series that converges for all magnitudes of the input [17].

The limitations of the power series representation of functions are circumvented by using orthogonal functions. In power series representation of functions the value of the series expansion of the function is strictly convergent for the values of the argument; therefore a region of convergence must be defined. In orthogonal function representation, the functions are required to converge in the mean square sense; therefore the convergence requirement is not imposed on the argument value but on the average of a product of functions. Wiener extended the notion of orthogonality to functionals and built a canonical model for nonlinear systems. This extension, known as the Wiener model, is discussed next.

B. THE WIENER MODEL FOR A DISCRETE TIME NONLINEAR SYSTEM

Weiner defined a class of systems suitable for his representation which has been called the *Wiener class*. A system that is a member of the Wiener class has the following properties:

1. The system is not "explosive." For a finite variance random input the system output also has finite variance.

2. The system does not have infinite memory. The present value of the system output must become asymptotically independent of the past values of the input.

1. G-Functional Representation of Nonlinear Systems

From the Volterra functionals, Wiener formed a set of orthogonal functionals which he called the G-functionals, since they are orthogonal if the input is a white Gaussian time function [22, 17]. The type of convergence of the orthogonal series is *convergence in the mean.* As a result, the class of nonlinear systems that can be represented by the Wiener G-functionals is larger than the class of systems describable by a Volterra series. Although Wiener developed and reported his results for continuous time, we shall state the results in discrete time as is more suitable to our purpose. Let us define the functional

$$\Gamma_p[h_p, h_{p-1(p)}, h_{p-2(p)}, \cdots, h_{0(p)}; w(n)] = h_{0(p)} + \sum_{i=1}^{p} \left[\sum_{k_1=0}^{\infty} \cdots \sum_{k_i=0}^{\infty} h_{i(p)}(k_1, k_2, \cdots, k_i) w(n-k_1) w(n-k_2) \cdots w(n-k_i) \right] \tag{3.13}$$

as the p^{th} degree *non-homogeneous* Volterra functional. Observe that Γ_p is a sum of homogeneous Volterra functionals of the type defined by (3.3). The p in parenthesis in the subscript of a kernel indicates that it is a member of a p^{th} degree non-homogeneous functional, and we define $h_{p(p)} = h_p$. Wiener developed the set of the orthogonal G-functionals by requiring that[1]

$$\mathcal{E}\left\{\mathcal{H}_m[w(n)]\Gamma_p[h_p, h_{p-1(p)}, h_{p-2(p)}, \cdots, h_{0(p)}; w(n)]\right\} = 0 \qquad \text{for } m < p \tag{3.14}$$

[1] Although the original results were formulated using temporal averages we will restate the results in terms of ensemble averages.

when the input $w(n)$ is white Gaussian noise with variance σ_0^2. In other words, Γ_p is orthogonal to any functional $\mathcal{H}_m$ of lower degree. The particular functionals Γ_p satisfying these conditions are called G-functionals and are denoted by $\mathcal{G}_p[g_p; w(n)]$. The quantity g_p is the leading kernel of the functional from which the kernels $g_{p-1(p)}, g_{p-2(p)}, \cdots, g_{0(p)}$ are derived. The symmetry properties of both the Volterra and the Wiener kernels and the average of the product of Gaussian processes result in the following properties:

1. For the even order G-functionals, the derived kernels of odd order are zero and the functional expression in the form of (3.13) has only even order kernels. An analogous statement applies to the odd order G-functionals for which the even order derived kernels are zero. The general expression for the Wiener G-functional is then given by

$$\mathcal{G}_p[g_p: w(n)] = \sum_{m=0}^{\lfloor \frac{p}{2} \rfloor} \mathbf{G}_{p-2m}[w(n)] \tag{3.15}$$

where $\lfloor \frac{p}{2} \rfloor$ represents the greatest integer less than or equal to $\frac{p}{2}$ and $\mathbf{G}_{p-2m}[w(n)]$ is the $(p-2m)^{th}$ degree homogeneous functional given by

$$\mathbf{G}_{p-2m}[w(n)] = \sum_{k_1=0}^{\infty} \cdots \sum_{k_{p-2m}=0}^{\infty} g_{p-2m(p)}(k_1, k_2, \cdots, k_{p-2m}) w(n-k_1) w(n-k_2) \cdots w(n-k_{p-2m}) \tag{3.16}$$

2. The kernels $g_{p-2m(p)}$ for $m = 0, 1, 2, \ldots, \lfloor \frac{p}{2} \rfloor$ are derived from the leading kernel g_p by the relation

$$g_{p-2m(p)}(k_1, k_2, \cdots, k_{p-2m}) = \frac{(-1)^m p! (\sigma_o^2)^m}{(p-2m)!(m)! 2^m} \sum_{\nu_1=0}^{\infty} \cdots \sum_{\nu_m=0}^{\infty} g_p(\nu_1, \nu_1, \cdots, \nu_m, \nu_m, k_1, k_2, \cdots, k_{p-2m}) \tag{3.17}$$

The coefficients of the integrals of the derived kernels of $\mathcal{G}_p$ are similar to the coefficients in the expression of the Hermite polynomial

$$\mathbf{H}_p(x) = \sum_{m=0}^{\lfloor \frac{p}{2} \rfloor} \frac{(-1)^m p!(\sigma_o^2)^m}{(p-2m)!(m)!2^m} x^{(p-2m)} \tag{3.18}$$

Both of these properties are shown in Appendix **B**.

In general, the Volterra functionals $\mathcal{H}_p$ and the Wiener functionals $\mathcal{G}_p$ of the same system are different. The Volterra representation of a general nonlinear system is

$$x(n) = \sum_{i=0}^{\infty} \mathcal{H}_i[w(n)] \tag{3.19}$$

while the Wiener representation of the same system is

$$x(n) = \sum_{i=0}^{\infty} \mathcal{G}_i[g_i; w(n)] \tag{3.20}$$

(In both cases the sum is possibly finite.) Since the Volterra functionals are homogeneous, the Volterra functional of some degree p in (3.19) is obtained by combining all the p^{th} degree kernels in (3.20). For a finite order system the Volterra functional is obtained by summing the leading and the derived functionals of the same degree along the columns of Table 1.1 while the G-functionals are formed by summing along the rows in the same table. The G-functional expansion of $\mathcal{H}_p$ is therefore

$$\mathcal{H}_p[w(n)] = \sum_{m=0}^{\infty} \mathbf{G}_{p(p+2m)}[g_{p(p+2m)}; w(n)] \tag{3.21}$$

and the kernels h_p and $g_{p(p+2m)}$ are related in an analogous way.

From the above considerations it is important to observe the following:

1. The Wiener functionals and kernels are functions of the power level of the input σ_o^2 while the Volterra functionals and kernels are invariant with the power level of the system input. Further, *the Wiener G-functionals are <u>not</u> orthogonal unless the input is white Gaussian with variance* σ_o^2.

G functional	homogeneous functionals						
$\mathcal{G}_0$	G_0						
$\mathcal{G}_1$		G_1					
$\mathcal{G}_2$	$G_{0(2)}$		G_2				
$\mathcal{G}_3$		$G_{1(3)}$		G_3			
$\mathcal{G}_4$	$G_{0(4)}$		$G_{2(4)}$		G_4		
$\mathcal{G}_5$		$G_{1(5)}$		$G_{3(5)}$		G_5	
$\vdots$		$\vdots$		$\vdots$		$\vdots$	
Volterra fuctionals:	$\mathcal{H}_0$	$\mathcal{H}_1$	$\mathcal{H}_2$	$\mathcal{H}_3$	$\mathcal{H}_4$	$\mathcal{H}_5$	$\cdots$

TABLE 3.1: The relation between the Wiener and Volterra functionals

2. The class of systems describable by the Volterra representation is a proper subset of the class of systems describable by the Wiener G-functionals. Since the Wiener G-functional series is an orthogonal series for the white Gaussian input, it does not require the restriction on the magnitude of the input to avoid the strict convergence problem as in the case of Volterra series. This means that a system may have a G-functional representation and not a Volterra series representation. Any system that can be represented by a Volterra series, however, can also have a G-functional representation.[1]

Before continuing the development of the Wiener theory, it is necessary to point out two important additional properties of the G-functionals. First, it follows from the general form of the functional and (3.14)-(3.17) that the G-functional is linear with respect to its kernel, i.e.,

$$\mathcal{G}_p[c_1 g_p + c_2 f_p; w(n)] = c_1 \mathcal{G}_p[g_p; w(n)] + c_2 \mathcal{G}_p[f_p; w(n)] \tag{3.22}$$

where c_1 and c_2 are any real or complex constants. Secondly, the expectation of the product of any two G-functionals (say for two different nonlinear systems) is given

[1]See [17] for further discussion.

by

$$\mathcal{E}\{\mathcal{G}_{p_1}[g_{p_1};w(n)]\mathcal{G}_{p_2}[f_{p_2};w(n)]\} = \begin{cases} 0 & \text{for } p_1 \neq p_2 \\ \gamma_{p_1}(\sigma_o) & \text{for } p_1 = p_2 \end{cases} \tag{3.23}$$

where

$$\gamma_{p_1}(\sigma_o) = p_1!\sigma_0^{2p_1} \sum_{k_1=0}^{\infty} \cdots \sum_{k_{p_1}=0}^{\infty} g_{p_1}(k_1,k_2,\cdots,k_{p_1}) f_{p_1}(k_1,k_2,\cdots,k_{p_1}) \tag{3.24}$$

The latter is equivalent to taking the expectation of the product of the two leading kernels only. The proof of these statements is also given in Appendix B.

2. Orthogonal Development of the G-Functional

It is shown (in Section C of this chapter) that the Wiener kernel g_p can be expanded in the p-dimensional discrete Laguerre series

$$g_p(k_1,k_2,\cdots,k_p) = \sum_{m_1=0}^{\infty} \cdots \sum_{m_p=0}^{\infty} c_{m_1 m_2 \cdots m_p} \lambda_{m_1}(k_1)\lambda_{m_2}(k_2)\cdots\lambda_{m_p}(k_p) \tag{3.25}$$

where $\lambda_m(k)$ is the m^{th} degree *discrete Laguerre function* and that the G-functional of $\lambda_i^{(p)}(k)$ can be correspondingly expanded in the form

$$\mathcal{G}_p[\lambda_i^{(p)};w(n)] = \sum_{i=0}^{\lfloor \frac{p}{2} \rfloor} \frac{(-1)^i p!(\sigma_o^2)^i}{(p-2i)!(i)!2^i} \left[\sum_{k=0}^{\infty} \lambda_i(k)w(n-k)\right]^{(p-2m)} \tag{3.26}$$

for white noise with variance σ_o^2. It is thus seen that $\mathcal{G}_p[g_p: w(n)]$ can be represented by sums of products of terms

$$y_i(n) = \sum_{k=0}^{\infty} \lambda_i(k)w(n-k) \tag{3.27}$$

where $y_i(n)$ is the output of a system whose impulse response is the i^{th} degree Laguerre function $\lambda_i(n)$ and whose input is $w(n)$. This output is also a zero-mean Gaussian process with variance σ_o^2. Thus from (3.26) it is seen that $\mathcal{G}_p$ can be represented by a set of linear systems that provide all of the "memory" followed by a nonlinear system *without* memory that acts to form sums of products of the $y_i(n)$.

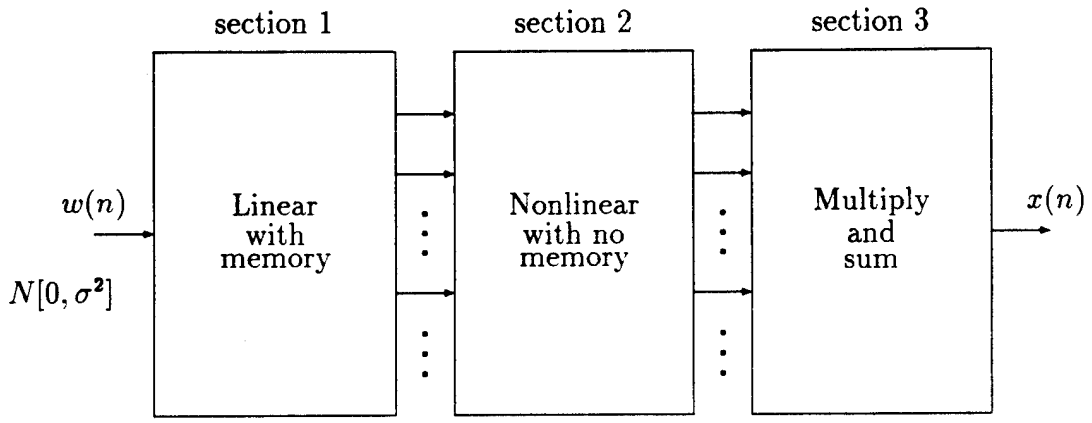

Figure 3.1: The Wiener model of nonlinear processes.

Further, the coefficients in (3.26) can be identified as those of the Hermite polynomials. Therefore (using the linearity property (3.22)) the G-functional is simply given by the equation

$$\mathcal{G}_p[g_p; w(n)] = \sum_{m_1=0}^{\infty} \cdots \sum_{m_p=0}^{\infty} c_{m_1 \cdots m_p} \prod_{j=1}^{N} \mathbf{H}_{p_j}(y_{m_j}(n)) \tag{3.28}$$

in which $m_{j_1} \neq m_{j_2}$ for $j_1 \neq j_2$ and $\sum_{j=1}^{N} p_j = p$. The Hermite polynomials are orthogonal and the expectation of the product of two polynomials is given by

$$\mathcal{E}\{\mathbf{H}_{p_1}(y_{m_1}(n))\mathbf{H}_{p_2}(y_{m_2}(n))\} = \delta_{p_1 p_2}\delta_{m_1 m_2} p_1! \sigma_o^{2p_1} \tag{3.29}$$

Since the Wiener representation of a nonlinear system is the sum of G-functionals, it follows that the system has the same form of representation as the individual $\mathcal{G}_p[g_p; w(n)]$. If the nonlinear system model is driven by the required white Gaussian noise, we have a rich representation for a large class of random processes. Fig. 3.1 thus represents what will be referred to as *Wiener model of nonlinear processes.* This model consists of three sections. Section 1 is a single-input multi-output linear

system with memory that can be represented as a bank of linear filters driven by the same input $w(n)$. The input $w(n)$ is a zero-mean white Gaussian process with variance σ_o^2. The impulse responses of the filters in section 1 are the discrete Laguerre functions;

Fig. 3.2 shows the structure of this section which, in general, has infinite length. Section 2 is a multi-input multi-output system with no memory, which can be repre-

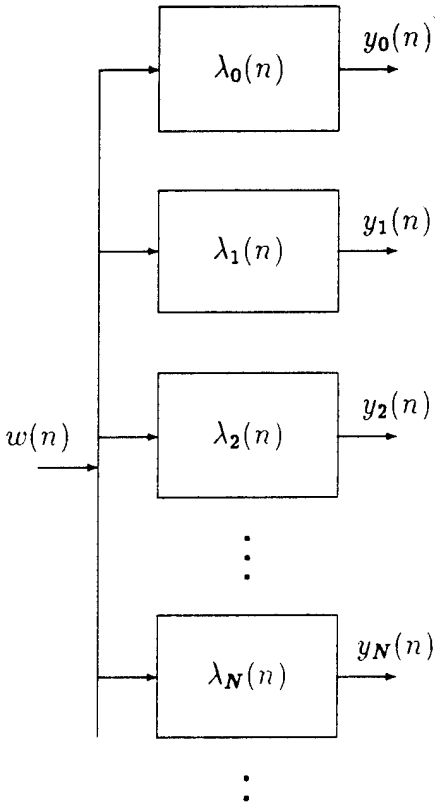

Figure 3.2: Section 1 of the Wiener model

sented by an infinite bank of identical memoryless nonlinear single-input multi-output blocks, each driven by one of the outputs of section 1. Each block of this section consists of parallel nonlinear functions described by the Hermite polynomials. Although

in the general model this section may also be infinite, processes with a finite degree of nonlinearity will require sections only up to some finite order N_N. The structure of one of the nonlinear blocks is shown in Fig. 3.3. Section 3 is a multi-input

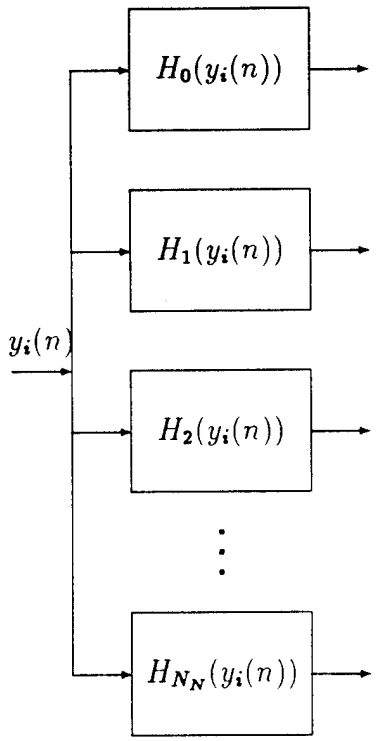

Figure 3.3: Section 2 of the general Wiener model.

single-output memoryless system whose inputs are the outputs of section 2. The model output is the weighted sum of certain products of the inputs. The products are formed by taking one and only one output from each block in section 2. Sections 1 and 2 are the same for all systems and therefore all processes in the Wiener class. The parameters, which are the weighting coefficients of section 3, characterize a particular process. In this case a stationary finite-moment random process can be

approximated by a finite size Wiener model [17]. The model size is given by :

1. The highest order, N_L, of the Laguerre functions in section 1. This section consists of the N_L+1 systems with impulse responses given by $\lambda_0(n), \lambda_1(n), \cdots, \lambda_{N_L}(n)$. It will be seen shortly that the order N_L roughly corresponds to the "length of memory" of the nonlinear system.

2. The highest order, N_N, of the Hermite polynomials in each of the nonlinear blocks of section 2. Therefore, each block consists of $N_N + 1$ parallel nonlinear memoryless systems described by the Hermite polynomials $\mathbf{H_0}(y_i(t)), \mathbf{H_1}(y_i(t))$, $\cdots, \mathbf{H}_{N_N}(y_i(t))$. As mentioned earlier, this corresponds to the degree of nonlinearity present in the process (quadratic, cubic, etc.).

Let us now consider a particular vector of indices $\boldsymbol{\alpha}_i$ of size $N_L + 1$ defined as

$$\boldsymbol{\alpha}_i = [\alpha_{i0}, \alpha_{i1}, \cdots, \alpha_{iN_L}] \tag{3.30}$$

with

$$\alpha_{ij} \in \{0, 1, \cdots, N_N\} \qquad \text{for } j = 0, 1, \cdots, N_L \tag{3.31}$$

Then the model output is given by

$$x(n) = \sum_{\Sigma[\boldsymbol{\alpha}_i]=0}^{N_N} c_{\boldsymbol{\alpha}_i} \prod_{j=0}^{N_L} \mathbf{H}_{\alpha_{ij}}(y_j(n)) \tag{3.32}$$

in which the summation is over all vectors $\boldsymbol{\alpha}_i$ such that the sum of the components is less than or equal to N_N. For a particular $\boldsymbol{\alpha}_i$ the term

$$\mathbf{Q}_{\boldsymbol{\alpha}_i}(n) = \prod_{j=0}^{N_L} \mathbf{H}_{\alpha_{ij}}(y_j(n)) \tag{3.33}$$

is a multidimensional Hermite polynomial in the outputs of the linear section. We can refer to this multidimensional Hermite polynomial as the *Q-polynomial.* Thus the model output can be written more concisely as

$$x(n) = \sum_{\Sigma[\boldsymbol{\alpha}_i]=0}^{N_N} c_{\boldsymbol{\alpha}_i} \mathbf{Q}_{\boldsymbol{\alpha}_i}(n) \tag{3.34}$$

The quantity $\Sigma[\boldsymbol{\alpha}_i]$ is the order of the functional represented by the Q-polynomial $\mathbf{Q}_{\boldsymbol{\alpha}_i}(n)$. The Q-polynomials are zero mean except for the one with index $\boldsymbol{\alpha}_0 = 0$ where the mean equals one. Then

$$\mathcal{E}\{x(n)\} = c_0 \tag{3.35}$$

C. DEVELOPMENT OF THE LINEAR SECTION OF THE DISCRETE WIENER MODEL

A discrete nonlinear system that is a member of the Wiener class of nonlinear sytems can be represented using the G-functional representation. The kernels of the G-functional have the property

$$\sum_{k_1=0}^{\infty}\sum_{k_2=0}^{\infty}\cdots\sum_{k_p=0}^{\infty} g_p^2(k_1, k_2, \cdots, k_p) < \infty \tag{3.36}$$

Therefore these kernels can be expanded in terms of a complete set of orthonormal functions. In this section we derive the expression for these functions which were shown to be the transfer functions of the bank of the linear systems that represent section 1 of the model. We also develop the orthogonal expansion of the G-functionals.

1. The Transfer Functions of the Linear Bank, the Discrete Laguerre Functions

The single-sided discrete time function $v_m(n)$ of order m is defined as

$$v_m(n) = \frac{n!}{m!(n-m)!}\rho^{n-m}u(n-m) \qquad |\rho| < 1 \tag{3.37}$$

where $u(n)$ is the unit step function. The z-transform of $v_m(n)$ is easily shown to be

$$V_m(z) = \frac{z}{(z-\rho)^{m+1}} \qquad \text{for } |z| > |\rho| \tag{3.38}$$

Since the functions $v_m(n)$ can be thought of as individual terms of a polynomial, it is resonable to expect that an arbitrary time function, say $\phi(n)$, could be represented in some way by a weighted sum of the $v_m(n)$. The problem is more tractable however if we instead use a set of *orthonormal* functions, derived from the $v_m(n)$, in the representation. These functions denoted by $\lambda_i(n)$ can be derived as follows.

The functions $\lambda_i(n)$ are formed by a linear combination

$$\lambda_i(n) = \sum_{m=0}^{i} c_{mi} v_m(n) \tag{3.39}$$

with z-transform

$$\Lambda_i(z) = \sum_{m=0}^{i} c_{mi} V_m(z) \tag{3.40}$$

The coefficients c_{mi} for $m = 0, 1, \cdots, i$ and $i = 0, 1, 2, 3, \cdots$ need to be determined such that

$$\sum_{n=0}^{\infty} \lambda_{i_1}(n)\lambda_{i_2}(n) = \delta_{i_1 i_2} \tag{3.41}$$

which is equivalent to

$$\frac{1}{2\pi j} \oint_{C1} z^{-1} \Lambda_{i_1}(z).\Lambda_{i_2}(z^{-1}) dz = \delta_{i_1 i_2} \tag{3.42}$$

where the contour $C1$ is the unit circle in the complex z-plane.

To satisfy the conditions, we proceed as follows. Substituting (3.38) in (3.40) we obtain

$$\begin{aligned} \Lambda_i(z) &= \sum_{m=0}^{i} c_{mi} \frac{z}{(z-\rho)^{m+1}} \\ &= \frac{c_{0i} z(z-\rho)^i + c_{1i} z(z-\rho)^{i-1} + \cdots + c_{ii} z}{(z-\rho)^{i+1}} \\ &= c_{0i} \frac{z P_i(z)}{(z-\rho)^{i+1}} \end{aligned} \tag{3.43}$$

where $P_i(z)$ is an i^{th} degree polynomial in z that can be factored as

$$P_i(z) = (z - z_1)(z - z_2) \cdots (z - z_i) \tag{3.44}$$

Then

$$\Lambda_i(z) = c_{0i}\frac{z(z-z_1)(z-z_2)\cdots(z-z_i)}{(z-\rho)^{i+1}} \tag{3.45}$$

is a rational polynomial function which has one zero at the origin, i zeros at $z_1, z_2, \cdots, z_i$, and $(i+1)$ poles at ρ. Now since

$$\Lambda_i(z^{-1}) = -c_{0i}\rho^{-(i+1)}\left(\prod_{m=1}^{i} z_m\right)\frac{(z-z_1^{-1})(z-z_2^{-1})\cdots(z-z_i^{-1})}{(z-\rho^{-1})^{i+1}} \tag{3.46}$$

it follows that

$$\begin{aligned}\frac{1}{2\pi j}\oint_{C1} z^{-1}\Lambda_{i_1}(z)\Lambda_{i_2}(z^{-1})dz &= -c_{0i_1}c_{0i_2}\rho^{-(i_2+1)}\left(\prod_{m=1}^{i_2} z_m\right)\\ &\times\frac{1}{2\pi j}\oint_{C1}\frac{(z-z_1)(z-z_2)\cdots(z-z_{i_1})(z-z_1^{-1})(z-z_2^{-1})\cdots(z-z_{i_2}^{-1})}{(z-\rho)^{i_1+1}(z-\rho^{-1})^{i_2+1}}dz\end{aligned} \tag{3.47}$$

which in general is not zero. If the coefficients $c_{0i}, c_{1i}, \cdots, c_{ii}$ in (3.43) are chosen however such that the zeros

$$z_1 = z_2 = \cdots = \rho^{-1}$$

then (3.45) becomes

$$\Lambda_i(z) = c_{0i}\frac{z(z-\rho^{-1})^i}{(z-\rho)^{i+1}} \qquad |z| > |\rho| \tag{3.48}$$

and (3.47) becomes

$$\frac{1}{2\pi j}\oint_{C1} z^{-1}\Lambda_{i_1}(z)\Lambda_{i_2}(z^{-1})dz = -c_{0i_1}c_{0i_2}\rho^{-2i_2-1}\frac{1}{2\pi j}\oint_{C1}\frac{(z-\rho^{-1})^{i_1-i_2-1}}{(z-\rho)^{i_1-i_2+1}}dz \tag{3.49}$$

If $i_1 < i_2$ the integrand does not have any poles inside the unit circle, which means that the integral equals to zero. If $i_1 > i_2$ then it has $i_1 - i_2 + 1$ poles at ρ but the integral is also equal to zero since

$$\frac{1}{(i_1-i_2)!}\frac{d^{i_1-i_2}}{dz^{i_1-i_2}}(z-\rho^{-1})^{i_1-i_2-1} = 0$$

Therefore the functions $\lambda_i(n)$ whose transform is given by (3.48) satisfy the orthogonality condition.

To normalize this set of orthogonal functions, set $i_1 = i_2 = i$ in (3.49) and equate to 1:

$$\begin{aligned}\frac{1}{2\pi j}\oint_{C1} z^{-1}\Lambda_i(z)\Lambda_i(z^{-1})dz &= -c_{0i}^2\rho^{-2i-1}\frac{1}{2\pi j}\oint_{C1}\frac{1}{(z-\rho^{-1})(z-\rho)}dz \\ &= -c_{0i}^2\frac{\rho^{-2i-1}}{(\rho-\rho^{-1})} \\ &= 1 \end{aligned} \tag{3.50}$$

Thus it follows that

$$c_{0i} = \rho^i\sqrt{1-\rho^2} \tag{3.51}$$

and

$$\Lambda_i(z) = \sqrt{1-\rho^2}\rho^i\frac{z(z-\rho^{-1})^i}{(z-\rho)^{i+1}} \qquad |z| > |\rho| \tag{3.52}$$

To obtain an expression for c_{mi} in (3.39), equation (3.52) is put in the form

$$\begin{aligned}\Lambda_i(z) &= \sqrt{1-\rho^2}\rho^i\frac{z(z-\rho^{-1})^i}{(z-\rho)^{i+1}} \\ &= \sqrt{1-\rho^2}\rho^i\frac{z}{(z-\rho)^{i+1}}\left[(z-\rho)+(\rho-\rho^{-1})\right]^i \\ &= \sqrt{1-\rho^2}\rho^i\frac{z}{(z-\rho)^{i+1}}\left[(z-\rho)-\rho^{-1}(1-\rho^2)\right]^i \\ &= \sqrt{1-\rho^2}\rho^i\frac{z}{(z-\rho)^{i+1}}\sum_{m=0}^{i}(-1)^m\rho^{-m}(1-\rho^2)^m\begin{pmatrix} i \\ m\end{pmatrix}(z-\rho)^{i-m} \\ &= \sum_{m=0}^{i}(-1)^m\rho^{i-m}(1-\rho^2)^{\frac{2m+1}{2}}\begin{pmatrix} i \\ m\end{pmatrix}\frac{z}{(z-\rho)^{m+1}} \\ &= \sum_{m=0}^{i}(-1)^m\rho^{i-m}(1-\rho^2)^{\frac{2m+1}{2}}\begin{pmatrix} i \\ m\end{pmatrix}V_m(z) \end{aligned} \tag{3.53}$$

Equating the coefficients of $V_m(z)$ in (3.40) and (3.53) we obtain

$$c_{mi} = (-1)^m\rho^{i-m}(1-\rho^2)^{\frac{2m+1}{2}}\begin{pmatrix} i \\ m\end{pmatrix} \tag{3.54}$$

Then substituting (3.37) and (3.54) in (3.39) we arrive at the explicit expression

$$\lambda_i(n) = \sum_{m=0}^{i}(-1)^m\rho^{i-m}(1-\rho^2)^{\frac{2m+1}{2}}\binom{i}{m}\frac{n!}{m!(n-m)!}\rho^{n-m}u(n-m)$$

$$= \sqrt{1-\rho^2}\sum_{m=0}^{i}(-1)^m\binom{i}{m}\binom{n}{m}(1-\rho^2)^m\rho^{i+n-2m}u(n-m) \quad (3.55)$$

This is the *discrete Laguerre function* $\lambda_i(n)$ which describes the impulse response of the i^{th} system in the bank of linear filters in the discrete Wiener model. The bank of linear filters comprising section 1 of the model is then represented by the impulse responses described by the discrete Laguerre functions.

Basic results on the discrete Laguerre functions can be found in [25, 26, 27]. This complete set of orthonormal functions has recently gained considerable interest in linear system applications. Among these applications are the problem of identification of causal stable systems [28, 29, 30], the design of adaptive filters [31] and the analysis and computation of local and running cross-correlation functions using Lagurre cross-correlation sequences [32, 31]. The expression (3.55), although formidable in the time domain, will be shown to have an exceptionally convenient interpretation in the transform domain which not only provides insight about properties of the discrete Laguerre functions, but also leads to simple procedures for their computation. However, we first show that the Wiener G-functionals can *indeed* be expanded in terms of the discrete Laguerre functions and therefore that the structure shown in Fig. 3.1 is an appropriate one.

2. Orthogonal Expansion of the G-Functionals

Given a set of orthogonal functions in one variable, a function of several variables can be expanded as a sum of products of these functions. Thus the Wiener kernel g_p can be expanded in the p-dimensional series

$$g_p(k_1, k_2, \cdots, k_p) = \sum_{m_1=0}^{\infty}\cdots\sum_{m_p=0}^{\infty} c_{m_1 m_2 \cdots m_p}\lambda_{m_1}(k_1)\lambda_{m_2}(k_2)\cdots\lambda_{m_p}(k_p) \qquad (3.56)$$

where $\lambda_m(k)$ is the m^{th} degree discrete Laguerre function and the coefficients $c_{m_1 m_2 \cdots m_p}$ are given by

$$c_{m_1 m_2 \cdots m_p} = \sum_{k_1=0}^{\infty} \cdots \sum_{k_p=0}^{\infty} g_p(k_1, k_2, \cdots, k_p) \lambda_{m_1}(k_1) \lambda_{m_2}(k_2) \cdots \lambda_{m_p}(k_p) \tag{3.57}$$

In this expansion the kernels remain symmetric with respect to the time arguments since the indices m_1 through m_p assume the same set of values from the the integers in the set $[0, \infty)$. By Applying the linearity property (3.22) and the expansion in (3.56), it can be seen that the p^{th} degree G-functional has an expansion of the form

$$\mathcal{G}_p[g_p; w(n)] = \sum_{m_1=0}^{\infty} \cdots \sum_{m_p=0}^{\infty} c_{m_1 m_2 \cdots m_p} \mathcal{G}_p^{\dagger}[\lambda_{m_1} \lambda_{m_2} \cdots \lambda_{m_p}; w(n)] \tag{3.58}$$

where $\mathcal{G}_p^{\dagger}$ is a nonhomogeneous functional with leading kernel $g_p^{\dagger} = \lambda_{m_1} \lambda_{m_2} \cdots \lambda_{m_p}$. The kernels of each term in this functional expansion may have the Laguerre functions repeated; therefore we write an individual kernel in the form

$$g_p^{\dagger} = \lambda_{m_1}^{p_1} \lambda_{m_2}^{p_2} \cdots \lambda_{m_N}^{p_N} \tag{3.59}$$

such that

$$\sum_{j=1}^{N} p_j = p \qquad N = 1, 2, \cdots \text{or } p \tag{3.60}$$

It is important to point out that this kernel $g_p^{\dagger}$ is not symmetric with respect to its arguments; but the sum (3.56) or (3.58) produces a kernel g_p which *is* symmetric. This happens because other kernels in the sum have time arguments that permute in the functions $\lambda_i^{(p_i)}$. To give an example, for $N = 2$ and $p_1 + p_2 = p$ the leading homogeneous functional $\mathbf{G}_p[g_p; w(n)]$ can be written as

$$
\begin{aligned}
&\mathbf{G}_p^{\dagger}[\lambda_{i_1}^{(p_1)}\lambda_{i_2}^{(p_2)};w(n)] = \\
&\sum_{k_1=0}^{\infty}\cdots\sum_{k_p=0}^{\infty}\lambda_{i_1}(k_1)\lambda_{i_1}(k_2)\cdots\lambda_{i_1}(k_{p_1})\lambda_{i_2}(k_{p_1+1})\lambda_{i_2}(k_{p_1+2})\cdots\lambda_{i_2}(k_p) \\
&\qquad\times w(n-k_1)w(n-k_2)\cdots w(n-k_p) \\
&=\sum_{k_1=0}^{\infty}\cdots\sum_{k_{p_1}=0}^{\infty}\lambda_{i_1}(k_1)\lambda_{i_1}(k_2)\cdots\lambda_{i_1}(k_{p_1})w(n-k_1)w(n-k_2)\cdots w(n-k_{p_1}) \\
&\times\sum_{k_1=0}^{\infty}\cdots\sum_{k_{p_2}=0}^{\infty}\lambda_{i_2}(k_1)\lambda_{i_2}(k_2)\cdots\lambda_{i_2}(k_{p_2})w(n-k_1)w(n-k_2)\cdots w(n-k_{p_2}) \\
&=\mathbf{G}_{p_1}^{\dagger}[\lambda_{i_1}^{(p_1)};w(n)]\mathbf{G}_{p_2}^{\dagger}[\lambda_{i_2}^{(p_2)};w(n)]
\end{aligned}
\tag{3.61}
$$

To keep the symmetry property we could repeat the expansion terms $(p_1+p_2)!$ times permuting the time arguments and divide their sum by $(p_1+p_2)!$. In this case however, each term would give us the same value, namely

$$\mathbf{G}_{p_1}[\lambda_{i_1}^{(p_1)};w(n)]\mathbf{G}_{p_2}[\lambda_{i_2}^{(p_2)};w(n)]$$

Then we would multiply again by $(p_1+p_2)!$ which leads to the same result if we assume that the kernels of the homogeneous functional are symmetric. In general, for any value of N

$$\mathbf{G}_p^{\dagger}[\lambda_{i_1}^{(p_1)}\lambda_{i_2}^{(p_2)}\cdots\lambda_{i_N}^{(p_N)};w(n)] = \mathbf{G}_{p_1}^{\dagger}[\lambda_{i_1}^{(p_1)};w(n)]\mathbf{G}_{p_2}^{\dagger}[\lambda_{i_2}^{(p_2)};w(n)]\cdots\mathbf{G}_{p_N}^{\dagger}[\lambda_{i_N}^{(p_N)};w(n)] \tag{3.62}$$

The derived kernels of the G-functional are related to the leading kernel by (3.14)-(3.17) and the functions λ_i are orthogonal i.e.

$$\sum_{k=0}^{\infty}\lambda_i(k)\lambda_j(k) = \delta_{ij} \tag{3.63}$$

In the above example, the derived homogeneous functionals are obtained by reducing the order of the kernel using (3.17), letting two time arguments be equal and summing over their value in the range $[0,\infty)$. Since the symmetry of the kernel includes terms of all possible permutations of the time arguments, we have three cases

- Summing the product $\lambda_{i_1}(k)\lambda_{i_1}(k)$ which equals one; therefore the exponent p_1 is reduced to $p_1 - 2$ while p_2 remains the same.
- Summing the product $\lambda_{i_2}(k)\lambda_{i_2}(k)$ which equals one; therefore the exponent p_2 is reduced to $p_2 - 2$ while p_1 remains the same.
- Summing the product $\lambda_{i_1}(k)\lambda_{i_2}(k)$ which equals zero from the orthogonality property.

