NASA-CR-204830

NAG1-1634

ELECTRICAL AND COMPUTER ENGINEERING
DEPARTMENT

CLEMSON UNIVERSITY

CLEMSON, SC 29634-0915

TR-121796-5604P

AN OVERDETERMINED SYSTEM FOR IMPROVED AUTOCORRELATION BASED SPECTRAL MOMENT ESTIMATOR PERFORMANCE

by

Byron M. Keel

Radar Systems Laboratory
Technical Report No. 22

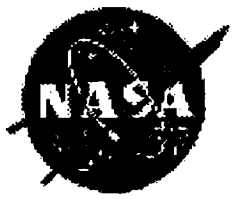
AN OVERDETERMINED SYSTEM FOR IMPROVED AUTOCORRELATION BASED SPECTRAL MOMENT ESTIMATOR PERFORMANCE

by

Byron M. Keel

Technical Report #22
December 17, 1996

Radar Systems Laboratory
Electrical and Computer Engineering Department
Clemson University
Clemson, SC 29634-0915



Laser Radar Sensor Development Signal Processing Studies
Grant NAG-1-1634
National Aeronautics and Space Administration
Langley Research Center
Hampton. VA 23665

ABSTRACT

Autocorrelation based spectral moment estimators are typically derived using the Fourier transform relationship between the power spectrum and the autocorrelation function along with using either an assumed form of the autocorrelation function, e.g., Gaussian, or a generic complex form and applying properties of the characteristic function. Passarelli has used a series expansion of the general complex autocorrelation function and has expressed the coefficients in terms of central moments of the power spectrum. A truncation of this series will produce a closed system of equations which can be solved for the central moments of interest.

The autocorrelation function at various lags is estimated from samples of the random process under observation. These estimates themselves are random variables and exhibit a bias and variance that is a function of the number of samples used in the estimates and the operational signal-to-noise ratio. This contributes to a degradation in performance of the moment estimators.

This dissertation investigates the use autocorrelation function estimates at higher order lags to reduce the bias and standard deviation in spectral moment estimates. In particular, Passarelli's series expansion is cast in terms of an overdetermined system to form a framework under which the application of additional autocorrelation function estimates at higher order lags can be defined and assessed. The solution of the overdetermined system is the least squares solution. Furthermore, an overdetermined system can be solved for any moment or moments of interest and is not tied to a particular form of the power spectrum or corresponding autocorrelation function.

As an application of this approach, autocorrelation based variance estimators are defined by a truncation of Passarelli's series expansion and applied to simulated Doppler weather radar returns which are characterized by a Gaussian shaped power

spectrum. The performance of the variance estimators determined from a closed system is shown to improve through the application of additional autocorrelation lags in an overdetermined system. This improvement is greater in the narrowband spectrum region where the information is spread over more lags of the autocorrelation function. The number of lags needed in the overdetermined system is a function of the spectral width, the number of terms in the series expansion, the number of samples used in estimating the autocorrelation function, and the signal-to-noise ratio. The overdetermined system provides a robustness to the chosen variance estimator by expanding the region of spectral widths and signal-to-noise ratios over which the estimator can perform as compared to the closed system.

ACKNOWLEDGMENTS

The author would like to thank his advisor, Dr. Ernest G. Baxa, Jr., for his technical guidance and personal support. The author would also like to thank the members of his committee, Dr. Frank M. Cholewinski, Dr. John J. Komo, and Dr. Robert J. Schalkoff, for their support in this effort.

The author would like to thank the National Aeronautics and Space Administration which supported this work under grant No. NAG-1-1634 associated with the Wake Vortex Lidar Project at Langley Research Center. Technical support under this grant has been provided by Mr. Grady J. Koch of the Aerospace Electronics Systems Division.

The author would like to thank the Georgia Tech Research Institute for its support in helping the author to complete his degree. In particular, the author would like to acknowledge the support of his past and present management, Admiral Richard H. Truly, Dr. Edward K. Reedy, Dr. Robert N. Trebits, Mr. Guy V. Morris, Dr. Mark A. Richards, and Mr. Trent G. Farill, for their support. The author would also like to thank Dr. Marvin N. Cohen and Mr. Melvin L. Belcher for their allowing the author to jump ship and return to school full time.

The author would like to take this opportunity to thank those family members who have made a significant impact in his life toward this degree. With his deepest heart felt gratitude and love, the author would like to thank his wife, Kelley, for her support and encouragement over the past nine years without whom none of this would have been possible. The author would like to also thank his three children, Katie, Kimberly, and Christopher, from whom the author stole precious time in order to complete this degree. The author looks forward to being a husband and father again. The author would also like to thank his parents, Murray and Willise Keel, who have sacrificed all their lives for the good of their children. In addition,

the author would like to thank his brother, Bruce, his grandparents, Shellie and Zelma Still, and his great aunt, Dalice Mae Birt, for their support. The author also humbly acknowledges the grace and support of his heavenly Father in all that he does.

TABLE OF CONTENTS

Table of Contents (Continued)

LIST OF FIGURES

List of Figures (Continued)

Page

List of Figures (Continued)

Page

List of Figures (Continued)

Page

List of Figures (Continued)

Page

List of Figures (Continued)

Page

List of Figures (Continued)

Page

CHAPTER I

INTRODUCTION

Problem Statement

The need to measure power spectrum parameters is found in many fields including meteorology, seismology, acoustics, and astronomy. The power spectrum is defined as the Fourier transform of the autocorrelation function associated with a stationary random process. Parameters associated with the power spectrum include, but are not limited to, the total power, the mean frequency, the frequency spread (variance), and the skewness. Measurements of these moments can be made in the frequency domain by applying discrete centroiding techniques to approximate the moment definitions. The Fourier based techniques require an estimate of the power spectrum using traditional or modern spectral estimation techniques. Often the amount of data to be processed is quite large, and this places a considerable burden on modern signal processing equipment to meet the real-time requirements.

Fortunately, for spectral moment estimates, Rummler [28, 20] has shown that information compression can be achieved through the autocorrelation function (ACF). Using the relationship between the characteristic function defined in probability theory and the Fourier transform relationship between the autocorrelation function and the power spectral density (PSD), Rummler developed the well-known pulse-pair mean and width estimators which require only estimates of the zeroth and first autocorrelation lags. Zrnic and others [37, 38, 40, 19, 31, 32] have compared the pulse-pair estimators to Fourier based estimators and established under what conditions each is optimum. Other autocorrelation based estimators can be derived from an assumed form of the PSD and associated autocorrelation function. In meteorological processing, the Doppler return (PSD) is often Gaussian shaped [9] with

a corresponding Gaussian shaped autocorrelation function. This fact has been applied to the development of an autocorrelation based variance estimator for Gaussian shaped spectra [38].

The pulse-pair estimators have been used extensively in many fielded systems, but Passarelli [26] has shown that the pulse-pair estimators are a subset of a much larger set of autocorrelation based moment estimators that can be derived from a McLauren series expansion of the complex autocorrelation function. Passarelli shows that the coefficients of the McLauren series expansion can be expressed in terms of moments about the mean of the power spectral density. This series expansion can be truncated and a closed system formed which can be solved for the moment or moments of interest.

The autocorrelation based spectral moment estimators are a function of the estimated autocorrelation function at various lags obtained from the observed random process. In general, there are two estimators used for estimating the autocorrelation function from an ergodic random process. One is unbiased but has a larger variance especially at the higher lags values, and the other is asymptotically unbiased and tends yield a lower variance in the estimate. The unbiased estimator is also known to yield an autocorrelation sequence estimate that may not reach its maximum at the zeroth lag. The variance in the autocorrelation lag estimates is a function of the number of samples used in the estimate and the signal-to-noise ratio.

It follows then that the autocorrelation based spectral moment estimator's performance will be a function of the quality of the autocorrelation lag estimates. In fielded systems, both the number of samples available for estimating the autocorrelation function at various lags and the signal-to-noise ratio of the system are driven by constraints of the physical environment and system requirements for detection and parameter estimation. Therefore, one may be limited in ones ability to improve the quality of the autocorrelation function estimate.

In measurement systems which allow for the estimation of the autocorrelation function at lags beyond the zeroth and first, the opportunity to use more lags may provide improvements in spectral moment estimator performance. The use of additional autocorrelation lags in the case of an overdetermined system has been used in autoregressive moving average (ARMA) spectral estimation to improve spectral resolution for a given (n,m) ARMA model [4]. In addition, higher order lags have been used in ARMA modeling for coefficient estimation [34, 5], and Bruzzone and Kaveh [3] have defined a relative information index to measure the information provided by the autocorrelation function at different lags under various signal-to-noise conditions. Based on the relative information index, a design criterion is defined for selecting those lags which contribute information to the spectral estimate. Also, the use of additional autocorrelation lags has been applied in the case of mean estimation through the poly-pulse-pair [35, 21] to reduce the variance in the estimate. The poly-pulse-pair estimator can easily be derived from Passarelli's expansion. In cases where the autocorrelation function is defined in closed form, such as the Gaussian shaped PSD, a fit of the measured data to the shape of the autocorrelation function has been applied [1, 27]. However, there are cases where the use of additional lags has proven to be of little merit. Srivastava and Jameson [33] have attempted to apply a poly-pulse-pair like approach to the autocorrelation based variance estimators derived for the case of an assumed Gaussian shaped PSD without success.

This work takes Passarelli's expansion which provides a mechanism for relating the central moments of the PSD to the autocorrelation function and cast the closed systems defined by a truncation of the series expansion into overdetermined systems. The additional lags in the overdetermined system are used to improve estimator performance through a reduction in estimator bias and standard deviation. Passarelli's autocorrelation function expansion is not based on an assumed form of the autocorrelation function (e.g., Gaussian shaped, etc.) which, therefore, allows this approach to be applied to any random process.