Thus the required functionals are obtained by just reducing the exponents two at a time and permuting all of these reductions among the λ's. In this way we arrive at the results mentioned in (3.26) for the G-functional in which $\mathcal{G}_p[g_p; w(n)]$ is represented by sums of products of terms $y_i(n)$ as given by (3.27).

3. Realization of Section 1 of the Wiener Model

Having shown that the Wiener G-functionals can indeed be expanded in terms of the discrete Laguerre functions. let us return to the form of these functions and the corresponding realization of section 1 in Fig. 3.1. Equation (3.52) can be put in the form

$$\Lambda_i(z) = \sqrt{1-\rho^2}\frac{z}{(z-\rho)}\left[\frac{\rho(z-\rho^{-1})}{(z-\rho)}\right]^i \tag{3.64}$$

or

$$\Lambda_i(z) = \Lambda_0(z)\left[\frac{\rho(z-\rho^{-1})}{(z-\rho)}\right]^i \tag{3.65}$$

for which

$$\Lambda_0(z) = \sqrt{1-\rho^2}\frac{z}{(z-\rho)} \tag{3.66}$$

Since for $|\rho| < 1$ the term in brackets is the transfer function of an allpass causal stable linear filter

$$H_{ap}(z) = \frac{\rho(z-\rho^{-1})}{(z-\rho)} = \frac{\rho - z^{-1}}{1-\rho z^{-1}} \tag{3.67}$$

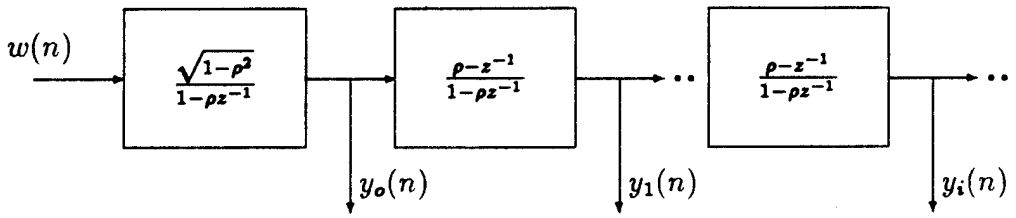

Figure 3.4: Realization of section 1 of the model.

it follows that $\Lambda_i(z)$ can be written as

$$\Lambda_i(z) = \Lambda_{i-1}(z)H_{ap}(z) \tag{3.68}$$

or

$$\Lambda_i(z) = \Lambda_0(z)H_{ap}^i(z) \tag{3.69}$$

Since section 1 of the model required a bank of the $\lambda_i(n)$ this means that section 1 can be realized by a linear stable causal system with a single pole at ρ followed by a train of cascaded allpass filters each with a transfer function $H_{ap}(z)$ as shown in Fig. 3.4. The frequency response of the leading Laguerre system $\Lambda_0(e^{j\omega})$ is shown in Fig. 3.5 for $\rho = 0.5$, while the frequency response of the allpass filter $H_{ap}(e^{j\omega})$ is shown in Fig. 3.6. The impulse response of the allpass filter is given by

$$h_{ap}(n) = \rho\delta(n) - (1-\rho^2)\rho^{(n-1)}u(n-1) \tag{3.70}$$

Beginning with

$$\lambda_0(k) = \sqrt{1-\rho^2}\rho^k u(k) \tag{3.71}$$

the remaining functions can be computed by the convolution form

$$\lambda_i(n) = \lambda_{i-1}(n) * h_{ap}(n) \tag{3.72}$$

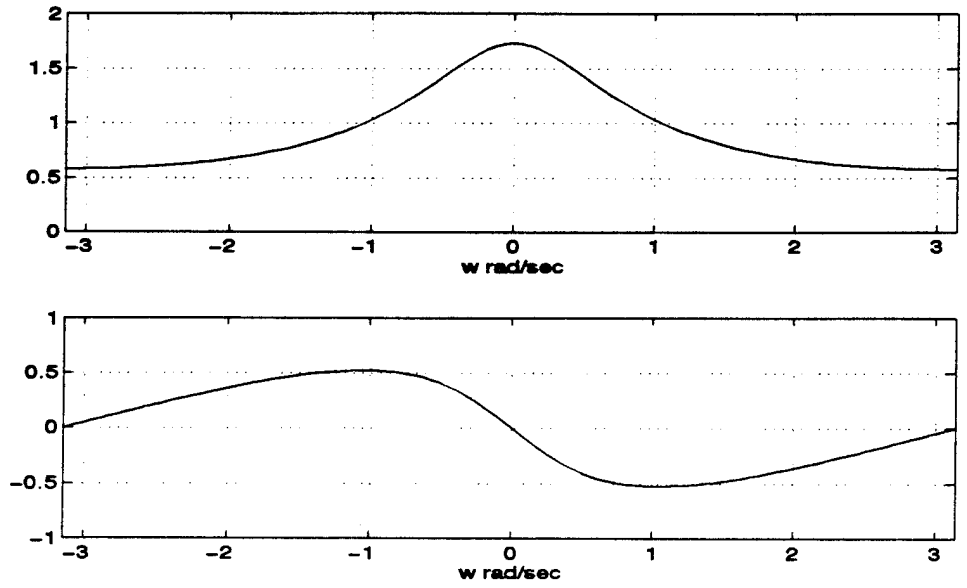

Figure 3.5: The leading Laguerre function $\lambda_0(\epsilon^{j\omega})$ for $\rho = 0.5$.

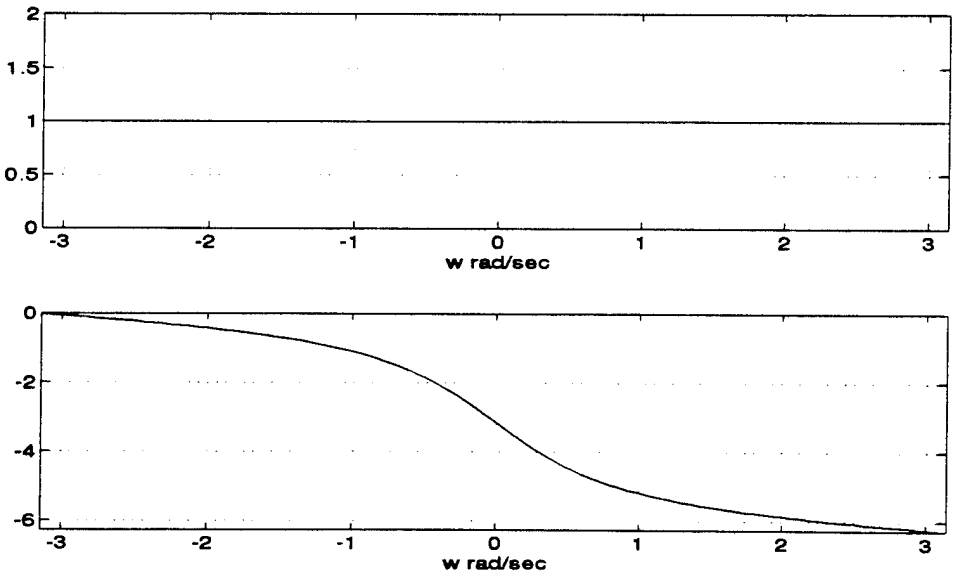

Figure 3.6: The allpass filter transfer function $H_{ap}(\epsilon^{j\omega})$ for $\rho = 0.5$.

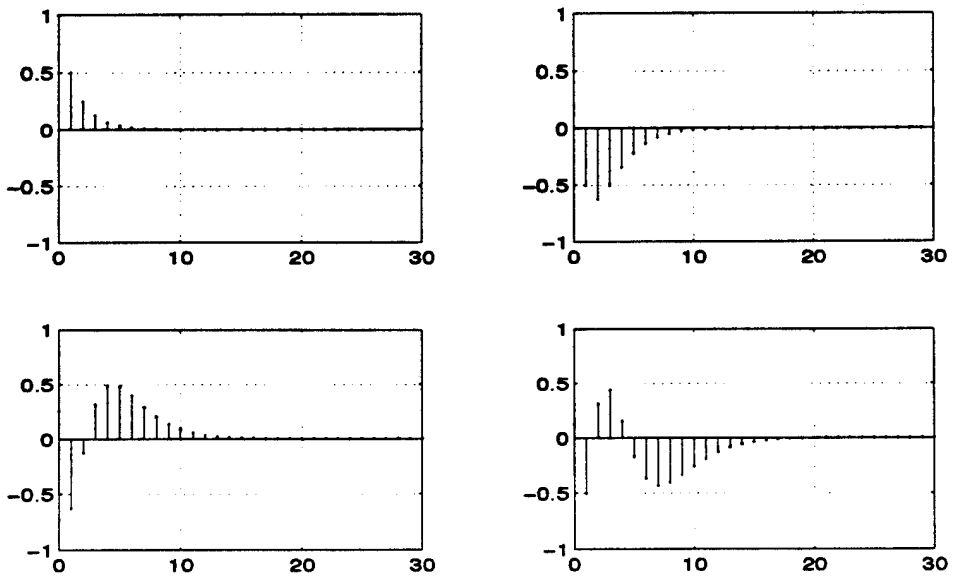

Figure 3.7: Example of the discrete Laguerre functions for $\rho = 0.5$.

or by the recurrence relation

$$\lambda_i(n) = \rho\lambda_i(n-1) + \rho\lambda_{i-1}(n) - \lambda_{i-1}(n-1) \tag{3.73}$$

An example of four discrete Laguerre functions for $\rho = 0.5$ is shown in Fig. 3.7. Note how the energy in the sequence tends to move away further from the origin as the order increases. Thus when these functions are convolved with an input waveform the higher order Laguerre filters tend to bring in more of the "past history" of the waveform.

D. CROSS-CORRELATION FUNCTIONS AND CROSS-SPECTRA OF THE OUTPUTS OF THE LAGUERRE SYSTEMS

An important part of this dissertation is to derive expressions for the higher order statistics (moments, cumulants, polyspectra) of the output $x(n)$ of the Wiener model. Given the structure of the model, these higher order statistics will involve cross-moments between the outputs of the linear systems comprising section 1. Since

these outputs are Gaussian and have zero mean, it is sufficient to describe their cross-correlation. It will be seen that because of the considerable structure brought to the problem by the use of the discrete Laguerre functions, these cross-correlation functions inherit a number of interesting properties. This in turn provides a basis for the organized computation of the Wiener model's higher-order statistics.

1. The Laguerre Cross-Correlation Sequences and Their Transforms

The cross-correlation function between the outputs y_{i_1} and y_{i_2} of the linear section is given by

$$\begin{aligned} R_{y_{i_1},y_{i_2}}(l) &= \mathcal{E}\{y_{i_1}(n)y_{i_2}(n+l)\} \\ &= \mathcal{E}\left\{\sum_{k_1=0}^{\infty}\lambda_{i_1}(k_1)w(n-k_1)\sum_{k_2=0}^{\infty}\lambda_{i_2}(k_2)w(n+l-k_2)\right\} \\ &= \sum_{k_1=0}^{\infty}\sum_{k_2=0}^{\infty}\lambda_{i_1}(k_1)\lambda_{i_2}(k_2)\mathcal{E}\{w(n-k_1)w(n+l-k_2)\} \end{aligned} \tag{3.74}$$

since $w(n)$ is white Gaussian with variance σ_0^2

$$\mathcal{E}\{w(n-k_1)w(n+l-k_2)\} = \sigma_0^2\delta(l+k_1-k_2) \tag{3.75}$$

and (3.74) becomes

$$\begin{aligned} R_{y_{i_1},y_{i_2}}(l) &= \sigma_0^2\sum_{k=0}^{\infty}\lambda_{i_1}(k)\lambda_{i_2}(l+k) \\ &= \sigma_0^2 r_{i_1,i_2}(l) \end{aligned} \tag{3.76}$$

where

$$r_{i_1,i_2}(l) = \sum_{k=0}^{\infty}\lambda_{i_1}(k)\lambda_{i_2}(l+k) \tag{3.77}$$

The quantity $r_{i_1,i_2}(l)$ will be called the *Laguerre cross-correlation sequence*. From well-known properties of the z-transform it can also be expressed as

$$r_{i_1,i_2}(l) = \frac{1}{2\pi j}\oint_{C1} z^{l-1}\Lambda_{i_2}(z)\Lambda_{i_1}(z^{-1})dz \tag{3.78}$$

The complex cross-spectral density function $S_{y_{i_1},y_{i_2}}(z)$ is the z-transform of the cross-correlation function $R_{y_{i_1},y_{i_2}}(l)$ and is then given by

$$S_{y_{i_1},y_{i_2}}(z) = \sigma_o^2 s_{i_1,i_2}(z) \tag{3.79}$$

where $s_{i_1,i_2}(z)$ is the z-transform of the Laguerre cross-correlation sequence

$$s_{i_1,i_2}(z) = \sum_{l=-\infty}^{\infty} r_{i_1,i_2}(l) z^{-l} \tag{3.80}$$

with

$$r_{i_1,i_2}(l) = \frac{1}{2\pi j} \oint_{C1} z^{l-1} s_{i_1,i_2}(z) dz \tag{3.81}$$

The quantity $s_{i_1,i_2}(z)$ will be referred to as the *Laguerre cross-spectral density function*. Then by comparing (3.78) and(3.80) and substituting (3.65) we have

$$\begin{aligned} s_{i_1,i_2}(z) &= \Lambda_{i_2}(z)\Lambda_{i_1}(z^{-1}) \\ &= (1-\rho^2)\frac{z}{(z-\rho)}\left[\frac{\rho(z-\rho^{-1})}{(z-\rho)}\right]^{i_2} \frac{z^{-1}}{(z^{-1}-\rho)}\left[\frac{\rho(z^{-1}-\rho^{-1})}{(z^{-1}-\rho)}\right]^{i_1} \\ &= -\frac{(1-\rho^2)}{\rho}\frac{z}{(z-\rho)(z-\rho^{-1})}\left[\frac{\rho(z-\rho^{-1})}{(z-\rho)}\right]^{i_2-i_1} \end{aligned} \tag{3.82}$$

This function has the interesting property that it depends only on the *difference* between the orders of the two Laguerre linear system outputs and not on the actual values of the orders. If the order difference is denoted by

$$d = i_2 - i_1 \tag{3.83}$$

the Laguerre cross-spectral density function is

$$s_d(z) = -\frac{(1-\rho^2)}{\rho}\frac{z}{(z-\rho)(z-\rho^{-1})}\left[\frac{\rho(z-\rho^{-1})}{(z-\rho)}\right]^{d} \tag{3.84}$$

and the complex cross-spectral density function of the two outputs is given by

$$S_{y_{i_1},y_{i_2}}(z) = \sigma_0^2 s_{i_2-i_1}(z) \tag{3.85}$$

By the same considerations, the Laguerre cross-correlation sequence is a function only of the difference $d = i_2 - i_1$ and can be denoted by

$$r_{i_1 i_2}(l) = r_{i_2 - i_1}(l) = r_d(l) \tag{3.86}$$

Therefore (3.76) can also be written as

$$R_{y_{i_1}, y_{i_2}}(l) = \sigma_0^2 r_{i_2 - i_1}(l) \tag{3.87}$$

Since the Laguerre cross-correlation sequences and their transforms depend only on the order difference d, it is interesting to examine their behavior for different values of this parameter. The leading cross-spectral density function for which $d = 0$ is the complex power spectral density function of the output of *any* of the Laguerre systems and is given by

$$S_y(z) = \sigma_0^2 s_0(z) = \frac{\sigma_0^2(1 - \rho^2)}{(z - \rho)(z^{-1} - \rho)} \tag{3.88}$$

The corresponding autocorrelation function is

$$R_y(l) = \sigma_0^2 r_0(l) = \sigma_0^2 \rho^{|l|} \tag{3.89}$$

Further (3.84) indicates that

$$\begin{aligned} s_d(z) &= s_{d-1}(z) H_{ap}(z) \\ &= s_0(z) H_{ap}^d(z) \end{aligned} \tag{3.90}$$

Fig. 3.8 shows a plot of the sequence $r_0(l)$ for $\rho = 0.5$ and the corresponding power spectral density function $S_0(e^{j\omega})$, which is $S_0(z)$ evaluated on the unit circle. Starting with $r_0(l)$ the Laguerre cross-correlation sequence can be derived by the convolution form

$$r_d(l) = r_{d-1}(l) * h_{ap}(l) \tag{3.91}$$

or by the recurrence relation

$$r_d(l) = \rho r_d(l - 1) + \rho r_{d-1}(l) - r_{d-1}(l - 1) \tag{3.92}$$

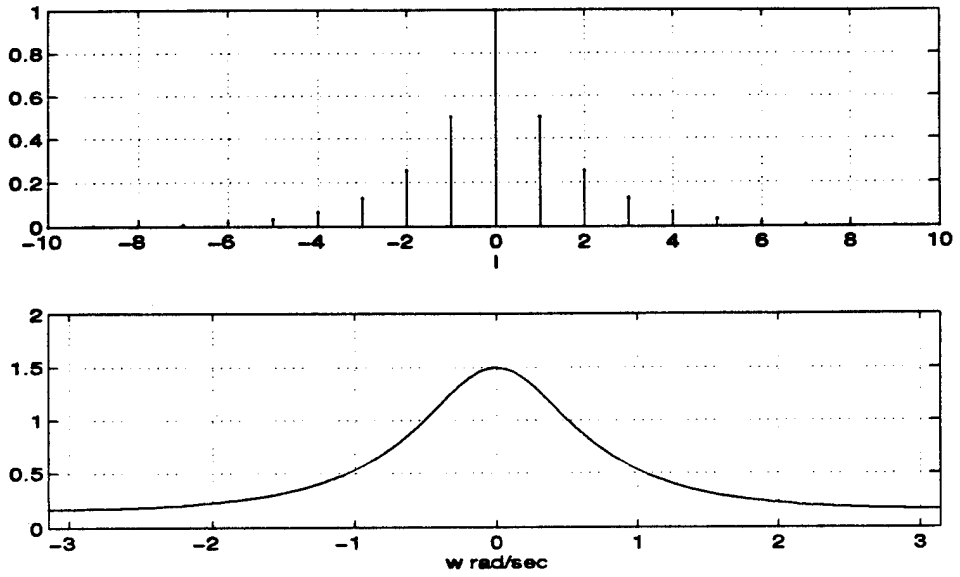

Figure 3.8: The Laguerre autocorrelation sequence $r_0(l)$ and power spectral function $s_0(e^{jw})$ for $\rho = 0.5$.

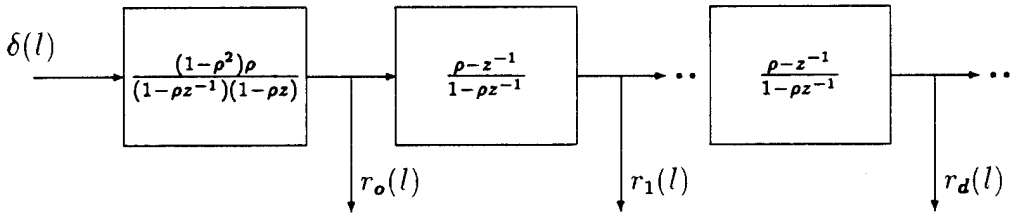

Figure 3.9: The generation of the Laguerre cross-correlation function.

Both of these relations follow from considerations similar to those in section C. Further, the sequence can be thought of as the "impulse response" of the "system" shown in Fig. 3.9. This picture does not represent a real system, but only a convenient way to show the relation between the Laguerre cross-correlstion sequences.

By substituting (3.84) in (3.81), $r_d(l)$ has the representation

$$\begin{aligned} r_d(l) &= \frac{1}{2\pi j}\oint_{C1} z^{l-1} s_d(z)dz \\ &= -(1-\rho^2)\rho^{d-1}\frac{1}{2\pi j}\oint_{C1}\frac{z^l(z-\rho^{-1})^{d-1}}{(z-\rho)^{d+1}}dz \end{aligned} \tag{3.93}$$

with the following characteristics:

1. For $d > 0$ and $l = 0$

$$r_d(0) = -(1-\rho^2)\rho^{d-1}\frac{1}{d!}\frac{d^d}{dz^d}(z-\rho)^{d-1} = 0 \tag{3.94}$$

(This is a restatement of the fact that the $y_i(n)$ are orthogonal.)

2. For $d > 0$ and $l < 0$ a change of variables $z = v^{-1}$ can be made in (3.93). Then

$$r_d(l) = -(1-\rho^2)\rho^{-(d+1)}\frac{1}{2\pi j}\oint_{C1}\frac{v^{-l}(v-\rho)^{d-1}}{(v-\rho^{-1})^{d+1}}dv = 0 \tag{3.95}$$

since the integrand has no poles inside the unit circle. In other words, for $d > 0$ the $r_d(l)$ are *strictly right-sided.*

3. For $d > 0$ and $l > 0$ (3.93) can be evaluated as

$$\begin{aligned} r_d(l) &= -(1-\rho^2)\rho^{d-1}\frac{1}{d!}\frac{d^d}{dz^d}[z^l(z-\rho^{-1})^{d-1}] \\ &= (1-\rho^2)\frac{\rho^{d+l}}{d}\sum_{m=1}^{\min[l,d]}\frac{(-1)^m(d+l-m)!}{(m-1)!(d-m)!(l-m)!}\rho^{-2m} \end{aligned} \tag{3.96}$$

which provides an explicit expression for $r_d(l)$.

We also note here that negative values of d need not be considered explicitly since from the definition (3.77) with $d = i_1 - i_2$

$$r_{-d}(l) = r_d(-l)$$

Characteristics 1 and 2 above imply that the cross-correlation function (not the auto-correlation) of the outputs of the Laguerre systems is zero if the output of a higher Laguerre order system is averaged with the output of a lower Laguerre order system at an equal or greater time lag. In other words. the output of a Laguerre system is orthogonal to (uncorrelated with) the future values of the output of the lower order Laguerre system. An example of the first four Laguerre cross-correlation sequences is shown in Fig. 3.10.

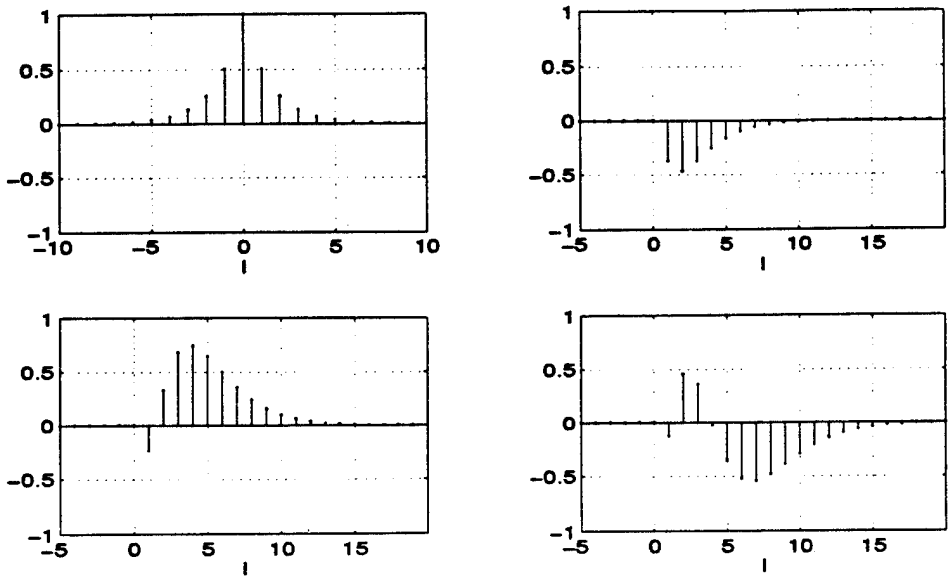

Figure 3.10: Laguerre cross-correlation functions for $\rho = 0.5$.

2. The Effect of the Parameter ρ on the System Characteristic

The above derivations of the transfer functions and cross-spectrum (impulse responses and cross-correlation) were carried out for real values of the system pole location ρ. The impulse response of any stable causal system can be expanded in terms of the complete orthonormal set of discrete Laguerre functions. The expansion generally requires an infinite number of these functions. When an impulse response is approximated as a finite weighted sum of Laguerre functions however, the error in this expansion is affected by the choice of the parameter ρ (pole location in the complex z-plane). Stated another way, for a given mean square approximation error, the minimum number of terms in the expansion is reduced when ρ is appropriately chosen. For positive real ρ the linear system has a lowpass characteristic. If the system is required to have a highpass characteristic the best choice of ρ is on the real negative axis. For a system with a bandpass characteristic the pole is best chosen as a complex value. This leads to a system with a complex output. However.

combining the response of two systems with parameters ρ and ρ^* (ρ conjugate) as shown in Fig. 3.11 produces a real output process with essentially the same statistical characteristics. This follows because conjugating ρ merely results in conjugation of the functionals $\mathcal{G}_p$ in the Wiener representation (see (3.20)). To obtain the real-

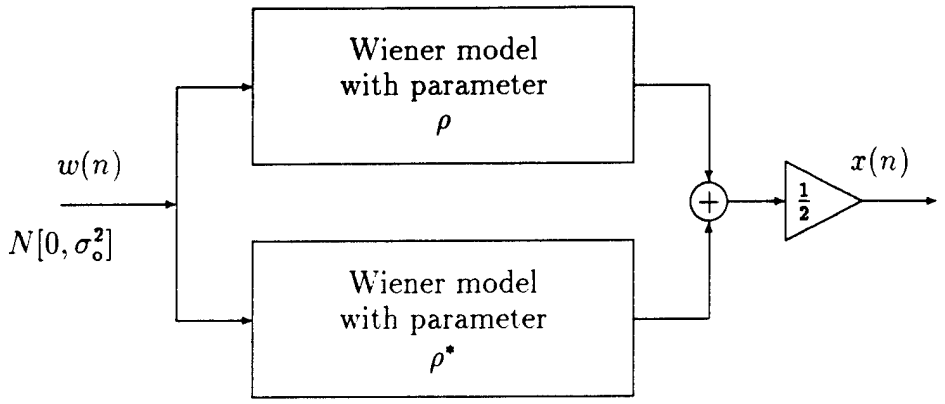

Figure 3.11: The nonlinear model for bandpass processes

valued output shown in Fig. 3.11 we can equivalently work with the top system and take the real part of its output. When a complex value of ρ is used, (3.66),(3.67) and (3.84) need to be modified slightly.

1. The leading Laguerre function becomes:

$$\Lambda_0(z) = \sqrt{1-|\rho|^2}\frac{z}{(z-\rho)} \tag{3.97}$$

2. The allpass filter transfer function becomes:

$$H_{ap}(z) = \frac{\rho^*(z-\rho^{*-1})}{(z-\rho)} = \frac{\rho^* - z^{-1}}{1-\rho z^{-1}} \tag{3.98}$$

3. The leading Laguerre cross-spectrum becomes:

$$s_0(z) = \frac{(1-|\rho|^2)}{(z-\rho)(z^{-1}-\rho^*)} \tag{3.99}$$

These have the frequency characteristics similar to those for real positive ρ except for a shift on the frequency axis equal to $\angle\rho$.

IV. CUMULANTS AND POLYSPECTRA OF THE WIENER MODEL OUTPUT

As has been mentioned earlier, our interest in the Wiener model is to represent a class of discrete random processes. The output of the Wiener model has been shown to be represented as the weighted sum of multinomials (3.34) known as the Q-polynomials. These multinomials are denoted by $\mathbf{Q}_{\boldsymbol{\alpha}_i}$ and each was shown in (3.33) to be formed as a specific product of the outputs of the memoryless nonlinear blocks of section 3 of the model. Since each of these outputs is represented by a Hermite polynomial formed from the output of a linear system driven by a zero-mean white Gaussian process, the Q-polynomial is a sum of functionals of a white Gaussian process similar to those discussed in Chapter II. Thus the procedures developed in that chapter to calculate the cross-moments and the cross-spectral functions for this type of functionals can be applied here to develop cumulants and polyspectra of processes represented by the Wiener model. Moreover, since the impulse responses of the linear filters in section 1 are the discrete Laguerre functions, the cross-correlation functions and cross-spectral functions of the outputs of the linear filters were shown to have special properties which result in a structured method of representing the required cumulants and polyspectra of the modeled process.

In the following section we present a detailed view of the Wiener model and define the Q-polynomials for a model of specific dimensions. In the next sections we formulate the process cumulants and develop an organized procedure for their evaluation. In the last section that we formulate expressions for the process polyspectra and develop a procedure for their evaluation.

A. THE DETAILED WIENER MODEL

We consider here a finite version of the Wiener model for discrete nonlinear systems presented in Chapter III defined by the two model dimensions N_L and N_N. The integer N_L is the highest order of discrete Laguerre functions representing the impulse responses of the linear filters comprising section 1 of the model; that is, section 1 consists of $N_L + 1$ filters with impulse responses given by the discrete Laguerre functions $\lambda_0(n), \lambda_1(n), \cdots, \lambda_{N_L}(n)$. The integer N_N is the highest degree of nonlinearity in each memoryless nonlinear block comprising section 2 of the model. N_N is therefore the order of the highest order Hermite polynomial in these blocks, which is the highest degree of nonlinearity in this representation. The complete structure of this model is shown in Fig. 4.1 for $N_L = N_N = 2$. In general if a system is used to model a process with nonlinearity degree p, the lower degrees of nonlinearity are also included in the representation. This follows since the system is represented using the Wiener G-functionals $\mathcal{G}_0, \mathcal{G}_1, \cdots, \mathcal{G}_{p-1}, \mathcal{G}_p$. The following rules are therefore necessary to define the Q-polynomials.

1. The kernels of the G-functionals are symmetric with respect to their time arguments. Therefore the expansion of the leading kernel in terms of a product of discrete Laguerre functions must include all the possible permutations of products of these functions with total sum of exponents equal to the order of the kernel. This means that the G-functional is a sum of terms each formed by one of the possible permutations of products of Hermite polynomials whose sum of orders is equal to the functional order. Therefore in the model representation *the p^{th} degree nonlinearity is represented by all the possible permutations of products of the outputs of section 2 that result in multinomials of degree p.*

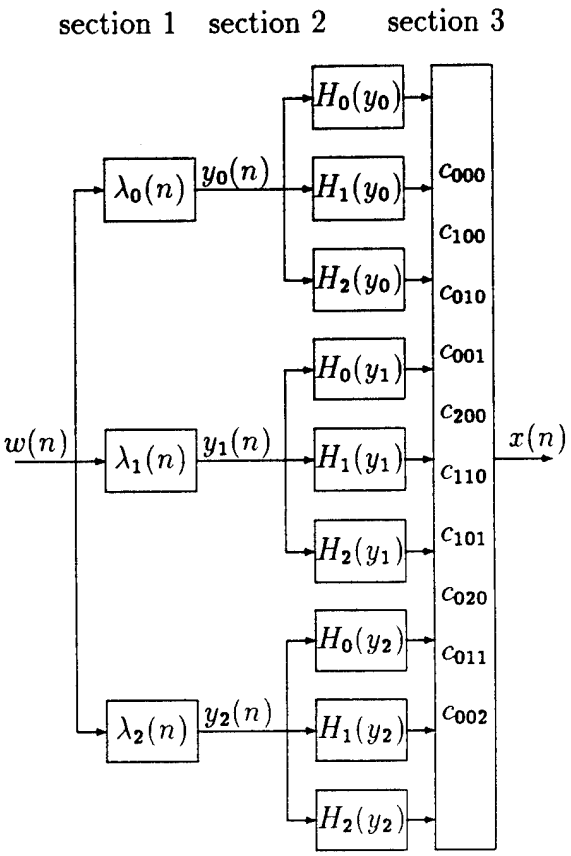

Figure 4.1: The detailed Wiener model for $N_L = 2$ and $N_N = 2$.

2. The relation between the derived kernels and the leading kernel of the G-functional implies that when a Hermite polynomial of one of the linear section outputs is included in the product, no other polynomials of the same output are included. Therefore *in each permutation of the products of the outputs of section 2, one and only one output is taken from each nonlinear block at a time.*

3. If the nonlinearity of the process is required to be of degree N_N, the expression for the nonlinear system output does not include functionals of order greater than N_N. This means that *the sum of the orders of Hermite polynomials in each term does not exceed* N_N.

Recall from Chapter III that the Q-polynomials, which replace the G-functional representation of the output of the model, are formed as follows:

1. Each polynomial is assigned a vector of indices

$$\boldsymbol{\alpha_i} = [\alpha_{i0}, \alpha_{i1}, \cdots, \alpha_{iN_L}]$$

The value of each component $\alpha_{ij} \in [0, 1, 2, \cdots, N_N]$ indicates the order of the Hermite polynomial of the corresponding output $y_j(n)$ that is included in the product, i.e.,

$$\begin{aligned} \mathbf{Q}_{\boldsymbol{\alpha_i}}(n) &= \mathbf{Q}_{\alpha_{i0}\alpha_{i1}\cdots\alpha_{iN_L}}(n) \\ &= \mathbf{H}_{\alpha_{i0}}(y_0(n))\mathbf{H}_{\alpha_{i1}}(y_1(n))\cdots\mathbf{H}_{\alpha_{iN_L}}(y_{N_L}(n)) \end{aligned} \tag{4.1}$$

2. The set of Q-polynomials are weighted and summed over all the possible values of the vectors $\boldsymbol{\alpha_i}$. To keep the symmetry property of the functional kernels, each α_{ij} assumes the values $0, 1, 2, \cdots, N_N$ for same vector $\boldsymbol{\alpha_i}$ in the set. The maximum degree of nonlinearity is maintained by keeping the sum of the components of $\boldsymbol{\alpha_i}$ less than or equal to N_N. Hence the Q-polynomial representation of the output for specific dimensions N_L and N_N is

$$x(n) = \sum_{\Sigma[\boldsymbol{\alpha_i}]=0}^{N_N} c_{\boldsymbol{\alpha_i}} \mathbf{Q}_{\boldsymbol{\alpha_i}}(n) \tag{4.2}$$

The quantity $\Sigma[\boldsymbol{\alpha_i}]$ (see (2.90)) is the order of the functional represented by the Q-polynomial $\mathbf{Q}_{\boldsymbol{\alpha_i}}(n)$. When $\Sigma[\boldsymbol{\alpha_i}] = 0$ we have the multinomial $\mathbf{Q_0}$ which is the product of all $\mathbf{H_0}(y_i(n))$. Since all $\mathbf{H_0}(y_i(n))$ are equal to 1 it follows that $\mathbf{Q_0}$ is also 1. This term represents the zero-order homogeneous functional, which is a constant value. The weighted sum of Q-polynomials for which $\Sigma[\boldsymbol{\alpha_i}] = 1$ represents the linear portion of the representation and is equal to the non-homogeneous functional of first order $\mathcal{G}_1[w(n)]$. Similarly the Q-polynomials for which $\Sigma[\boldsymbol{\alpha_i}] = 2$ represent

the quadratic part $\mathcal{G}_2[w(n)]$, and so on. In our development of the cumulant and polyspectra expressions we will need to define a procedure for ordering the indices of the polynomials. We choose the following rules of ordering because they result in an organized structure of the arrays of the values of cross-moments and cross-spectra of the Q-polynomials required in our formulation. The indices are ordered using the vector notation

$$\boldsymbol{\alpha}_i > \boldsymbol{\alpha}_j \tag{4.3}$$

with the following interpretation:

1. $\boldsymbol{\alpha}_i > \boldsymbol{\alpha}_j$ if $\Sigma[\boldsymbol{\alpha}_i] > \Sigma[\boldsymbol{\alpha}_j]$. In other words, the former is greater than the latter if it represents a functional of order greater than that represented by the latter.