To assess the framework defined by the overdetermined system for improving moment estimator performance, the variance estimator is chosen for evaluation. The square root of the variance is defined as a width estimate in meteorological signal processing and is used to measure the turbulence associated with an event. Autocorrelation based moment estimators are applied in many meteorological measurement systems. Pulsed Doppler weather radars allow meteorologists to measure rainfall rates (proportional to estimated average return power), average velocity (estimated mean Doppler shift), degree of turbulence (a function of the estimated variance in the Doppler shift), and other physical attributes associated with such phenomena as windshear, tornados, thunderstorms, wake vortices, and other natural or man-induced atmospheric conditions. But with the ranging and Doppler extraction capabilities associated with pulsed Doppler radars over large volumes of space, comes the requirement to process large quantities of data in real-time. The autocorrelation based moment estimators allow one to process the data in real-time without the need to perform the transformation from the autocorrelation domain to the frequency domain. In addition, the Gaussian shaped power spectrum and corresponding autocorrelation function are found to model the power spectrum for many signals including the Doppler return from meteorological events. Therefore, in evaluating the overdetermined variance estimators simulated Doppler weather returns will be used to measure performance.

Contribution to the Field

This work contributes to the field by defining a framework in which to improve the performance of autocorrelation based spectral moment estimators through the inclusion of estimates of the autocorrelation function at higher order lags. The framework is defined as an overdetermined system in terms of a truncation of Passarelli's series expansion. It is shown that the solution of the overdetermined system in terms of Passarelli's expansion yields a least squares solution. The overdetermined system is shown to reduce estimator bias and standard deviation for variance

estimators defined by Passarelli's expansion in the narrowband case and assuming a Gaussian shaped spectrum. Performance bounds are defined for several of the overdetermined variance estimators and related to observed performance. The overdetermined system is robust in that it extends the operating region of variance estimators over a larger region of spectral widths and signal-to-noise ratios. It is shown that the number of lags used in the overdetermined system is a function of the quality of the autocorrelation function estimate (based on the number of samples used in the estimate and the signal-to-noise ratio), the number of terms in the series expansion, and the spectral width.

Organization

This dissertation is organized into five chapters. The current chapter serves to describe the problems addressed and the resulting contributions to the field. Chapter 2 defines autoregressive moving average spectral estimation techniques. This chapter also defines the overdetermined Yule-Walker equations for an AR model and illustrates how they are used to improve spectral estimates. This chapter, therefore, serves as a foundation for casting Passarelli's series expansion in an overdetermined system. Chapter 3 is an overview of Fourier and autocorrelation based spectral moment estimators. This chapter also includes a derivation of Passarelli's autocorrelation based moment estimators and serves as the basic structure from which to build the overdetermined system. Chapter 4 defines the overdetermined system for Passarelli's series expansion and shows how the additional lags can be used to improve estimator performance. Chapter 5 discusses the key results of this dissertation and defines potential future work in this area.

CHAPTER II

ARMA SPECTRAL ESTIMATION

Introduction

This chapter is an overview of autoregressive moving average (ARMA) spectral estimation. The ARMA parameter estimation techniques discussed in this chapter include the overdetermined Yule-Walker equations which are used to increase the amount of information extracted from the autocorrelation function estimate. The application of an overdetermined system in ARMA spectral estimation serves as a basis for extending the overdetermined system to autocorrelation based spectral moment estimators in order to improve estimator performance.

In Appendix A, the Fourier based methods of spectral estimation are defined. These methods assume a realization of N samples of an ergodic random process from which lags $-N \leq k \leq N$ of the autocorrelation function can be estimated. It is observed that these methods show a tendency toward a large variance in the spectral estimate. The Blackman-Tukey method offers a reduction in the variance of the spectral estimate by windowing the autocorrelation estimate. However, windowing in Fourier based techniques reduces the frequency resolution of the estimator and biases the estimate. ARMA spectral estimation techniques provide a means for increasing the spectral resolution while using fewer autocorrelation lags and, thereby, reducing the variance of the estimate.

The spectral factorization property, in Appendix A, states that the power spectrum of a random process can be viewed as the power spectrum of the output of a stable and causal linear system driven by white noise. For the special case where $P_{XX}(z)$ can be expressed as a rational polynomial function in z

$$P_{XX}(z) = \sigma^2 \frac{N(z)}{D(z)}\,, \tag{1}$$

Equation 1 can be expressed as

$$P_{XX}(z) = \sigma^2 \frac{A(z)}{B(z)} \frac{A^*(\frac{1}{z^*})}{B^*(\frac{1}{z^*})} \tag{2}$$

where the linear system's transfer function is defined by

$$H(z) = \frac{A(z)}{B(z)} . \tag{3}$$

This assumed form of the power spectral density lends itself to a class of models termed the autoregressive moving average (ARMA) models. The parameters associated with these models may be obtained from the autocorrelation function estimate as described in the following sections. Two important differences arise when comparing Fourier based spectral estimation techniques and ARMA modeling. The first difference is that the ARMA models do not assume a finite length autocorrelation function as implied by the windowing in the Fourier based methods. This extension of the autocorrelation function leads to spectral estimators which exhibit higher spectral resolution over Fourier based methods. The second difference is that the order of the ARMA models determines the number of autocorrelation lags required to estimate the power spectral density. The number of autocorrelation lags required is usually much lower than that for Fourier based methods.

ARMA Modeling

Theory

ARMA modeling techniques have been successfully used to estimate the PSD and associated parameters in radar signal processing [11], in speech signal processing [18], and in other fields where the signal of interest may be characterized as a random process. The ARMA model consists of a linear system driven by a white noise source, $u(n)$, as shown in Figure 1. The transfer function, $H(z)$, for an ARMA process is expressed as

$$H(z) = \frac{b_0 z^M + b_1 z^{M-1} + \cdots + b_{M-1} z + b_M}{z^N + a_1 z^{N-1} + \cdots + a_{N-1} z + a_N} = \frac{B(z)}{A(z)} . \tag{4}$$

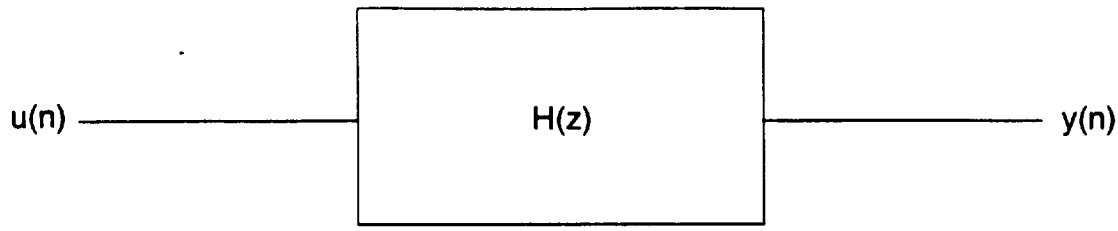


Figure 1. A pictorial of an autoregressive moving average model.

This is a rational transfer function which contains, in general, both poles and zeros. The order of the numerator and denominator is determined by the process being modeled.

The ARMA model can be subdivided into two models having distinct, desirable qualities. The first model is an all-zero model termed the moving average (MA) model. The transfer function for the moving average model is

$$H(z) = b_0 z^M + b_1 z^{M-1} + \cdots + b_{M-1} z + b_M \tag{5}$$

and the corresponding difference equation is

$$y(n) = b_0 u(n) + b_1 u(n-1) + \cdots + b_{M-1} u(n-M+1) + b_M u(n-M) \tag{6}$$

which is a finite impulse response (FIR) filter. The MA model is used to model random processes whose PSD's are smooth or have well defined valleys. The second model is an all-pole model termed the autoregressive (AR) model. The transfer function for the AR model can be expressed as

$$H(z) = \frac{z^N}{z^N + a_1 z^{N-1} + \cdots + a_{N-1} z + a_N} \tag{7}$$

and the corresponding difference equation is

$$y(n) = a_1 y(n-1) + \cdots + a_{N-1} y(n-N+1) + a_N y(n-N) + u(n) \tag{8}$$

which is an infinite impulse response (IIR) filter. The AR model is used to model random processes whose PSD's contain well defined peaks. The question of which

model to use in representing a random process is partially answered by Kolmogorov and Wold. Kolmogorov [16] and Wold [36] have shown that both an infinite order AR(∞) model and an infinite order MA(∞) model can be used to represent any order ARMA model. Therefore, one is free to use higher order AR or MA models to approximate any random process.

As defined in Appendix A, the z-transform of the second order statistics for the input and output of a linear system are related to the system transfer function by

$$P_{YY}(z) = H(z)\, H^*\left(\frac{1}{z^*}\right) P_{uu}(z) \ . \tag{9}$$

Now, if $H(z)$ is evaluated at $z = \exp(j2\pi f)$, then the associated PSD is

$$P_{ARMA}(f) = \sigma^2 \left|\frac{B(f)}{A(f)}\right|^2 \tag{10}$$

where σ^2 is the noise power associated with the input white noise process, and $A(f)$ and $B(f)$ are the transfer function polynomials evaluated at $z = \exp(j2\pi f)$. The PSD for a MA process can be expressed as

$$P_{MA}(f) = \sigma^2 |B(f)|^2 \tag{11}$$

and, the PSD for a AR process can be expressed as

$$P_{AR}(f) = \sigma^2 \left|\frac{1}{A(f)}\right|^2 \ . \tag{12}$$

An explicit relationship between the ARMA model parameters and the second-order statistics (the ACF) can be obtained. Letting

$$P_{YY}(z)\, A(z) = \frac{B^*\left(\frac{1}{z^*}\right)}{A^*\left(\frac{1}{z^*}\right)} B(z)\, \sigma^2 \tag{13}$$

then

$$P_{YY}(z)\, A(z) = H^*\left(\frac{1}{z^*}\right) B(z)\, \sigma^2 \tag{14}$$

and taking the inverse z-transform of the left side of Equation 14 yields

$$r_{YY}[k] \star a_k = \sum_{l=0}^{N} a_l r_{YY}[k-l] \ . \tag{15}$$

Now, taking the inverse z-transform of the right side yields

$$Z^{-1}\left[H^*\left(\frac{1}{z^*}\right)B(z)\,\sigma^2\right] = \sigma^2 \sum_{l=0}^{M} b_l h^*[l-k] \ . \tag{16}$$