2. If $\Sigma[\boldsymbol{\alpha}_i] = \Sigma[\boldsymbol{\alpha}_j]$, then $\boldsymbol{\alpha}_i > \boldsymbol{\alpha}_j$ if the partial sums (below) satisfy

$$\sum_{k=m}^{N_L} \alpha_{ik} > \sum_{k=m}^{N_L} \alpha_{jk} \qquad \exists m \in [0, 1, 2, \cdots, N_L] \tag{4.4}$$

In fact (4.4) defines the entire ordering since it includes the case $m = 0$. This means that if the Laguerre systems are divided into two groups, one vector of indices is considered greater than the other one if the group of the higher order Laguerre systems is represented by Hermite polynomials that have a greater sum of orders. For example if $N_L = 2$ and $N_N = 2$ the output $x(n)$ is comprised of the weighted sum of ten Q-polynomials that have indices $\boldsymbol{\alpha}_i$ with the following descending order:

$$\begin{aligned}
&002 > 011 > 020 > 101 > 110 > 200 \quad \text{second order (quadratic) terms} \\
&001 > 010 > 100 \quad \text{first order (linear) terms} \\
&000 \quad \text{zeroth order (constant) term}
\end{aligned} \tag{4.5}$$

N_L/N_N	1	2	3	4
1	3	6	10	15
2	4	10	20	35
3	5	15	35	70
4	6	21	56	126
5	7	28	84	210

TABLE 4.1: Number of section 3 parameters $c_{\boldsymbol{\alpha}_i}$ for a set of model dimensions.

In this example $x(n)$ can be represented as

$$x(n) = \sum_{\boldsymbol{\alpha}_i=000}^{002} c_{\boldsymbol{\alpha}_i} \mathbf{Q}_{\boldsymbol{\alpha}_i} \tag{4.6}$$

where the coefficients $c_{\boldsymbol{\alpha}_i}$ are the parameters that define section 3 of the model. The number of coefficients N depends only on the value of the model dimensions N_L and N_N and increases with both dimensions, as shown in Table 4.1.

B. CUMULANTS OF THE OUTPUT OF THE WIENER MODEL

If the output of the Wiener model is represented as a sum of Q-polynomials (see (3.34)) then the k^{th} order output cumulant can be written as

$$\begin{aligned} C_x^{(k)}(l_1, l_2, \cdots, l_{k-1}) &= \sum_{\boldsymbol{\alpha}_1} \sum_{\boldsymbol{\alpha}_2} \cdots \sum_{\boldsymbol{\alpha}_k} c_{\boldsymbol{\alpha}_1} c_{\boldsymbol{\alpha}_2} \cdots c_{\boldsymbol{\alpha}_k} \\ &\quad \times \text{cum}(\mathbf{Q}_{\boldsymbol{\alpha}_1}(n), \mathbf{Q}_{\boldsymbol{\alpha}_2}(n+l_1), \cdots, \mathbf{Q}_{\boldsymbol{\alpha}_k}(n+l_{k-1})) \end{aligned} \tag{4.7}$$

To be notationally correct we should denote the $\boldsymbol{\alpha}$'s by $\boldsymbol{\alpha}_{i_1}, \cdots, \boldsymbol{\alpha}_{i_k}$ but this notation gets extremely cumbersome in the development that follows. Therefore $\boldsymbol{\alpha}_1$ for example should be interpreted as a generic index vector not simply the first one in

the set, and the sum over $\boldsymbol{\alpha}_1$ is the sum over all such index vectors. We can write (4.7) more simply as

$$C_{\boldsymbol{x}}^{(k)}(\boldsymbol{l}) = \sum_{\boldsymbol{\alpha}} c_{\boldsymbol{\alpha}} C_{\boldsymbol{\alpha}}^{(k)}(\boldsymbol{l}) \tag{4.8}$$

where

$$C_{\boldsymbol{\alpha}}^{(k)}(\boldsymbol{l}) = \mathrm{cum}\left(\mathbf{Q}_{\boldsymbol{\alpha}_1}(n), \mathbf{Q}_{\boldsymbol{\alpha}_2}(n+l_1), \cdots, \mathbf{Q}_{\boldsymbol{\alpha}_k}(n+l_{k-1})\right)$$

and $\boldsymbol{\alpha}$ is a concatination of the vectors $\boldsymbol{\alpha}_i$, that is

$$\boldsymbol{\alpha} = [\boldsymbol{\alpha}_1^T, \boldsymbol{\alpha}_2^T, \cdots, \boldsymbol{\alpha}_k^T]^T \tag{4.9}$$

while $c_{\boldsymbol{\alpha}}$ is the product of the coefficients

$$c_{\boldsymbol{\alpha}} = c_{\boldsymbol{\alpha}_1} c_{\boldsymbol{\alpha}_2} \cdots c_{\boldsymbol{\alpha}_k} \tag{4.10}$$

and the sum in (4.8) is over all the possible permutation of the indices.

The number of terms in (4.7) would seem to be equal to the number of the coefficients $c_{\boldsymbol{\alpha}}$ raised to the power k, which is a huge number even for low dimensional models, and increases exponentially as the order of the cumulant increases. Fortunately this is not the case because the cumulant $C_{\boldsymbol{\alpha}}^{(k)}(\boldsymbol{l})$ of Q-polynomials has properties that make its value identically zero for many well-defined combinations of the indices. In the following, we define the combinations of indices for which the terms in (4.7) are *not* equal to zero. Moreover from the combination of indices we develop a formula to compute the value of the cumulant.

To obtain an expression for the k^{th}-order output cumulant of $C_{\boldsymbol{x}}^{(k)}(\boldsymbol{l})$ we need to develop an expression for the k^{th}-order cross-cumulant of the Q-polynomials $C_{\boldsymbol{\alpha}}^{(k)}(\boldsymbol{l})$. For these cumulants we need to compute the expectation of the product of the Q-polynomials (moments) and substitute in the moment-cumulant relations described in Chapter II. To develop the expression for these expectations we expand the Q-polynomials as the product of Hermite polynomials. These are in turn functions of

the Gaussian outputs y_i of the linear filters. The values of the components of the vectors of indices $\boldsymbol{\alpha}$ together with the time lag arguments of the Q-polynomials can be used to determine if in a product of Laguerre cross-correlation sequences there exists at least one which is zero. This principle is used to construct a set of sufficient conditions on the vectors of indices such that the cumulant of the Q-polynomials is not identically zero. Then we can apply the procedure developed in Chapter III to find the desired expression. The analysis is presented in three subsections. First we form a general expression for the moments of Q-polynomials. Next we show by example how we simplify the general expression and eliminate most of the terms. Finally we present the general efficient procedure for computation of the cumulants.

1. General Expression for Moments of Q-Polynomials

For a model of dimensions N_L and N_N the cross-moment of the Q-polynomials can be expressed as

$$\begin{aligned}\mu^{(k)}_{\boldsymbol{\alpha}}(\boldsymbol{l}) &= \mathcal{E}\left\{\mathbf{Q}_{\boldsymbol{\alpha}_1}(n)\mathbf{Q}_{\boldsymbol{\alpha}_2}(n+l_1)\cdots\mathbf{Q}_{\boldsymbol{\alpha}_k}(n+l_{k-1})\right\} \\ &= \sum_{\mathbf{m}=\mathbf{0}}^{\lfloor\frac{\boldsymbol{\alpha}}{2}\rfloor} \frac{(-1)^{\Sigma[\mathbf{m}]}\boldsymbol{\alpha}!}{\mathbf{m}!(\boldsymbol{\alpha}-2\mathbf{m})!}\left(\frac{\sigma_o^2}{2}\right)^{\Sigma[\mathbf{m}]}\eta^{(k)}_{\boldsymbol{\alpha}-2\mathbf{m}}(\boldsymbol{l})\end{aligned} \tag{4.11}$$

where $\eta^{(k)}_{\boldsymbol{\alpha}-2\mathbf{m}}(\boldsymbol{l})$ is defined as the moment

$$\eta^{(k)}_{\boldsymbol{\alpha}-2\mathbf{m}}(\boldsymbol{l}) = \mathcal{E}\left\{\mathbf{Y}^{\boldsymbol{\alpha}_1-2\mathbf{m}_1}(n)\mathbf{Y}^{\boldsymbol{\alpha}_2-2\mathbf{m}_2}(n+l_1)\cdots\mathbf{Y}^{\boldsymbol{\alpha}_k-2\mathbf{m}_k}(n+l_{k-1})\right\} \tag{4.12}$$

and where the following vector notation is used

$$\mathbf{m} = [\mathbf{m}_1^T, \mathbf{m}_2^T, \cdots, \mathbf{m}_k^T]^T$$

$$\mathbf{m}_i = [m_{i0}, m_{i1}, \cdots, m_{iN_L}^T]^T$$

$$\sum_{\mathbf{m}=\mathbf{0}}^{\lfloor \frac{\boldsymbol{\alpha}}{2} \rfloor} = \sum_{m_{00}=0}^{\lfloor \frac{\alpha_{01}}{2} \rfloor} \sum_{m_{01}=0}^{\lfloor \frac{\alpha_{01}}{2} \rfloor} \cdots \sum_{m_{kN_L}=0}^{\lfloor \frac{\alpha_{kN_L}}{2} \rfloor}$$

$$\mathbf{Y}^{\boldsymbol{\alpha}_i - 2\mathbf{m}_i}(n) = y_0^{\alpha_{i0}-2m_{i0}}(n) y_1^{\alpha_{i1}-2m_{i1}}(n) \cdots y_{N_L}^{\alpha_{iN_L}-2m_{iN_L}}(n) \qquad (4.13)$$

We seek here to simplify this general expression, to outline its specific structure, and to show which terms need to be computed and which do not. The cross-moment of the Q-polynomials has been expanded as a weighted sum of expectations of products of Gaussian variables $\eta_{\boldsymbol{\alpha}}^{(k)}$. The weights in this summation are products of the coefficients of the Hermite polynomials. Since the sum of the exponents in each cross-moment $\eta_{\boldsymbol{\alpha}-2\mathbf{m}}$ is

$$\Sigma[\boldsymbol{\alpha} - 2\mathbf{m}] = \Sigma[\boldsymbol{\alpha}] - 2\Sigma[\mathbf{m}]$$

then the sum of exponents is even if $\Sigma[\boldsymbol{\alpha}]$ is even. Thus a necessary condition for the cross-moments of the Q-polynomials to be non-zero is that the total sum of indices of the polynomials $\Sigma[\boldsymbol{\alpha}]$ is even. The value of the cumulant in (4.8) is therefore the sum of *the terms with indices* $\boldsymbol{\alpha}$ *for which each* $\Sigma[\boldsymbol{\alpha}]$ *is even.* In this case the value of the cross-moment $\eta_{\boldsymbol{\alpha}}^{(k)}$ is obtained by following the procedure described in Chapter II. The symmetry properties of the cumulant functions of real stationary random processes are utilized in this respect to calculate the cumulants in the non-redundant region defined by

$$0 \geq l_1 \geq l_2 \cdots \geq l_{k-1} \qquad (4.14)$$

This choice ensures that the time lag arguments of the submatrices of $\mathbf{C}_y^{(m)}$ defined as $\mathbf{R}(n)$ in (2.88) are non-negative. The entries of these matrices are the cross-correlation functions of the outputs of the Laguerre filters comprising section 1 of

the model and this gives the matrices specific important structure. Recall that the Laguerre cross-correlation functions have the following properties:

1. The index d of each cross-corelation function is a single integer representing the difference between the Laguerre orders i and j of the correlated outputs $(d = j - i)$.

2. The cross-correlation functions (not the autocorrelation functions) are strictly single sided; for positive values of d the function $r_d(l)$ is non-zero only for strictly positive values of the time lag l.

It follows from the above that the diagonal $(N_L + 1) \times (N_L + 1)$ blocks denoted $\mathbf{R}(0)$ in (2.87) are the identity matrices because for $l = 0$

$$r_d(0) = \begin{cases} 1 & \text{for } d = 0 \\ 0 & \text{for } d \neq 0 \end{cases} \tag{4.15}$$

For cumulant values in the non-redundant region, the off-diagonal blocks above the main diagonal are upper triangular and Toeplitz because for $d > 0$ and $l > 0$ $r_d(-l)$ is identically zero and $r_d(l)$ in general is not zero. The blocks have the form

$$\mathbf{R}(l) = \begin{bmatrix} r_0(l) & r_1(l) & r_2(l) & \cdots & r_{N_L}(l) \\ 0 & r_0(l) & r_1(l) & \cdots & r_{N_L-1}(l) \\ 0 & 0 & r_0(l) & \cdots & r_{N_L-2}(l) \\ \vdots & \vdots & \vdots & \vdots & \\ 0 & 0 & 0 & \cdots & r_0(l) \end{bmatrix}; \quad l > 0 \tag{4.16}$$

Since the overall matrix $\mathbf{C}_y^{(k)}(\boldsymbol{l})$ is symmetric the off diagonal blocks below the main diagonal are lower triangular. Then for the case of the k^{th}-order cumulant of the output of the Wiener model the matrix of cross-correlation sequences becomes

$$\mathbf{C}_y(\boldsymbol{l}) = \begin{bmatrix} \mathbf{R}(0) & \mathbf{R}(l_1) & \mathbf{R}(l_2) & \cdots & \mathbf{R}(l_{k-1}) \\ \mathbf{R}^T(l_1) & \mathbf{R}(0) & \mathbf{R}(l_2 - l_1) & \cdots & \mathbf{R}(l_{k-1} - l_1) \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \mathbf{R}^T(l_{k-1} - l_1) & \mathbf{R}^T(l_{k-1} - l_2) & \mathbf{R}^T(l_{k-1} - l_3) & \cdots & \mathbf{R}(0) \end{bmatrix} \tag{4.17}$$

where the $\mathbf{R}(0)$ blocks, as previously noted, are the identity matrix.

The corresponding $k(N_L+1) \times k(N_L+1)$ multiplicity matrix $\mathbf{M}$ is constructed such that no permutation resulting in a zero-valued term is allowed. Since the moment expression is a sum of the products of system cross-correlation sequences, the value of any term including a zero-valued entry of $\mathbf{C}_y$ is zero. In any permutation we must keep the multiplicity of a zero-valued entry of $\mathbf{C}_y$ equal to zero. Then an entry of $\mathbf{M}$ is zero if the corresponding entry of $\mathbf{C}_y$ is zero. This means that the multiplicity matrix $\mathbf{M}$ has the same structure as $\mathbf{C}_y$, i.e.,

1. The matrix $\mathbf{M}$ is partitioned into $k \times k$ square blocks each of size $(N_L+1) \times (N_L+1)$.

2. The diagonal blocks are also diagonal containing *non-negative even-valued entries* along the diagonal.

3. The upper non-diagonal blocks are upper triangular and the matrix is symmetric.

4. The values of the entries of $\mathbf{M}$ in the last column and in the last row or along the diagonal of the lower right block are dependent on the values of the preceding entries in the columns or rows. This dependence is required to maintain the condition that the sum along any row or column is equal to the corresponding exponent $\alpha_{ij} - 2m_{ij}$. We write this more simply using the reduction notation as

$$\Sigma[\mathbf{M}] = \boldsymbol{\alpha} - 2\mathbf{m}$$

As an example, consider computing the second order cumulant of a model with dimension $N_L = 2$. The matrix of the Laguerre cross-correlation sequences has

the form

$$\mathbf{C}_{\boldsymbol{y}}(l) = \begin{bmatrix} 1 & 0 & 0 & r_0(l) & r_1(l) & r_2(l) \\ 0 & 1 & 0 & 0 & r_0(l) & r_1(l) \\ 0 & 0 & 1 & 0 & 0 & r_0(l) \\ r_0(l) & 0 & 0 & 1 & 0 & 0 \\ r_1(l) & r_0(l) & 0 & 0 & 1 & 0 \\ r_2(l) & r_1(l) & r_0(l) & 0 & 0 & 1 \end{bmatrix} \tag{4.18}$$

where the above mentioned properties are evident. The associated multiplicity matrix $\mathbf{M}$ has a similar form and must satisfy the condition

$$\underbrace{\begin{bmatrix} 2i_1 & 0 & 0 & i_2 & i_3 & i_4 \\ 0 & 2i_5 & 0 & 0 & i_6 & i_7 \\ 0 & 0 & 2i_8 & 0 & 0 & i_9 \\ i_2 & 0 & 0 & 2i_{10} & 0 & 0 \\ i_3 & i_6 & 0 & 0 & 2i_{11} & 0 \\ i_4 & i_7 & i_9 & 0 & 0 & 2i_{12} \end{bmatrix}}_{\mathbf{M}} \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} \alpha_{10} - 2m_{10} \\ \alpha_{11} - 2m_{11} \\ \alpha_{12} - 2m_{12} \\ \alpha_{20} - 2m_{20} \\ \alpha_{21} - 2m_{21} \\ \alpha_{22} - 2m_{22} \end{bmatrix} \tag{4.19}$$

Note that the vector on the right hand side represents the exponents of y_i in a specific term in the expression (4.11), and that the moment $\eta^{(2)}_{\boldsymbol{\alpha}-2\mathbf{m}}(l)$ is expressed using (2.92) as

$$\begin{aligned} &\eta^{(k)}_{\boldsymbol{\alpha}-2\mathbf{m}}(\boldsymbol{l}) = \\ &= (\sigma_o^2)^{\frac{\Sigma[\boldsymbol{\alpha}]-2\Sigma[\mathbf{m}]}{2}} (\boldsymbol{\alpha}_1 - 2\mathrm{m}_1)!(\boldsymbol{\alpha}_2 - 2\mathrm{m}_2)! \\ &\sum_{\Sigma[\mathbf{M}]=\boldsymbol{\alpha}-2\boldsymbol{m}} \frac{1}{2^{\frac{\mathrm{tr}(\mathbf{M})}{2}}} \prod_{j_1=1}^{2(2+1)} \frac{(C_{\boldsymbol{y}}(j_1,j_1))^{\frac{M(j_1,j_1)}{2}}}{(\frac{M(j_1,j_1)}{2})!} \prod_{j_2=j_1+1}^{2(2+1)} \frac{(C_{\boldsymbol{y}}(j_1,j_2))^{M(j_1,j_2)}}{M(j_1,j_2)!} \end{aligned} \tag{4.20}$$

In this example the entries of $\mathbf{M}$ are divided into two groups; the first is that of the independent entries $\{i_1, i_2, i_3, i_5, i_6, i_8\}$ and the second is that of the dependent entries $\{i_4, i_7, i_9, i_{10}, i_{11}, i_{12}\}$. The summation

$$\Sigma[\mathrm{M}] = \boldsymbol{\alpha} - 2\mathrm{m}$$

is actually a multiple summation over the non-negative integer values of the independent entries while the dependent ones are determined from condition (4.19). In

general, for a system of any size N_L, the value of the k^{th}-order cross-moment is given by

$$\eta^{(k)}_{\boldsymbol{\alpha}-2\mathbf{m}}(\boldsymbol{l}) = (\sigma_o^2)^{\frac{\Sigma[\boldsymbol{\alpha}-2\mathbf{m}]}{2}}(\boldsymbol{\alpha}_1 - 2\mathbf{m}_1)!(\boldsymbol{\alpha}_2 - 2\mathbf{m}_2)!\cdots(\boldsymbol{\alpha}_k - 2\mathbf{m}_k)!$$
$$\times \sum_{\Sigma[\mathbf{M}]=\boldsymbol{\alpha}-2\mathbf{m}} \frac{1}{2^{\frac{tr(\mathbf{M})}{2}}} \prod_{j_1=1}^{k(N_L+1)} \frac{(C_y(j_1,j_1))^{\frac{M(j_1,j_1)}{2}}}{(\frac{M(j_1,j_1)}{2})!} \prod_{j_2=j_1+1}^{k(N_L+1)} \frac{(C_y(j_1,j_2))^{M(j_1,j_2)}}{M(j_1,j_2)!} \tag{4.21}$$

Now we are ready to consider the cross-moment of the Q-polynomials which we denote by $\mu^{(k)}_{\boldsymbol{\alpha}}$. A necessary condition for this statistic to be non-zero is that $\Sigma[\boldsymbol{\alpha}]$ is an even value and the permutations are arranged so that the specific structure of the multiplicity matrix $\mathbf{M}$ is kept as described above. In this case the value of the cros-moment is calculated by substituting (4.21) in (4.11) to obtain

$$\mu^{(k)}_{\boldsymbol{\alpha}}(\boldsymbol{l})$$
$$= \sum_{\mathbf{m}_1=0}^{\lfloor\frac{\boldsymbol{\alpha}_1}{2}\rfloor} \cdots \sum_{\mathbf{m}_k=0}^{\lfloor\frac{\boldsymbol{\alpha}_k}{2}\rfloor} \frac{(-1)^{\Sigma[\mathbf{m}]}\boldsymbol{\alpha}_1!\boldsymbol{\alpha}_2!\cdots\boldsymbol{\alpha}_k!}{\mathbf{m}_1!\mathbf{m}_2!\cdots\mathbf{m}_k!(\boldsymbol{\alpha}_1 - 2\mathbf{m}_1)!(\boldsymbol{\alpha}_2 - 2\mathbf{m}_2)!\cdots(\boldsymbol{\alpha}_k - 2\mathbf{m}_k)!} \left(\frac{\sigma_o^2}{2}\right)^{\Sigma[\mathbf{m}]}$$

$$\times(\sigma_o^2)^{\frac{\Sigma[\boldsymbol{\alpha}_i-2\mathbf{m}_i]}{2}}(\boldsymbol{\alpha}_1 - 2\mathbf{m}_1)!(\boldsymbol{\alpha}_2 - 2\mathbf{m}_2)!\cdots(\boldsymbol{\alpha}_k - 2\mathbf{m}_k)!$$
$$\times \sum_{\Sigma[\mathbf{M}]=\boldsymbol{\alpha}-2\mathbf{m}} \frac{1}{2^{\frac{tr(\mathbf{M})}{2}}} \prod_{j_1=1}^{k(N_L+1)} \frac{(C_y(j_1,j_1))^{\frac{M(j_1,j_1)}{2}}}{(\frac{M(j_1,j_1)}{2})!} \prod_{j_2=j_1+1}^{k(N_L+1)} \frac{(C_y(j_1,j_2))^{M(j_1,j_2)}}{M(j_1,j_2)!} \tag{4.22}$$

Notice that since the first summations are over the terms $\mathbf{m_i}$, the dependent entries of the multiplicity matrix change to maintain the condition $\Sigma[\mathbf{M}] = \boldsymbol{\alpha} - 2\mathbf{m}$. The last equation can be simplified to

$$\mu^{(k)}_{\boldsymbol{\alpha}}(\boldsymbol{l})$$
$$= (\sigma_o^2)^{\frac{\Sigma[\boldsymbol{\alpha}]}{2}}\boldsymbol{\alpha}_1!\boldsymbol{\alpha}_2!\cdots\boldsymbol{\alpha}_k! \sum_{\mathbf{m}_1=0}^{\lfloor\frac{\boldsymbol{\alpha}_1}{2}\rfloor}\sum_{\mathbf{m}_2=0}^{\lfloor\frac{\boldsymbol{\alpha}_2}{2}\rfloor} \cdots \sum_{\mathbf{m}_k=0}^{\lfloor\frac{\boldsymbol{\alpha}_k}{2}\rfloor} \frac{(-1)^{\Sigma[\mathbf{m}]}}{\mathbf{m}_1!\mathbf{m}_2!\cdots\mathbf{m}_k!(2)^{\Sigma[\mathbf{m}]}}$$
$$\times \sum_{\Sigma[\mathbf{M}]=\boldsymbol{\alpha}-2\mathbf{m}} \frac{1}{2^{\frac{tr(\mathbf{M})}{2}}} \prod_{j_1=1}^{k(N_L+1)} \frac{(C_y(j_1,j_1))^{\frac{M_{\mathbf{m}}(j_1,j_1)}{2}}}{(\frac{M(j_1,j_1)}{2})!} \prod_{j_2=j_1+1}^{k(N_L+1)} \frac{(C_y(j_1,j_2))^{M(j_1,j_2)}}{M(j_1,j_2)!} \tag{4.23}$$

It appears that the general expression (4.23) for the cross-moment of Q-polynomials is extremely complex. However it will be shown that considerable simplifications are possible and that when moments are combined to produce cumulants, even further simplifications arise. These combine to make the overall computation quite managable.

2. Rules for Computation of Moments of Q-polynomials

The key to reducing (4.23) to a simple form is the following. Although we have specified some necessary conditions for the terms in (4.23) to be non-zero, these conditions are not sufficient. In fact, if we proceed in calculating the cumulant of the Q-polynomial according to the arrangement in (4.23) many computations are *not necessary* due to the following:

1. For a large number of configurations of the values of the first summation parameters $\mathbf{m}_i$ there is no multiplicity matrix $\mathbf{M}$ with non-negative integer entries that satisfies the the condition $\Sigma[\mathbf{M}] = \boldsymbol{\alpha} - 2\mathbf{m}$. In these cases the corresponding term is equal to zero.

2. In many other configurations there exsits a matrix $\mathbf{M}$ satisfying the above condition but the resulting terms sum to zero. Therefore the value of the moment is also zero.

This results in a tremendous simplification of the resulting moment expressions and a corresponding savings in computation. For instance, in the case of the second-order moment of Q-polynomials of our example, (4.23) reduces to the considerably simpler expression

$$\mu_{\boldsymbol{\alpha}}^{(2)}(l) = (\sigma_o^2)^{\frac{\Sigma[\boldsymbol{\alpha}]}{2}} \boldsymbol{\alpha}_1!\boldsymbol{\alpha}_2! \sum_{i_3=0}^{\alpha_{10}-\alpha_{20}} \frac{r_0^{i_2+i_6+i_9}(l) r_1^{i_3+i_7}(l) r_2^{i_4}(l)}{i_2!i_3!i_4!i_7!i_9!} \tag{4.24}$$

where the quantities i_2, i_4, i_6, i_7 and i_9 depend upon the value of i_3 as described in Appendix C.

In Appendix C we examine the conditions for which we can avoid computing the terms in the Q-polynomial given by (4.11) which are identically zero. This leads to the following principles:

1. The non-zero terms in the expression for the cross-moment of Q-polynomials come only from the expectation of the product of the first terms in the Hermite polynomials corresponding to $\mathbf{m} = \mathbf{0}$. This means that the Q-polynomial cross-moment expression (4.11) reduces to a *single term* and that the computational effort is correspondingly reduced by a large factor.

2. This expectation is taken in a special way such that the pairing permutations do not include any permutation that would result in the term $r_0(0)$ as a factor. This corresponds to setting the values of the diagonal entries of the multiplicity matrix $\mathbf{M}$ equal to zero.

Although the Q-polynomial expression (4.23) is reduced to one term only, this term is also investigated to determine the relations between the components of the index vectors $\boldsymbol{\alpha}_i$ such that this term itself is not identically zero. The relations between the index vector components depend upon the order of the moment; therefore we develop these relations for each moment order below.

a. The First Order Moment (Mean) of the Q-polynomials

The procedure described above can be applied to compute the value of the first order moment. In this case, according to the principles just cited, we take the moment of the product of only the first terms of the Hermite polynomials without allowing the pairing that results in having $r_0(0)$ as one of the quantities within the

product. In this case however, the matrix of cross-correlation sequences $\mathbf{C}_{\boldsymbol{y}}$ in (4.17) is the upper-left block denoted by $\mathbf{R}(0)$ for which all of the off-diagonal entries are zero and the diagonal entries are $r_{\mathbf{0}}(0) = 1$. Also the $\mathbf{M}$ matrix is diagonal with even valued diagonal elements. If the sum of the components of the vector $\boldsymbol{\alpha}$ is odd then it is clear that the mean is zero. For any $\boldsymbol{\alpha}$ with even sum of components, the diagonal elements of $\mathbf{M}$ must satisfy

$$\Sigma[\mathbf{M}] = \boldsymbol{\alpha}$$

This means that if $\boldsymbol{\alpha} \neq \mathbf{0}$ the diagonal elements are not equal to zero and each of the individual components α_{ij} of the vector $\boldsymbol{\alpha}$ must be even. In other words, when expanding the Q-polynomial using the Hermite polynomials and taking the expectation, the sum of the resulting terms has the form

$$(1-1)^{\frac{\alpha_{i0}}{2}}(1-1)^{\frac{\alpha_{i1}}{2}}\cdots(1-1)^{\frac{\alpha_{iN_L}}{2}} \tag{4.25}$$

which equals zero except for $\boldsymbol{\alpha} = \mathbf{0}$. In that case

$$\mu_{\mathbf{0}}^{(1)} = \mathcal{E}\{Q_{\mathbf{0}}(n)\} = 1$$

Let us now apply this result to compute the mean of the process $x(n)$ which is given by

$$\begin{aligned} \mathcal{E}\{x(n)\} &= \mathcal{E}\left\{\sum_{\boldsymbol{\alpha}_i} c_{\boldsymbol{\alpha}_i}\mathbf{Q}_{\boldsymbol{\alpha}_i}(n)\right\} \\ &= \sum_{\boldsymbol{\alpha}_i} c_{\boldsymbol{\alpha}_i}\mathcal{E}\{\mathbf{Q}_{\boldsymbol{\alpha}_i}(n)\} \end{aligned} \tag{4.26}$$

All the terms under summation are zero except the first one for which

$$\mathcal{E}\{x(n)\} = c_{\mathbf{0}}\mathcal{E}\{Q_{\mathbf{0}}\} = c_{\mathbf{0}} \tag{4.27}$$

Therefore $c_{\mathbf{0}}$ is equal to the mean of the process. Since the cumulants of a random process do not change when a constant is added to or subtracted from the process

we will henceforth assume that c_0 is equal to zero and thus that $x(n)$ has mean zero. If c_0 is not zero, this affects only the mean of the process and not the higher order cumulants.

b. Second-Order Moment

Let us now examine the moment

$$\mu_{\boldsymbol{\alpha}}^{(2)} = \mathcal{E}\{\mathbf{Q}_{\boldsymbol{\alpha}_1}(n)\mathbf{Q}_{\boldsymbol{\alpha}_2}(n+l)\} \qquad \text{for } l \geq 0$$

with

$$\boldsymbol{\alpha}_i = [\alpha_{i0}, \alpha_{i1}, \alpha_{i2}, \cdots, \alpha_{i(N_L-2)}, \alpha_{i(N_L-1)}, \alpha_{iN_L}]^T \tag{4.28}$$

and

$$\boldsymbol{\alpha} = [\boldsymbol{\alpha}_1^T, \boldsymbol{\alpha}_2^T]^T \tag{4.29}$$

From the principles stated above (see section IV.2) this moment is calculated as the expectation of the product of only the first term of each Hermite polynomial of order corresponding to the individual components of the two index vectors. Thus this moment becomes

$$\mathcal{E}\{\mathbf{Q}_{\boldsymbol{\alpha}_1}(n)\mathbf{Q}_{\boldsymbol{\alpha}_2}(n+l)\} = \mathcal{E}_H\{y_0^{\alpha_{10}}(n)y_1^{\alpha_{11}}(n)y_2^{\alpha_{12}}(n)\cdots y_{N_L}^{\alpha_{1N_L}}(n)y_0^{\alpha_{20}}(n+l)y_1^{\alpha_{21}}(n+l)\cdots y_{N_L}^{\alpha_{2N_L}}(n+l)\} \tag{4.30}$$

where the operator $\mathcal{E}_H$ denotes the expectation obtained by summing up all the possible pairing permutations *excluding* those that result in one or more instances $r_0(0)$. Since

$$\sigma_o^2 r_0(0) = \mathcal{E}\{y_i(n)y_i(n)\} = \mathcal{E}\{y_i(n+l)y_i(n+l)\}$$

this means that all of the $y_j(n)$ or $y_j(n+l)$ of the first term of a Hermite polynomial must be paired with those from *other* polynomials. Note also that the pairing with

other y_j that have the same time argument is equal to zero because of the property $r_d(0) = 0$ for $d \neq 0$. Finally, note that if is l is positive the variable $y_i(n)$ of some Laguerre order can only be paired with another variable $y_j(n)$ with Laguerre order $j > i$. This follows because $r_d(l) = r_{j-i}(l) = 0$ for $j \leq i$. Therefore, let us arrange the components of the index vectors according to the Laguerre order and the time index as follows

$$\begin{array}{rccccccc} & y_0 & y_1 & y_2 & \cdots & y_{N_L-2} & y_{N_L-1} & y_{N_L} \\ n+l: & \alpha_{20} & \alpha_{21} & \alpha_{22} & \cdots & \alpha_{2(N_L-2)} & \alpha_{2(N_L-1)} & \alpha_{2N_L} \\ n \ : & \alpha_{10} & \alpha_{11} & \alpha_{12} & \cdots & \alpha_{1(N_L-2)} & \alpha_{1(N_L-1)} & \alpha_{1N_L} \end{array}$$

To obtain a non-zero result $y_{N_L}(n)$ can be paired with $y_{N_L}(n+l)$ only. Therefore the exponent of the latter must be greater than or equal to the exponent of the former, i.e.,

$$\alpha_{2N_L} \geq \alpha_{1N_L} \tag{4.31}$$

Also since $y_{N_L-1}(n)$ can be paired only with $y_{N_L-1}(n+l)$ or $y_{N_L}(n+l)$ to obtain a nonzero result, the sum of the exponents of the two outputs with highest Laguerre orders must satisfy

$$\alpha_{2N_L} + \alpha_{2(N_L-1)} \geq \alpha_{1N_L} + \alpha_{1(N_L-1)} \tag{4.32}$$

Continuing this argument leads to the set of conditions

$$\begin{aligned} &\alpha_{2N_L} + \alpha_{2(N_L-1)} + \alpha_{2(N_L-2)} \geq \alpha_{1N_L} + \alpha_{1(N_L-1)} + \alpha_{1(N_L-2)} \\ &\alpha_{2N_L} + \alpha_{2(N_L-1)} + \alpha_{2(N_L-2)} + \alpha_{2(N_L-3)} \geq \alpha_{1N_L} + \alpha_{1(N_L-1)} + \alpha_{1(N_L-2)} + \alpha_{1(N_L-3)} \\ &\qquad\qquad\qquad\qquad \vdots \\ &\alpha_{2N_L} + \alpha_{2(N_L-1)} + \cdots + \alpha_{22} + \alpha_{21} \geq \alpha_{1N_L} + \alpha_{1(N_L-1)} + \cdots + \alpha_{12} + \alpha_{11} \\ &\alpha_{2N_L} + \alpha_{2(N_L-1)} + \cdots + \alpha_{22} + \alpha_{21} + \alpha_{20} \geq \alpha_{1N_L} + \alpha_{1(N_L-1)} + \cdots + \alpha_{12} + \alpha_{11} + \alpha_{10} \end{aligned} \tag{4.33}$$

Now let the vector $\boldsymbol{\beta}_i^q$ be defined as the vector composed of the rightmost m components of the vector $\boldsymbol{\alpha}_i$.

$$\boldsymbol{\beta}_i^q = [\alpha_{i(N_L-q+2)}, \cdots, \alpha_{iN_L}, \alpha_{i(N_L+1)}] \tag{4.34}$$

This means that $\boldsymbol{\beta}_i^q$ represents the exponents of the outputs of the q systems with highest Laguerre orders. With this definition, (4.31)-(4.33) can be summarized as:

$$\boldsymbol{\beta}_2^q \geq \boldsymbol{\beta}_1^q \qquad \text{for } q = 1, 2, \cdots, N_L + 1 \tag{4.35}$$

where the last relation for $q = N_L + 1$, implies that the total sum of the components in $\boldsymbol{\alpha}_2$ is greater than the total sum of the components in $\boldsymbol{\alpha}_1$.

Now, let us consider the relation of the index components starting with the outputs of the lowest Laguerre orders. The output $y_0(n+l)$ can be paired with $y_0(n)$ only. Therefore

$$\alpha_{10} \geq \alpha_{20} \tag{4.36}$$

The output $y_1(n+l)$ can be paired with $y_0(n)$ or $y_1(n)$ only, which implies that

$$\alpha_{10} + \alpha_{11} \geq \alpha_{20} + \alpha_{21} \tag{4.37}$$

By similar consideration the following relations must hold:

$$\begin{aligned}
&\alpha_{10} + \alpha_{11} + \alpha_{12)} \geq \alpha_{20} + \alpha_{21} + \alpha_{22} \\
&\alpha_{10} + \alpha_{11} + \alpha_{12} + \alpha_{13} \geq \alpha_{20} + \alpha_{21} + \alpha_{22} + \alpha_{23} \\
&\qquad\qquad\qquad\qquad \vdots \\
&\alpha_{10} + \alpha_{11} + \cdots + \alpha_{1(N_L-2} + \alpha_{1(N_L-1)} \geq \alpha_{20} + \alpha_{21} + \cdots + \alpha_{2(N_L-2)} + \alpha_{2(N_L-1)} \\
&\alpha_{10} + \alpha_{11} + \cdots + \alpha_{1(N_L-1)} + \alpha_{1N_L} \geq \alpha_{20} + \alpha_{21} + \cdots + \alpha_{2(N_L-1)} + \alpha_{2N_L}
\end{aligned} \tag{4.38}$$

If the vector $\boldsymbol{\gamma}_i^q$ is defined to be composed of the leftmost q components of $\boldsymbol{\alpha}_i$ (representing the exponents of the outputs of the q systems with lowest Laguerre orders),

$$\boldsymbol{\gamma}_i^q = [\alpha_{i0}, \alpha_{i1}, \cdots, \alpha_{i(q-1)}] \tag{4.39}$$

then the relations (4.36)-(4.38) can be summarized as

$$\boldsymbol{\gamma}_2^q \leq \boldsymbol{\gamma}_1^q \qquad \text{for } q = 1, 2, \cdots, N_L + 1 \tag{4.40}$$

The last relation in (4.40) (for $q = N_L + 1$) means that $\Sigma[\boldsymbol{\alpha}_1]$ is greater than $\Sigma[\boldsymbol{\alpha}_2]$. This and the last relation in (4.35) (for $m = N_L+1$) are not true unless $\Sigma[\boldsymbol{\alpha}_1] = \Sigma[\boldsymbol{\alpha}_2]$. This implies that:

1. $\Sigma[\boldsymbol{\alpha}]$ is even where $\boldsymbol{\alpha} = [\boldsymbol{\alpha}_1{}^T, \boldsymbol{\alpha}_2{}^T]^T$.

2. Since $\boldsymbol{\alpha}_i = [\boldsymbol{\gamma}_i^{N_L+1-q} \quad \boldsymbol{\beta}_i^q]$ for any q the relations (4.40) are true if and only if the relations (4.35) are true.

From this condition we also observe that if the Q-polynomials in the expression (4.2) are divided into groups according to the degree of nonlinearity they represent (i.e. linear, quadratic, cubic, etc.), the necessary and sufficient conditions for the second-order cross-cumulant of two Q-polynomials to be not identically zero are:

1. The two polynomials belong to the *same* group of nonlinearity.

2. The relations (4.33) or (4.35) hold.

c. Third-Order Moment

The relations between the vectors of indices $\boldsymbol{\alpha}_1$, $\boldsymbol{\alpha}_2$ and $\boldsymbol{\alpha}_3$ such that

$$\mu_{\boldsymbol{\alpha}}^{(3)}(l_1, l_2) = \mathcal{E}\{\mathbf{Q}_{\boldsymbol{\alpha}_1}(n)\mathbf{Q}_{\boldsymbol{\alpha}_2}(n + l_1)\mathbf{Q}_{\boldsymbol{\alpha}_3}(n + l_2)\} \neq 0 \qquad \text{for } l_2 \geq l_1 \geq 0 \tag{4.41}$$

are developed as follows. The components of each vector are arranged according to descending time argument value and increasing Laguerre order of the corresponding output:

$$\begin{array}{lccccccc} & y_0 & y_1 & y_2 & \cdots & y_{N_L-2} & y_{N_L-1} & y_{N_L} \\ n+l_2: & \alpha_{30} & \alpha_{31} & \alpha_{32} & \cdots & \alpha_{3(N_L-2)} & \alpha_{3(N_L-1)} & \alpha_{3N_L} \\ n+l_1: & \alpha_{20} & \alpha_{21} & \alpha_{22} & \cdots & \alpha_{2(N_L-2)} & \alpha_{2(N_L-1)} & \alpha_{2N_L} \\ n: & \alpha_{10} & \alpha_{11} & \alpha_{12} & \cdots & \alpha_{1(N_L-2)} & \alpha_{1(N_L-1)} & \alpha_{1N_L} \end{array}$$

We have two conditions on these variables. The first condition follows from the condition on the expectation of the product of Gaussian variables; namely $\Sigma[\boldsymbol{\alpha}]$ is even. The second condition is obtained if we examine the conditions on the values of the individual components in each vector. We find the relations

$$\boldsymbol{\beta}_3^q + \boldsymbol{\beta}_2^q \geq \boldsymbol{\beta}_1^q \qquad \text{for } q = 1, 2, \cdots, N_L + 1 \tag{4.42}$$

$$\boldsymbol{\gamma}_1^q + \boldsymbol{\gamma}_2^q \geq \boldsymbol{\gamma}_3^q \qquad \text{for } q = 1, 2, \cdots, N_L + 1 \tag{4.43}$$

$$\Sigma[\boldsymbol{\alpha}_1] + \boldsymbol{\beta}_3^q \geq \boldsymbol{\beta}_2^q \qquad \text{for } q = 1, 2, \cdots, N_L + 1 \tag{4.44}$$

where the vectors $\boldsymbol{\beta}_i^q$ and $\boldsymbol{\gamma}_i^q$ have the same definition as before. For $q = N_L + 1$ equations (4.42), (4.43) and (4.44) yield the following relations:

$$\begin{aligned} \Sigma[\boldsymbol{\alpha}_1] &\leq \Sigma[\boldsymbol{\alpha}_2] + \Sigma[\boldsymbol{\alpha}_3] \\ \Sigma[\boldsymbol{\alpha}_3] &\leq \Sigma[\boldsymbol{\alpha}_1] + \Sigma[\boldsymbol{\alpha}_2] \\ \Sigma[\boldsymbol{\alpha}_2] &\leq \Sigma[\boldsymbol{\alpha}_3] + \Sigma[\boldsymbol{\alpha}_1] \end{aligned} \tag{4.45}$$

If these three relations are true one of the relations in (4.42), (4.43) and (4.44) becomes redundant. Therefore necessary and sufficient conditions for the values of the components in the three vectors of exponents to ensure that the value of the moment is not identically zero are :

1. The sum of all the components of the three vectors $\Sigma[\boldsymbol{\alpha}]$ is even.

2. The sum of the components of any two vectors is greater than or equal to the sum of the components in the third.

3. The relations (4.42) and (4.43) are satisfied.

Similar relations can be developed between the vectors of indices of the Q-polynomials for computing the cross-moment of any order. The set of indices in the expression of the cumulant of the output of the model is generated then examined to exclude the terms that do not satisfy these relations. The value of the cumulant is then obtained by calculating the rest of the terms because their values are not identically zero. In general, the non-zero valued Q-polynomial cross-moment in (4.11) is then given by

$$\mu_{\boldsymbol{\alpha}}^{(k)}(\boldsymbol{l}) = (\sigma_o^2)^{\frac{\Sigma[\boldsymbol{\alpha}]}{2}} \boldsymbol{\alpha}_1!\boldsymbol{\alpha}_2!\cdots\boldsymbol{\alpha}_k! \sum_{\Sigma[\mathbf{M}]=\boldsymbol{\alpha}} \prod_j \frac{(r_j(l_j))^{i_j}}{i_j!} \tag{4.46}$$

where the first summation is over the possible multiplicity matrix permutations. The product is over all the nonzero-valued entries i_j of the multiplicity matrix and r_j is the corresponding Laguerre cross-correlation with time argument l_j.

3. Computation of Cumulants of Q-Polynomials and Model Output

To obtain an expression for the k^{th}-order cumulant of the output $x(n)$ as described by (4.7) we need to develop an expression for the k^{th}-order cross-cumulant of the Q-polynomials

$$C_{\boldsymbol{\alpha}}^{(k)}(l_1, l_2, \cdots, l_{k-1})$$

Since the Q-polynomials are all zero-mean except $\mathbf{Q_0}$ (which has a mean equal to one) the mean of the model output was shown to be given by the value of the coefficient c_0 (see (4.27)). In our development of the model output cumulants we assume that this coefficient equals zero so the model output henceforth has zero mean. Consequently the Q-polynomial representation of the model output does not include $\mathbf{Q}_{\boldsymbol{\alpha}_0}$. The expressions of the model output cumulants and the Q-polynomial cumulants are

given by (2.14). In this case the Q-polynomial cumulants of orders 2 to 4 have the expression

$$
\begin{aligned}
C_{\boldsymbol{\alpha}}^{(2)}(l) &= \mathcal{E}\left\{\mathbf{Q}_{\boldsymbol{\alpha}_1}(n)\mathbf{Q}_{\boldsymbol{\alpha}_2}(n+l)\right\} \\
C_{\boldsymbol{\alpha}}^{(3)}(\boldsymbol{l}) &= \mathcal{E}\left\{\mathbf{Q}_{\boldsymbol{\alpha}_1}(n)\mathbf{Q}_{\boldsymbol{\alpha}_2}(n+l_1)\mathbf{Q}_{\boldsymbol{\alpha}_3}(n+l_2)\right\} \\
C_{\boldsymbol{\alpha}}^{(4)}(\boldsymbol{l}) &= \mathcal{E}\left\{\mathbf{Q}_{\boldsymbol{\alpha}_1}(n)\mathbf{Q}_{\boldsymbol{\alpha}_2}(n+l_1)\mathbf{Q}_{\boldsymbol{\alpha}_3}(n+l_2)\mathbf{Q}_{\boldsymbol{\alpha}_3}(n+l_3)\right\} \\
&\quad -\mathcal{E}\left\{\mathbf{Q}_{\boldsymbol{\alpha}_1}(n)\mathbf{Q}_{\boldsymbol{\alpha}_2}(n+l_1)\right\}\mathcal{E}\left\{\mathbf{Q}_{\boldsymbol{\alpha}_3}(n+l_2)\mathbf{Q}_{\boldsymbol{\alpha}_4}(n+l_3)\right\} \\
&\quad -\mathcal{E}\left\{\mathbf{Q}_{\boldsymbol{\alpha}_1}(n)\mathbf{Q}_{\boldsymbol{\alpha}_3}(n+l_2)\right\}\mathcal{E}\left\{\mathbf{Q}_{\boldsymbol{\alpha}_2}(n+l_1)\mathbf{Q}_{\boldsymbol{\alpha}_4}(n+l_3)\right\} \\
&\quad -\mathcal{E}\left\{\mathbf{Q}_{\boldsymbol{\alpha}_1}(n)\mathbf{Q}_{\boldsymbol{\alpha}_4}(n+l_3)\right\}\mathcal{E}\left\{\mathbf{Q}_{\boldsymbol{\alpha}_2}(n+l_1)\mathbf{Q}_{\boldsymbol{\alpha}_3}(n+l_2)\right\}
\end{aligned}
\tag{4.47}
$$

It can be seen that the second- and third-order cumulants of the Q-polynomials are given directly by their respective moments. As for the fourth-order cumulant, we still can compute it directly using the same procedure we have developed to compute the moment. Since the last three terms cancel with similar terms that arise in the computation of the first term, we need only to compute the moment using the procedure developed above and set a condition so that the common terms which cancel are not computed. This condition is imposed on the permutations simply by setting terms of the multiplicity matrix $\mathbf{M}$ to zero. For the fourth-order moment this matrix has 4×4 blocks and we do our computations on the upper trianglar parts of the six off-diagonal blocks. The entries of this matrix take on values as described before such that at least three of the six blocks have entries which are not all equal to zero.

The computation of the model output cumulant can be summarized as follows:

1. From the model dimensions N_L and N_N construct a vector of indices $\boldsymbol{\alpha}$ the partitions of which are the Q-polynomial indices $\boldsymbol{\alpha}_i$.

2. For computation of a cumulant of order k, test all the possible combinations of k indices $\boldsymbol{\alpha} = [\boldsymbol{\alpha}_{i_1}{}^T \ \boldsymbol{\alpha}_{i_2}{}^T \ \cdots \boldsymbol{\alpha}_{i_k}{}^T]^T$ to determine those that satisfy the conditions for a nonzero-valued cross-moment.

3. For each of the combinations that give nonzero-valued moments use (4.46) to obtain the value of the cumulant by applying the proper condition on the permutations of the matrix of multiplicity.

4. From the values of the coefficients $c_{\boldsymbol{\alpha}_i}$ the output cumulant is given by (4.7).

C. POLYSPECTRA OF THE OUTPUT OF WIENER MODEL

The polyspectra of the output of the Wiener model are the Fourier transforms of the corresponding cumulants of this output. Due to the linearity of the Fourier transform these polyspectra are the weighted sum of the polyspectra of the individual Q-polynomials in the expression (4.2) with the same coefficients $c_{\boldsymbol{\alpha}}$. Although it might at first seem that the procedures developed in the previous section could be adopted directly to compute the polyspectra, the regions of symmetry of these quantities and other considerations make the problem sufficiently complicated that it is better to develop the spectral procedures on their own. The resulting expressions and procedures for computation are analogous to those for the cumulants but differ in important details.

Since the Q-polynomials are functionals of Gaussian random variables $y_i(n)$, the expressions derived in Chapter II Section C.2 can be applied to the compute cross-polyspectra of the Q-polynomials. The second-order system cross-spectral functions, which are the entries of the matrix $\boldsymbol{\mathcal{S}}$ defined in Chapter II, are shown in Chapter III to be the Laguerre cross-spectral functions $s_{\boldsymbol{d}}(\omega)$. In the following development we

make use of the functions

$$s_d(z) = \sum_{l=-\infty}^{\infty} r_d(l) z^{-l} \tag{4.48}$$

which are used in Chapter III to derive the Laguerre cross-correlation sequences. These functions, when evaluated on the unit circle, are the Laguerre cross-spectral functions[1]. We also make use of the orthogonality of Q-polynomials, which can be detected from the vectors of indices $\boldsymbol{\alpha}$, to represent the polyspectra as a sum of only the terms that are not guaranteed to be identically zero.

Since the Q-polynomial moments of orders higher than third are not equal to their cumulants (see (4.47)), their Fourier transform is different than the polyspectra. We shall call the Fourier tranforms of these moments the *Q-spectral functions* and denote the k^{th}-order Q-spectral function as $S_Q^{\boldsymbol{\alpha}}(\boldsymbol{\omega})$. The Q-spectral function has the form

$$\begin{aligned} S_Q^{\boldsymbol{\alpha}}(\boldsymbol{\omega}) &= S_Q^{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_k}(\omega_1,\omega_2,\cdots,\omega_{k-1}) \\ &= (\sigma_o^2)^{\frac{\Sigma[\boldsymbol{\alpha}]}{2}} \boldsymbol{\alpha}_1!\boldsymbol{\alpha}_2!\cdots\boldsymbol{\alpha}_k! \sum_{\mathbf{m}_1=\mathbf{0}}^{\lfloor\frac{\boldsymbol{\alpha}_1}{2}\rfloor} \sum_{\mathbf{m}_2=\mathbf{0}}^{\lfloor\frac{\boldsymbol{\alpha}_2}{2}\rfloor} \cdots \sum_{\mathbf{m}_k=\mathbf{0}}^{\lfloor\frac{\boldsymbol{\alpha}_k}{2}\rfloor} \frac{(-1)^{\Sigma[\mathbf{m}]}}{\mathbf{m}_1!\mathbf{m}_2!\cdots\mathbf{m}_k!(2)^{\Sigma[\mathbf{m}]}} \\ &\sum_{\Sigma[\mathbf{M}]=\boldsymbol{\alpha}-2\mathbf{m}} 2^{-\frac{\mathrm{tr}(\mathbf{M})}{2}} \left(\prod_{j_1=1}^{k(N_L+1)} \frac{1}{(\frac{M(j_1,j_1)}{2})!} \prod_{j_2=j_1+1}^{k(N_L+1)} \frac{1}{M(j_1,j_2)!} \right) S_{conv}^{\boldsymbol{\alpha}-2\mathbf{m}}(\boldsymbol{\omega}) \end{aligned} \tag{4.49}$$

where $\mathbf{M}$ is the multiplicity matrix, which depends on the summation parameters $\mathbf{m}$ through the relation $\Sigma[\mathbf{M}] = \boldsymbol{\alpha} - 2\mathbf{m}$. As described in Chapter II, the quantity $S_{conv}^{\boldsymbol{\alpha}-2\mathbf{m}}$ is formed as the product of the two quantities :

1. The product of values of each quantity in the block matrix $\mathbf{R}(0)$ that is included in the pairing permutation.

[1]We note as before, that the use of the term $s_d(\omega)$ rather than $s_d(e^{j\omega})$ is slightly abusive but results in some simplification of notation.

2. The multiple convolution of the system cross-spectral functions $s_{\boldsymbol{d}}(z)$ that are included in the pairing permutation. These were defined in Chapter II as the entries of the off-diagonal blocks $\mathbf{S}$ of the matrix $\boldsymbol{\mathcal{S}}$ given by (2.106).

Since the matrix $\mathbf{M}$ determines the multiplicity of each of the above quantities, the entries of this matrix need to be set to zero whenever the inclusion of the corresponding spectral functions results in a term $S_{\mathbf{conv}}^{\boldsymbol{\alpha}-2\mathbf{m}}(\boldsymbol{\omega})$ which is identically zero.

We have noted earlier (see Section B.1) that the diagonal blocks of the multiplicity matrix are diagonal. Since each of the off-diagonal blocks $\mathbf{S}^{\omega}$ of $\boldsymbol{\mathcal{S}}$ is Hermitian symmetric, if the upper triangle has entries that are expressed as $s_{\boldsymbol{d}}(\omega)$ for $d = 0, 1, \cdots, N_L$, then the lower triangle entries are given by $s_{\boldsymbol{d}}(-\omega)$. The corresponding z-transforms are $s_{\boldsymbol{d}}(z)$ and $s_{\boldsymbol{d}}(z^{-1})$. For purpose of this discussion let us denote a block of the multiplicity matrix corresponding to a block S^{ω} in the spectral matrix as $\mathbf{M}^{\omega}$. Now if the pairing permutation results in a combination of entries from the both the upper triangle and the lower triangle then at least one of the convolutions takes the form

$$s_{\boldsymbol{d}_1}(z) * s_{\boldsymbol{d}_2}(z^{-1}) = \frac{1}{2\pi j} \oint_{C1} v^{-1} s_{\boldsymbol{d}_1}\left(\frac{z}{v}\right) s_{\boldsymbol{d}_2}(v) dv \tag{4.50}$$

If we substitue the value of $s_{\boldsymbol{d}}(z)$ from (3.84) we get

$$s_{\boldsymbol{d}_1}(z) * s_{\boldsymbol{d}_2}(z^{-1}) = \frac{(1-\rho^2)^2}{\rho^2} \rho^{-(\boldsymbol{d}_1+\boldsymbol{d}_2)} z \frac{1}{2\pi j} \oint_{C1} \frac{(v - z\rho)^{\boldsymbol{d}_1 - 1}}{(v - z\rho^{-1})^{\boldsymbol{d}_1+1}} \frac{(v-\rho)^{\boldsymbol{d}_2-1}}{(v-\rho^{-1})^{\boldsymbol{d}_2+1}} dv \tag{4.51}$$

which is identically zero because there are no poles inside the unit circle. Therefore if the values of the polyspectra are computed in the non-redundant regions for which

$$0 \leq \omega_1 \leq \omega_2 \cdots \leq \omega_{k-1} \tag{4.52}$$

the off-diagonal blocks $\mathbf{M}^{\omega}$ of the multiplicity matrix $\mathbf{M}$ must be upper triangular to exclude the permutations for which $s_{\boldsymbol{d}}(z^{-1})$ is included. Then the multiplicity matrix

in this procedure is exactly the same as that used for computing the moments. Since the terms in the lower trianglular portion are never used, the block $\mathbf{S}^{\omega}$ can be replaced by one which is upper triangular and Toeplitz:

$$\mathbf{S}^{\omega} = \begin{bmatrix} s_0(\omega) & s_1(\omega) & s_2(\omega) & \cdots & s_{N_L}(\omega) \\ 0 & s_0(\omega) & s_1(\omega) & \cdots & s_{N_L-1}(\omega) \\ 0 & 0 & s_0(\omega) & \cdots & s_{N_L-2}(\omega) \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & s_0(\omega) \end{bmatrix} \tag{4.53}$$

Notice that a single subscript $d = j - i$ is used for the system cross-spectral functions rather than the double index used in Chapter II because the Laguerre cross-spectral functions depend only on the difference between the two Laguerre system orders i and j.

In the development in Appendix C of the relations between the vectors of indices for which the value of the moment is not identically zero the results were based upon excluding permutations that contain the terms $r_0(0)$. These are the diagonal entries of the matrix $\mathbf{C}_y$ in (4.17) and result in a common factor that equals zero. Since these functions also appear in the polyspectra expressions under similar conditions, the same factorization exists with the same zero values. This means that terms of the Q-spectral functions are identically zero in exactly the same cases for which the moments of the Q-polynomials are identically zero. The relations between the vectors of indices developed in the previous section are therefore the same for the Q-spectral functions. Thus the expressions (4.49) simplifies in the same way as the expression (4.23) did for the time domain computation.

Let us now consider the remaining part of the problem, namely computing the quantities $S_{conv}^{\boldsymbol{\alpha}-2\mathbf{m}}(\boldsymbol{\omega})$. It will be seen that, like in other calculations involving the Wiener model, there is abundant structure present so that the form of the result can be predicted from an analysis presented here and the actual computation can be implemented by a table lookup scheme.

1. Multiple Convolutions of the Laguerre System Cross-Spectral Functions

The Q-spectral functions given by (4.49) are obtained by performing multiple convolutions on the system cross-spectral functions $s_{\boldsymbol{d}}(z)$ with given multiplicity to compute the frequency-dependent quantities $S^{\boldsymbol{\alpha}}_{\boldsymbol{conv}}(\boldsymbol{\omega})$. Convolving the spectral functions by directly applying the convolution formula (4.50) is a tedious operation especially when this needs to be repeated many times. Fortunately however, we can use the characteristic of the Laguerre cross-spectral functions to develop a procedure to obtain the value of a multiple convolution algebraically. If a system cross-spectral function has a multiplicity m, this quantity equals the sum of the entries along the corresponding diagonals, principal and non-principal, in the block $\mathbf{M}^{\omega}$ of the multiplicity matrix $\mathbf{M}$. This means that the contribution of this function $s_{\boldsymbol{d}}(z)$ to the quantity $S^{\boldsymbol{\alpha}}_{\boldsymbol{conv}}$ is the result of convolution with itself $(m-1)$ times. The result is then convolved with the contributions of the other functions. Let us use the notation $s_{\boldsymbol{d}}(z;\rho)$ to denote the a complex cross-spectral density function with parameter ρ and define $s^{\boldsymbol{m}}_{\boldsymbol{d}}(z;\rho)$ as

$$\begin{aligned} s^{(m)}_{\boldsymbol{d}}(z;\rho) &= \underbrace{s_{\boldsymbol{d}}(z;\rho) * s_{\boldsymbol{d}}(z;\rho) * s_{\boldsymbol{d}}(z;\rho) * \cdots * s_{\boldsymbol{d}}(z;\rho)}_{m-1 \text{ convolutions}} \\ &= s_{\boldsymbol{d}}(z;\rho) * s^{(m-1)}_{\boldsymbol{d}}(z;\rho) \end{aligned} \tag{4.54}$$

Using this notation we can also define

$$\begin{aligned} s^{(0)}_{\boldsymbol{d}}(z;\rho) &= \delta(z) \\ s^{(1)}_{\boldsymbol{d}}(z;\rho) &= s_{\boldsymbol{d}}(z;\rho) \end{aligned} \tag{4.55}$$

We consider various special cases below.

a. Multiple Self-Convolution of the Complex Spectral Density Function $s_0(z;\rho)$

When the sum of the entries along the principal diagonal of the block $\mathbf{M}^{\omega}$ is non-zero then the system complex spectral density function $s_0(z;\rho)$ has a contribution to the required spectral function S^{α}_{conv}. This contribution is obtained by performing a number of successive convoutions. The Laguerre system complex spectral density function has the form

$$s_0(z;\rho) = -\frac{(1-\rho^2)}{\rho}\frac{z}{(z-\rho)(z-\rho^{-1})} \tag{4.56}$$

The first convolution is given by

$$\begin{aligned} s_d^{(2)}(z;\rho) &= s_0(z;\rho) * s_0(z;\rho) \\ &= \frac{(1-\rho^2)^2}{\rho^2} z \frac{1}{2\pi j}\oint_{C1} \frac{v}{(v-z\rho)(v-z\rho^{-1})(v-\rho)(v-\rho^{-1})} dv \end{aligned} \tag{4.57}$$

Since $|z| < |\rho^{-1}|$ it follows that $|z\rho| < 1|$. Therefore the integrand has two poles inside the unit circle, at locations ρ and $z\rho$, and by evaluating residues the result of this convolution is found to be

$$\begin{aligned} s_0^{(2)}(z;\rho) &= -\frac{(1-\rho^4)}{\rho^2}\frac{z}{(z-\rho^2)(z-\rho^{-2})} \\ &= -\frac{(1-(\rho^2)^2)}{\rho^2}\frac{z}{(z-\rho^2)(z-\rho^{-2})} \\ &= s_0(z;\rho^2) \end{aligned} \tag{4.58}$$

It is seen that the two poles of the complex power spectral density function are moved from the locations ρ and ρ^{-1} to ρ^2 and ρ^{-2} respectively as a result of one convolution. One more convolution can be similarly shown to result in the quantity

$$\begin{aligned} s_0^{(3)}(z;\rho) &= s_0(z;\rho) * s_0^2(z;\rho) = s_0(z;\rho) * s_0(z;\rho^2) \\ &= -\frac{(1-(\rho^3)^2)}{\rho^3}\frac{z}{(z-\rho^3)(z-\rho^{-3})} \\ &= s_0(z;\rho^3) \end{aligned} \tag{4.59}$$

Continuing in this way we can find that

$$s_0^{(m)}(z;\rho) = s_0(z;\rho^m) \tag{4.60}$$

Thus the result of $m-1$ convolutions on m functions $s_0(z;\rho)$ is equal to the same function with the parameter ρ raised to the m power; its poles are moved from ρ and ρ^{-1} to ρ^m and ρ^{-m}.

b. Convolutions of up to Three Laguerre Cross-Spectral Density Functions $s_{d_1}(z;\rho)$, $s_{d_2}(z;\rho)$, $s_{d_3}(z;\rho)$

For the other Laguerre spectral functions $s_{d_1}(z;\rho)$ for which $d_1 > 0$ the sum of the elements along the corresponding non-principal diagonal of $\mathbf{M}^{\omega}$ determines the number of convolutions to be performed. Let us first consider the convolution of $s_{d_1}(z;\rho)$ with itself:

$$\begin{aligned} s_{d_1}^{(2)}(z;\rho) &= s_{d_1}(z;\rho) * s_{d_1}(z;\rho) \\ &= \frac{1}{2\pi j}\oint_{C1} v^{-1} s_{d_1}(v) s_{d_1}\left(\frac{z}{v}\right) dv \\ &= z(1-\rho^2)^2 \frac{1}{2\pi j}\oint_{C1} \frac{v(v-\rho^{-1})^{(d_1-1)}(v-z\rho)^{(d_1-1)}}{(v-\rho)^{(d_1+1)}(v-z\rho^{-1})^{(d_1+1)}} dv \end{aligned} \tag{4.61}$$

Since the integrand has (d_1+1) poles inside the unit circle at location ρ the value of this convolution is given by

$$\begin{aligned} s_{d_1}^{(2)}(z;\rho) &= z(1-\rho^2)^2 \frac{1}{d_1!}\frac{d^{d_1}}{dv^{d_1}}\left[\frac{v(v-\rho^{-1})^{(d_1-1)}(v-z\rho)^{(d_1-1)}}{(v-z\rho^{-1})^{(d_1+1)}}\right]_{v=\rho} \\ &= z(1-\rho^2)^2 \underbrace{\sum_{m_1=0}^{1}\sum_{m_2=0}^{(d_1-1)}\sum_{m_3=0}^{(d_1-m_1-m_2)}}_{m_4=d_1-m_1-m_2-m_3} \frac{1}{m_1!m_2!m_3!m_4!}\frac{1}{(1-m_1)!}\rho^{1-m_1} \\ &\quad \times \frac{(d_1-1)!}{(d_1-1-m_2)!}(\rho-\rho^{-1})^{d_1-1-m_2}\frac{(d_1-1)!}{(d_1-1-m_3)!}(\rho-z\rho)^{d_1-1-m_3} \\ &\quad \times (-1)^{m_4}\frac{(d_1+m_4)!}{d_1!}\frac{1}{(\rho-z\rho^{-1})^{(d_1+1+m_4)}} \end{aligned}$$

$$= z(1-\rho^2)^2 \sum_{m_1=0}^{1} \sum_{m_2=0}^{(d_1-1)} \binom{1}{m_1} \binom{d_1-1}{m_2} \rho^{(1-m_1)} (\rho-\rho^{-1})^{(d_1-1-m_2)}$$
$$\times \underbrace{\sum_{m_3=0}^{(n-m_1-m_2)}}_{m_4=n-m_1-m_2-m_3} \binom{d_1-1}{m_3} \binom{d_1+m_4}{d_1} \frac{(-1)^{m_4}(\rho-z\rho)^{(d_1-1-m_3)}}{(\rho-z\rho^{-1})^{(d_1+1+m_4)}} \tag{4.62}$$

Since m_1 can take on values of only 0 and 1, the double summation over m_1 and m_2 can be simplified. In particular, the term

$$\binom{1}{m_1}$$

is equal to 1, and with a change of variables

$$m_1 + m_2 = k' \tag{4.63}$$

k' assumes values either in the range $[1, d_1]$ (for $m_1 = 1$) or the range $[0, d_1 - 1]$ (for $m_1 = 0$). Therefore (4.62) becomes

$$s_{d_1}^{(2)}(z;\rho) = -(1-\rho^2) \sum_{k'=0}^{d_1} \sum_{m_2=k'-1}^{\min[k', d_1-1]} (-1)^{d_1-m_2} \binom{d_1-1}{m_2} \rho^{-2(d_1-m_2)} (1-\rho^2)^{d_1-m_2}$$
$$\times \underbrace{\sum_{m_3=0}^{(d_1-k')}}_{m_3 \in [0, d_1-1]} (-1)^{m_3} \binom{d_1-1}{m_3} \binom{2d_1-k'-m_3}{d_1} \rho^{2(2d_1+1-k'-m_3)}$$
$$\times \frac{z(z-1)^{(d_1-1-m_3)}}{(z-\rho^2)^{(2d_1+1-k'-m_3)}} \tag{4.64}$$

Now let us define

$$W_{d_1}^{\prime(1)}(k') = \sum_{m_2=k'-1}^{\min[k', d_1-1]} (-1)^{d_1-m_2} \binom{d_1-1}{m_2} \rho^{-2(d_1-m_2)} (1-\rho^2)^{d_1-m_2} \tag{4.65}$$

Since $m_2 = 0$ for $k' = 0$, this quantity has the value

$$W_{d_1}^{\prime(1)}(0) = (-1)^{d_1} \rho^{-2d_1} (1-\rho^2)^{d_1} \tag{4.66}$$

Similarly since $m_2 = d_1 - 1$, for $k' = d_1$ we have

$$W_{d_1}^{\prime(1)}(d_1) = -\rho^{-2}(1-\rho^2) \tag{4.67}$$

and for $k' \in [1, d_1 - 1]$

$$\begin{aligned} W_{d_1}^{\prime(1)}(k') &= (-1)^{d_1-k'+}\begin{pmatrix} d_1-1 \\ k'-1 \end{pmatrix}\rho^{-2(d_1-k'+1)}(1-\rho^2)^{d_1-k'+1} \\ &\quad +(-1)^{d_1-k'}\begin{pmatrix} d_1-1 \\ k' \end{pmatrix}\rho^{-2(d_1-k')}(1-\rho^2)^{d_1-k'} \\ &= (-1)^{d_1-k'}\begin{pmatrix} d_1 \\ k' \end{pmatrix}(\rho^{-2}-1)^{d_1-k'}[1-\frac{k'}{d_1}\rho^{-2}] \end{aligned} \tag{4.68}$$

Notice that if we let $k' = 0$ and $k' = d_1$ in (4.68) we obtain the relations (4.66) and (4.67) respectively. Therefore (4.68) is the required expression for $W_{d_1}^{\prime(1)}(k')$ for all $k' \in [0, d_1]$. We then substitute $k = d_1 - k'$ and define $W_{d_1}^{(1)}(k)$ as

$$W_{d_1}^{(1)}(k) = (-1)^k\begin{pmatrix} d_1 \\ k \end{pmatrix}(\rho^{-2}-1)^k\left[1-\frac{d_1-k}{d_1}\rho^{-2}\right] \tag{4.69}$$

Hence (4.64) becomes

$$\begin{aligned} s_{d_1}^{(2)}(z;\rho) &= -(1-\rho^2)\sum_{k=0}^{d_1} W_{d_1}^{(1)}(k)\sum_{m_3=0}^{\min[k,d_1-1]}(-1)^{m_3}\begin{pmatrix} d_1-1 \\ m_3 \end{pmatrix}\begin{pmatrix} d_1+k-m_3 \\ d_1 \end{pmatrix} \\ &\quad \times\rho^{2(d_1+1+k-m_3)}\frac{z(z-1)^{(d_1-1-m_3)}}{(z-\rho^2)^{(d_1+1+k-m_3)}} \end{aligned} \tag{4.70}$$

The expression (4.70) reveals important characteristics of the convolution of the two Laguerre cross-spectral functions of the same order. This convolution can be obtained algebraically as a sum of spectral functions that have one zero at the origin and multiple other zeros at $z = 1$. The poles of these functions are at $z = \rho^2$. The number of these functions depends on the orders n of the cross-spectral functions being convolved and they differ only in the multiplicities of their poles and zeros. We will see presently that similar properties hold for multiple convolutions of the Laguerre cross-spectral functions of the same order and convolutions of Laguerre cross-specrtal functions of different orders.

Now let us define the spectral functions embeded in (4.70) that result from the first convolution as

$$\mathcal{X}^{(1)}(z;p,q) = \rho^{2p}\frac{z(z-1)^q}{(z-\rho^2)^p} \tag{4.71}$$

Then (4.70) can be written as

$$s^{(2)}_{d_1}(z;\rho) = \sum_{k=0}^{d_1}\sum_{m_3=0}^{\min[k,d_1-1]} b^{(1)}(k,m_3)\mathcal{X}^{(1)}(z;p^{(1)}(k,m_3),q^{(1)}(m_3)) \tag{4.72}$$

where

$$\begin{aligned} b^{(1)}(k,m_3) &= (-1)^{m_3+1}\begin{pmatrix} d_1-1 \\ m_3 \end{pmatrix}\begin{pmatrix} d_1+k-m_3 \\ d_1 \end{pmatrix}(1-\rho^2)W^{(1)}_{d_1}(k) \\ p^{(1)}(k,m_3) &= d_1+2+k+q^{(1)}(m_3) \\ q^{(1)}(m_3) &= d_1-1-m_3 \end{aligned} \tag{4.73}$$

with $m_3 = 0,1,\cdots,k$ and $k = 0,1,2\cdots,d_1$. Therefore the pole and zero multiplicities have a range of values given by

$$\begin{aligned} p^{(1)} &= d_1+1, d_1+2,\cdots,d_1+2-q^{(1)} \\ q^{(1)} &= 0,1,2,\cdots,d_1-1 \end{aligned} \tag{4.74}$$

The procedure we have employed so far is useful for the other cases to be considered in this section and the results will be analogous. In particular, suppose it is desired to form the spectral term $s^{(2)}_{d_1,d_2}(z) = s_{d_1}(z) * s_{d_2}(z)$. We can assume with no loss of generality that $d_2 < d_1$. In this case we obtain the same results except that the limits on m_3 in (4.72) change from $m_3 \in [0, d_1-1]$ to $0 \le m_3 \le \min[d_1-1, d_2-k]$ and the ranges of the pole and zero multiplicities in (4.74) change to

$$\begin{aligned} p^{(1)} &= d_1+1, d_1+2,\cdots,d_1+2+q^{(1)} \\ q^{(1)} &\in [d_1-d_2-1, d_1-1] \end{aligned} \tag{4.75}$$

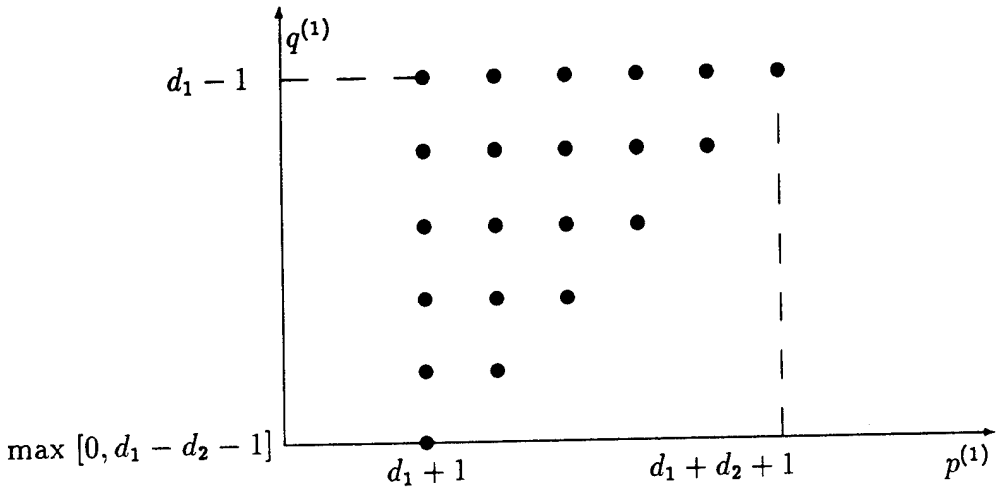

Figure 4.2: The multiplicities of the poles and zeros in $s_{d_1}(z;\rho) * s_{d_2}(z;\rho)$.