Therefore, the ARMA(M,N) process can be expressed as

$$r_{YY}[k] = \begin{cases} -\sum_{l=1}^{N} a_l r_{YY}[k-l] + \sigma^2 \sum_{l=0}^{M-k} h^*[l]\,b[l+k] & \text{for } k = 0, 1, \ldots, M \\ -\sum_{l=1}^{N} a_l r_{YY}[k-l] & \text{for } k \geq M+1 \end{cases} \tag{17}$$

where $h(n)$ is a causal sequence. Equation 17 results in a set of nonlinear equations to be solved for the ARMA parameters (σ^2, a_k for $k = 1, \ldots, N$, and b_k for $k = 0, \ldots, M$). For the MA process, let $a_k = \delta(k)$ and $h(n) = b_n$ which yields

$$r_{YY}[k] = \begin{cases} \sigma^2 \sum_{l=0}^{N-k} b^*[l]\,b[l+k] & \text{for } k = 0, 1, \ldots, M \\ 0 & \text{for } k \geq M+1 \end{cases} \ . \tag{18}$$

Again, this is a set of nonlinear equations that must be solved for the MA parameters (σ^2, b_k for $k = 0, \ldots, M$). For the AR process, let $b_k = \delta(k)$ which yields

$$r_{YY}[k] = \begin{cases} -\sum_{l=1}^{N} a_l r_{YY}[k-l] + \sigma^2 & \text{for } k = 0 \\ -\sum_{l=1}^{N} a_l r_{YY}[k-l] & \text{for } k \geq 1 \end{cases} \ . \tag{19}$$

Equation 19 is termed the Yule-Walker equations. Note that this is a set of linear equations which can be written in matrix form as

$$\begin{bmatrix} r_{YY}[0] & r_{YY}[-1] & \cdots & r_{YY}[-(N-1)] \\ r_{YY}[1] & r_{YY}[0] & \cdots & r_{YY}[-(N-2)] \\ \vdots & \vdots & \ddots & \vdots \\ r_{YY}[N-1] & r_{YY}[N-2] & \cdots & r_{YY}[0] \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_N \end{bmatrix} = - \begin{bmatrix} r_{YY}[1] \\ r_{YY}[2] \\ \vdots \\ r_{YY}[N] \end{bmatrix} \ . \tag{20}$$

Also, note that for a complex WSS random process the correlation sequence is conjugate symmetric, $r(-k) = r^*(k)$. Therefore, the correlation matrix in Equation 20 is Hermitian symmetric and Toeplitz since all the diagonal elements are equal. This set of equations can be extended to include the white noise variance where

$$\begin{bmatrix} r_{YY}[0] & r_{YY}[-1] & \cdots & r_{YY}[-N] \\ r_{YY}[1] & r_{YY}[0] & \cdots & r_{YY}[-(N-1)] \\ \vdots & \vdots & \ddots & \vdots \\ r_{YY}[N] & r_{YY}[N-1] & \cdots & r_{YY}[0] \end{bmatrix} \begin{bmatrix} 1 \\ a_1 \\ \vdots \\ a_N \end{bmatrix} = \begin{bmatrix} \sigma^2 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \ . \tag{21}$$

Of the three models, the AR, with its set of linear equations relating the ACF to the models parameters, is the simplest to solve. The AR model is used extensively in the literature and is used in lieu of an ARMA or MA model due to the difficulty in estimating the zero locations. In addition, for most physical systems, the poles contain most of the relevant information (harmonic location, spectrum width, etc.). Also, as stated previously, the fact that the AR(∞) is capable of representing any ARMA or MA model supports the use an AR model of higher order over the MA and ARMA given the difficulties in estimating the zero locations.

The Yule-Walker equations resulting from the AR model can be solved using common matrix analysis techniques such as Gaussian elimination or by exploiting the properties of the Hermitian Toeplitz matrix. Gaussian elimination techniques require on the order of $O(N^3)$ operations. Levison and Durbin [10] have developed an algorithm based on the Hermitian and Toeplitz nature of the Yule-Walker equations which allows for the solution of the Yule-Walker equations in $O(N^2)$ operations. Given the Yule-Walker equations in Equation 21, they can be modified to form

$$\begin{bmatrix} r_{YY}[0] & r^*_{YY}[1] & \cdots & r^*_{YY}[N] \\ r_{YY}[1] & r_{YY}[0] & \cdots & r^*_{YY}[(N-1)] \\ \vdots & \vdots & \ddots & \vdots \\ r_{YY}[N] & r_{YY}[N-1] & \cdots & r_{YY}[0] \\ r_{YY}[N+1] & r_{YY}[N] & \cdots & r_{YY}[0] \end{bmatrix} \begin{bmatrix} 1 \\ a_1^N \\ \vdots \\ a_N^N \\ 0 \end{bmatrix} = \begin{bmatrix} \sigma_N^2 \\ 0 \\ \vdots \\ 0 \\ \xi_N \end{bmatrix} . \tag{22}$$

Given the Hermitian Toeplitz nature of the Yule-Walker equations, Equation 22 can be written in an equivalent form as

$$\begin{bmatrix} r_{YY}[0] & r_{YY}[1] & \cdots & r_{YY}[N+1] \\ r^*_{YY}[1] & r_{YY}[0] & \cdots & r_{YY}[(N)] \\ \vdots & \vdots & \ddots & \vdots \\ r^*_{YY}[N] & r^*_{YY}[N-1] & \cdots & r_{YY}[1] \\ r^*_{YY}[N+1] & r^*_{YY}[N] & \cdots & r_{YY}[0] \end{bmatrix} \begin{bmatrix} 0 \\ a_N^N \\ \vdots \\ a_{N-1}^N \\ 1 \end{bmatrix} = \begin{bmatrix} \xi_N \\ 0 \\ \vdots \\ 0 \\ \sigma_N^2 \end{bmatrix} . \tag{23}$$

Taking the complex conjugate of Equation 23 and scaling it by a complex number κ yields

$$\kappa_{N+1}\begin{bmatrix} r_{YY}[0] & r^*_{YY}[1] & \cdots & r^*_{YY}[N+1] \\ r_{YY}[1] & r_{YY}[0] & \cdots & r^*_{YY}[(N)] \\ \vdots & \vdots & \ddots & \vdots \\ r_{YY}[N] & r_{YY}[N-1] & \cdots & r^*_{YY}[1] \\ r_{YY}[N+1] & r_{YY}[N] & \cdots & r_{YY}[0] \end{bmatrix}\begin{bmatrix} 0 \\ a_N^N \\ \vdots \\ a_{N-1}^N \\ 1 \end{bmatrix}^* = \kappa_{N+1}\begin{bmatrix} \xi_N \\ 0 \\ \vdots \\ 0 \\ \sigma_N^2 \end{bmatrix}^* \tag{24}$$

The complex number κ is often termed the reflection coefficient. Adding Equations 21 and 24 yields

$$R_{N+1}\left\{\begin{bmatrix} 1 \\ a_1^N \\ \vdots \\ a_N^N \\ 0 \end{bmatrix} + \kappa_{N+1}\begin{bmatrix} 0 \\ a_N^N \\ \vdots \\ a_1^N \\ 1 \end{bmatrix}^*\right\} = \begin{bmatrix} \sigma_N^2 \\ 0 \\ \vdots \\ 0 \\ \xi_N \end{bmatrix} + \kappa_{N+1}\begin{bmatrix} \xi_N \\ 0 \\ \vdots \\ 0 \\ \sigma_N^2 \end{bmatrix}^* \tag{25}$$

where

$$R_{N+1} = \begin{bmatrix} r_{YY}[0] & r^*_{YY}[1] & \cdots & r^*_{YY}[N+1] \\ r_{YY}[1] & r_{YY}[0] & \cdots & r^*_{YY}[(N)] \\ \vdots & \vdots & \ddots & \vdots \\ r_{YY}[N] & r_{YY}[N-1] & \cdots & r^*_{YY}[1] \\ r_{YY}[N+1] & r_{YY}[N] & \cdots & r_{YY}[0] \end{bmatrix}. \tag{26}$$

Now, letting

$$\kappa_{N+1} = -\frac{\xi_N}{\sigma_N^2}. \tag{27}$$

Equation 25 can be written as

$$R_{N+1}\,\underline{a}_{N+1} = \sigma^2_{N+1}[1\ 0\ \ldots\ 0]^T \tag{28}$$

where

$$\underline{a}_{N+1} = \begin{bmatrix} 1 \\ a_1^{N+1} \\ \vdots \\ a_{N-1}^{N+1} \\ a_N^{N+1} \end{bmatrix} = \begin{bmatrix} 1 \\ a_1^N \\ \vdots \\ a_N^N \\ 0 \end{bmatrix} + \kappa_{N+1}\begin{bmatrix} 0 \\ a_N^N \\ \vdots \\ a_1^N \\ 1 \end{bmatrix}^* \tag{29}$$

and

$$\sigma^2_{N+1} = \sigma_N^2\left[1 - |\kappa_{N+1}|^2\right]. \tag{30}$$

Table I. Levison Durbin Recursion

1. Initialize the routine.
   (a) $a_0^0 = 1$
   (b) $\sigma_0^2 = r_{XX}(0)$
2. 2. for i $= 0, 1, \ldots, N-1$
   (a) $\xi_i = r_{XX}(i+1) + \sum_{k=1}^{i} a_k^i r_{XX}(i-k+1)$
   (b) $\kappa_{i+1} = -\frac{\xi_i}{\sigma_i^2}$
   (c) For $k = 1, 2, \ldots, i, \quad a_k^{i+1} = a_k^i + \kappa_{i+1} a_{i-k+1}^{i*}$
   (d) $a_{i+1}^{i+1} = \kappa_{i+1}$
   (e) $\sigma_{i+1}^2 = \sigma_{N+1}^2 = \sigma_N^2 [1 - |\kappa_{N+1}|^2]$

The result of this exercise is a recursive algorithm for obtaining the AR filter coefficients of order N+1 using the coefficients at order N. The algorithm is given in Table I.

There are many sub-optimal approaches in the literature for estimating the AR parameters. This list includes the Yule-Walker equations, the Burg algorithm, and the least squares approach just to name a few. There are also sequential methods based on the time series data which include the gradient adaptive least mean squares algorithm and the recursive least squares. With all the attention given to the area of spectral estimation over the past 10-20 years, there are several excellent text describing these different techniques [14, 29]. This dissertation will not explore the different techniques but will concentrate on the Yule-Walker equations which directly employ an estimate of the autocorrelation matrix.

Model Order Determination

Model order determination is a critical factor in representing a random process. Too low an order yields a smoothed spectrum, and too high an order may yield spurious peaks. Most algorithms, which attempt to estimate the model order from an observed data set, focus on observing a minimum in the estimated white noise variance, $\hat{\sigma}^2$. Akaike [29] has proposed two methods for estimating the order of a process. The first method, termed the final prediction error method (FPE), estimates the order based on the minimization of the following equation

$$FPE(k) = \frac{N+k}{N-k}\hat{\sigma}^2 \tag{31}$$

where N is the number of samples collected and k is the current order of the model used to represent the process. A second method proposed by Akaike for estimating the model order is termed the Akaike information criterion (AIC) and is based on minimizing the following equation

$$AIC(k) = N \ln\left(\hat{\sigma}^2\right) + 2k\,. \tag{32}$$

High Resolution

ARMA modeling offers an enhanced capability over Fourier based methods in resolving closely spaced signals. These techniques are often termed "high resolution" spectral estimators. ARMA modeling is able to enhance its resolution capability through the fact that the ARMA model does not window the autocorrelation function. Equation 17 does not assume the autocorrelation function goes to zero after a specified number of lags as assumed in the Fourier based methods. In addition, the lack of any implied window function prevents the normal sidelobes introduced by a window function in the spectrum.

Effect of White Noise

In most physical systems, the random process under observation contains signal plus additive observation noise. The observation noise should not be confused with

the white noise source used to drive the ARMA model. In modeling a random process, an appropriate model must be chosen which can characterize the total process including the observation noise. Exclusion of the observation noise from the model will result in a biased estimate. Since AR models are often used to characterize a random process, it is of benefit to observe what happens to an AR process when white observation noise is added. Given an AR process with $P_{AR}(z)$ and additive independent white noise with its corresponding $P_w(z) = \sigma_w^2$, the resulting z-transform description can be written as

$$P_{YY}(z) = P_{AR}(z) + \sigma_w^2 \tag{33}$$

or

$$P_{YY}(z) = \frac{\sigma^2}{|A(z)|^2} + \sigma_w^2 \tag{34}$$

and adding the two terms yields

$$P_{YY}(z) = \frac{\sigma^2 + \sigma_w^2 |A(z)|^2}{|A(z)|^2} = P_{ARMA}(z) \ . \tag{35}$$

The addition of white observation noise has changed the AR process to an ARMA process[12]. From Equation 35, as the signal-to-noise ratio (SNR) increases, the zeros of the system move to the poles of the system resulting in a flat or smoothed spectrum. Kay [14] has given four methods of removing the effects of the observation noise in AR modeling. The four methods are:

1. use an ARMA model in lieu of an AR model to better represent the process;
2. use a filter to reduce the SNR;
3. adjust the model parameters to compensate for the noise [13]; and
4. use a higher order AR model to better approximate the ARMA process.

The Autocorrelation Function

Introduction

The AR model serves as a compact tool for relating properties of an autocorrelation function to the power spectrum. This section will define the Fourier uncertainty principle and use a first order AR model to illustrate the property. This section will also discuss white and colored noise properties and methods for removing the noise from the ACF before estimating the PSD.

The Uncertainty Principle

The uncertainty principle [2, 23] of the Fourier transform states that the duration-bandwidth product $\Delta\tau\,\Delta\omega$ must meet the following constraint

$$\Delta\tau\,\Delta\omega \geq \frac{1}{2} \tag{36}$$

with equality for the Gaussian distribution function. The uncertainty principle will be used in Chapters III and IV to explain performance issues related to the autocorrelation based moment estimators.

First Order AR Model Example

An ARMA model consists a set of difference equations which can be solved to yield the general form of the autocorrelation function for that process. For a real first order AR(1) process, the Yule-Walker equations can be written as

$$r_{XX}(k) = -a_1 r_{XX}(k-1) + \sigma^2 \quad \text{for } k = 0 \tag{37}$$

$$r_{XX}(k) = -a_1 r_{XX}(k-1) \quad \text{for } k \neq 0\,. \tag{38}$$

Solving the difference equations for $r_{XX}(k)$ yields

$$r_{XX}(k) = \frac{\sigma^2}{1-a_1^2}(-a_1)^{|k|}\,. \tag{39}$$

Since a_1 is required to be inside the unit circle for a stable system, the resulting autocorrelation function is a decaying function ($|a_1| \leq 1$). Also, if a_1 is positive,

the function oscillates between negative and positive values, and if a_1 is negative, the function is positive for all k. Examples of both cases are given in Figures 2 and 3. The corresponding PSD can written as

$$P_{XX}(\omega) = \frac{\sigma^2}{|1 + a_1 \exp(-j\omega)|^2} \tag{40}$$

since the AR process is the output of linear system driven by white noise

$$P_{XX}(z) = \sigma^2 H(z) H^*(\frac{1}{z^*}) \tag{41}$$

and the system transfer function is defined as

$$H(z) = \frac{z}{z + a_1} \, . \tag{42}$$

Given the AR(1) PSD in Equation 40, if the coefficient is negative, then the PSD represents a lowpass process. The bandwidth of the lowpass process is determined by the distance of the pole in Equation 42 from the unit circle. The bandwidth dependence is illustrated in Figures 4 and 5 for the cases where $a_1 = 0.9$ and $a_1 = 0.5$, respectively. The closer the pole is to the unit circle the narrower the bandwidth. The corresponding autocorrelation functions are given in Figures 6 and 7. In comparing the autocorrelation functions to their corresponding PSD's, an inverse relationship is seen between the duration in the correlation domain and the bandwidth in the frequency domain. These examples imply that narrowband processes have autocorrelation functions that decay slowly compared to wideband processes. This can be explained through the Fourier uncertainty principle.

The Noise Autocorrelation Function

Any real measurement system has an associated noise source or sources due to the random motion of electrons within the components forming the system. In addition to noise sources within the measurement system, there are other sources of contamination which are usually labeled as clutter and are a part of the measured signal. The clutter is usually an undesired portion of the measured signal

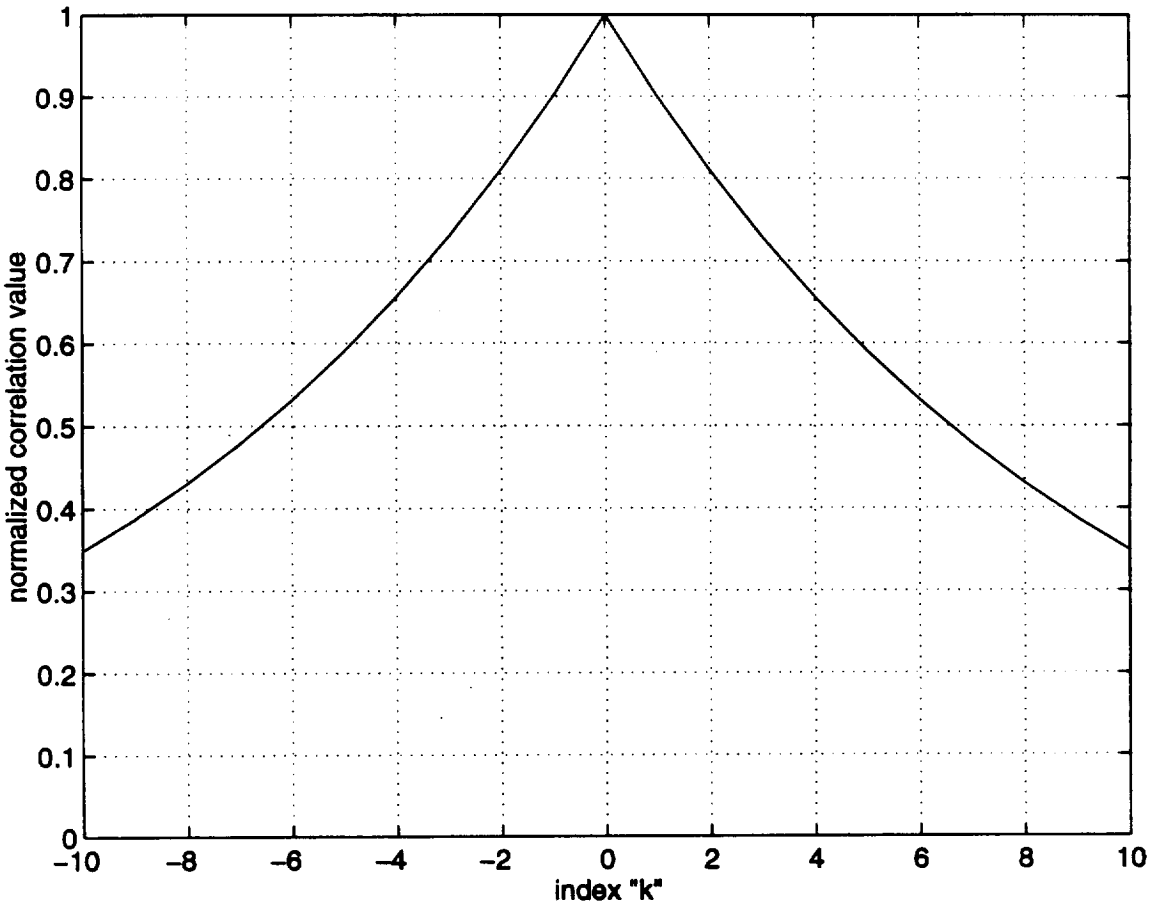


Figure 2. The autocorrelation function associated with an AR(1) process with $a_1 = -0.9$.