Fig. 4.2 shows the values assumed by $p^{(1)}$ and $q^{(1)}$ in (4.72) when convolving $s_{d_1}(z;\rho)$ and $s_{d_2}(z)$ for $d_2 \leq d_1$. From this plot we note that the minimum value of the difference $p^{(1)} - q^{(1)}$ is given by

$$\begin{aligned} \min(p^{(1)} - q^{(1)}) &= \min(p^{(1)}) - \max(q^{(1)}) \\ &= (d_1 + 1) - (d_1 - 1) = 2 \end{aligned} \tag{4.76}$$

Now let us consider what happens when the spectral term $s^{(2)}_{d_1,d_2}(z) = s_{d_1}(z) * s_{d_2}(z)$ is convolved with yet another cross-spectral function $s_{d_3}(z)$. We will show that the general structure of the problem observed so far is maintained. The convolution of these cross-spectral functions can be written as

$$s_{d_3}(z) * s^{(2)}_{d_1,d_2}(z) = \sum_{k=0}^{d_1} \sum_{m_3=0}^{\min[k,d_1-1]} b^{(1)}(k, m_3) s_{d_3}(z;\rho) * \mathcal{X}^{(1)}(z; p^{(1)}, q^{(1)}) \tag{4.77}$$

where each term under the summation involves a convolution given by

$$
\begin{aligned}
s_{d_3}(z;\rho) * \mathcal{X}^{(1)}(z;p^{(1)},q^{(1)}) &= \frac{1}{2\pi j}\oint_{C_1} v^{-1} s_{d_3}(z;\rho)\mathcal{X}^{(1)}(\frac{z}{v};p^{(1)},q^{(1)})dv \\
&= z(-1)^{p^{(1)}-q^{(1)}+1}\rho^{d_3}\frac{(1-\rho^2)}{\rho} \\
&\quad \times \frac{1}{2\pi j}\oint_{C_1} \frac{v^{p^{(1)}-q^{(1)}-1}(v-\rho^{-1})^{d_3-1}(v-z)^{q^{(1)}}}{(v-\rho)^{d_3+1}(v-z\rho^{-2})^{p^{(1)}}}dv
\end{aligned}
\tag{4.78}
$$

Since (4.76) establishes that $p^{(1)} - q^{(1)} \geq 1$, the integrand has at least one zero at the origin and $d_3 + 1$ poles inside the unit circle at ρ. Thus we can find an expression for the convolution (4.78) similar to that of (4.72), namely

$$
s_{d_3}(z;\rho) * \mathcal{X}^{(1)}(z;p^{(1)},q^{(1)}) = \sum_{k=0}^{d_3}\sum_{m_3=0}^{\min[q^{(1)},d_3-k]} b^{(2)}(k,m_3)\mathcal{X}^{(2)}(z;p^{(2)}(k,m_3),q^{(2)}(m_3))
\tag{4.79}
$$

where the function $\mathcal{X}^{(2)}$ is defined as

$$
\mathcal{X}^{(2)}(z;p,q) = \frac{\rho^{3p}}{\rho^q}\frac{z(z-\rho)^q}{(z-\rho^3)^p}
\tag{4.80}
$$

and where the remaining quantities in (4.79) are given by

$$
\begin{aligned}
p^{(2)} &= p^{(1)} + d_3 - k - m_3 \\
q^{(2)} &= q^{(1)} - m_3 \\
b^{(2)}(k,m_3) &= (-1)^{m_3}\begin{pmatrix} q^{(1)} \\ m_3 \end{pmatrix}\begin{pmatrix} p^{(1)}-1+d_3-k-m_3 \\ d_3-k-m_3 \end{pmatrix} W_{d_3}^{(2)}(k,p^{(1)}-q^{(1)}) \\
W_{d_3}^{(2)}(k,t) &= \sum_{m_2=k-t+1}^{\min[k,d_3-1]} \rho^{(d_3-1)}\begin{pmatrix} d_3 \\ k \end{pmatrix}(\rho^{-2}-1)^{(d_3-m_2)}\begin{pmatrix} d_3-m_2 \\ m \end{pmatrix} \\
&\quad \times \frac{\begin{pmatrix} k \\ m_2 \end{pmatrix}\begin{pmatrix} d_3-k \\ d_3-m_2 \end{pmatrix}}{\begin{pmatrix} t-1+m_2-k \\ t-1 \end{pmatrix}}
\end{aligned}
\tag{4.81}
$$

The function $\mathcal{X}^{(2)}$ has one zero at the origin, $q^{(2)}$ zeros at ρ. and $p^{(2)}$ poles at ρ^3. The limits of summations in (4.79) determine the ranges of $p^{(2)}$ and $q^{(2)}$ or equivalently

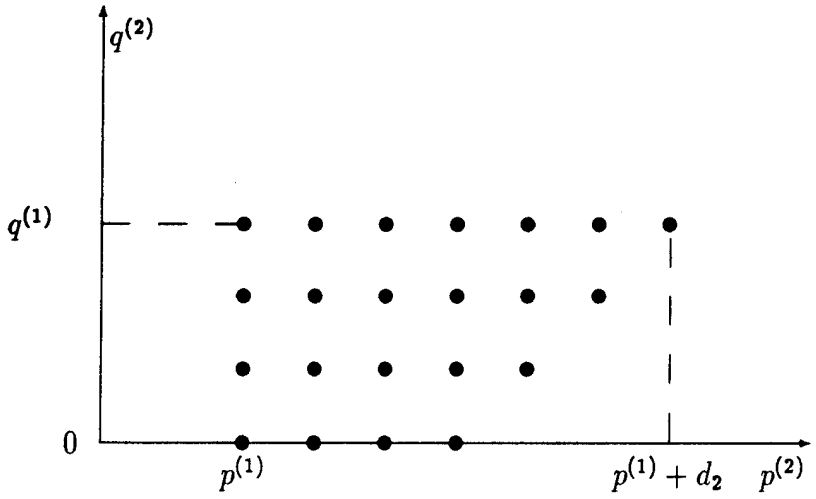

Figure 4.3: The range of the functions $\mathcal{X}^{(2)}$ as a result of $s_{d_3}(z;\rho) * \mathcal{X}^{(1)}$.

the set of spectral functions $\mathcal{X}^{(2)}$. For each value of $p^{(1)}$ and $q^{(1)}$ we have

$$\begin{aligned} p^{(2)} &= p^{(1)}, p^{(1)}+1, \cdots, p^{(1)}+d_3+q^{(2)}-q^{(1)} \\ q^{(2)} &= 0, 1.2, \cdots, q^{(1)} \end{aligned} \tag{4.82}$$

Fig. 4.3 shows the region of pole and zero multiplicities of the functions $\mathcal{X}^{(2)}$ that result from convolving $s_{d_3}(z)$ with a single function $\mathcal{X}^{(1)}$ in (4.77). To define all the functions $\mathcal{X}^{(2)}$ that result from the convolution we combine (4.74) and (4.82). Then the entire set of spectral functions $\mathcal{X}^{(2)}$ has pole and zero multiplicities in the region defined by

$$\begin{aligned} p^{(2)} &= d_1+1, d_1+2.\cdots, d_1+d_3+2+q^{(2)} \\ q^{(2)} &= 0.1.2.\cdots.d_1-1 \end{aligned} \tag{4.83}$$

This region is shown in Fig. 4.4.

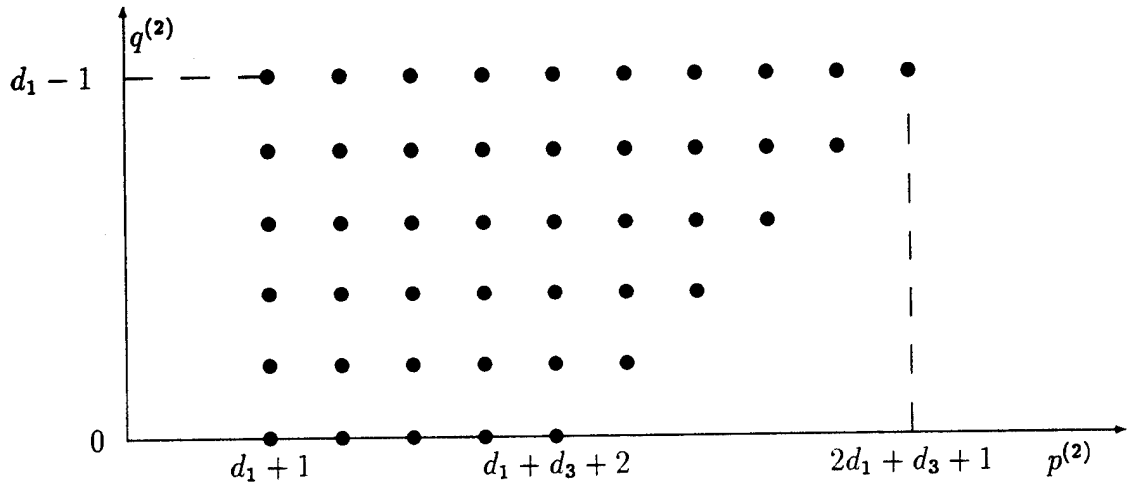

Figure 4.4: The range of the functions $\mathcal{X}^{(2)}$ as a result of $s_{d_3}(z) * s^{(2)}_{d_1}(z)$.

c. Convolution of Any Number of Laguerre Cross-Spectral Functions

To generalize the results of the previous subsections, let us assume that we need to perform convolutions of the Laguerre cross-spectral functions of orders $d_1, d_2, \cdots, d_{\kappa+1}$ such that

$$d_1 \geq d_2 \geq \cdots \geq d_{\kappa+1} \geq 0$$

We can denote this by

$$s^{(\kappa+1)}_{\mathbf{d}}(z) = \underbrace{s_{d_1}(z) * s_{d_2}(z) * s_{d_3}(z) * \cdots * s_{d_{\kappa+1}(z)}}_{\kappa \text{ convolutions}} \tag{4.84}$$

According to the above discussions we could convolve $s_{d_1}(z)$ with $s_{d_2}(z)$ then convolve the resulting spectral function with $s_{d_3}(z)$ and so on until all the required κ convolutions are performed. In so doing we can benefit from the above relations to establish an efficient procedure to obtain the result. We proceed by induction.

Let us define the spectral function $\mathcal{X}^{(\nu)}(z;p,q)$ as one of the spectral functions that results from performing ν convolutions. This spectral function has the general form

$$\begin{aligned}\mathcal{X}^{(\nu)}(z;p,q) &= \frac{\rho^{p(\nu+1)}}{\rho^{q(\nu-1)}}\frac{z(z-\rho^{\nu-1})^q}{(z-\rho^{\nu+1})^p} \\ &= \frac{z(z\rho^{-(\nu-1)}-1)^q}{(z\rho^{-(\nu+1)}-1)^p} \end{aligned} \tag{4.85}$$

It is seen in the previous subsections that when a function of this form is convolved with one of the Laguerre cross-spectral density functions, a weighted sum of spectral functions of the same form results. Thus the convolution of $s_d(z;\rho)$ with $\mathcal{X}^{(\nu)}(z;p^{(\nu)},q^{(\nu)})$ produces an expression that can be written as

$$\begin{aligned}&s_d(z;\rho) * \mathcal{X}^{(\nu)}(z;p^{(\nu)},q^{(\nu)}) \\ &= \sum_{k=0}^{d}\sum_{m_3=0}^{\min[q^{(\nu)},d-k]} b^{(\nu+1)}(k,m_3)\mathcal{X}^{(\nu+1)}(z;p^{(\nu+1)}(k,m_3),q^{(\nu+1)}(m_3))\end{aligned} \tag{4.86}$$

where

$$\begin{aligned}p^{(\nu+1)} &= p^{(\nu)} + d - k - m_3 \\ q^{(\nu+1)} &= q^{(\nu)} - m_3 \\ b^{(\nu+1)}(k,m_3) &= (-1)^{m_3}\binom{q^{(\nu)}}{m_3}\binom{p^{(\nu)}-1+d-k-m_3}{d-k-m_3} W_d^{(\nu+1)}(k,p^{(\nu)}-q^{(\nu)}) \\ W_d^{(\nu+1)}(k,t) &= \sum_{m_2=k-t+1}^{\min[k,d-1]} (\rho^{-2}-1)^{(d-m_2)}\left(\frac{d-m_2}{d}\right) \\ &\qquad \times \frac{\binom{k}{m_2}\binom{d-k}{d-m_2}}{\binom{t-1+m_2-k}{t-1}}\end{aligned} \tag{4.87}$$

where the difference between the pole and zero multiplicities in $\mathcal{X}^{(\nu)}$ satisfies

$$p^{(\nu)} - q^{(\nu)} \geq 2$$

The set of spectral functions $\mathcal{X}^{(\nu+1)}$ that result in the convolution (4.86) differ only in the multiplicities $p^{(\nu+1)}$ and $q^{(\nu+1)}$ which are determined by the limits of summation in (4.86). The ranges of $p^{(\nu+1)}$ and $q^{(\nu+1)}$ or equivalently the spectral functions $\mathcal{X}^{(\nu+1)}$ are obtained from the order d of the Laguerre cross-spectral function $s_d(z)$ and the pole and zero multiplicities of $\mathcal{X}^{(\nu)}$ as follows

$$p^{(\nu+1)} = p^{(\nu)}, p^{(\nu)}+1, \cdots, p^{(\nu)}+d+q^{(\nu+1)}-q^{(\nu)}$$
$$q^{(\nu+1)} = 0, 1, 2, \cdots, q^{(\nu)} \tag{4.88}$$

Now let us consider in detail the entire convolution (4.84). This can be obtained by following a set of successive steps that give the value of this convolution in an organized algebraic procedure. The value of this convolution is a sum of spectral functions $\mathcal{X}^{(\alpha)}$ of the form

$$s_{\mathbf{d}}^{(\kappa+1)}(z) = \sum_i c_i^{(\kappa)} \mathcal{X}^{(\kappa)}(z; p_i, q_i) \tag{4.89}$$

where

$$\mathcal{X}^{(\kappa)}(z; p_i, q_i) = \frac{\rho^{p_i(\kappa+1)}}{\rho^{q_i(\kappa-1)}} \frac{z(z-\rho^{\kappa-1})^{q_i}}{(z-\rho^{\kappa+1})^{p_i}} \tag{4.90}$$

It is clear that *the set of zero and pole locations are determined simply from the* <u>*number*</u> *of convolutions*. The value of the convolution is obtained by determining the values of p_i, q_i and c_i based on the orders of the Laguerre spectral functions $d_1, d_2, \cdots, d_{\kappa+1}$. We can make the following observations:

1. If this convolution were performed successively then as a result of the first convolutions, $s_{d_1}(z) * s_{d_2}(z)$ we would have pole and zero multiplicities determined by a plot similar to that of Fig. 4.2. Each of the resulting functions $\mathcal{X}^{(1)}(z)$ is then convolved with $s_{d_3}(z)$. The pole zero multiplicities in the resulting spectral function $\mathcal{X}^{(2)}(z)$ are obtained by extending those in Fig. 4.2 using the extension map described by Fig. 4.3 for the Laguerre spectral function order

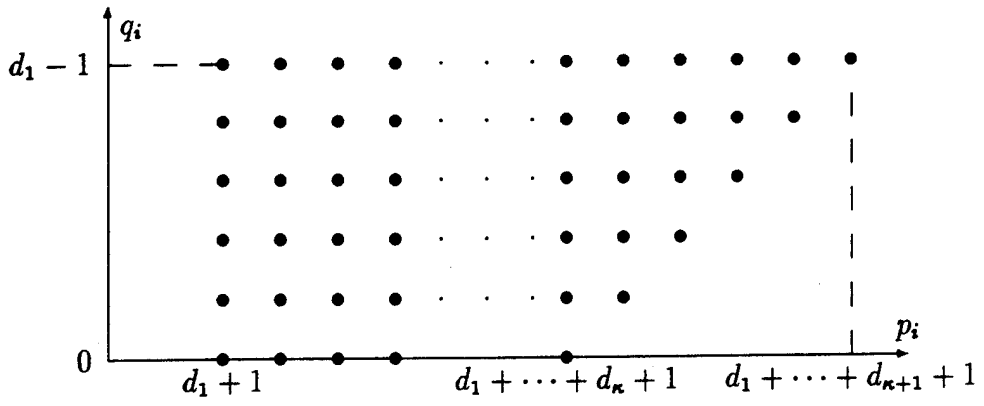

Figure 4.5: The zero and pole multiplicities in $s_{\mathbf{d}}^{\kappa+1}(z)$.

m_3. This results in a range of the pole and zero multiplicities described by a new plot similar to that of Fig. 4.4. This procedure is continued until the last extension map corresponding to the Laguerre spectral function order $m_{\kappa+1}$. We can obtain the same result directly however by taking the region of p_i and q_i in (4.89) to be as follows:

- The maximum value of zero multiplicity is $d_1 - 1$.
- The minimum value of zero multiplicity is $\max[0. d_1 - d_2 - \cdots - d_{\kappa+1}]$.
- The maximum value of pole multiplicity is $d_1 + d_2 + \cdots + d_{\kappa+1} + 1$.
- the minimum value of pole multiplicity is $d_1 + 1$.

We can also use a plot similar to that in Fig. 4.5.

2. The coefficients $c_i^{(\kappa)}$ in (4.89) can also be obtained by following a recursive operation. Let us assume that after the first ν convolutions a set of coefficients $c_i^{(\nu)}$ were computed and associated with the corresponding p_i and q_i. For the $\nu + 1^{st}$ convolution, involving $s_{d_{\nu+2}}$, the extension map in Fig. 4.3 for $d_{\nu+2}$ is used to determine the multiplicity pairs $[p^{(\nu+1)}, q^{(\nu+1)}]$ that result from the pairs $[p^{(\nu)}, q^{(\nu)}]$. Then if from the extension map the $\nu + 1^{st}$ convolution results in generating $[p_j^{(\nu+1)}, q_j^{(\nu+1)}]$ from $[p_i^{(\nu)}, q_i^{(\nu)}]$ then the coefficient $c_i^{(\nu)}$ is modified

to $c_j^{(\nu+1)}$ using the multiplier $b^{(\nu+1)}$ defined in (4.87). This multiplier is now written as by $b^{(\nu+1)}(i,j)$ to indicate that it corresponds to the j^{th} multiplicity pair generated from the i^{th} pair after $\nu+1$ convolutions. Thus the coefficient $c_j^{(\nu+1)}$ is updated according to

$$c_j^{(\nu+1)} = c_j^{(\nu)} + b^{(\nu+1)}(i,j)c_i^{(\nu)} \tag{4.91}$$

Finally, for each value of the vectors of indices $\boldsymbol{\alpha}$, the Q-spectral function can be obtained as follows:

1. The multiple convolution of the Laguerre spectral functions which are the entries of one block $\mathbf{S}^{\omega}$ of the matrix $\boldsymbol{\mathcal{S}}$ given in (2.106) is obtained as a weighted sum of the spectral functions $\mathcal{X}^{(\alpha)}(z;p_i,q_i)$ in the form of (4.89). The poles of these spectral functions are located at $\rho^{(\nu+1)}$ while the zeros are located at $\rho^{(\nu-1)}$ where ν equals the sum of the entries of the block $\mathbf{M}^{\omega}$ of $\mathbf{M}$ that corresponds to the block $\mathbf{S}^{\omega}$ of $\boldsymbol{\mathcal{S}}$.

2. The range of the pole multiplicities p_i, the zero multiplicity q_i, and the coefficients c_i in (4.89) are determined as described above from the orders of the Laguerre cross-spectral functions in $\mathbf{S}^{\omega}$ for which the corresponding entries in $\mathbf{M}^{\omega}$ have non-zero values.

d. Convolution of $s_0^{(m)}(z;\rho)$ with $s_{\mathbf{d}}^{(\kappa+1)}(z;\rho)$

To obtain the value of the multiple convolution of the Laguerre cross-spectral functions in the block $\mathbf{S}^{\omega}$ we first convolve the Laguerre cross-spectral function $s_0(z;\rho)$ a number of times m_0 corresponding to the sum of the entries on the principal diagonal of the block $\mathbf{M}^{\omega}$ of the multiplicity matrix $\mathbf{M}$. The result of this self convolution is given by (4.60). We also perform the convolution on the rest

of the Laguerre cross-spectral functions (residing on the non-principal diagonals of $\mathbf{S}^{\omega}$ the value of this convolution is given by (4.89). Thus the value of the required convolution is then given by

$$\begin{aligned} S^{\omega}_{\mathbf{conv}} &= s_0^{(m_0)}(z;\rho) * s_{\mathbf{d}}^{(\kappa+1)}(z) \\ &= \sum_i c_i^{\kappa} s_0^{(m_0)}(z;\rho) * \mathcal{X}^{(\kappa)}(z;p_i,q_i) \end{aligned} \tag{4.92}$$

The terms under summation in this expression are obtained by convolving $s_0^{(m_0)}(z;\rho)$ with each of the spectral functions $\mathcal{X}^{(\kappa)}(z;p_i,q_i)$. This convolution has the form

$$\begin{aligned} s_0^{(m_0)}(z;\rho) * \mathcal{X}^{(\kappa)}(z;p_i,q_i) &= \frac{1}{2\pi j}\oint_{C_1} v^{-1} s_0^{(m_0)}(v;\rho)\mathcal{X}^{(\kappa)}\left(\frac{z}{v};p_i,q_i\right)dv \\ &= -z\frac{(1-(\rho^{m_0})^2)}{\rho^{m_0}}\frac{1}{2\pi j}\oint_{C_1}\frac{(-1)^{p_i-q_i}v^{(p_i-q_i-1)}(v-z\rho^{-(\kappa-1)})^{q_i}}{(v-\rho^{m_0})(v-\rho^{-m_0})(v-z\rho^{-(\kappa+1)})^{p_i}}dv \end{aligned} \tag{4.93}$$

Since the integrand has only one pole inside the unit circle at location ρ^{m_0}, the value of the convolution in (4.93) is easily obtained as

$$s_0^{(m_0)}(z;\rho) * \mathcal{X}^{(\kappa)}(z;p_i,q_i) = \frac{\rho^{p_i(m_0+\kappa+1)}}{\rho^{q_i(m_0+\kappa-1)}}\frac{z(z-\rho^{m_0+\kappa-1})^{q_i}}{(z-\rho^{m_0+\kappa+1})^{p_i}} \tag{4.94}$$

Notice that this convolution results in one term only. Moreover the pole and zero multiplicities of $\mathcal{X}^{(\kappa)}$ do not change and the value of this convolution can be easily obtained by just moving the pole and the zero from $\rho^{\kappa+1}$ and $\rho^{\kappa-1}$ to $\rho^{m_0+\kappa+1}$ and $\rho^{m_0+\kappa-1}$ respectively. More precisely from (4.94) and the definition (4.85) we have

$$s_0^{(m_0)}(z;\rho) * \mathcal{X}^{(\kappa)}(z;p_i,q_i) = \mathcal{X}^{(\kappa+m_0)}(z;p_i,q_i) \tag{4.95}$$

V. MODELING A NON-GAUSSIAN RANDOM PROCESS USING THE WIENER MODEL

In Chapter IV we develop the procedure for computing the cumulant and polyspectrum (of any order) of the output of the Wiener model. These higher-order statistics are shown to be functions of the model parameters defined in Chapter III. The first such model parameter is σ_o^2 the variance of the zero-mean white Gaussian process $w(n)$ that drives the system. This parameter specifies the set of Hermite polynomials that describe the memoryless nonlinearities represented by section 2 of the model. The second parameter, ρ, defines the pole and zero locations of the transfer functions of the linear filters comprising section 1 of the model. The impulse responses of these filters are the discrete Laguerre functions and the outputs are zero-mean Gaussian processes each with variance σ_o^2. Thus. the auto- and cross-correlation functions of the outputs of these filters (and in fact all of the higher-order statistics) are completely specified by σ_o^2 and ρ. The remaining parameters are the coefficients $c_{\boldsymbol{\alpha}}$ in the Q-polynomial representation of the output of the model (3.34). According to our definition of the indices represented by the vector $\boldsymbol{\alpha}$, these parameters are uniquely determined once the model dimensions N_L and N_N are specified. Thus for a given set of model dimensions the number of the model parameters is finite and the expressions of the model output statistics can be obtained.

In this chapter we demonstrate the application of the Wiener model in representing some non-Gaussian discrete random processes that fail to meet the conditions of linearity. The parameters of the discrete Wiener model are determined such that the model output statistics match the modeled process statistics up to some given order over a specific region of support.

A. NONLINEARITY DETECTION IN RANDOM PROCESSES

The statistics of a linear random process either in the time domain (cumulants) or in the frequency domain (polyspectra) have characteristics that nonlinear processes do not share. A significant body of work has been reported to detect and characterize the nonlinearities in time series [4, 33, 5].

In [34] Lawrance and Lewis (1987) present a class of stationary random processes $x(n)$ for which the autocovariance function, $C_x^{(2)}(l)$, has an autoregression structure, i.e., this function satisfies the p^{th} order Yule-Walker autoregression

$$C_x^{(2)}(l) = a_1 C_x^{(2)}(l-1) + a_2 C_x^{(2)}(l-2) + \cdots + a_p C_x^{(2)}(l-p) \tag{5.1}$$

where a_i are appropriate real-valued constants. This class of processes includes (among others) two subclasses of stationary random processses with mean μ that have an autoregerssive representation. The first is the class of processes for which the linear conditional expectation has the form

$$\mathcal{E}[x(n) - \mu | x(n-1), x(n-2), \cdots, x(n-p)] = a_1(x(n-1) - \mu) + a_2(x(n-2) - \mu) + \cdots + a_p(x(n-p) - \mu) \tag{5.2}$$

Processes in this class may be nonlinear. The second class, which is a subset of the class satisfying (5.2), is that of linear stationary random processes with mean μ which satisfy

$$x(n) - \mu = a_1(x(n-1) - \mu) + a_2(x(n-2) - \mu) + \cdots + a_p(x(n-p) - \mu) + w(n) \tag{5.3}$$

where $w(n)$ are independent and identically-distributed (i.i.d.) and the coefficients a_i are real-valued constants satisfying the condition that the linear autoregressive model represented by (5.3) is stable.

The order p linear autoregressive residual $\epsilon^{(p)}(n)$ is given by the process $x(n)$ after subtracting its best linear least-square predictor in terms of $x(n-1)$, $x(n-2), \cdots, x(n-p)$, i.e.,

$$\epsilon^{(p)}(n) = x(n) - \mu - a_1(x(n-1)-\mu) - a_2(x(n-2)-\mu) - \cdots - a_p(x(n-p)-\mu) \quad (5.4)$$

Lawrance and Lewis [34] based the principle of detecting the nonlinearity in processes with autoregressive correlation structure (5.1) on the following:

1. For processes satisfying (5.1) the linear autoregressive residuals, $\epsilon^{(p)}$, are uncorrelated.

2. For linear processes satisfying (5.3) the linear AR residuals $\epsilon^{(p)}$ are not only uncorrelated but they are also *independent*.

According to this distinction, the higher-order dependency of the residuals is investigated by considering the two cross-correlation quantities :

$$\begin{aligned} &\mathrm{Corr}[(x(n)-\mu)^2, \epsilon^{(p)}(n+l)] \\ &\mathrm{Corr}[(x(n)-\mu), \left(\epsilon^{(p)}(n+l)\right)^2] \end{aligned} \quad (5.5)$$

for $l = 0, \pm 1, \pm 2, \cdots$. If a process satisfying (5.1) is linear (i.e satisfying (5.3)) then the two cross-correlation functions in (5.5) are zero for all $l \neq 0$. This test is a useful one, but it is based on a particular class of processes defined by (5.1).

A more general test for nonlinearity was formulated by Rao and Gabr (1980) [35] and later extended by Hinich (1982) [36]. This test is based on the fact that if a process $x(n)$ is linear (in mean square), then there exists a sequence $h(n)$ such that $x(n)$ is represented by

$$x(n) = \sum_{k=0}^{\infty} h(k)w(n-k) \quad (5.6)$$

where $w(n)$ are i.i.d. and the sequence $h(n)$ is square summable . In this case, given the process bispectrum $S_x^{(3)}(\omega_1, \omega_2)$ and power spectrum $S_x^{(2)}(\omega)$, it is easily shown that the ratio

$$\frac{|S_x^{(3)}(\omega_1, \omega_2)|^2}{S_x^{(2)}(\omega_1) S_x^{(2)}(\omega_2) S_x^{(2)}(\omega_1 + \omega_2)} \tag{5.7}$$

is a constant independent of ω_1 and ω_2. In practical applications estimates of the process bispectrum and power spectrum functions are used to test for its linearity.

The linearity test of Hinich actually follows from the fact that the k^{th}-order polyspectrum of a process represented by (5.6) is has the form

$$S_x^{(k)}(\omega_1, \omega_2, \cdots, \omega_{k-1}) = \beta_k H(\omega_1) H(\omega_2) \cdots H(\omega_{k-1}) H(-\omega_1 - \omega_2 - \cdots - \omega_{k-1}) \tag{5.8}$$

where β_k is a nonzero constant and $H(\omega)$ is the Fourier transform of $h(n)$. In other words, linear processes have the characteristic that their polyspectra are *factorable*: this is a generalization of the correlation-based innovations representation of random processes. The representation (2.43) is possible if the process *power spectral density function* is factorable, i.e., if it satisfies the Paley-Wiener condition. For a rational power spectral density, the procedure developed by Wiener [1] can be used to determine the poles and zeros of the linear system transfer function. The factorization of polyspectra however is not as simple.

Tekalp and Erdem (1989) [13] established necessary and sufficient conditions for the existence of a stable linear time-invariant system driven by a non-Gaussian k^{th}-order stationary white process such that the system output k^{th}-order polyspectrum matches a given one. The procedure is outlined briefly below.

The k^{th}-order *complex cepstrum* of a k^{th}-order stationary random process $x(n)$ is defined as the multidimensional inverse z-transform of the logarithm of the complex polyspectrum $S_x^{(k)}(z_1, z_2, \cdots, z_{k-1})$

$$c_x^{(k)}(l_1, l_2, \cdots, l_{k-1}) = \mathcal{Z}^{-1}\left\{\log[S_x^{(k)}(z_1, z_2, \cdots, z_{k-1})]\right\} \tag{5.9}$$

The region of convergence is assumed to contain the k-dimensional unit circle $|z_1| = |z_2| = \cdots = |z_k| = 1$, and the integration is performed over that contour. For a discrete random process that has the representation (5.6) its k^{th}-order complex cepstrum is given by

$$c_x^{(k)}(l_1, l_2, \cdots, l_{k-1}) = c_w^{(k)}(l_1, l_2, \cdots, l_{k-1}) + \underbrace{g(-l_i) \prod_{\substack{j=1 \\ j \neq i}}^{k-1} \delta(l_j - l_i)}_{\text{for any } i} + \sum_{i=1}^{k-1} g(l_i) \prod_{\substack{j=1 \\ j \neq i}}^{k-1} \delta(l_j) \tag{5.10}$$

where $\delta(l)$ is the unit sample function and

$$g(l) = \mathcal{Z}^{-1}\{\log[H(z)]\} \tag{5.11}$$

From (2.43) the complex cepstrum of $x(n)$ is also given by

$$c_x^{(k)}(l_1, l_2, \cdots, l_{k-1}) = \log[\beta_k] \prod_{j=1}^{k-1}(l_j) + \underbrace{g(-l_i) \prod_{\substack{j=1 \\ j \neq i}}^{k-1} \delta(l_j - l_i)}_{\text{for any } i} + \sum_{i=1}^{k-1} g(l_i) \prod_{\substack{j=1 \\ j \neq i}}^{k-1} \delta(l_j) \tag{5.12}$$

Equation (5.12) implies that the region of support of the complex cepstrum of $x(n)$ is the union of the regions $\mathcal{R}_i$ defined as

$$\mathcal{R}_i = \{(l_1, l_2, \cdots, l_{k-1}) : -\infty < l_i < \infty, l_j = 0 \text{ for } j = 1, 2, \cdots, k-1 \text{ and } j \neq i,$$

and

$$i = 1, 2, \cdots, k-1 \tag{5.13}$$

$$\mathcal{R}_k = \{(l_1, l_2, \cdots, l_{k-1}) : -\infty < l_1 = l_2 = \cdots = l_{k-1} < \infty\} \tag{5.14}$$

Tekalp and Erdem proved that a necessary and sufficient condition for the process $x(n)$ to have the linear representation (5.6) is that its complex cepstrum is zero outside the region of support

$$\bigcup_{i=1}^{k} \mathcal{R}_i$$

If this condition is not satisfied the polyspectrum is not factorable and there does not exist a linear time-invariant system which when driven by a white process gives the required output polyspectrum. In this case a nonlinear model needs to be considered.

The likelihood of encountering a factorable polyspectrum has also been studied by Tekalp and Erdem [12]. They proved that the subset of factorable polyspectra has measure zero in the set of polynomial polyspectra that correspond to finite-support cumulants. Thus from a practical point of view the linear non-Gaussian model can be used only as an approximation when this is justified.

B. THE DISCRETE STOCHASTIC PROCESSES REPRESENTABLE BY THE WIENER MODEL

The structure of the discrete Wiener model (in Chapter III) together with the analysis of the model's cumulants and polyspectra (Chapter IV) provide the following characterization:

1. The linear stage of the model, section 1, is represented by a bank of causal, stable linear filters whose impulse responses are the discrete Laguerre functions. The transfer functions of these filters, $\Lambda_i(z)$, are analytical functions with no zeros on the unit circle. The outputs of this stage are zero-mean Gaussian random variables and their cross-correlation functions are given by the Laguerre cross-correlation sequences $r_d(l)$. Thus, the complex cross-spectral functions of these outputs are given by the Laguerre cross-spectral functions $s_d(z)$ which are also analytical functions with no zeros on the unit circle.

2. The nonlinearities in the model are represented by the Q-polynomials which are a wieghted sum of multinomials of the outputs of section 1. Since the model output is a sum of products of correlated Gaussian variables, the output cumulant is computable as a specific sum of products of the Laguerre correlation

sequences $r_d(l)$. The output polyspectrum is given by a sum of multiple convolutions of the Laguerre complex cross-spectral functions $s_d(z)$. This convolution was shown to be representable as a sum of spectral functions $\mathcal{X}^{(\alpha)}(z_i; p_j, q_j)$ which are also analytical on the multidimensional unit circle.

While sufficient conditions that a random process must satisfy to be representable by the Wiener model are not known, certain necessary conditions are evident from the above characterization:

1. The cumulants of any order of a processes to be representable by the Wiener model *must asymptotically approach zero* as any of the time lag arguments increases. This follows because output cumulants of any order of the model are obtained as products of the Laguerre cross-correlation sequences $r_d(l)$ and all these sequences asymptotically approach zero as their lag arguments increase.

2. Unless the polyspectrum is zero everywhere, as in the case of a Gaussian process, a process to be represented by the Wiener model must not be *bandlimited*. a bandlimited process is here defined as one for which at least one polyspectrum of order k is zero over a nonzero-volume in the k-dimensional subspace that defines its region of support.