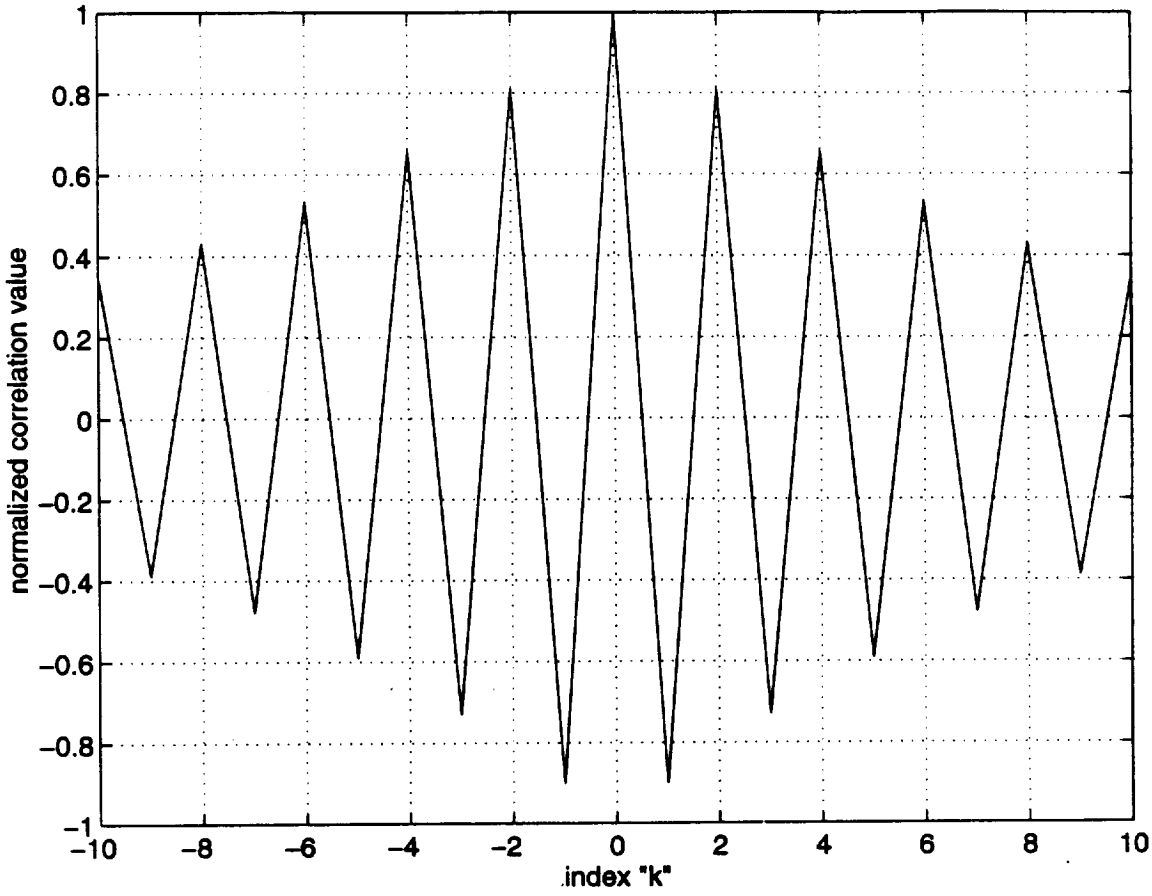


Figure 3. The autocorrelation function associated with an AR(1) process with $a_1 = 0.9$.

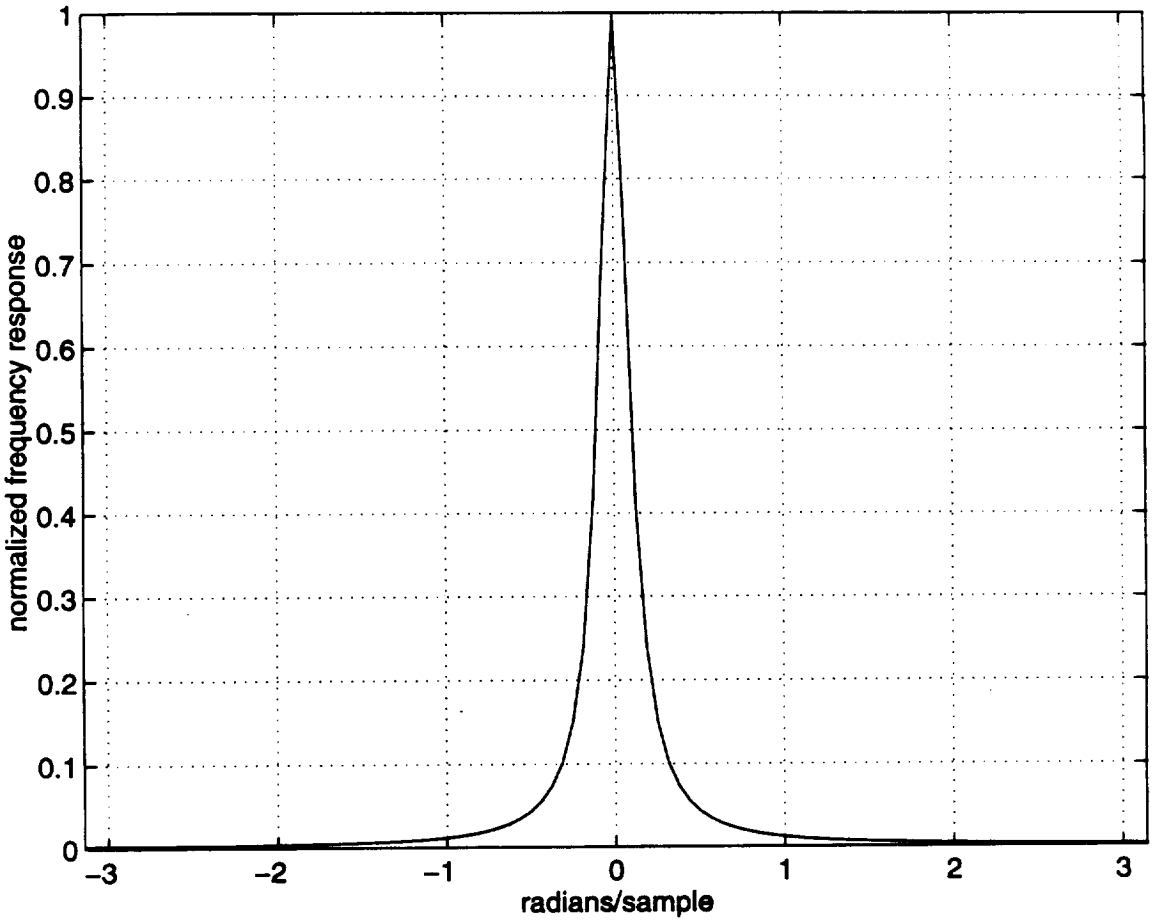


Figure 4. The power spectral density associated with an AR(1) process with $a_1 = -0.9$.

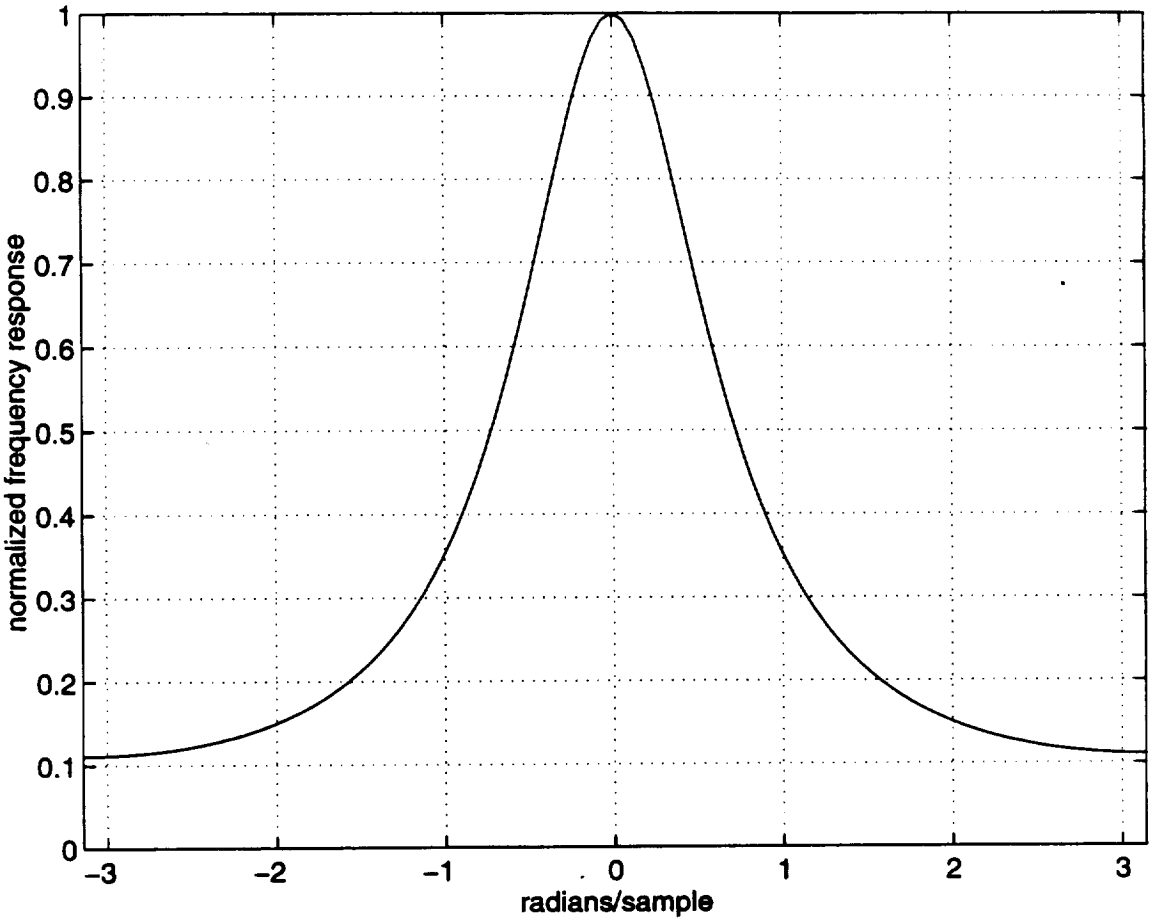


Figure 5. The power spectral density associated with an AR(1) process with $a_1 = -0.5$.

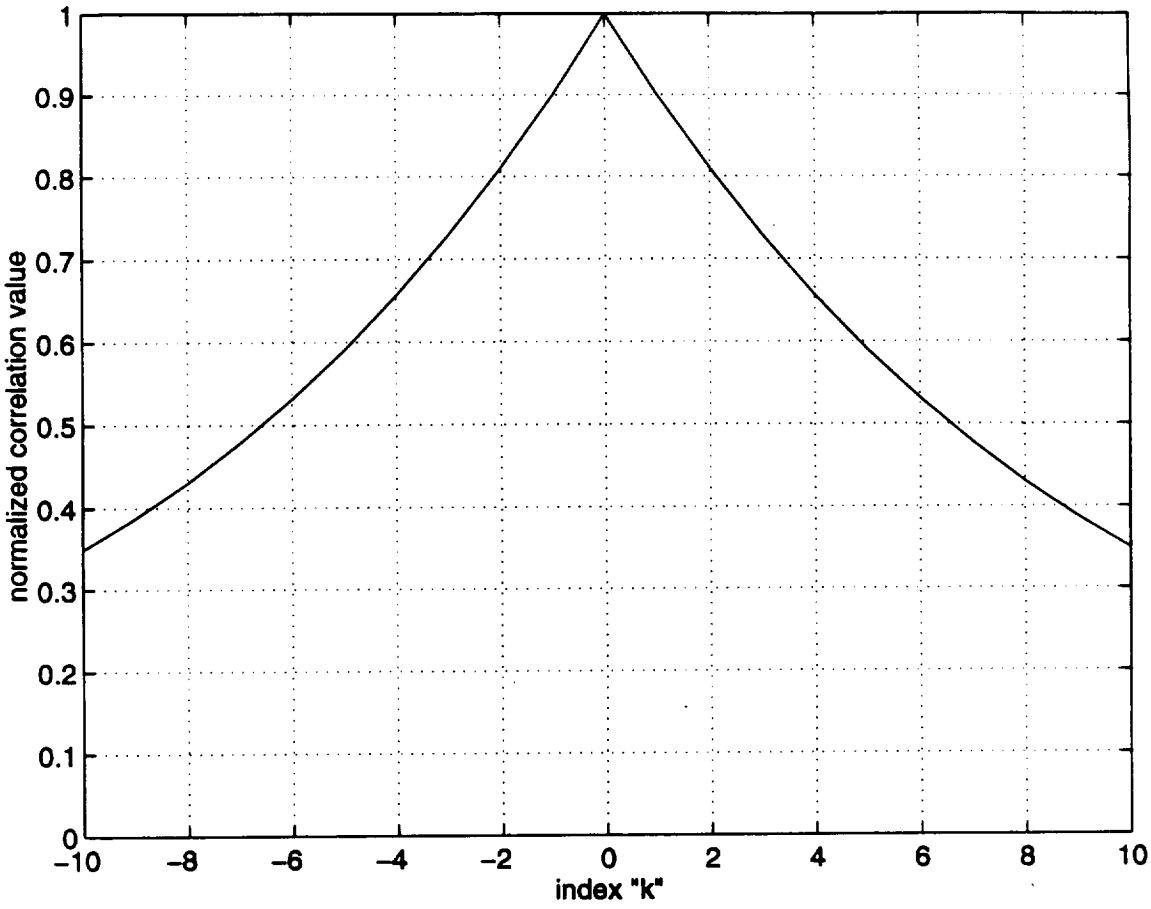


Figure 6. The autocorrelation function associated with an AR(1) process with $a_1 = -0.9$.

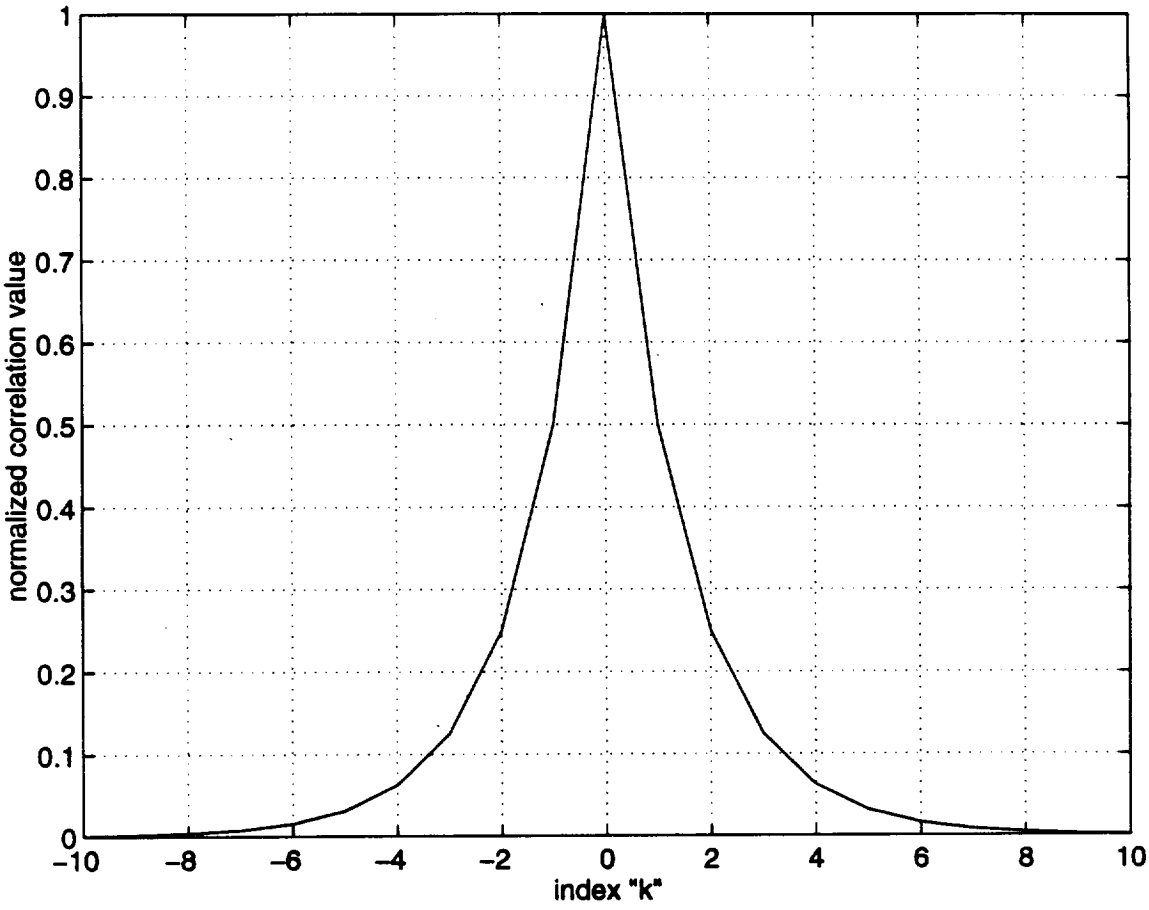


Figure 7. The autocorrelation function associated with an AR(1) process with $a_1 = -0.5$.

and may cause a bias in parameters extracted from the measurement. A priori knowledge of the noise or clutter signal aids in designing methods for noise or clutter rejection/suppression. In the case where little knowledge is known about the noise source, matched filters based on the transmit signal are used to enhance signal detection.

In physical systems, the noise is often additive and statistically independent of the "signal". A random process, Y, containing signal, X, plus additive independent noise, N, will yield an autocorrelation function which is the sum of the individual autocorrelation functions

$$r_{YY}(k) = r_{XX}(k) + r_{NN}(k) \; . \tag{43}$$

Since the Fourier transform is a linear operator, the PSD for the total process is

$$P_{YY}(\omega) = P_{XX}(\omega) + P_{NN}(\omega) \; . \tag{44}$$

A special case exist when the PSD of the noise is constant for all ω

$$P_{NN}(\omega) = \sigma^2 \;\; \text{for all } \omega \; . \tag{45}$$

This is termed a white noise process. The resulting autocorrelation function is a delta function at lag $k = 0$

$$r_{NN}(k) = \sigma^2 \, \delta(k) \; . \tag{46}$$

In this case, all the information about the noise is contained in the zeroth autocorrelation lag. In traditional and modern spectral estimation techniques where estimates of the autocorrelation function are used to estimate the power spectral density, the noise contained in the zeroth autocorrelation lag estimate will bias the power spectrum estimate of the signal.

Compensation for or removal of the noise bias is possible using correlation subtraction techniques. In the case of white noise, an estimate of the noise power, $\hat{r}_{NN}(0)$, is subtracted from the total power to obtain an estimate of the signal power

$$\hat{r}_{XX}(0) = \hat{r}_{YY}(0) - \hat{r}_{NN}(0) \; . \tag{47}$$

Estimates of the noise power are obtained during periods when the signal of interest is known to be absent. The white noise case is an optimum noise suppression scenario in that all the noise power is contained in the zeroth autocorrelation lag. However, there are other cases where the noise process is non-white (or colored).

Colored noise contains correlated noise samples which result in non-zero values in the ACF away from lag zero. Correlation subtraction routines can also be used in this case when a priori knowledge of the noise source is available. An example is the case where a white noise source has been passed through a known noise suppression filter to improve the SNR.

Information Relative to Lag Index

The Yule-Walker equations, as defined in Equation 19, represent a system of equations with N+1 unknowns and potentially an infinite number of equations. In practice, however, the first N+1 equations starting with $k = 0$ are assumed when referring to the Yule-Walker equations. Considering the general case, the Yule-Walker equations represent a set of N+1 unknowns and M equations ($M \geq (N+1)$) which may be used to form an overdetermined set of equations ($M > (N+1)$).