3. The process to be represented by the Wiener model must not be periodic or almost periodic [37]. Since the model polyspectrum is represented with spectral functions that have the same form and all have multiple poles at the same location, these spectral functions do not have any singularities on the multidimensional unit circle. Therefore these functions cannot produce any line spectral components.

C. DISCRETE PROCESS MODELING BY CUMULANT MATCHING

For a given model dimensions N_L and N_N the output of the Wiener model is given by

$$x(n) = \sum_{\boldsymbol{\alpha}_i} \mathbf{Q}\boldsymbol{\alpha}_i \tag{5.15}$$

and the k^{th}-order output cumulant is given by

$$C_x^{(k)}(\boldsymbol{l}) = \sum_{\boldsymbol{\alpha}} c_{\boldsymbol{\alpha}} C_{\boldsymbol{\alpha}}(\boldsymbol{l}) \tag{5.16}$$

where each $\boldsymbol{\alpha}$ vector is a concatenation of the k $\boldsymbol{\alpha}_i$ vectors and the summation is only over those combinations for which the $C_x^{(k)}$ are non-zero. Let us define the vector $\mathbf{c}$ as the vector of N coefficients appearing in (5.15) and ordered with indices in ascending order, where order for the indices $\boldsymbol{\alpha}_i$ is defined by (4.3). Thus

$$\mathbf{c} = [c_{\boldsymbol{\alpha}_1}, c_{\boldsymbol{\alpha}_2}, \cdots, c_{\boldsymbol{\alpha}_N}]^T \tag{5.17}$$

where here the variables $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots$ and $\boldsymbol{\alpha}_N$ represent *specific* vector of indices of the N terms that appear in (5.15). Further define the vector of cumulants

$$\mathbf{c}_x^{(k)}(\boldsymbol{l}) = [C_{\boldsymbol{\alpha}^{(1)}}^{(k)}(\boldsymbol{l})\ C_{\boldsymbol{\alpha}^{(2)}}^{(k)}(\boldsymbol{l})\ \cdots\ C_{\boldsymbol{\alpha}^{(k\times N)}}^{(k)}(\boldsymbol{l})]^T \tag{5.18}$$

where the variables $\boldsymbol{\alpha}^{(1)}, \boldsymbol{\alpha}^{(2)}, \cdots, \boldsymbol{\alpha}^{(k\times N_L)}$ represent the *paritcular* larger vectors of indices $\boldsymbol{\alpha}$ that appear in (5.16). Then (5.16) can be written using matrix notation as

$$C_x^{(k)}(\boldsymbol{l}) = (\mathbf{c}_x^{(k)}(\boldsymbol{l}))^T \mathbf{c}^{\otimes k} \tag{5.19}$$

where

$$\mathbf{c}^{\otimes k} = \underbrace{\mathbf{c} \otimes \mathbf{c} \otimes \cdots \otimes \mathbf{c}}_{k \text{ times}} \tag{5.20}$$

and $\otimes$ denotes the direct (Kronecker) product of vectors.

If a set of specific lag values $\boldsymbol{l}^1, \boldsymbol{l}^2, \cdots, \boldsymbol{l}^M$ are now chosen, then a vector of these values can be formed as

$$\mathbf{X}_m = \mathcal{C}^{(k)} \mathbf{c}^{\otimes k} \tag{5.21}$$

where

$$\mathcal{C}^{(k)} = \begin{bmatrix} \mathbf{c}_x^{(k)T}(\boldsymbol{l}^1) \\ \mathbf{c}_x^{(k)T}(\boldsymbol{l}^2) \\ \vdots \\ \mathbf{c}_x^{(k)T}(\boldsymbol{l}^M) \end{bmatrix} \tag{5.22}$$

The entries in each row of the matrix $\mathcal{C}^{(k)}$ are the Q-polynomial cross-cumulants which are functions of the parameter ρ and the input variance σ_o^2. The rows differ in their time lag vector.

Let us observe that although the quantities σ_o^2 and ρ are parameters of the system, the theory provides a model using any finite positive real value for σ_o^2 and any ρ satisfying $|\rho| < 1$. Although the variance σ_o^2 is essential for ensuring the orthogonality of the Hermite polynomials, once its value is chosen, the scale of the output can be adjusted by appropriate scaling of the $c_{\boldsymbol{\alpha}_i}$ parameters. Therefore in our modeling, we fix the value of σ_o^2 at 1 and do not attempt to adjust it.

Although in theory the model could be applied with any fixed value of ρ with magnitude less than 1, in practice, because the model dimension N_L is finite, the value of ρ has a significant effect on the error involved in matching a given set of higher-order statistics. Therefore, when developing the Wiener model for a random process, we include both the parameter ρ and the coefficients $c_{\boldsymbol{\alpha}_i}$ in spite of the fact that there could be some redundancy.

1. Model Parameter Identification Using the Extended Kalman Filter Algorithm

Given an appropriate stochastic process represented by a set of measurement data, we seek to model it as the output of a Wiener discrete nonlinear system. In this

case we estimate the process cumulants over defined regions of support and arrange these estimates in a vector $\mathbf{X}_p$. The vector $\mathbf{X}_m$ in (5.21) is then defined over the same region of support. Let us define the modeling error $\mathbf{E}$ as

$$\mathbf{E} = \mathbf{X}_p - \mathbf{X}_m \tag{5.23}$$

We seek to determine the values of the model parameters that minimize the squared magnitude $\mathbf{E}^{*T}\mathbf{E}$. Since the components of the vector $\mathbf{X}_m$ are nonlinear functions of the model parameters, a technique for solving a nonlinear optimization problem is required. The extended Kalman filter algorithm was used because it was found to converge to a solution faster than other techniques tested (e.g., the steepest descend method). During the optimization we need to know the cumulant vector $\mathbf{X}_m$ and its gradient with respect to the model parameters; therefore it is necessary to have an efficient method to compute these arrays. In Appendix D we describe a procedure that enables one to compute both the cumulant vector and its gradient in a type of look-up table that saves signficant computation and storage. The extended Kalman filter algorithm, which is used in the optimization, is also described in Appendix **D**.

2. Simulation Results

In this section we present the results of two experimental examples. In the first example the procedure developed in this research is applied to model a synthetically generated set of data. In the second example the procedure is applied to a data set obtained from a biological measurement record. We refer to these data sets as the "original data." In both examples the vector $\mathbf{X}_p$ that represents the original data cumulant values consists of the estimates made from the data of both the second- and third-order cumulants in the nonredundant regions of support. The vector $\mathbf{X}_m$ representing the model output cumulants is constructed for the same regions of support. Note that cumulant values used in this vector are computed using

the expressions developed in the dissertation, i.e., they are not estimated from the model output. The extended Kalman filter algorithm is then applied to identify the model parameters minimizing the squared magnitude of the modeling error. After the model parameters are obtained, they are used to generate a data sequence from the Wiener model using these parameters, i.e., the model is driven by a computer generated white Gaussian sequence to produce data. The cumulants of the generated output are then estimated and compared to the cumulants estimated from the original data.

In this section we also compare the generated sequence to the original data record. This is repeated for some different model sizes, i.e., some different values of N_L and N_N.

a. Synthetically Generated Data

In this example a discrete Wiener model was constructed with dimensions $N_L = N_N = 2$. The poles and zeros of the corresponding transfer functions are defined by the parameter ρ which is chosen to be 0.65 in this simulation. Section 3 in this example is defined by the values of nine coefficients $c_{\boldsymbol{\alpha}}$ not including $c_{\mathbf{0}}$. A zero-mean unit-variance Gaussian sequence of 1024 time points was applied to the model to generate the random process. The second- and third-order cumulants of the random process were estimated as shown in Fig. 5.1. In the following experiments models of different size were used to model the generated data sequence. The Laguerre dimension N_L was given values in the range $\{1, 2, 3, 4, 5\}$ while the nonlinearity dimension N_N was given the values $\{2, 3, 4\}$ for each value of N_L. From the results of these experiments we could form the following conclusions:

1. For all the model dimensions, the estimated value of the parameter ρ was very close to the original value namely in the region [.635..671].

2. None of the trials resulted in values of the components of the vector **c** relatively close to the original values although in all cases the cumulant matching error was acceptably small ($< 5\%$).

3. Although the cumulant matching was acceptable, the plots of the model output sequence showed temporal variations judged to be similar to those of the original sequence for the case of quadratic models only. For higher degrees of nonlinearity the sequence had temporal variations that appeared to be different from the original data.

The results for $\{N_L = 1, N_N = 2\}$ and $\{N_L = 2, N_N = 2\}$ are shown here to demonstrate the efficiency of the Wiener model in modeling nonlinear data sequences. The resulting vectors of cumulants are compared with that of the original data in Fig. 5.2. The output sequences of the different size models is compared to the original data sequence in Fig. 5.3.

b. Biological Measurement Data

In this example we attempted to model a data set obtained from a record of biological measurements. This data set is one that has been made available to researchers in the field of time series analysis by the Santa Fe Institute [38]. The complete data record contains 34,000 points of the heart beat rate, chest volume, and blood oxygen concentration together with the EEG state of a sleeping patient in the sleep laboratory of the Beth Israel Hospital in Boston. One segment of the blood oxygen concentration data was chosen while the patient was in the same EEG state for a sufficiently long period of time (around 2500 points). A plot of the selected data segment is shown in Fig. 5.4, while estimates of its second- and third-order cumulants are shown in Fig. 5.5. This data set was modeled with the same range of dimensions as in the first example. The model parameters were determined such

that samples of the second- and third-order cumulants of both the model output and the data set match in the least square sense. The results of these trials showed the following:

1. The estimated values of the parameter ρ were very close together in all the trials and approximately equal to 0.95. This indicates that the linear part of the model has a low-pass frequency response.

2. The vector of coefficients **c** was considerably diverse among all the trials although the cumulant matching was acceptable.

3. The resulting data plots were comparable to the original data plot for quadratic models but unlike the original data for higher degrees of nonlinearity.

The cumulant vectors of the data set and the model output are compared in Fig. 5.6. The details of the data and the model outputs for different model sizes are shown in Fig. 5.7.

c. Analysis of Results

The results of the above examples indicate some characteristics of the procedure used for discrete random process modeling. These characteristics can be summarized as follows:

1. The linear part of the model adjusts rapidly to the frequency content of the modeled process. The filter bandwidth is properly adapted to the input process bandwidth as shown by the estimated value of the parameter ρ.

2. The solution obtained for the coefficients of section 3 of the model is not unique. This can be explained by the sparseness of the matrices constructed from the theoretical output cumulants. The resulting nonlinear equations to be solved for

these coefficients can be partitioned into a group of uncoupled sets of nonlinear equations, each with a separate subset of these coefficients. Each of these sets of equations has more than one solution. One solution from each set provides a problem solution that exactly gives the same result as any other combination of the different solutions of the different sets.

3. Over estimation of the model nonlinearity order may give an acceptable cumulant matching result while the time variations of the data sets may appear to be far different. This seems to be illustrative of the fact that for a highly nonlinear model second- and third-order cumulants are insufficient to characterize the process.

The method employed here to match cumulants of a given time series are not the only possible methods and the results summarized above are as much dependent on the method as they are dependent on the model. We hope that other experiments and methods will later be investigated.

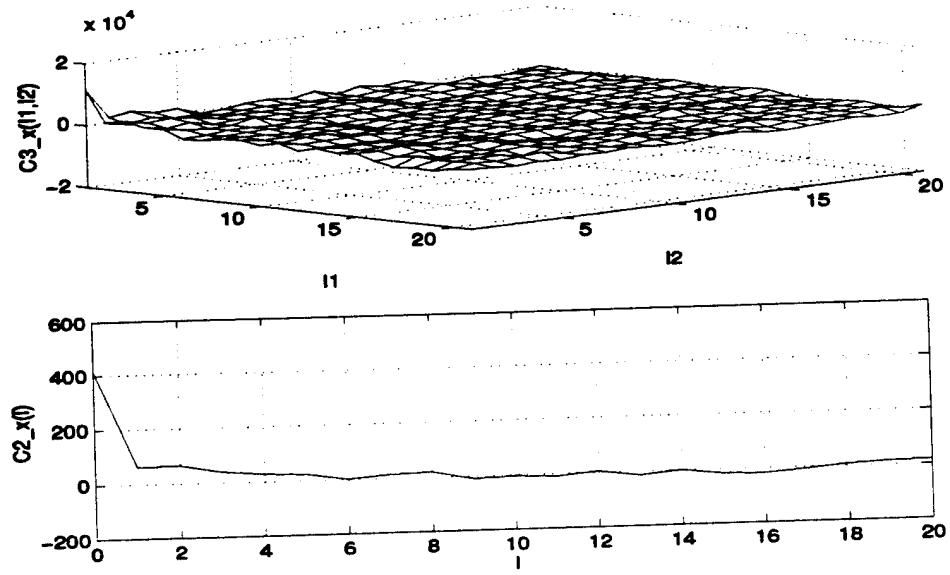

Figure 5.1: Plots of the estimated cumulants $C_x^{(3)}(l_1, l_2)$ and $C_x^{(2)}(l)$ for synthetically generated data.

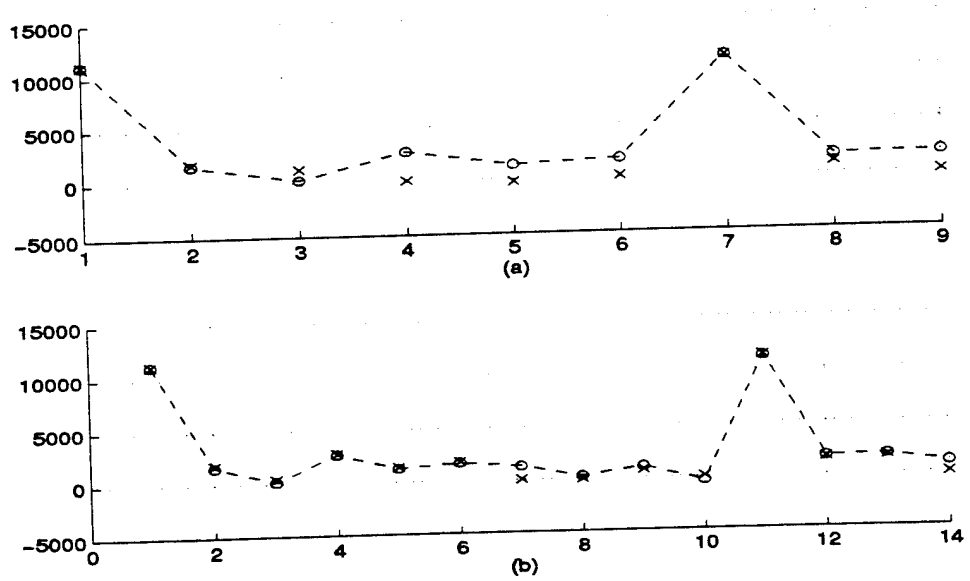

Figure 5.2: The results of cumulant matching for the first example; "o" = original set "x"=model output; (a): $N_L = 1, N_N = 2$ (b): $N_L = 2, N_N = 2$.

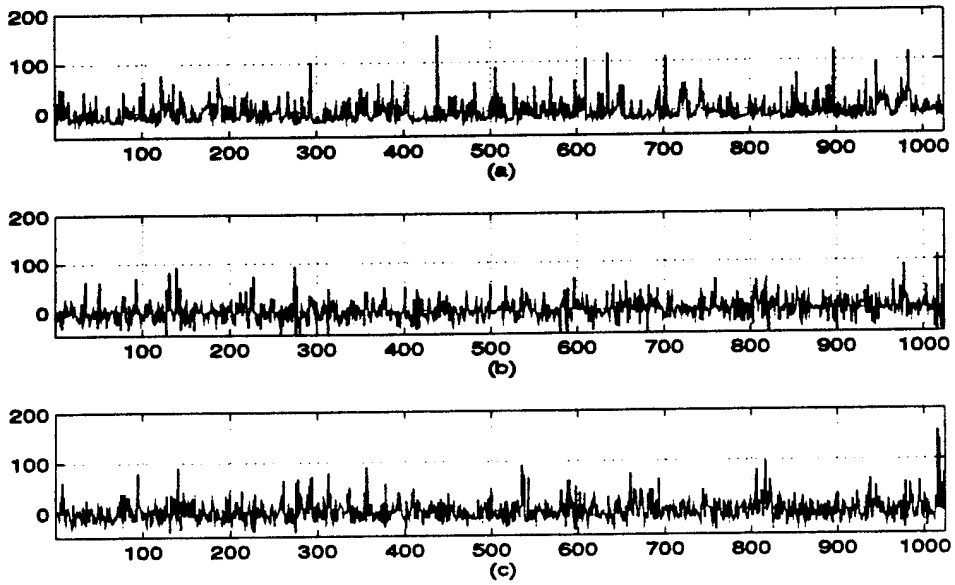

Figure 5.3: Example 1 data (a): Original data sequence (b): Model output for $N_L = 1, N_N = 2$ (c): Model output for $N_L = 1, N_N = 2$.

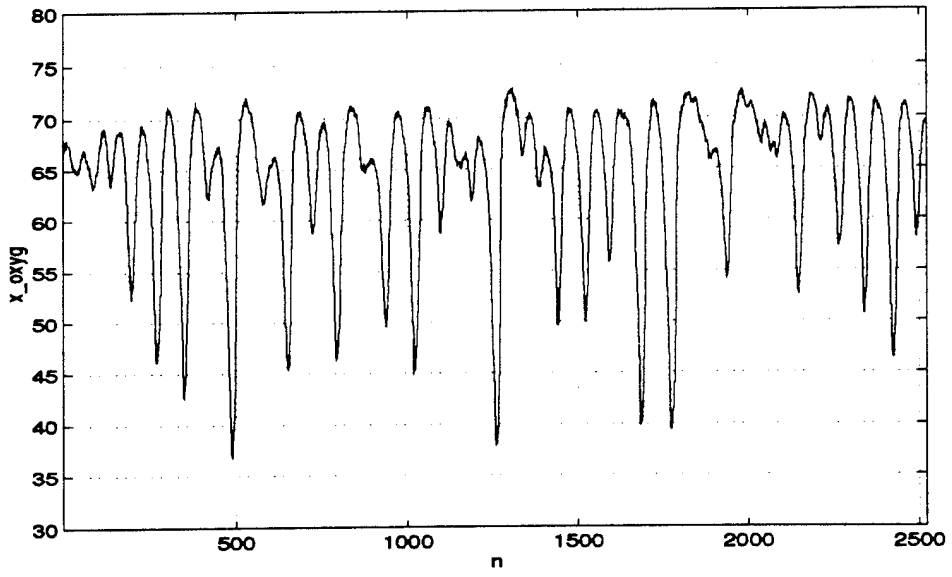

Figure 5.4: Example 2 data : a segment of the data set representing a patient's blood oxygen concentration.

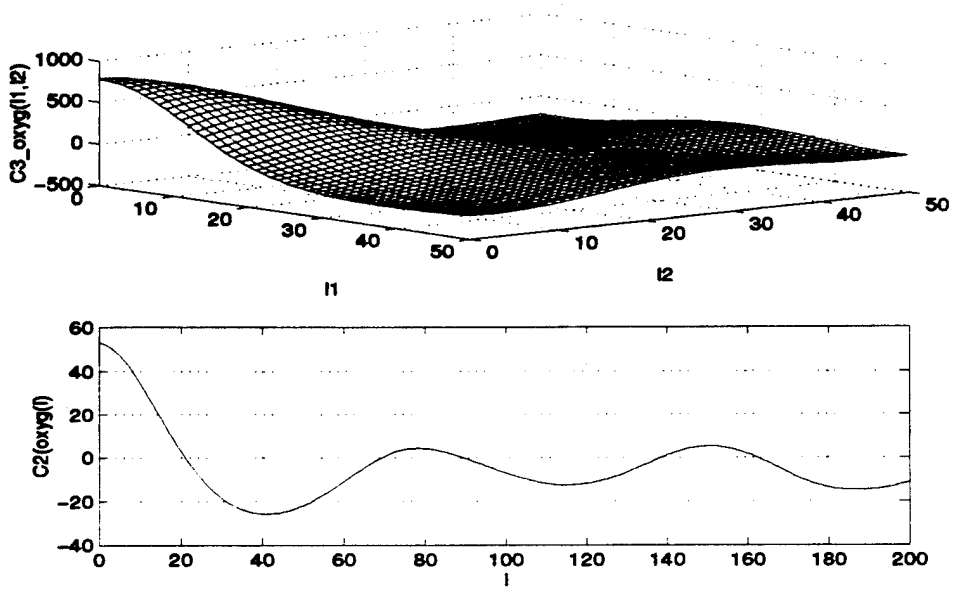

Figure 5.5: Example 2 data (a): Estimate of data third-order cumulant (b): Estimate of the covariance function.

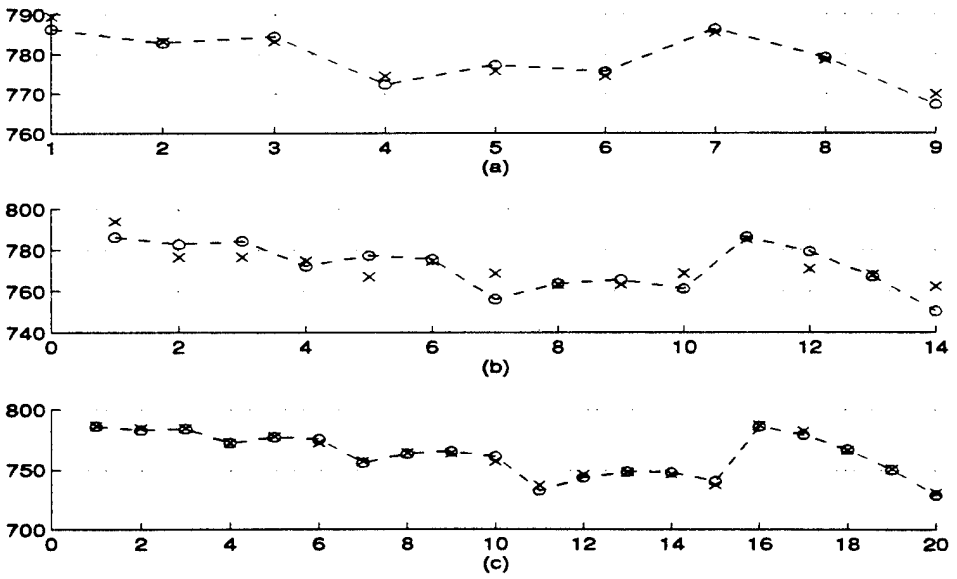

Figure 5.6: Cumulant vectors comparison for example 2, "o" = original set "x"=model output:(a): $N_L = 1, N_N = 2$. (b): $N_L = 2, N_N = 2$. (c): $N_L = 3, N_N = 2$.

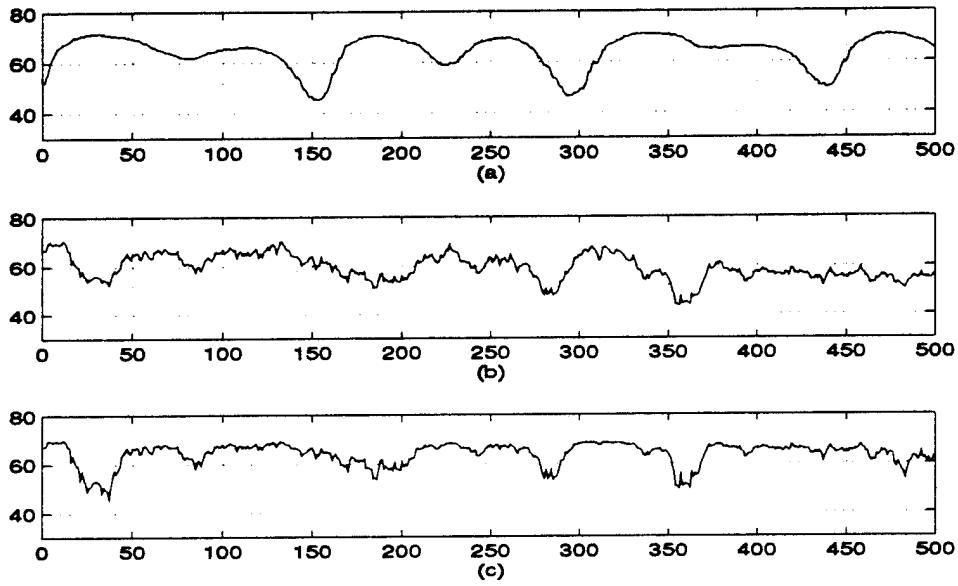

Figure 5.7: Example 2 results; (a) Original data set. (a):Model output for $N_L = 1, N_N = 2$. (b):model output for $N_L = 2, N_N = 2$.

VII. CONCLUSION

The canonical representation of nonlinear systems developed by Norbert Wiener to circumvent the problems of their series-like representation has been used for many years in parameter identification and characterization for such nonlinear systems. The applications that have been reported in this regard have been primarily for continuous time systems and are based on second-order statistics. The kernels representing the system are obtained by computing the cross-correlation between the model output and its driving input.

The research presented in this dissertation is focused on the use of the Wiener nonlinear system model to represent a general class of discrete stochastic processes and, in particular on:

1. Development of the model for discrete nonlinear systems and discrete random processes.

2. Analysis of the higher-order statistics of the model output, which can represent a broad class of discrete random processes.

3. Application of the theory and the results to model empirical data.

The Wiener model is comprised of three cascaded stages. The first stage is a bank of orthogonal linear filters driven by the the same white Gaussian process. The second stage consists of sets of identical memoryless nonlinear modules which form orthogonal multinomials of the outputs of the linear stage. The model output is then represented as a weighted sum of such multinomials. This output is characterized by the summation weights which are the constant parameters of the model final stage.

The higher-order statistics are generally very demanding in computation and in data size especially if they are involved in nonlinear system applications. The Wiener model is shown in this dissertation to be tractible in such applications because of its inherent orthogonality properties. Since the model input is white Gaussian, the orthogonality of its components results in expressing the auto- and cross-correlation functions of the outputs of the linear stage and generally all the model output cumulants in terms of the cross-correlation sequences between the different order Laguerre functions. These cross-correlation sequences depend only on the difference between the orders of the correlated Laguerre functions. As a result of this property, a very small number of these sequences is needed to formulate the model output statistics of any order. Another important characteristic of these functions is that (except for the autocorrelation sequence) they are single-sided, which makes most of the terms in the general expression of the output statistics identically zero.

We have exploited the model characteristics in developing a procedure that minimizes the computation and storage required to obtain the value of the higher-order cumulants of the model output. We have also developed the computation of these cumulants such that most of the effort is done off-line only once and stored in look-up tables. These tables can be used in any application of this model that requires computing the higher order cumulants of its output.

In the frequency domain we introduced a method that reduces the cost of computing the model output polyspectra. The computation of such multidimensional spectral functions involves a large number of integral convolutions. We first proved that, similar to the cumulant functions, the size of this computation is greatly reduced due to the model structure. Moreover, we showed that such convolution computations can be performed algebraically in a structured recursive procedure.

In the last chapter of this dissertation we discuss the representability of random processes using the Wiener model. The polyspectra of such processes are necessarily analytic and not bandlimited. Two examples of modeling processes satisfying these conditions are presented. These examples demonstrate the viability of our approach. We hope that further experiments of this type, possibly using other methods of matching will be conducted in the future.

The development of this structured model still has areas to be covered and questions to be answered. In analyzing the characteristic of discrete random processes that can be represented using the Wiener model we provided some necessary conditions. Sufficient conditions are yet to be found and the class of those processes representable by the Wiener model needs to be defined.

The development of a procedure to compute the model output polyspectra can be done following the results of the time domain development. The algebraic formulation of computing the multiple convolution of Laguerre cross-spectral functions can be directly applied to obtain the power spectral density and higher-order polyspectra of the model output. The use of these results to predict model behavior and for synthesis of processes with given polyspectral characteristics has yet to be explored however.

The spectral functions denoted by $\mathcal{X}^{(\alpha)}(z; p_i, q_i)$ have multiple poles and zeros at locations $\rho^{\alpha+1}$ and $\rho^{\alpha-1}$ respectively. These poles and zeros move towards the origin as α increases. This means that as the number of convolutions required to represent the polyspectra increases, these spectral functions approach the form z^{-k}. Since the number of convolutions is proportional to the degree of nonlinearity, in highly nonlinear terms the spectral functions can be approximated by z^{-k} which greatly simplifies the representation. In effect, these terms can be approximated by a weighted sum of functions of magnitude one with linear phase. Even if the

nonlinearity is not high we can extend the number of Laguerre filters to compensate for a reduction in the magnitude of ρ to justify this approximation. These alternatives need to be analyzed and the effect of each on the the model output representation needs to be evaluated.

In summary, the work reported in this dissertation may be just the introduction to a whole new body of work yet to be explored. We have demonstrated the feasibilty of the Wiener model as a generic representation for a large class of random processes. This model expands in a natural way the much more restricted class of linear Gaussian processes that are currently so well understood. The utility and practicality of the model result from important structure foreseen by Wiener in its development and the recent new interest in higher-order statistics that has undergone development in only the last decade. Although our initial exploration of this new area leaves many questions unanswered, we hope that the problems posed here will continue to be addressed and lead to a significant improvement in our ability to understand and model the signals and other data that are encountered in the real world.

APPENDIX A

EXPECTATION OF THE PRODUCT OF JOINTLY GAUSSIAN RANDOM VARIABLES

Let $\mathbf{y} = [y_1, y_2, \cdots, y_k]$ be a vector of real zero-mean jointly Gaussian random variables. The expectation of the product of these variables raised to some powers has the form

$$\mathcal{E}\{y_1^{\nu_1} y_2^{\nu_2} \cdots y_k^{\nu_k}\} \tag{A.1}$$

The value of this moment is zero if the sum of the exponents $\sum_{i=1}^{k} \nu_i$ is odd. Let us outline a method here for computing this moment when the sum of the exponents is even. The value of the moment can be computed by generalizing equation (2.34) to

$$\begin{aligned} \mathrm{M}_{\mathbf{y}}^{(\nu_1+\nu_2+\cdots+\nu_k)} &= \mathcal{E}\{y_1^{\nu_1} y_2^{\nu_2} \cdots y_k^{\nu_k}\} \\ &= \mathcal{E}\{\underbrace{y_1 y_1 \cdots y_1}_{\nu_1} \underbrace{y_2 y_2 \cdots y_2}_{\nu_2} \cdots \underbrace{y_k y_k \cdots y_k}_{\nu_k}\} \\ &= \sum \prod [\mathcal{E}\{y_{j_1} y_{j_2}\}]^{q_{j_1 j_2}} \end{aligned} \tag{A.2}$$

where the exponent $q_{j_1 j_2}$ means that y_{j_1} and y_{j_2} are paired $q_{j_1 j_2}$ times and the summation is *over all the possible permutations.* Since more than one permutation can result in the same value of the term under summation (A.2) can be put in the form

$$\mathrm{M}_{\mathbf{y}}^{(\nu_1+\nu_2+\cdots+\nu_k)} = \sum_{p} c_p \prod [\mathcal{E}\{y_{j_1} y_{j_2}\}]^{q_{j_1 j_2}} \tag{A.3}$$

where in this case the summation is *over all the distinct permutations* and the coefficient c_p is the number of non-distinct permutations that have the same value. The multiplication is *over all the distinct pairing* of the y_j's with each pair raised to a

power that equals the number of its repetitions in the pairing configuration. In all cases the relation

$$\sum_{j_1,j_2} q_{j_1 j_2} = \sum_{i=1}^{k} \nu_i \tag{A.4}$$

must hold for each term under the summation. The procedure to find the value of this moment begins by constructing the $k \times k$ correlation matrix $\mathbf{C}$

$$\mathbf{C} = \begin{bmatrix} C(1,1) & C(1,2) & C(1,3) & \cdots & C(1,k) \\ C(2,1) & C(2,2) & C(2,3) & \cdots & C(2,k) \\ C(3,1) & C(3,2) & C(3,3) & \cdots & C(3,k) \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ C(k,1) & C(k,2) & C(k,3) & \cdots & C(k,k) \end{bmatrix} \tag{A.5}$$

where

$$C(i,j) = \mathcal{E}\{y_i y_j\} \tag{A.6}$$

and the same size matrix $\mathbf{M}$ of *non-negative integer* entries such that

$$\mathbf{M} = \begin{bmatrix} 2i_1 & i_2 & i_3 & \cdots & i_k \\ i_{k+1} & 2i_{k+2} & i_{k+3} & \cdots & i_{2k} \\ i_{2k+1} & i_{2k+2} & 2i_{2k+3} & \cdots & i_{3k} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ i_{(k-1)k+1} & i_{(k-1)k+2} & i_{(k-1)k+3} & \cdots & 2i_{k^2} \end{bmatrix} \tag{A.7}$$

Note that for clarity in the following discussion it is prefered to use k^2 linearly indexed variables i_j to represent the entries of $\mathbf{M}$ rather than variables with dual (row,column) indices. We also force the diagonal entries to assume even values for reasons that are explained in a moment.

The entries of the multiplicity matrix $\mathbf{M}$ are used to determine the exponents $q_{j_1 j_2}$ in (A.3). Since, in general, some of the exponents of the random variables y_i in (A.2) are greater than one, there are some permutations that result in pairing a random varaible y_i with y_j more than once. This results in having the quantity $C(i,j)$ appearing in the product in (A.3) a number of times equal to the number of the resulting identical pairs of y_i and y_j. The value of the off-diagonal entry $\mathrm{M}(i,j)$

is the number of y_i paired with the same number of y_j to give $C(i,j)^{\mathbf{M}(i,j)}$. The diagonal entry $\mathrm{M}(j,j)$ in row and column j is the number of the elements taken from y_j that are paired together and the result gives $C(j,j)^{\frac{\mathbf{M}(j,j)}{2}}$. Therefore the values of the diagonal entries must be even.

At this point it is required to compute the number of permutations to perform the pairing for a specific configuration of the entries of the matrix **M**. We start from the first row and examine all the possible values that the entries of this row can assume. Then we examine the other rows moving downwards in order. When a row is examined its entries are examined in order from left to right. Therefore starting with the upper left entry $\mathrm{M}(1,1) = 2i_1$ for $i_1 = 0,1,2,\cdots,\lfloor\frac{\nu_1}{2}\rfloor$ the number of permutations is obtained by computing the number of ways of choosing $2i_1$ elements out of ν_1 multiplied by the number of pairing permutations of $2i_1$ elements. This yields

$$\begin{aligned} N_{per}^{(11)} &= \frac{\nu_1!}{2i_1!(\nu_1-2i_1)!} \times \frac{2i_1}{i_1!2^{i_1}} \\ &= \frac{\nu_1!}{i_1!(\nu_1-2i_1)!2^{i_1}} \end{aligned} \tag{A.8}$$

Since the sum of the elements in the first row of M equals ν_1 then for each value of i_1 the value of $\mathrm{M}(1,2) = i_2 \leq \nu_1 - 2i_1$. Also the sum of the elements in the second column must equal ν_2 which means that $i_2 \leq \nu_2$. Therefore the multiplicity matrix entry $\mathrm{M}(1,2)$ equals i_2 such that $i_2 = 0,1,2,\cdots,\min\{\nu_1-2i_1,\nu_2\}$. The number of permutations in this case is given by the number of ways of choosing i_2 elements out of $\nu_1 - 2i_1$ multiplied by the number of taking the same number i_2 of the y_2 out of ν_2 then multiplied by the value $i_2!$ corresponding to the number of pairing permutations. That is,

$$N_{per}^{(12)} = \frac{(\nu_1-2i_1)!}{i_2!(\nu_1-2i_1-i_2)!} \times \frac{\nu_2!}{i_2!(\nu_2-i_2)!} \times i_2!$$

$$= \frac{(\nu_1 - 2i_1)!\nu_2!}{i_2!(\nu_1 - 2i_1 - i_2)!(\nu_2 - i_2)!} \tag{A.9}$$

Similarly the other entries of the first row of $\mathbf{M}$ follow as

$$\begin{aligned} N_{per}^{(1j)} &= \frac{(\nu_1 - \sum_{k=1}^{j-1} \mathrm{M}(1,k))!}{i_j!(\nu_1 - \sum_{k=1}^{j} \mathrm{M}(1,k))!} \times \frac{\nu_j!}{i_j!(\nu_2 - i_j)!} \times i_j! \\ &= \frac{(\nu_1 - \sum_{k=1}^{j-1} \mathrm{M}(1,k))!\nu_j!}{i_j!(\nu_1 - \sum_{k=1}^{j} \mathrm{M}(1,k))!(\nu_j - i_j)!} \end{aligned} \tag{A.10}$$

and the total number of permutations of a specific configuration of the elements of the first row equals the product of the numbers of the above permutations,i.e.,

$$N_{per}^{(1)} = \frac{\nu_1!\nu_2!\cdots\nu_k!}{2^{i_1} i_1!i_2!\cdots i_k!(\nu_1 - 2i_1 - i_2 \cdots - i_k)!(\nu_2 - i_2)!\cdots(\nu_k - i_k)!} \tag{A.11}$$

For the second row we start with $\mathrm{M}(2,1) = i_{k+1}$ and

$$i_{k+1} = 0, 1, 2, \cdots, \min\{(\nu_1 - 2i_1 - i_2 \cdots - i_k), (\nu_2 - i_2)\}$$

The number of possible permutations is

$$\begin{aligned} N_{per}^{(21)} &= \frac{(\nu_1 - 2i_1 - i_2 \cdots - i_k)!}{i_{k+1}!(\nu_1 - 2i_1 - i_2 \cdots - i_k - i_{k+1})!} \times \frac{(\nu_2 - i_2)!}{i_{k+1}!(\nu_2 - i_2 - i_{k+1})!} \times i_{k+1}! \\ &= \frac{(\nu_1 - 2i_1 - i_2 \cdots - i_k)!(\nu_2 - i_2)!}{i_{k+1}!(\nu_1 - 2i_1 - i_2 \cdots - i_k - i_{k+1})!(\nu_2 - i_2 - i_{k+1})!} \end{aligned} \tag{A.12}$$

For the entry $\mathrm{M}(2,2) = 2i_{k+2}$ with $i_{k+2} = 0, 1, 2, \cdots, \lfloor \frac{\nu_2 - i_2 - i_{k+1}}{2} \rfloor$ the number of permutations is

$$\begin{aligned} N_{per}^{(22)} &= \frac{(\nu_2 - i_2 - i_{k+1})!}{2i_{k+2}!(\nu_2 - i_2 - i_{k+1} - 2i_{k+2})!} \times \frac{2i_{k+2}!}{i_{k+2}!2^{i_{k+2}}} \\ &= \frac{(\nu_2 - i_2 - i_{k+1})!}{2^{i_{k+2}} i_{k+2}!(\nu_2 - i_2 - i_{k+1} - 2i_{k+2})!} \end{aligned} \tag{A.13}$$

Proceeding with the rest of the entries of the matrix $\mathbf{M}$, we can examine all the possible values of the entries and compute the number of permutations of each configuration noting that in all cases the sum of the entries of both the j^{th} row and the j^{th} column must equal ν_j. The coefficient c_p in (A.3) is given by

$$c_p = \prod_{i=1}^{k} \prod_{j=1}^{k} N_{per}(ij) \tag{A.14}$$

Due to the cancellation of of terms that appear in both the denominator and the nominator, the value of the moment is given by

$$M_{\mathbf{y}}^{(\nu_1+\nu_2+\cdots+\nu_k)} = \nu_1!\nu_2!\cdots\nu_k!\sum \frac{1}{2^{\frac{\mathrm{tr}(I)}{2}}} \prod_{j_1=1}^{k} \frac{(C(j_1,j_1))^{\frac{\mathbf{M}(j_1,j_1)}{2}}}{(\frac{\mathbf{M}(j_1,j_1)}{2})!} \prod_{j_2\neq j_1} \frac{(C(j_1,j_2))^{\mathbf{M}(j_1,j_2)}}{(\mathrm{M}(j_1,j_2))!} \tag{A.15}$$

where the summation is over all possible configurations of the matrix $\mathbf{M}$ that satisfy the above mentioned condition on the sum of the entries along the rows and the columns. We can write this condition as

$$\begin{aligned} \sum_j \mathrm{M}(i,j) &= \nu_i \\ \sum_i \mathrm{M}(i,j) &= \nu_j \end{aligned} \tag{A.16}$$

The first product corresponds to the product of the diagonal entries of the correlation matrix $\mathbf{C}$ (the autocorrelation) while the second corresponds to the off-diagonal ones (the cross-correlation).

APPENDIX B

THE WIENER G-FUNCTIONALS FOR DISCRETE NONLINEAR SYSTEMS

A. THE RELATION BETWEEN THE DERIVED WIENER KERNELS AND THE LEADING KERNEL

Since the Wiener Functional $\mathcal{G}_p[g_p, g_{p-1(p)}, \cdots, g_{0(p)}; w(n)]$ is required to be orthogonal to any Volterra kernel of order less than p when the input is Gaussian,(3.14), the set of p equations

$$\mathcal{E}\left\{\mathcal{H}_i[h_i; w(n)]\mathcal{G}_p[g_p, g_{p-1(p)}, \cdots, g_{0(p)}; w(n)]\right\} = 0 \qquad \text{for } i = 0, 1, 2, \cdots, p-1 \quad \text{(B.1)}$$

are used to determine the relation between the derived kernels and the leading kernel. Equation (B.1) can be put in the form

$$\mathcal{E}\left\{\mathcal{H}_i[h_i; w(n)]\mathcal{G}_p[g_p, g_{p-1(p)}, \cdots, g_{0(p)}; w(n)]\right\} = \mathcal{E}\left\{\mathcal{H}_i[h_i; w(n)]\sum_{j=0}^{p}\mathbf{G}_j[g_j(p); w(n)]\right\}$$

$$= \sum_{j=0}^{p}\mathcal{E}\left\{\mathcal{H}_i[h_i; w(n)]\mathbf{G}_j[g_j(p); w(n)]\right\} = 0 \qquad \text{for } i = 0, 1, 2, \cdots, p-1 \qquad \text{(B.2)}$$

where $\mathbf{G}_j[g_j(p); w(n)]$ is the j^{th} homogeneous functional corresponding to the leading kernel g_p for $j = p$ or one of the derived kernels otherwise. Each of the terms under the summation has the form

$$\mathcal{E}\left\{\sum_{k_1=0}^{\infty}\sum_{k_2=0}^{\infty}\cdots\sum_{k_{i+j}=0}^{\infty}h_i(k_1, k_2, \cdots, k_i)g_{j(p)}(k_{i+1}, \cdots, k_{i+j})w(n-k_1)\cdots w(n-k_{i+j})\right\}$$

$$= \sum_{k_1=0}^{\infty}\sum_{k_2=0}^{\infty}\cdots\sum_{k_{i+j}=0}^{\infty}h_i(k_1, k_2, \cdots, k_i)g_{j(p)}(k_{i+1}, k_{i+2}, \cdots, k_{i+j})$$

$$\times\mathcal{E}\left\{w(n-k_1)w(n-k_2)\cdots w(n-k_{i+j})\right\} \qquad \text{(B.3)}$$

in which the expectation of the product of Gaussian variables is involved. The result is that the value of this term is zero if the number of the Gaussian variables $i+j$ is odd. This means that if the Volterra kernel h_i has even order i, then (B.2) will have terms that correspond to g-kernels with even orders only. Likewise, if h_i is odd (B.2) will contain only odd order g-kernels. Since p may be even or odd, g_p may appear in either collection of terms. Therefore, for simplicity in the following development p is assumed to be odd but the result also applies to the case of p even.

Starting with the expectation of the product of the G-functional with the zeroth order Volterra kernel we have

$$\begin{aligned}
&\mathcal{E}\Bigg\{ h_0 g_{0(p)} + \sum_{k_1=0}^{\infty}\sum_{k_2=0}^{\infty} h_0 g_{2(p)}(k_1,k_2) w(n-k_1)w(n-k_2) \\
&+ \sum_{k_1=0}^{\infty}\sum_{k_2=0}^{\infty}\sum_{k_3=0}^{\infty}\sum_{k_4=0}^{\infty} h_0 g_{4(p)}(k_1,k_2,k_3,k_4) w(n-k_1)w(n-k_2)w(n-k_3)w(n-k_4) \\
&+ \cdots + \sum_{k_1=0}^{\infty}\cdots\sum_{k_{p-1}=0}^{\infty} h_0 g_{p-1(p)}(k_1,\cdots,k_{p-1}) w(n-k_1)\cdots w(n-k_{p-1}) \Bigg\} = 0
\end{aligned} \tag{B.4}$$

Then applying the properties of the expectation of the product of Gaussian variables and the symmetry properties of the Volterra and Wiener kernels with respect to their arguments produces

$$\begin{aligned}
h_0 \Bigg[& g_{0(p)} + \sigma_o^2 \sum_{k=0}^{\infty} g_{2(p)}(k,k) + 3\sigma_o^4 \sum_{k_1=0}^{\infty}\sum_{k_2=0}^{\infty} g_{4(p)}(k_1,k_1,k_2,k_2) + \cdots \\
&+ \frac{(p-1)!}{(\frac{p-1}{2})!2^{(\frac{p-1}{2})}} \sigma_o^{(p-1)} \sum_{k_1=0}^{\infty}\cdots\sum_{k_{\frac{p-1}{2}}=0}^{\infty} g_{p-1(p)}(k_1,k_1,k_2,k_2,\cdots,k_{\frac{p-1}{2}},k_{\frac{p-1}{2}}) \Bigg] = 0
\end{aligned} \tag{B.5}$$

Since this is required to be true for any h_0, the zeroth order kernel relation is

$$g_{0(p)} + \sigma_o^2 \sum_{k=0}^{\infty} g_{2(p)}(k,k) + 3\sigma_o^4 \sum_{k_1=0}^{\infty}\sum_{k_2=0}^{\infty} g_{4(p)}(k_1,k_1,k_2,k_2) + \cdots$$
$$+ \frac{(p-1)!}{(\frac{p-1}{2})!2^{(\frac{p-1}{2})}} \sigma_o^{(p-1)} \sum_{k_1=0}^{\infty} \cdots \sum_{k_{\frac{p-1}{2}}=0}^{\infty} g_{p-1(p)}(k_1,k_1,k_2,k_2,\cdots,k_{\frac{p-1}{2}},k_{\frac{p-1}{2}}) = 0 \tag{B.6}$$

Now we proceed through the Volterra kernels of even order. Letting $i = 2$ in (B.2) we have

$$\mathcal{E}\left\{ \sum_{k_1=0}^{\infty}\sum_{k_2=0}^{\infty} h_2(k_1,k_2) g_{0(p)} w(n-k_1)w(n-k_2) \right.$$
$$+ \sum_{k_1=0}^{\infty}\sum_{k_2=0}^{\infty}\sum_{k_3=0}^{\infty}\sum_{k_4=0}^{\infty} h_2(k_1,k_2) g_{2(p)}(k_3.k_4) w(n-k_1)w(n-k_2)w(n-k_3)w(n-k_4)$$
$$+ \sum_{k_1=0}^{\infty}\sum_{k_2=0}^{\infty}\cdots\sum_{k_6=0}^{\infty} h_2(k_1,k_2) g_{4(p)}(k_3.k_4,k_5,k_6) w(n-k_1)w(n-k_2)\cdots w(n-k_6)$$
$$\vdots$$
$$\left. + \sum_{k_1=0}^{\infty}\cdots\sum_{k_{p+1}=0}^{\infty} h_2(k_1,k_2) g_{p-1(p)}(k_3.k_4.\cdots,k_{p+1}) w(n-k_1)\cdots w(n-k_{p+1}) \right\} = 0 \tag{B.7}$$

Now, let us divide the pairing permutations into two sets the first of which is such that $w(n-k_1)$ and $w(n-k_2)$ are paired together and the second such that they are not. We have

$$\sigma_o^2 \sum_{k=0}^{\infty} h_2(k,k) \Bigg[\quad g_{0(p)} + \sigma_o^2 \sum_{k_1=0}^{\infty} g_{2(p)}(k_1,k_1) + 3\sigma_o^4 \sum_{k_1=0}^{\infty}\sum_{k_2=0}^{\infty} g_{4(p)}(k_1,k_1,k_2,k_2)$$

$$+ \quad \cdots + \frac{(p-1)!}{(\frac{p-1}{2})!2^{(\frac{p-1}{2})}} \sigma_o^{(p-1)} \sum_{k_1=0}^{\infty} \cdots \sum_{k_{\frac{p-1}{2}}=0}^{\infty} g_{p-1(p)}(k_1,k_1,k_2,k_2,\cdots,k_{\frac{p-1}{2}},k_{\frac{p-1}{2}}) \Bigg]$$

$$+ \quad 2\sigma_o^4 \sum_{k_1=0}^{\infty}\sum_{k_2=0}^{\infty} h_2(k_1,k_2) \Bigg[\quad g_{2(p)}(k_1,k_2) + 6 \sum_{k_3=0}^{\infty} g_{4(p)}(k_1,k_2,k_3,k_3) + \cdots$$

$$+ \quad \sigma_o^{p-3} \frac{(p-1)!}{2!(\frac{p-3}{2})!2^{(\frac{p-3}{2})}} \sum_{k_3=0}^{\infty} \cdots \sum_{k_{\frac{p+1}{2}}=0}^{\infty} g_{p-1(p)}(k_1,k_2,k_3,k_3,\cdots,k_{\frac{p+1}{2}},k_{\frac{p+1}{2}}) \Bigg] = 0 \tag{B.8}$$

From (B.6) the first term (in brackets) is identically zero. Therefore, the second term must equal zero for all values of $h_2(k_1,k_2)$; this provides the second even-kernel relation

$$g_{2(p)}(k_1,k_2) + 6\sum_{k=0}^{\infty} g_{4(p)}(k_1,k_2,k,k) + \cdots$$
$$+\sigma_o^{p-3} \frac{(p-1)!}{2!(\frac{p-3}{2})!2^{(\frac{p-3}{2})}} \sum_{k_3=0}^{\infty} \cdots \sum_{k_{\frac{p+1}{2}}=0}^{\infty} g_{p-1(p)}(k_1,k_2,k_3,k_3,\cdots,k_{\frac{p+1}{2}},k_{\frac{p+1}{2}}) = 0 \tag{B.9}$$

Let us take this one step further by letting $i = 4$ in (B.3) and dividing the pairing permutation into three sets. In the first set all of the time arguments in $h_4(k_1,k_2,k_3,k_4)$ are paired together and are not paired with any of the g-kernel arguments. In the second set only two of the arguments are paired together and each one of the other two is paired with one argument of the g-kernels, if possible. In the third set each of the arguments is paired with one of the arguments of the g-kernels,

if possible. By doing this, the value of the expectation we obtain is

$$
\begin{aligned}
& 3\sigma_o^4 \sum_{k_1=0}^{\infty}\sum_{k_2=0}^{\infty} h_4(k_1,k_1,k_2,k_2)\Bigg[\quad g_{0(p)} + \sigma_o^2 \sum_{k=0}^{\infty} g_{2(p)}(k,k) \\
+\ & 3\sigma_o^4 \sum_{k_3=0}^{\infty}\sum_{k_4=0}^{\infty} g_{4(p)}(k_3,k_3,k_4,k_4) + \cdots \\
+\ & \sigma_o^{(p-1)} \frac{(p-1)!}{(\frac{p-1}{2})!2^{(\frac{p-1}{2})}} \sum_{k_3=0}^{\infty}\cdots\sum_{k_{\frac{p+3}{2}}=0}^{\infty} g_{p-1(p)}(k_3,k_3,k_4,k_4,\cdots,k_{\frac{p+3}{2}},k_{\frac{p+3}{2}})\Bigg] \\
& +12\sigma_o^6 \sum_{k_1=0}^{\infty}\sum_{k_2=0}^{\infty}\sum_{k_3=0}^{\infty} h_4(k_1,k_2,k_3,k_3)\Bigg[\quad g_2(k_1,k_2,) \\
+\ & 6\sigma_o^2 \sum_{k_4=0}^{\infty} g_{4(p)}(k_1,k_2,k_4,k_4) + \cdots + \sigma_o^{p-3} \frac{(p-1)!}{2!(\frac{p-3}{2})!2^{(\frac{p-3}{2})}} \frac{(p-1)!}{2!(\frac{p-3}{2})!2^{(\frac{p-3}{2})}} \\
& \times \sum_{k_4=0}^{\infty}\cdots\sum_{k_{\frac{p+3}{2}}=0}^{\infty} g_{p-1(p)}(k_1,k_2,k_4,k_4,\cdots,k_{\frac{p+3}{2}},k_{\frac{p+3}{2}})\Bigg] \\
+\ & 24\sigma_o^8 \sum_{k_1=0}^{\infty}\sum_{k_2=0}^{\infty}\sum_{k_3=0}^{\infty}\sum_{k_4=0}^{\infty} h_4(k_1,k_2,k_3,k_4)\Bigg[\quad g_{4(p)}(k_1,k_2,k_3,k_4) + \cdots \\
& + \sigma_o^{p-5} \frac{(p-1)!}{4!(\frac{p-5}{2})!2^{(\frac{p-5}{2})}} \sum_{k_5=0}^{\infty}\cdots\sum_{k_{\frac{p+3}{2}}=0}^{\infty} g_{p-1(p)}(k_1,k_2,k_3,k_4,k_5,k_5,\cdots,k_{\frac{p+3}{2}},k_{\frac{p+3}{2}})\Bigg] \\
& = 0
\end{aligned}
\tag{B.10}
$$

This is similar to the previous case; the first and the second terms (in brackets) are identically zero from (B.6) and (B.9), which implies that the third term must be equal to zero for all $h_4(k_1,k_2,k_3,k_4)$. This leads to the third even-kernel relation

$$
\begin{aligned}
& g_{4(p)}(k_1,k_2,k_3,k_4) + 15\sigma_o^2 \sum_{k_5=0}^{\infty} g_{6(p)}(k_1,k_2,k_3,k_4,k_5,k_5) + \cdots \\
+\ & \sigma_o^{p-5} \frac{(p-1)!}{4!(\frac{p-5}{2})!2^{(\frac{p-5}{2})}} \sum_{k_5=0}^{\infty}\cdots\sum_{k_{\frac{p+3}{2}}=0}^{\infty} g_{p-1(p)}(k_1,k_2,k_3,k_4,k_5,k_5,\cdots,k_{\frac{p+3}{2}},k_{\frac{p+3}{2}}) = 0
\end{aligned}
\tag{B.11}
$$

By proceeding in the same manner up to $i = p - 1$ we obtain

$$\left[\frac{p-3}{2} \text{ terms equal to zero from above equations}\right] +$$
$$\sigma_o^{p-1}(p-1)! \sum_{k_1=0}^{\infty} \cdots \sum_{k_{p-1}=0}^{\infty} h_{p-1}(k_1, k_2, \cdots, k_{p-1}) g_{p-1(p)}(k_1, k_2, \cdots, k_{p-1}) = 0 \tag{B.12}$$

which implies that

$$\sum_{k_{p-1}=0}^{\infty} h_{p-1}(k_1, k_2, \cdots, k_{p-1}) g_{p-1(p)}(k_1, k_2, \cdots, k_{p-1}) = 0 \tag{B.13}$$

for all $h_{p-1}(k_1, k_2, \cdots, k_{p-1})$. This relation is satisfied only if

$$g_{p-1(p)}(k_1, k_2, \cdots, k_{p-1}) = 0 \tag{B.14}$$

which is the last even-kernel relation we need. If we substitute (B.14) in the relation just preceding it we obtain

$$g_{p-3(p)}(k_1, k_2, \cdots, k_{p-3}) = 0 \tag{B.15}$$

By continuing this backward substitution in all the above even-kernel relations (B.6)-(B.14) we find that *all the even-order derived kernels are equal to zero* .

To obtain the relation between the odd-order kernels we perform the expectation of the product of the G-functional and odd-order Volterra functionals. Making use of the properties of the average of the product of Gaussian variables and the symmetry properties of both the Volterra and the Wiener kernels, we can derive the following relations. Starting with the first order Volterra kernel we obtain

$$\sigma_o^2 \sum_{k_1=0}^{\infty} h_1(k_1) \Bigg[\quad g_{1(p)}(k_1) + 3\sigma_o^2 \sum_{k2=0}^{\infty} g_{3(p)}(k_1, k_2, k_2)$$

$$+ \quad 15\sigma_o^4 \sum_{k_2=0}^{\infty} \sum_{k_3=0}^{\infty} g_{5(p)}(k_1, k_2, k_2, k_3, k_3) + \cdots$$

$$+ \quad p \frac{(p-1)!}{(\frac{p-1}{2})! 2^{(\frac{p-1}{2})}} \sigma_o^{(p-1)} \sum_{k_2=0}^{\infty} \cdots \sum_{k_{\frac{p+1}{2}}=0}^{\infty} g_p(k_1, k_2, k_2, k_3, k_3, \cdots, k_{\frac{p+1}{2}}, k_{\frac{p+1}{2}}) \Bigg] = 0 \tag{B.16}$$

Since this must hold for all $h_1(k)$ then the first odd-kernel relation is

$$g_{1(p)}(k_1) + 3\sigma_o^2 \sum_{k2=0}^{\infty} g_{3(p)}(k_1, k_2, k_2) + 15\sigma_o^4 \sum_{k_2=0}^{\infty} \sum_{k_3=0}^{\infty} g_{5(p)}(k_1, k_2, k_2, k_3, k_3) + \cdots$$
$$+ p \frac{(p-1)!}{(\frac{p-1}{2})! 2^{(\frac{p-1}{2})}} \sigma_o^{(p-1)} \sum_{k_2=0}^{\infty} \cdots \sum_{k_{\frac{p+1}{2}}=0}^{\infty} g_p(k_1, k_2, k_2, k_3, k_3, \cdots, k_{\frac{p+1}{2}}, k_{\frac{p+1}{2}}) = 0 \tag{B.17}$$

Using the the third-order Volterra kernel we have

$$3\sigma_o^4 \sum_{k_1=0}^{\infty} \sum_{k_2=0}^{\infty} h_3(k_1, k_1, k_2) \Bigg[\quad g_{1(p)}(k_2) + 3\sigma_o^2 \sum_{k3=0}^{\infty} g_{3(p)}(k_2, k_3, k_3)$$

$$+ \quad 15\sigma_o^4 \sum_{k_3=0}^{\infty} \sum_{k_4=0}^{\infty} g_{5(p)}(k_2, k_3, k_3, k_4, k_4) + \cdots$$

$$+ \quad p \frac{(p-1)!}{(\frac{p-1}{2})! 2^{(\frac{p-1}{2})}} \sigma_o^{(p-1)} \sum_{k_3=0}^{\infty} \cdots \sum_{k_{\frac{p+3}{2}}=0}^{\infty} g_p(k_2, k_3, k_3, k_4, k_4, \cdots, k_{\frac{p+3}{2}}, k_{\frac{p+3}{2}}) \Bigg]$$

$$+ \quad 3!\sigma_o^6 \sum_{k_1=0}^{\infty} \sum_{k_2=0}^{\infty} \sum_{k_3=0}^{\infty} h_3(k_1, k_2, k_3) \Bigg[\quad g_{3(p)}(k_1, k_2, k_3)$$

$$+ \quad 60\sigma_o^2 \sum_{k4=0}^{\infty} g_{5(p)}(k_1, k_2, k_3, k_4, k_4) + \cdots$$

$$+ \quad \frac{(p)!}{(\frac{p-3}{2})! 2^{(\frac{p-3}{2})}} \sigma_o^{(p-3)} \sum_{k_4=0}^{\infty} \cdots \sum_{k_{\frac{p+3}{2}}=0}^{\infty} g_p(k_1, k_2, k_3, k_4, k_4, \cdots, k_{\frac{p+3}{2}}, k_{\frac{p+3}{2}}) \Bigg]$$

$$= 0 \tag{B.18}$$

Since from (B.17) the first term (in brackets) is identically zero, the second term is required to be equal to zero for all h_3. This implies that

$$g_{3(p)}(k_1,k_2,k_3)+60\sigma_o^2\sum_{k4=0}^{\infty}g_{5(p)}(k_1,k_2,k_3,k_4,k_4)+\cdots$$
$$+\frac{(p)!}{(\frac{p-3}{2})!2^{(\frac{p-3}{2})}}\sigma_o^{(p-3)}\sum_{k_4=0}^{\infty}\cdots\sum_{k_{\frac{p+3}{2}}=0}^{\infty}g_p(k_1,k_2,k_3,k_4,k_4,\cdots,k_{\frac{p+3}{2}},k_{\frac{p+3}{2}})=0 \tag{B.19}$$

This operation is continued until the expectation of the product of the G-functional and the last two Volterra kernels h_{p-4} and h_{p-2} is performed to yield the last two odd-kernel relations:

$$g_{p-4(p)}(k_1,k_2,\cdots,k_{p-4})+\frac{(p-2)!}{2!(p-4)!}\sigma_o^2\sum_{k_{p-3}=0}^{\infty}g_{p-2(p)}(k_1,k_2,\cdots,k_{p-3},k_{p-3})$$
$$+\quad\frac{p!}{2!(p-4)!2^2}\sigma_o^4\sum_{k_{p-3}=0}^{\infty}\sum_{k_{p-2}=0}^{\infty}g_p(k_1,k_2,\cdots,k_{p-3},k_{p-3},k_{p-2},k_{p-2})=0 \tag{B.20}$$

and

$$g_{p-2(p)}(k_1,k_2,\cdots,k_{p-2})+\frac{p!}{(p-2)!}\frac{\sigma_o^2}{2}\sum_{k_{p-1}=0}^{\infty}g_p(k_1,k_2,\cdots,k_{p-1},k_{p-1})=0 \tag{B.21}$$

This relation is equivalent to

$$g_{p-2(p)}(k_1,k_2,\cdots,k_{p-2})=-\frac{p!}{1!(p-2)!}\left(\frac{\sigma_o^2}{2}\right)^1\sum_{k_{p-1}=0}^{\infty}g_p(k_1,k_2,\cdots,k_{p-1},k_{p-1}) \tag{B.22}$$

Now substituting this in (B.20) and rearranging terms we find

$$g_{p-4(p)}(k_1,k_2,\cdots,k_{p-4})=$$
$$-\quad\frac{(p-2)!}{2!(p-4)!}\sigma_o^2\times-\frac{p!}{1!(p-2)!}\left(\frac{\sigma_o^2}{2}\right)^1\sum_{k_{p-3}=0}^{\infty}\sum_{k_{p-2}=0}^{\infty}g_p(k_1,k_2,\cdots,k_{p-3},k_{p-3},k_{p-2},k_{p-2})$$
$$-\frac{p!}{2!(p-4)!2^2}\sigma_o^4\sum_{k_{p-3}=0}^{\infty}\sum_{k_{p-2}=0}^{\infty}g_p(k_1,k_2,\cdots,k_{p-3},k_{p-3},k_{p-2},k_{p-2})$$
$$=\frac{p!}{2!(p-4)!}\left(\frac{\sigma_o^2}{2}\right)^2\sum_{k_{p-3}=0}^{\infty}\sum_{k_{p-2}=0}^{\infty}g_p(k_1,k_2,\cdots,k_{p-4},k_{p-3},k_{p-3},k_{p-2},k_{p-2}) \tag{B.23}$$

Finally substituting this in the preceding relations we obtain

$$
\begin{aligned}
& g_{p-6(p)}(k_1, k_2, \cdots, k_{p-6}) = \\
& - \frac{p!}{3!(p-6)!}\left(\frac{\sigma_o^2}{2}\right)^3 \sum_{k_{p-5}=0}^{\infty}\sum_{k_{p-4}=0}^{\infty}\sum_{k_{p-3}=0}^{\infty} g_p(k_1, k_2, \cdots, k_{p-5}, k_{p-5}, \cdots, k_{p-3}, k_{p-3}) \\
& \qquad\qquad \vdots \\
& g_{3(p)}(k_1, k_2, k_3) = \\
& (-1)^{\frac{p-3}{2}} \frac{p!}{(\frac{p-3}{2})!(3)!}\left(\frac{\sigma_o^2}{2}\right)^{\frac{p-3}{2}} \sum_{k_4=0}^{\infty} \cdots \sum_{k_{\frac{p-3}{2}}=0}^{\infty} g_p(k_1, k_2, k3, k_4, k_4, \cdots, k_{\frac{p-3}{2}}, k_{\frac{p-3}{2}}) \\
& g_{1(p)}(k_1) = \\
& (-1)^{\frac{p-1}{2}} \frac{p!}{(\frac{p-1}{2})!(1)!}\left(\frac{\sigma_o^2}{2}\right)^{\frac{p-1}{2}} \sum_{k_2=0}^{\infty} \cdots \sum_{k_{\frac{p-1}{2}}=0}^{\infty} g_p(k_1, k_2, k2, \cdots, k_{\frac{p-1}{2}}, k_{\frac{p-1}{2}}) \qquad \text{(B.24)}
\end{aligned}
$$

Similar results can be obtained if p is assumed to have an even value. The relation between the even-order derived kernels and the leading kernel is identical to those obtained for odd p.

B. AVERAGE OF THE PRODUCT OF TWO DISCRETE G-FUNCTIONALS

In this section we present the proof of the orthogonality of the discrete G-functionals, i.e., we show that

$$\mathcal{E}\{\mathcal{G}_{p_1}[g_{p_1}; w(n)]\mathcal{G}_{p_2}[f_{p_2}; w(n)]\} \qquad \text{(B.25)}$$

is zero for $p_1 \neq p_2$. We show how to develop an expression for the value of the average for $p_1 = p_2 = p$ and give that expression.

From the definition of the G-functional. $\mathcal{G}_p[g_p; w(n)]$ is orthogonal to any homogeneous functional of degree less than p. When $p_1 \neq p_2$ (assuming $p_1 > p_2$) we

expand $\mathcal{G}_{p_2}[f_{p_2}; w(n)]$ using (3.14) and substitute its value in the product to obtain

$$\begin{aligned}\mathcal{E}\{\mathcal{G}_{p_1}[g_{p_1}; w(n)]\mathcal{G}_{p_2}[f_{p_2}; w(n)]\} &= \mathcal{E}\left\{\sum_{m=0}^{\lfloor \frac{p_2}{2} \rfloor} \mathcal{G}_{p_1}[g_{p_1}; w(n)]\mathbf{G}_{p_2-2m}[f_{p_2-2m}; w(n)]\right\} \\ &= \sum_{m=0}^{\lfloor \frac{p_2}{2} \rfloor} \mathcal{E}\{\mathcal{G}_{p_1}[g_{p_1}; w(n)]\mathbf{G}_{p_2-2m}[f_{p_2-2m}; w(n)]\}\end{aligned} \tag{B.26}$$

Each term in the summation is the expectation of the product of the G-functional and a homogeneous functional with degree less than p_1. Thus from the definition of the G-functional, each term in the summation is equal to zero. This proves the first part of the property.

When $p_1 = p_2 = p$ all the terms under summation in (B.26) except the first (for $m = 0$) are equal to zero. In this case (B.26) reduces to

$$\mathcal{E}\{\mathcal{G}_p[g_p; w(n)]\mathcal{G}_p[f_p; w(n)]\} = \mathcal{E}\{\mathcal{G}_p[g_p; w(n)]\mathbf{G}_p[f_p; w(n)]\} \tag{B.27}$$

Now for the sake of clarity let us assume that p is odd and develop the value of the expectation. In this case

$$\begin{aligned}&\mathcal{E}\{\mathcal{G}_p[g_p; w(n)]\mathbf{G}_p[f_p; w(n)]\} \\ &= \mathcal{E}\Bigg\{ \sum_{k_1=0}^{\infty} \cdots \sum_{k_p=0}^{\infty} f_p(k_1, \cdots, k_p) w(n-k_1) \cdots w(n-k_p) \\ &\times \Bigg[\sum_{k_{p+1}=0}^{\infty} g_{1(p)}(k_{p+1}) w(n-k_{p+1}) \\ &+ \sum_{k_{p+1}=0}^{\infty} \sum_{k_{p+2}=0}^{\infty} \sum_{k_{p+3}=0}^{\infty} g_{3(p)}(k_{p+1}, k_{p+2}, k_{p+3}) w(n-k_{p+1}) w(n-k_{p+2}) w(n-k_{p+3}) \\ &+ \cdots + \sum_{k_{p+1}=0}^{\infty} \cdots \sum_{k_{2p}=0}^{\infty} g_p(k_{p+1}, \cdots, k_{2p}) w(n-k_{p+1}) \cdots w(n-k_{2p}) \Bigg] \Bigg\}\end{aligned} \tag{B.28}$$

Since the next steps are algebraically tedious but otherwise straightforward let us just give an outline of those steps. By switching the order of the expectation and

summation we involve the expectation of the product of Gaussian variables. Then the development is continued in the same way as we derived the relation between the kernels. The pairing permutation is divided into sets. In the first set only one time argument of the f_p kernel is set equal to one argument in each of the g kernels and the rest are paired together. In the second set each of three time arguments of the f_p kernel is set equal to one of the arguments of the g kernels, if possible, and the rest are paired together. This continues until in the last set each of the arguments of the f_p kernel is set equal to one argument of the g kernels. Since for the last set there is not enough arguments except in the leading kernel g_p we have

$$
\begin{aligned}
&\mathcal{E}\left\{\mathcal{G}_p[g_p; w(n)]\mathbf{G}_p[f_p; w(n)]\right\} = \\
&p\frac{(p-1)!}{(\frac{p-1}{2})!2^{\frac{p-1}{2}}}\sigma_o^{p+1}\sum_{k_1=0}^{\infty}\cdots\sum_{k_{\frac{p+1}{2}}}^{\infty} f_p(k_1,k_2,k_2,\cdots,k_{\frac{p+1}{2}},k_{\frac{p+1}{2}}) \\
\times\ &\left[g_{1(p)}(k_1)+3\sigma_o^2\sum_{k_{\frac{p+1}{2}+1}}^{\infty} g_{3(p)}(k_1,k_{\frac{p+1}{2}+1},k_{\frac{p+1}{2}+1})\right. \\
+\ &\cdots+p\frac{(p-1)!}{(\frac{p-1}{2})!2^{\frac{p-1}{2}}}\sigma_o^{p-1}\sum_{k_{\frac{p+1}{2}+1}}^{\infty}\cdots\sum_{k_p}^{\infty} g_p(k_1,k_{\frac{p+1}{2}+1},k_{\frac{p+1}{2}+1},\cdots,k_p,k_p)\Bigg] \\
+\ &\frac{p!}{(\frac{p-3}{2})!2^{\frac{p-3}{2}}}\sigma_o^{p+3}\sum_{k_1=0}^{\infty}\cdots\sum_{k_{\frac{p+3}{2}}}^{\infty} f_p(k_1,k_2,k_3,k_4,k_4,\cdots,k_{\frac{p+3}{2}},k_{\frac{p+3}{2}}) \\
\times\ &\left[g_{3(p)}(k_1,k_2,k_3)+60\sigma_o^2\sum_{k_{\frac{p+3}{2}+1}}^{\infty} g_{5(p)}(k_1,k_2,k_3,k_{\frac{p+3}{2}+1},k_{\frac{p+3}{2}+1})\right. \\
+\ &\cdots+\frac{p!}{3!(\frac{p-3}{2})!2^{\frac{p-3}{2}}}\sigma_o^{p-3}\sum_{k_{\frac{p+3}{2}+1}}^{\infty}\cdots\sum_{k_p}^{\infty} g_p(k_1,k_2,k_3,\frac{p+3}{2}+1,k_{\frac{p+3}{2}+1},\cdots,k_p,k_p)\Bigg] \\
&\qquad\vdots \\
+\ &\frac{p!}{2!}\sigma_o^{2p-2}\sum_{k_1=0}^{\infty}\cdots\sum_{k_{p-2}}^{\infty} f_p(k_1,\cdots,k_{p-2},k_{p-1},k_{p-1}) \\
\times\ &\left[g_{p-2(p)}(k_1,\cdots,k_{p-2})+\frac{p!}{(p-2)!2!}\sigma_o^2\sum_{k_{p-1}}^{\infty} g_p(k_1,\cdots,k_{p-2},k_{p-1},k_{p-1})\right] \\
+\ &p!\sigma_o^{2p}\sum_{k_1=0}^{\infty}\cdots\sum_{k_p}^{\infty} f_p(k_1,k_2,\cdots,k_p)g_p(k_1,k_2,\cdots,k_p) \qquad \text{(B.29)}
\end{aligned}
$$

All the sum of terms between brackets were shown to be equal to zero in (B.6)–(B.22).