Bruzzone and Kaveh [3] have shown that the number of autocorrelation lags, M, required in estimating the parameters of an AR model of fix order N, is a function of the information contained in each autocorrelation lag. As a measure of the information contained in a set of autocorrelation lags, Bruzzone and Kaveh have defined a relative information index (RII), R,

$$R = \frac{|I_r(\theta)|}{|I_y(\theta)|} \tag{48}$$

where $I_r(\theta)$ is Fisher's information matrix for the pole positions based on the autocorrelation estimates, and $I_y(\theta)$ is Fisher's information matrix for the pole positions based on the observed data set. The range of R is $0 \leq R \leq 1$ with equality of one indicating that the autocorrelation estimates are sufficient in estimating the pole

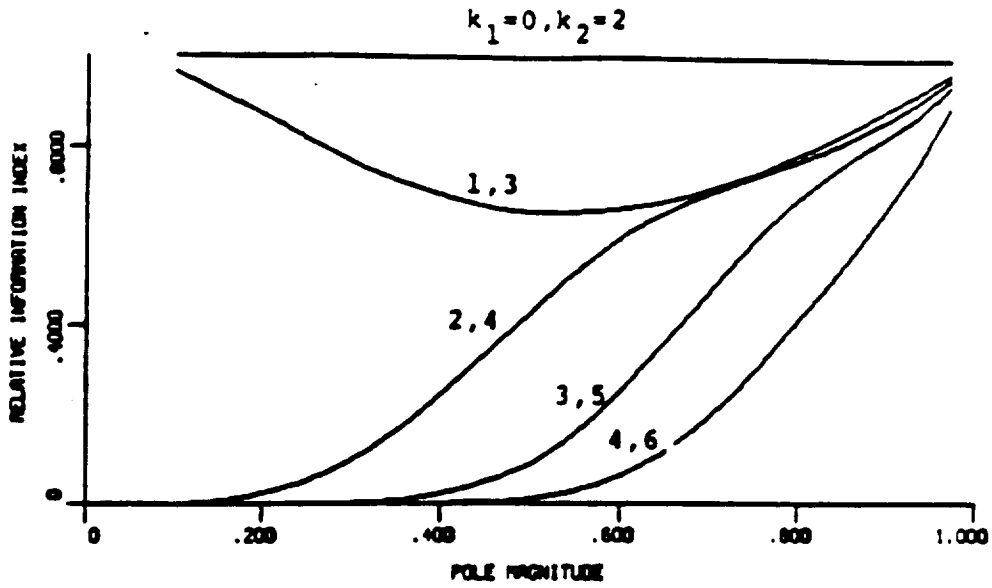


Figure 8. A plot of the RII for a AR(2) process with poles located at $\theta = \pm 0.25\pi$ as a function of the pole magnitudes and the autocorrelation lags $\{k_1, k_2\}$ used in the estimate. (Source: ©1984 IEEE, Bruzzone and Kaveh, "Informaiton Tradeoffs in Using the Sample Autocorrelation Function in ARMA Parameter Estimation", IEEE Trans. on ASSP, August 1984.)

positions. Bruzzone and Kaveh illustrate several examples for an AR process. A discussion of some of these examples is presented here in order to illustrate Bruzzone and Kaveh's conclusions.

The first example is an AR(2) process with the poles located at angles of $\theta = \pm 0.25\pi$. Figure 8 shows the relative information index as a function of pole magnitude and as a function of the autocorrelation lags $\{r_{YY}(k),\ k_1 \leq k_2\}$ used to compute the pole locations. As seen in Figure 8, wideband processes, ($|P| \leq 0.8$), require the inclusion of the first few autocorrelation lags in order to accurately estimate the pole locations.

The second example is an AR(2) process with additive independent white noise. Figure 9 is a plot of the relative information index as a function of SNR and pole radius, $|P|$, when lags zero through four are used. As seen in Figure 9, the relative

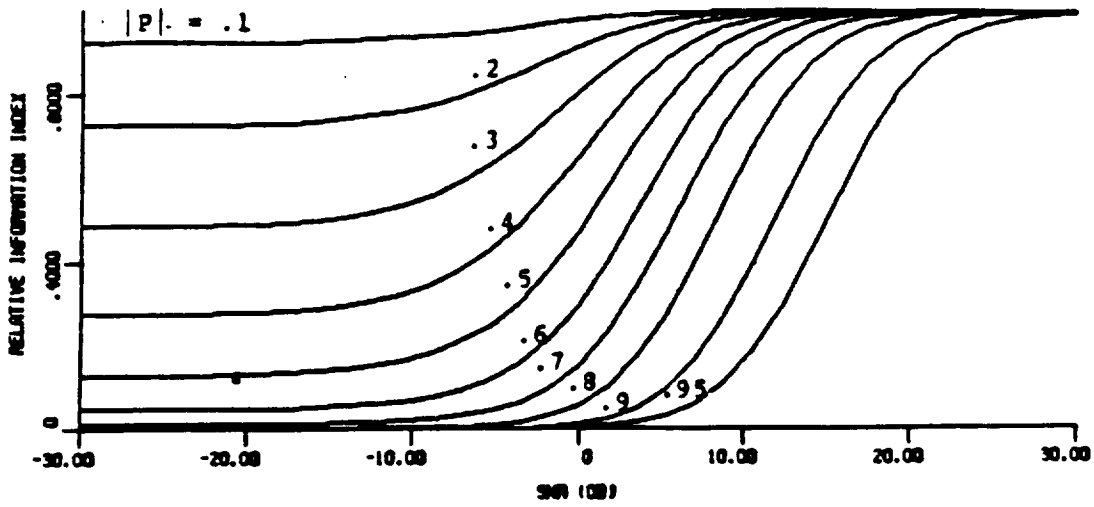


Figure 9. A plot of the RII for the case of a AR(2) model with poles located at $\theta = \pm 0.25\pi$ and additive white noise as a function of SNR and pole magnitude $| P |$. The AR(2) model estimate is made using lags zero through four. (Source: ©1984 IEEE, Bruzzone and Kaveh, "Informaiton Tradeoffs in Using the Sample Autocorrelation Function in ARMA Parameter Estimation", IEEE Trans. on ASSP, August 1984.)

information index falls off quickly as a function of SNR in the case of a narrowband processes ($|P| \geq 0.8$). In a similar example containing additive white noise and using an AR(2) process, the relative information index in Figure 10 is plotted as a function of the number of lags, ($k_0 = 0$ *to* k_2), used in the estimate and the pole locations in π-radians for a fixed radius of 0.95. This example illustrates that the higher lags of a narrowband process contain additional information which can be used in the spectral estimate.

In these examples, the white noise is concentrated in the zeroth lag. In selecting the autocorrelation lags to use in the AR model estimate, one might avoid using the zeroth lag entirely to reduce the noise bias. Figure 11 contains plots of the relative information index computed for a AR(2), with poles located at a radius of 0.95 and an angle of 0.25π radians, as a function of five autocorrelation lags with different initial indices. For a SNR above approximately 18 dB, the inclusion of

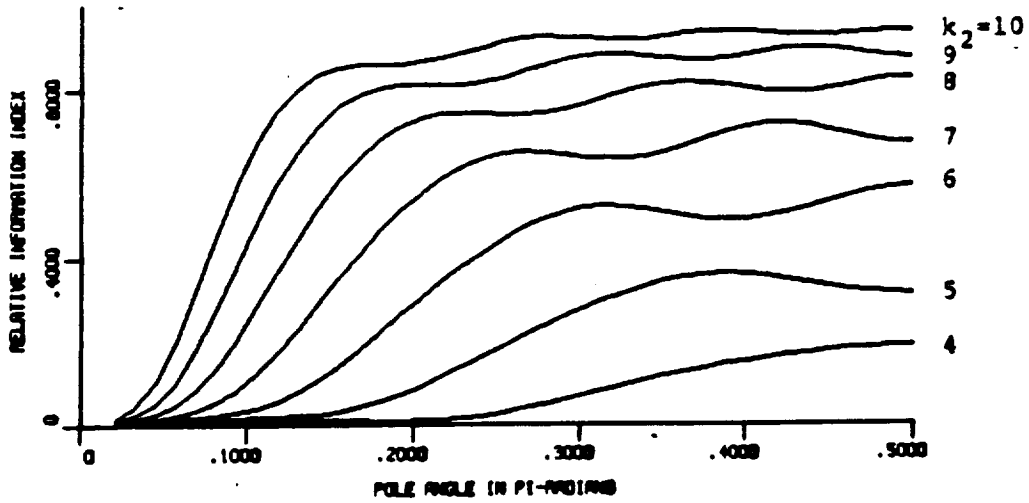


Figure 10. A plot of the RII for the AR(2) model with a pole magnitude of 0.95 (narrowband) as a function of the number of lags used in the estimate and the pole angle in radians. (Source: ©1984 IEEE, Bruzzone and Kaveh, "Informaiton Tradeoffs in Using the Sample Autocorrelation Function in ARMA Parameter Estimation", IEEE Trans. on ASSP, August 1984.)

the zeroth lag, even though it contains the noise power, is optimum for obtaining a good spectral estimate. However, when the SNR drops below 18 dB, the estimator is optimum when using the lags starting at index $k = 1$. Also, note that the latter lags starting as late as index four still contain some information about the process in the narrowband example.

For the wide-band case, it can be shown that most of the information is contained in the initial lags. Figure 12 is a plot of the relative information index for a AR(2) with poles located at a radius of 0.4 and an angle of 0.25π radians. Again, the set of autocorrelation lags used in the estimate consists of five lags with different initial indices. As seen in Figure 12, the latter lags contain little information about the wide-band process. The zeroth lag provides a great deal of information about the wideband process provided the SNR is greater than 5 dB. At lower SNR, the noise power dominates the zeroth lag causing a bias in information. However, only

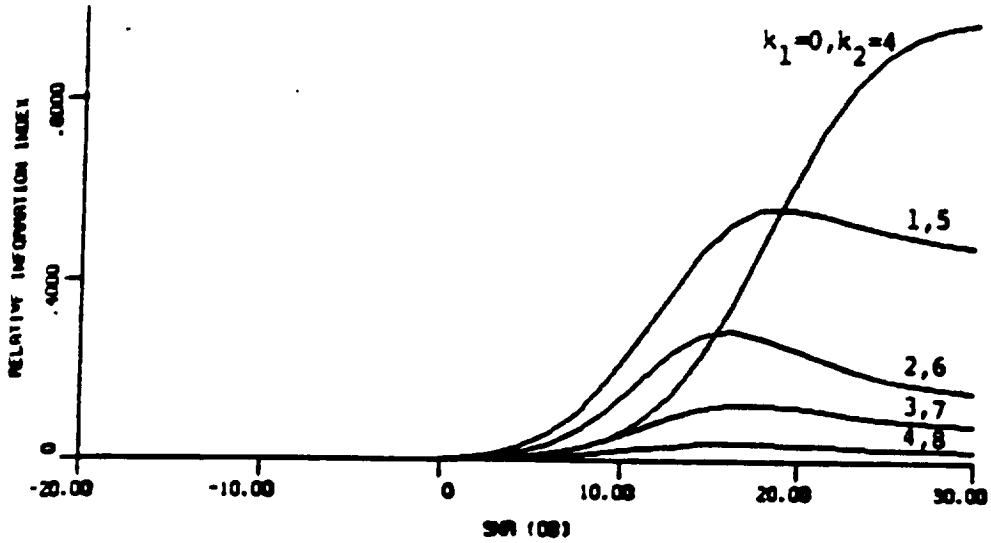


Figure 11. A plot of the RII for the case of a AR(2) process with poles located at $\theta = \pm 0.25\pi$ and a magnitude of 0.95 plotted as a function of the SNR and lags $\{k_1, k_2\}$ used in the estimate. (Source: ©1984 IEEE, Bruzzone and Kaveh, "Informaiton Tradeoffs in Using the Sample Autocorrelation Function in ARMA Parameter Estimation", IEEE Trans. on ASSP, August 1984.)

the zeroth lag contains the noise power, and therefore, lags one through five can be used at the lower SNR.