Therefore

$$
\mathcal{E}\left\{\mathcal{G}_p[g_p; w(n)]\mathcal{G}_p[f_p; w(n)]\right\} = p!\sigma_o^{2p}\sum_{k_1=0}^{\infty}\cdots\sum_{k_p}^{\infty} f_p(k_1,k_2,\cdots,k_p)g_p(k_1,k_2,\cdots,k_p) \qquad \text{(B.30)}
$$

An identical result is obtained for even values of p.

APPENDIX C

COMPUTATION OF THE NON-ZERO VALUED Q-POLYNOMIAL CROSS-MOMENT

The general expression for Q-polynomial cross-moment is given by (4.23), which is repeated here for convenience:

$$\mu_{\boldsymbol{\alpha}}^{(k)}(\boldsymbol{l}) = (\sigma_o^2)^{\frac{\Sigma[\boldsymbol{\alpha}]}{2}} \boldsymbol{\alpha}_1!\boldsymbol{\alpha}_2!\cdots\boldsymbol{\alpha}_k! \sum_{\mathbf{m}_1=0}^{\lfloor\frac{\boldsymbol{\alpha}_1}{2}\rfloor} \sum_{\mathbf{m}_2=0}^{\lfloor\frac{\boldsymbol{\alpha}_2}{2}\rfloor} \cdots \sum_{\mathbf{m}_k=0}^{\lfloor\frac{\boldsymbol{\alpha}_k}{2}\rfloor} \frac{(-1)^{\Sigma[\mathbf{m}]}}{\mathbf{m}_1!\mathbf{m}_2!\cdots\mathbf{m}_k!(2)^{\Sigma[\mathbf{m}]}}$$
$$\times \sum_{\Sigma[\mathbf{M}]=\boldsymbol{\alpha}-2\mathbf{m}} \frac{1}{2^{\frac{tr(\mathbf{M})}{2}}} \prod_{j_1=1}^{k(N_L+1)} \frac{(C_y(j_1,j_1))^{\frac{M_{\mathbf{m}}(j_1,j_1)}{2}}}{(\frac{M(j_1,j_1)}{2})!} \prod_{j_2=j_1+1}^{k(N_L+1)} \frac{(C_y(j_1,j_2))^{M(j_1,j_2)}}{M(j_1,j_2)!} \tag{C.1}$$

The goal here is to develop conditions for those terms in the expression which are zero or sum to zero since those terms will not need to be calculated. The development of sufficient conditions for the general case is very lengthy and requires more space than is feasible in this exposition. Instead, let us demonstrate these cases to derive the sufficient conditions for the example given in Chapter IV and show that the result can be generalized.

We begin by specializing (C.1) to the case of $k = 2$ and rearranging the terms for our example to write

$$\mu_{\boldsymbol{\alpha}}^{(2)}(l) = (\sigma_o^2)^{\frac{\Sigma[\boldsymbol{\alpha}]}{2}} \boldsymbol{\alpha}_1!\boldsymbol{\alpha}_2!$$
$$\times \sum_{m_1=0}^{t_1} \frac{(-1)^{m_1}}{m_1!2^{m_1}} \underbrace{\sum_{i_1=0}^{\lfloor\frac{\alpha_{10}-2m_1}{2}\rfloor} \sum_{i_2=0}^{\alpha_{10}-2i_1} \sum_{i_3=0}^{\alpha_{10}-2i_1-i_2}}_{i_4=\alpha_{10}-2m_1-2i_1-i_2-i_3\geq 0} \frac{r_0^{i_2}(l)r_1^{i_3}(l)r_2^{i_4}(l)}{i_1!i_2!i_3!i_4!2^{i_1}} \underbrace{\sum_{m_4=0}^{t_4} \frac{(-1)^{m_4}}{m_4!2^{m_4}} \frac{1}{i_{10}!2^{i_{10}}}}_{C}$$
$$\times \sum_{m_2=0}^{t_2} \frac{(-1)^{m_2}}{m_2!2^{m_2}} \underbrace{\sum_{i_5=0}^{\lfloor\frac{\alpha_{11}-2m_2}{2}\rfloor} \sum_{i_6=0}^{\alpha_{11}-2i_5}}_{i_7=\alpha_{11}-2m_2-2i_5-i_6\geq 0} \frac{r_0^{i_6}(l)r_1^{i_7}(l)}{i_5!i_6!i_7!2^{i_5}} \underbrace{\sum_{m_5=0}^{t_5} \frac{(-1)^{m_5}}{m_5!2^{m_5}} \frac{1}{i_{11}!2^{i_{11}}}}_{B}$$
$$\times \sum_{m_3=0}^{t_3} \frac{(-1)^{m_3}}{m_3!2^{m_3}} \underbrace{\sum_{i_8=0}^{\lfloor\frac{\alpha_{12}-2m_2}{2}\rfloor}}_{i_9=\alpha_{12}-2m_3-2i_8\geq 0} \frac{r_0^{i_9}(l)}{i_8!i_9!2^{i_8}} \underbrace{\sum_{m_6=0}^{t_6} \frac{(-1)^{m_6}}{m_6!2^{m_6}} \frac{1}{i_{12}!2^{i_{12}}}}_{A} \qquad \text{(C.2)}$$

In this arrangement the values of the m_j and the i_j are interdependent due to the condition (4.19). Therefore when we rearrange the summation the upper limits designated by t_1, t_2, t_3, t_4, t_5 and t_6 must be determined so that none of the i_j assumes a negative value and results in a rejected configuration of the matrix $\mathbf{M}$. When we consider these modified summations in detail there are many well defined cases in which terms clearly sum to zero. As we show later, eliminating these cases avoids the unnecessary computations which represent most of the terms in the expression.

Starting with the innermost summation in (C.2) marked A, let us proceed to determine the upper limit. From the condition specified by the last row of $\mathbf{M}$ in (4.19) we have

$$i_{12} = \frac{\alpha_{22} - i_4 - i_7 - i_9}{2} - m_6 \qquad \text{(C.3)}$$

The following restrictions need to be imposed on the relations between the summation parameters:

1. The quantity $\alpha_{22} - i_4 - i_7 - i_9$ must have an even value so that i_{12} is an integer.

2. Since i_4, i_7 and i_9 are non-negative integers and sum to a value less than α_{22} then

$$\frac{\alpha_{22} - i_4 - i_7 - i_9}{2} \leq \frac{\alpha_{22}}{2} \tag{C.4}$$

3. Since i_{12} must be non-negative

$$m_6 \leq \frac{\alpha_{22} - i_4 - i_7 - i_9}{2} \tag{C.5}$$

This means that we must set the upper limit of the summation t_6 to the value on the right of (C.5) to prevent i_{12} from assuming negative values, i.e.,

$$t_6 = \frac{\alpha_{22} - i_4 - i_7 - i_9}{2} \tag{C.6}$$

With these conditions the term A in the expression becomes

$$\frac{1}{2^{t_6} t_6!} \sum_{m_6=0}^{t_6} \frac{(-1)^{m_6} t_6!}{m_6!(t_6 - m_6)!} = \frac{(1-1)^{t_6}}{2^{t_6} t_6!} \tag{C.7}$$

which is zero for all non-zero values of t_6. This means that for some permutations the corresponding terms in the cumulant expression are summations of quantities that have a common factor of zero. Then to save the wasted effort of computing these terms we set the relation between the summation parameters to avoid these cases. This means that we let $t_6 = 0$, i.e..

$$\alpha_{22} - i_4 - i_7 - i_9 = 0 \tag{C.8}$$

which implies (using C.3) that

$$i_{12} = -m_6 \tag{C.9}$$

Since both i_{12} and m_6 must be non-negative integers the result is

$$\boxed{m_6 = i_{12} = 0} \tag{C.10}$$

and the term A in (C.2) equals one. This important relation means that because m_6 equals zero we consider only the first term of the Hermite polynomial $H_2(y_2(n+l))$.

Because i_{12} equals zero, the pairing of the y_i does not contain the value $\mathcal{E}\{y_2(n+l)y_2(n+l)\} = r_0(0)$ as a multiplier.

Now, from (C.8) and the relation between i_8 and i_9 given in the third row of (4.19) we have

$$\begin{aligned} i_8 &= \frac{\alpha_{12} - 2m_3 - i_9}{2} \\ &= \frac{\alpha_{12} - \alpha_{22} + i_4 + i_7}{2} - m_3 \end{aligned} \tag{C.11}$$

and the summation over i_8 disappears because i_8 is forced to have a specific value. In addition, the upper limit of summation over m_3 must be restricted to the cases that have non-negative integer values of i_8. Therefore, letting

$$t_3 = \frac{\alpha_{12} - \alpha_{22} + i_4 + i_7}{2}$$

the last line in (C.2) starting with the summation over i_3 is

$$\begin{aligned} &\sum_{m_3=0}^{t3} \frac{(-1)^{m_3}}{m_3! 2^{m_3}} \frac{r_0^{\alpha_{22}-i_4-i_7}(l)}{(t_3 - m_3)!(\alpha_{22} - i_4 - i_7)! 2^{t_3 - m_3}} \\ &= \frac{r_0^{\alpha_{22}-i_4-i_7}(l)}{(\alpha_{22} - i_4 - i_7)! t_3! 2^{t_3}} \sum_{m_3=0}^{t3} \frac{(-1)^{m_3} t_3!}{m_3!(t_3 - m_3)!} \\ &= \frac{r_0^{\alpha_{22}-i_4-i_7}(l)}{(\alpha_{22} - i_4 - i_7)!} \frac{(1-1)^{t_3}}{t_3! 2^{t_3}} \end{aligned} \tag{C.12}$$

Similar to the above discussion this term equals zero for all non-zero values of t_3. Hence the computational effort is avoided if we let $t_3 = 0$, which results in

$$\boxed{m_3 = i_8 = 0} \tag{C.13}$$

and

$$\alpha_{12} = \alpha_{22} - i_4 - i_7 \tag{C.14}$$

The entire term starting with the summation on m_3 (last line of (C.2)) then equals

$$\frac{r_0^{\alpha_{12}}(l)}{\alpha_{12}!} \tag{C.15}$$

The term marked B in the expression (C.2) is treated in the same way to obtain

$$\frac{1}{2^{t_5}t_5!}\sum_{m_5=0}^{t_5}\frac{(-1)^{m_5}t_5!}{m_5!(t_5-m_5)!} \tag{C.16}$$

where

$$t_5 = \frac{\alpha_{21}-i_3-i_6}{2} \tag{C.17}$$

This has a non-zero value for $t_5 = 0$, resulting in

$$\boxed{m_5 = i_{11} = 0} \tag{C.18}$$

and

$$i_6 = \alpha_{21} - i_3 \tag{C.19}$$

which implies that i_6 has a specific value for each value of i_3, and the summation over i_6 disappears. From (C.14) i_7 has a specific value given by

$$i_7 = \alpha_{22} - \alpha_{12} - i_4 \tag{C.20}$$

and the condition given in the second row of (4.19) implies that

$$i_7 = \alpha_{11} - 2m_2 - 2i_5 - i_6 \tag{C.21}$$

Therefore from the above

$$i_5 = \frac{\alpha_{11}+\alpha_{12}-\alpha_{21}-\alpha_{22}+i_3+i_4}{2} - m_2 \tag{C.22}$$

and the double summation over i_5 and i_6 disappears. The term starting with the summation over m_2 (second to last line in (C.2)) reduces to

$$\frac{r_0^{\alpha_{21}-i_3}(l)r_1^{\alpha_{12}+\alpha_{22}-\alpha_{12}-i_4}(l)}{(\alpha_{21}-i_3)!(\alpha_{12}+\alpha_{22}-\alpha_{12}-i_4)!t_2!2^{t_2}}\sum_{m_2=0}^{t_2}\frac{(-1)^{m_2}t_2!}{m_2!(t_2-m_2)!} \tag{C.23}$$

where

$$t_2 = \frac{\alpha_{11}+\alpha_{12}-\alpha_{21}-\alpha_{22}+i_3+i_4}{2} \tag{C.24}$$

As in the previous cases (C.23) sums to the binomial term $(1-1)^{t_2}$ which requires that $t_2 = 0$ and results in the relations

$$\boxed{m_2 = i_5 = 0} \tag{C.25}$$

and

$$i_4 + i_3 = \alpha_{21} + \alpha_{22} - \alpha_{11} - \alpha_{12} \tag{C.26}$$

Consequently the term represented by the second line of (C.2) becomes

$$\frac{r_0^{\alpha_{21}-i_3}(l) r_1^{\alpha_{22}-\alpha_{12}+i_4}(l)}{(\alpha_{21}-i_3)!(\alpha_{22}-\alpha_{12}+i_4)!} \tag{C.27}$$

In a similar manner the term marked C in (C.2) is put in the form

$$\frac{1}{2^{t_4} t_4!} \sum_{m_4=0}^{t_4} \frac{(-1)^{m_4} t_4!}{m_4!(t_4-m_4)!} \tag{C.28}$$

where

$$t_4 = \frac{\alpha_{20} - i_2}{2} \tag{C.29}$$

and summed to the binomial $(1-1)^{t_4}$. Then it is required that $t_4 = 0$ which yields

$$\boxed{m_4 = i_{10} = 0} \tag{C.30}$$

Then from (C.26) and the relation

$$i_4 = \alpha_{10} - 2m_1 - 2i_1 - i_2 - i_3 \tag{C.31}$$

we obtain

$$\begin{aligned} i_1 &= \frac{\alpha_{10} + \alpha_{11} + \alpha_{12} - \alpha_{20} - \alpha_{21} - \alpha_{22}}{2} - m_1 \\ i_2 &= \alpha_{20} \\ i_4 &= \alpha_{11} + \alpha_{22} - \alpha_{12} - \alpha_{21} - i_3 \end{aligned} \tag{C.32}$$

To guarantee that i_1 is a non-negative integer, we must have

$$t_1 = \frac{\alpha_{10} + \alpha_{11} + \alpha_{12} - \alpha_{20} - \alpha_{21} - \alpha_{22}}{2} \geq m_1 \tag{C.33}$$

The total term then becomes

$$\sum_{m_1=0}^{t_1} \frac{(-1)^{m_1}}{m_1! 2^{m_1}} \sum_{i_3+i_4=\alpha_{11}+\alpha_{22}-\alpha_{12}-\alpha_{21}} \frac{r_0^{\alpha_{20}}(l) r_1^{i_3}(l) r_2^{i_4}(l)}{(t_1-m_1)!\alpha_{20}! i_3! i_4! 2^{t_1-m_1}}$$

$$= \sum_{i_3+i_4=\alpha_{11}+\alpha_{22}-\alpha_{12}-\alpha_{21}} \frac{r_0^{\alpha_{20}}(l) r_1^{i_3}(l) r_2^{i_4}(l)}{\alpha_{20}! i_3! i_4! t_1! 2^{t_1}} \sum_{m_1=0}^{t_1} \frac{(-1)^{m_1} t_1!}{m_1!(t_1-m_1)!} \qquad \text{(C.34)}$$

which is only non-zero if $t_1 = 0$. This means that

$$\boxed{m_1 = i_1 = 0} \qquad \text{(C.35)}$$

and

$$\alpha_{10} + \alpha_{11} + \alpha_{12} = \alpha_{20} + \alpha_{21} + \alpha_{22} \qquad \text{(C.36)}$$

$$i_3 + i_4 = \alpha_{10} - \alpha_{20} \qquad \text{(C.37)}$$

By combining all of these results, the value of the Q-polynomial cumulant in this example becomes

$$\mu_{\boldsymbol{\alpha}}^{(2)}(\boldsymbol{l}) = (\sigma_o^2)^{\frac{\Sigma[\boldsymbol{\alpha}]}{2}} \boldsymbol{\alpha}_1! \boldsymbol{\alpha}_2! \sum_{i_3=0}^{\alpha_{10}-\alpha_{20}} \frac{r_0^{i_2+i_6+i_9}(l) r_1^{i_3+i_7}(l) r_2^{i_4}(l)}{i_2! i_3! i_4! i_7! i_9!} \qquad \text{(C.38)}$$

where the quantities i_2, i_4, i_6, i_7 and i_9 depend upon the value of i_3 as follows

$$\begin{aligned} i_2 &= \alpha_{20} \\ i_4 &= \alpha_{10} - \alpha_{20} - i_3 \\ i_6 &= \alpha_{21} - i_3 \\ i_7 &= \alpha_{20} + \alpha_{22} - \alpha_{10} - \alpha_{12} + i_3 \\ i_9 &= \alpha_{12} \end{aligned} \qquad \text{(C.39)}$$

In general to compute only the non-zero terms in the Q-polynomial cross-moment we need to maintain the following:

1. The Q-polynomial cross-moment is computed by computing the expectation of the product of *only the first terms* in the Hermite polynomials in the expansion. This corresponds to setting $\mathbf{m} = \mathbf{0}$ in (4.23).

2. This expectation is taken in a special way such that the pairing permutations do not include those that result in one or more autocorrelation functions of any of the outputs at zero lag, $r_0(0)$. This corresponds to setting the values of the diagonal entries of the multiplicity matrix $\mathbf{M}$ equal to zero.

APPENDIX D

COMPUTATIONAL METHODS AND OPTIMIZATION USING THE EXTENDED KALMAN FILTER

In Chapter V we are concerned with finding a model, through an optimization procedure, that matches the a set of cumulant values computed from data. Since the optimization is an iterative procedure, it is necessary to have an efficient way of to compute the cumulants in (5.21). Also, since the optimization uses gradient information, it is necessary to have an efficient way to compute the derivatives of the cumulants with respect to the model parameters. The analysis presented in Chapter IV provides the means for this efficient computation.

A. PROCEDURE FOR COMPUTING THE CUMULANT VECTOR $\mathbf{X}_m$

Since the Q-polynomial cumulants are computed directly from the vectors of indices, the structure of the arrays in (5.21) are the same for all the models that have the same dimensions N_L and N_N. Therefore for every pair of dimensions $[N_L, N_N]$ a table-like array of numerical values can be computed and stored in a library. This array contains all the information needed in (4.21). Each row in this array corresponds to one term in (4.21) and has $3 + k(k-1)/2$ fields defined as follows:

1. The first field has the index i of the k coefficients $c_{\boldsymbol{\alpha}}$, i.e., it specifies one element in $\mathbf{c}^{\otimes k}$.

2. The second field is the power of σ_o^2, which equals half the sum of all the components of the vectors of indices $\boldsymbol{\alpha}$.

3. The third field is a numerical value that results from dividing the product of the factorials of the components of the vectors of indices by the factorials of the nonzero entries of the multiplicity matrix $\mathbf{M}$.

4. Each of the last $k(k-1)/2$ fields has size $N_L + 1$ and corresponds to one block $\mathbf{R}(l)$ in the matrix $\mathbf{C}_y(\boldsymbol{l})$ in (4.17) with l either one of the components of $\mathbf{l}$ or a difference between two of them. The entries of each of these fields are the multiplicity of the corresponding Laguerre cross-correlation sequence $r_d(l)$.

According to (5.21) this array would be N^k long where the number of model coefficients, N, increases rapidly with the model dimensions N_L and N_N. This means that a slight change in the model dimensions or in the cumulant order would increase the computational cost dramatically. Also to compute the value of the cross-correlation sequence $r_d(l)$ every time it is needed using (3.96) is very costly, especially when the parameter ρ is changed as it is in every iteration of the optimization procedure. Fortunately, as we have seen, the model has structure which dramatically reduces the computational cost. Recall that:

1. The cumulants of the Q-polynomials are identically zero for most of the combinations of the vectors of indices $\boldsymbol{\alpha}$. This can be detected by running a simple test on $\boldsymbol{\alpha}$ to determine whether the corresponding term is identically zero or not. Using this test, presented in Chapter IV, greatly reduces the length of the stored array. This results in a significant saving of both storage and computation. Tables D.1 and D.2 show the actual length of this array compared to the quantity N^k for the cases of second and third-order cumulants respectively.

N_L/N_N	1	2	3	4
1	$\frac{4}{4}$	$\frac{13}{25}$	$\frac{29}{81}$	$\frac{54}{196}$
2	$\frac{9}{9}$	$\frac{45}{81}$	$\frac{145}{361}$	$\frac{370}{1156}$
3	$\frac{16}{16}$	$\frac{116}{196}$	$\frac{516}{1156}$	$\frac{1741}{4761}$
4	$\frac{25}{25}$	$\frac{250}{400}$		
5	$\frac{36}{36}$	$\frac{477}{729}$		

TABLE D.1: The ratio of the non-zero terms in the second-order cumulant function

N_L/N_N	1	2	3	4
1		$\frac{42}{125}$	$\frac{225}{729}$	$\frac{824}{2774}$
2		$\frac{213}{729}$	$\frac{1731}{6859}$	$\frac{10060}{39304}$
3		$\frac{759}{2744}$	$\frac{8746}{39304}$	$\frac{8746}{328509}$
4		$\frac{2171}{8000}$		
5		$\frac{5327}{19683}$		

TABLE D.2: The ratio of the non-zero terms in the third-order cumulant function

2. Both ρ and $\boldsymbol{l}$ are needed to compute the cross-correlation sequences $r_d(l_i)$. Since the maximum value of d equals N_L and the maximum value of l_i is the maximum component in $\boldsymbol{l}$, another array of the cross-correlation sequences can be formed such that its $(i,j)^{th}$ entry is $r_i(j)$ for $i = 0,1,2,\cdots,N_L$ and $j = 0,1,2,\cdots,\max\{\boldsymbol{l}\}$. The i^{th} row corresoponds to the sequence of order i while the j^{th} column corresponds to the value of the sequences at lag value j. Instead of using (3.96) to compute each $r_i(j)$ we use the recursion relation (3.92)

$$r_i(j) = \rho r_i(j-1) + \rho r_{i-1}(j) - r_{i-1}(j-1)$$

In this case we need to compute the values of $r_i(j)$ along the first row and the first column only and then compute the rest of the values recursively. The values in the first row are simply

$$r_0(j) = \rho^j \tag{D.1}$$

while the entries of the first column are given by

$$r_i(0) = \begin{cases} 1 & \text{for } i = 0 \\ 0 & \text{otherwise} \end{cases} \tag{D.2}$$

B. PROCEDURE FOR COMPUTING THE GRADIENT MATRIX H

In the optimization algorithm to be discussed shortly, it is necessary to compute the gradients of the vector $\mathbf{X}_m$ with respect to the model parameters. The gradient matrix $\mathbf{H}$ is thus defined as

$$\mathbf{H} = [\nabla_{\mathbf{c}}\mathbf{X}_m \;\; \nabla_{\sigma_o}\mathbf{X}_m \;\; \nabla_{\rho}\mathbf{X}_m] \tag{D.3}$$

Although this matrix can be formed from (5.21) it is preferable to avoid the unnecessary computations by utilizing the same information stored and used to compute

the vector $\mathbf{X}_m$. Therefore each of the three partitions of $\mathbf{H}$ is obtained by modifying specific fields in the cumulant data array and using this modified array to compute the gradients.

From (5.21) the gradient with respect to the vector $\mathbf{c}$ can be obtained by

$$\nabla_{\mathbf{c}}\mathbf{X}_m = \mathcal{C}^{(k)}\nabla_{\mathbf{c}}\mathbf{c}^{\otimes k} \tag{D.4}$$

Then the matrix $\mathcal{C}^{(k)}$ does not need to be recomputed; only the information obtained from the first field of the array of cumulant data is used to form another table with the same length. Each row of this table has N elements specifying the value of gradient of this term with respect to each of the model coefficients $c_{\boldsymbol{\alpha}}$. If a coefficient $c_{\boldsymbol{\alpha}}$ is not included in the product of coefficients then the corresponding entry in the table is zero. If the coefficient is included then the corresponding value in the table is

$$\nabla_{c_{\boldsymbol{\alpha}_i}} c_{\boldsymbol{\alpha}_i}^p c_{\boldsymbol{\alpha}_j}\cdots c_{\boldsymbol{\alpha}_k} = p c_{\boldsymbol{\alpha}_i}^{p-1} c_{\boldsymbol{\alpha}_j}\cdots c_{\boldsymbol{\alpha}_k} \tag{D.5}$$

Thus we need only to multiply the corresponding term by a value equal to the coefficient multiplicity in the first field then reduce the multiplicity by one.

Since a change in ρ changes terms in the cumulant expression that have the form

$$r_{d_1}^{\nu_1}(l_1) r_{d_2}^{\nu_2}(l_2)\cdots r_{d_p}^{\nu_p}(l_p)$$

the gradient of this quantity with respect to ρ is given by

$$r_{d_1}^{\nu_1}(l_1) r_{d_2}^{\nu_2}(l_2)\cdots r_{d_p}^{\nu_p}(l_p)\left[\nu_1\frac{r'_{d_1}(l_1)}{r_{d_1}(l_1)} + \nu_2\frac{r'_{d_2}(l_2)}{r_{d_2}(l_2)} + \cdots + \nu_p\frac{r'_{d_p}(l_p)}{r_{d_p}(l_p)}\right]$$

Notice that the derivative of each term is obtained by multiplying this term by

$$\left[\nu_1\frac{r'_{d_1}(l_1)}{r_{d_1}(l_1)} + \nu_2\frac{r'_{d_2}(l_2)}{r_{d_2}(l_2)} + \cdots + \nu_p\frac{r'_{d_p}(l_p)}{r_{d_p}(l_p)}\right]$$

where the ν_i are obtained from the original data array and the cross-correlation sequence values have already been computed and stored for a given range of orders

and time lags. Hence it remains to compute the values of $r'_d(l)$ for the same range. We can use (3.92) to obtain a recursion for the derivatives of the correlation sequences:

$$r'_i(j) = \rho r'_i(j-1) + \rho r'_{i-1}(j) - r'_{i-1}(j-1) + \frac{r_i(j) + r_{i-1}(j-1)}{\rho} \tag{D.6}$$

then precompute these values and store them in a separate array. In this array the entries of the first column, $r'_i(0)$ are all zeros and the first row is given by

$$r'_0(j) = j\rho^{j-1}$$

C. EXTENDED KALMAN FILTER ALGORITHM

For the extended Kalman filter optimization procedure the model parameters are arranged in a vector form

$$\zeta = [\mathbf{c}^T \quad \rho]^T \tag{D.7}$$

Then the recursive estimation of the value of the vector ζ that minimizes the quantity E^TE is proceeds according to the structure of the Kalman Filter technique [39]. Assuming that at the i^{th} recursion we have the value $\hat{\zeta}_i$ and the predection of $\hat{\zeta}_{i+1}$ is linear function of $\hat{\zeta}_i$ we can build our algorithm as follows

$$\begin{aligned}
\hat{\zeta}_{i/i} &= \hat{\zeta}_{i/i-1} + K_i[\mathbf{X}_p - \mathbf{X}_m(\hat{\zeta}_{i/i-1})] \\
\hat{\zeta}_{i+1/i} &= \hat{\zeta}_{i/i} \\
K_i &= \mathbf{P}_{i/i-1}\mathbf{H}_i\mathbf{W}_i^{-1} \\
\mathbf{W}_i &= \mathbf{H}_i^T\mathbf{P}_{i/i-1}\mathbf{H}_i + \mathbf{I} \\
\mathbf{P}_{i/i} &= \mathbf{P}_{i/i-1} - \mathbf{P}_{i/i-1}\mathbf{H}_iW_i^{-1}\mathbf{H}_i^T\mathbf{P}_{i/i-1} \\
\mathbf{P}_{i+1/i} &= \mathbf{P}_{i/i}
\end{aligned} \tag{D.8}$$

Where in this notation the matrix $\mathbf{H}_i$ is the gradient matrix in (D.3) computed for the model parameters vector $\hat{\zeta}_{i/i}$. We start the algorithm with proper initialization of ζ and $\mathbf{P}$.

LIST OF REFERENCES

[1] N. Wiener, *Extrapolation, Interpolation and Smoothing of Stationary Time Series*, John Wiley, New York, 1949.

[2] H. Wold, *The Analysis of Stationary Time Series*, Almquist and Wicksell, Uppsala, Sweden, 2nd ed. edition, 1954.

[3] A. N. Kolmogorov, "Interpolation and Extrapolation of Stationary Random Sequences," *Bulletin of the Academy of Sciences, USSR*, 5:3–14, 1941, (In Russian with a short form in German).

[4] Howell Tong, *Nonlinear Time Series, A Dynamical System Approach*, Oxford University Press, New York, 1980.

[5] David Guy and Keh-Shin Lii, "Nonlinear Discrimination within the Higher-Order Cumulant Structure," Proc. of Moments and Signal Processing Conf., Naval Postgraduate School, Monterey, California, March 1992.

[6] J. K. Tugnait, "Identification of Nonminimum Phase Linear Stochastic Systems," *Automatica*, 22, pp. 457–464, July 1987.

[7] J. K. Tugnait, "Identification of Linear Stochastic Systems via Second- and Fourth-order Cumulant Matching," *IEEE Trans. Inform. Theory*, IT-33, pp. 393–407, May 1987.

[8] G. B. Giannakis and Jerry M. Mendel, "Cumulant-Based Order Determination of Non-Gaussian ARMA Model," *IEEE Trans. Acoustics, Speech and Signal Processing*, 38(8), 1990.

[9] Jose A. R. Fonollosa and J. Vidal, "System Identification Using a Linear Combination of Cumulant Slices," *IEEE Trans. Signal Processing*, 41(7), pp. 2405–2412, July 1993.

[10] K. S. Lii and M. Rosenblatt, "Deconvolution and Estimation of Transfer Function Phase and Coefficient for Non-Gaussian Linear Processes," *Ann. Stat.*, 10(4),pp. 1195–1208, 1982.

[11] R. Pan and C. L. Nikias, "The Complex Cepstrum of Higher-Order Cumulants and Nonminimum Phase System Identification," *IEEE Trans. Acoust. Speech and Signal Processing*, 36, pp. 186–205, Feb. 1988.

[12] A. Tanju Erdem and A. Murat Tekalp, "On the Measure of the Set of Factorizable Polynomial Bispectra," *IEEE Trans. Acoustics, Speech and Signal Processing*, 38(9), pp. 1637–1639, September 1990.

[13] A. Murat Tekalp and A. Tanju Erdem, "Higher-Order Spectrum Factorization in One and Two Dimensions with Application in Signal Modeling and Nonminimum Phase System Identification," *IEEE Trans. Acoustics, Speech and Signal Processing*, 37(10), pp. 1537–1549, October 1989.

[14] James G. Stevens, *An Investigation of Multivariate Adapative Regression Splines for Modeling and Analysis of Univariate and Semi-Multivariate Time Series System*, PhD thesis, Naval Postgraduate School, Monterey, California, 1991.

[15] P. A. W. Lewis and J. G. Stevens, "Nonlinear Modeling of Time Series Using Multivariate Adaptive Regression Splines (MARS)." *J. Am. Stat. Assoc.*, 87. pp. 864–877. 1991.

[16] P. A. W. Lewis, B. K. Ray, and J. G. Stevens, "Modeling Time Series by Using Multivariate Adaptive Regression Splines (MARS)," In *Time Series Prediction*, May 1992.

[17] Martin Schetzen, *The Volterra and Wiener Theories of Nonlinear Systems*, John Wiley & Sons, New York, 1980.

[18] M. B. Priestley, *Nonlinear and Non-stationary Time Series Analysis*, Academic Press,Inc., San Diego, 1988.

[19] Y. S. Cho, S. B. Kim, and E. J. Powers. A Digital Technique to Estimate Second-Order Distortion Using Higher-Order Coherence Spectra, *IEEE Trans. Signal Processing*, 40(5), May 1992.

[20] S. W. Nam, S. B. Kim, and E. J. Powers, "On the Identification of a Third-Order Volterra Nonlinear System Using a Frequency-Domain Block RLS Adaptive Algorithm," In *IEEE ICASSP'90*, volume 5, pp. 2306–2309, April 1990.

[21] John Tsimbson and Kenneth V. Lever, "Measurement , Composition and Inversion of Orthogonal Representation of Memoryless Nonlinearities," In *IEEE Signal Processing Workshop on Higher-Order Statistics*, June 1993.

[22] Norbert Wiener, *Nonlinear Problems in Random Theory*, John Wiley & Sons, New York, 1958.

[23] Jerry M. Mendel, "Tutorial on Higher-Order Statistics (Spectra) in Signal Processing and System Theory," *Proceeding of the IEEE*, 79(3), March 1991.

[24] Murray Rosenblatt, *Stationary Sequences and Random Fields*. Birkhauser, Boston, 1985.

[25] Gabor Szego, *Orthonormal Polynomials*, American Mathematical Society, New York, 1939.

[26] Morris J. Gottlieb, "Concerning Some Polynomials Orthogonal on a Finite or Enumerable Set of Points," *Amer. J. Math.*, 60, pp. 453–458, 1938.

[27] H. Stahl and Vilmos Totik, *General Orthogonal Polynomials*, Cambridge University Press, New York, 1992.

[28] P. M. Makila, "Approximation of Stable Systems by Laguerre Filters," *Automatica*", 26(2), pp. 333–345, 1990.

[29] B. Wahlberg, "System Identification Using Laguerre Filters," *IEEE Transactions on Automatic Control*, 36(5), May 1991.

[30] C. Zervos, P. R. Belanger, and G. A. Dumont, "On PID Controllers Tuning using Orthonormal Series Identification." *Automatica*, 24, pp. 346–352, 1987.

[31] A. C. Den Brinke, "Laguerre-Domain Adaptive Filters," *IEEE Transactions on Signal Processing*, 42(4), April 1992.

[32] A. C. Den Brinker, "The Discrete Time Running Cross-Correlation Function Estimated from the Laguerre Spectra." In *Signal Processing VI:Theories and Applications*, Elsevier Science Publishers, British Vancouver, 1992.

[33] Chrysostomoms L. Nikias and Athina P. Petropulu, *Higher Order Spectra Analysis: A Nonlinear Signal Processing Framework*, Prentice Hall, 1993.

[34] A. J. Lawrance and Peter A. W. Lewis, "Higher-Order Residual Analysis for Nonlinear Time Series with Autoregressive Correlation Structure," *International Statistical Review*, 55(1), 1987.

[35] T. Subba Rao and M. M. Gabr, "A Test for Linearity of Stationary Time Series," *J. Time Series Analysis*, 1, pp. 145–158, 1980.

[36] M. Hinich, "Testing for Gaussianity and Linearity of a Stationary Time Series," *J. Time Series Analysis*, 3, pp. 169–176, 1982.

[37] C. W. Therrien, *Discrete Random Signals and Statistical Signal Processing*, Prentice Hall, New Jersey, 1992.

[38] A. S. Weigend and N. A. Gershenfeld, Time Series Prediction, In *Proceeding of NATO Advanced Research on Comparative Time Series Analysis*, May 1992.

[39] B. D. O Anderson and John B. Moore, *Optimal Filtering*, Prentice-Hall, New Jersey, 1979.

INITIAL DISTRIBUTION LIST

		No. Copies
1.	Defense Information Center Cameron Station Alexandria, VA 22304-6145	2
2.	Library Code 52 Naval Postgraduate School Monterey, CA 93943-5101	2
3.	Chairman, Code EC/Mw Department of Electrical and Computer Engineering Naval Postgraduate School Monterey, CA 93943-5121	1
4.	Professor Charles W. Therrien, Code EC/Ti Department of Electrical and Computer Engineering Naval Postgraduate School Monterey, CA 93943-5121	4
5.	Professor Roberto Cristi, Code EC/Cx Department of Electrical and Computer Engineering Naval Postgraduate School Monterey, CA 93943-5121	1
6.	Professor Murali Tummala, Code EC/Tu Department of Electrical and Computer Engineering Naval Postgraduate School Monterey, CA 93943-5121	1
7.	Professor H. H. Loomis, Jr., Code EC/Lm Department of Electrical and Computer Engineering Naval Postgraduate School Monterey, CA 93943-5121	1
8.	Professor Amr Zaky, Code CS/Za Department of Computer Science Naval Postgraduate School Monterey, CA 93943-5118	1

		No. Copies
9.	Professor Peter Lewis, Code OR/Lw Department of Operations Research Naval Postgraduate School Monterey, CA 93943-5219	1
10.	Military Technical College Kobry-Elkoba Cairo, Egypt	2
11.	Col. Attala I. Hashad Armament Authority Research Branch Cairo, Egypt	4