Based on their analysis, Bruzzone and Kaveh have developed a criterion for selecting autocorrelation lags for use in AR model estimation. The following is a summary of some of the critical issues defined by Bruzzone and Kaveh.

1. The zeroth and first autocorrelation lags are important in both the wide and narrow band cases and provide a substantial portion of the available information.

2. Fewer lags are needed for wideband processes.

3. In the case of narrowband processes, estimators using a large number of autocorrelation lags is suggested.

Cadzow [4] has also investigated using additional lags of the autocorrelation function in estimating ARMA model parameters. Cadzow's paper treats the overdetermined case of the Yule-Walker equations and shows that additional lags can be

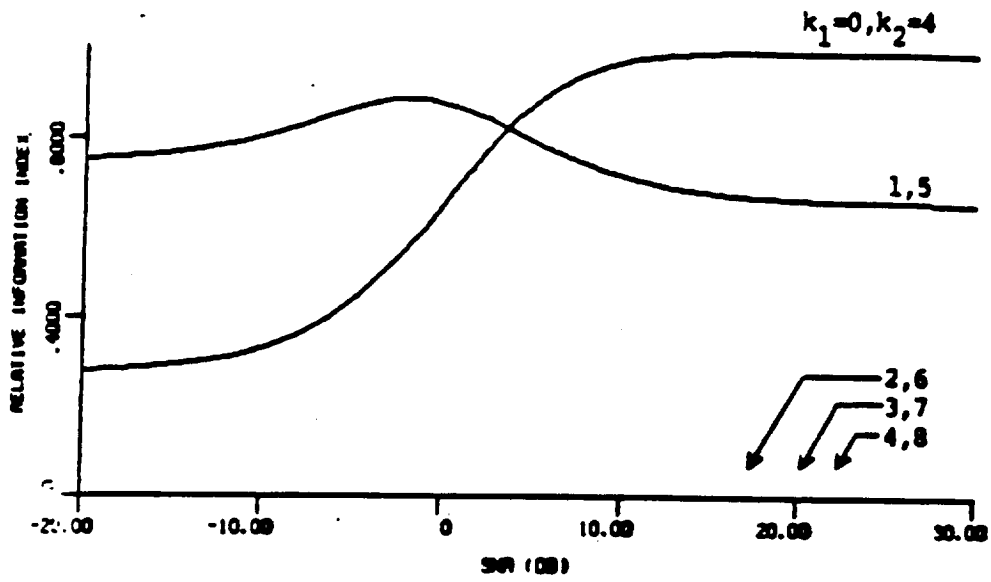


Figure 12. A plot of the RII index for the case of an AR(2) process with a magnitude of 0.4 (wideband) and an angle of $\theta = \pm 0.25\pi$, as a function of SNR and lags $\{k_1, k_2\}$ used in the estimate. (Source: ©1984 IEEE, Bruzzone and Kaveh, "Informaiton Tradeoffs in Using the Sample Autocorrelation Function in ARMA Parameter Estimation", IEEE Trans. on ASSP, August 1984.)

used to improve estimator performance. In order to show how the additional lags could be used to improve the spectral estimate, Cadzow simulated ten realizations of a signal consisting of two sinusoids embedded in white noise with normalized frequencies of 0.2 and 0.215. Figure 13 is a plot of the spectral estimate using a 20th order AR model. The large model order is used since the additive white noise mandates an ARMA model. Note that the two sinusoids are not distinguishable. In Figure 14, the model order has increased to twenty-four, and the two sinusoids are starting to separate. Cadzow demonstrates in Figure 15 that an AR model of size 20 using 50 equations in the overdetermined system yields a better estimate of the two sinusoid locations than the higher order AR model using the basic Yule-Walker equations.

In this section, a case was made for the use of additional autocorrelation lags in AR modeling. In Chapter IV, this overdetermined approach will be extended to autocorrelation based moment estimators in order to improve estimator performance.

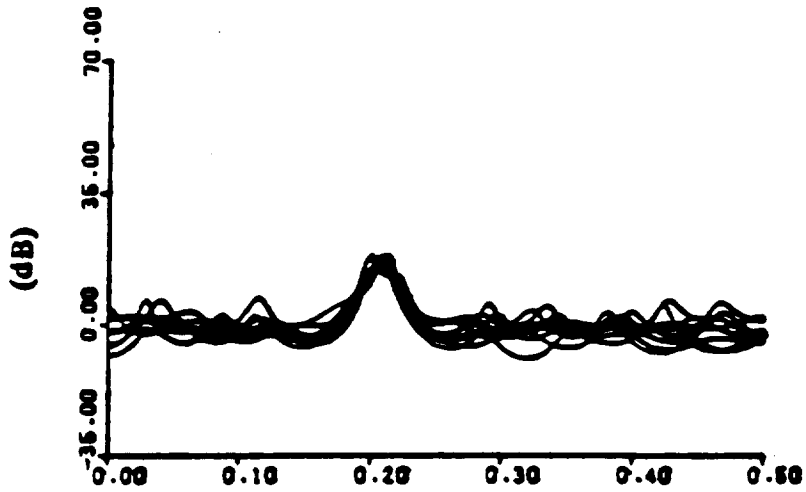


Figure 13. Ten realizations of a 20th order AR power spectrum estimate for two sinusoids embedded in white noise. (Source: ©1982 IEEE, Cadzow, "Spectral Estimation: An Overdetermined Rational Model Equation Approach", Proceedings of the IEEE, September 1982.)

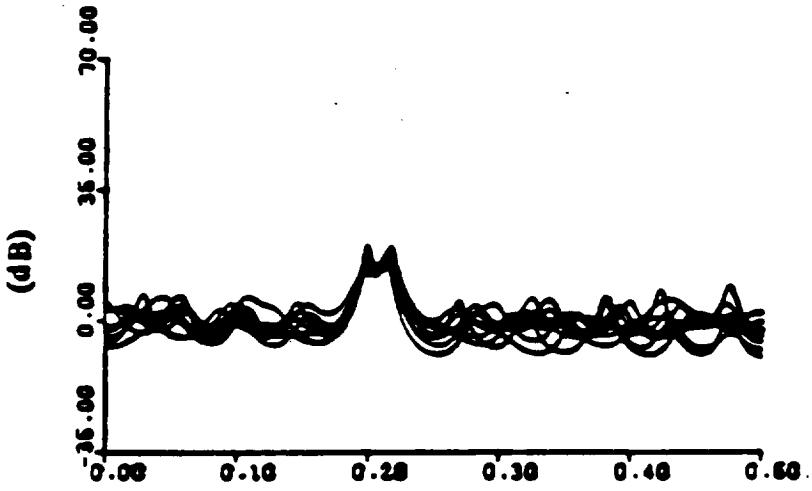


Figure 14. Ten realizations of a 24th order AR power spectrum estimate for two sinusoids embedded in white noise. (Source: ©1982 IEEE, Cadzow, "Spectral Estimation: An Overdetermined Rational Model Equation Approach", Proceedings of the IEEE, September 1982.)

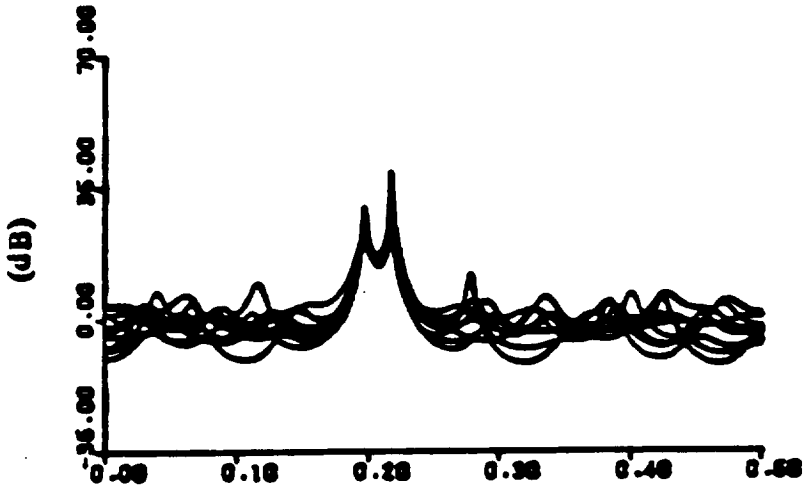


Figure 15. Ten realizations of a 20th order overdetermined AR power spectrum estimate based on 50 equations for two sinusoids embedded in white noise. (Source: ©1982 IEEE, Cadzow, "Spectral Estimation: An Overdetermined Rational Model Equation Approach", Proceedings of the IEEE, September 1982.)

CHAPTER III

SPECTRAL MOMENT ESTIMATORS

Introduction

This chapter starts by relating moments of the power spectral density to moments defined for a probability density function. This chapter then reviews Fourier and autocorrelation based techniques for estimating moments of the PSD. The Fourier based techniques attempt to approximate the moment integral using discrete centroiding techniques given a traditional or modern power spectrum estimate. The autocorrelation based techniques include the commonly applied pulse-pair mean and variance estimators and Passarelli's autocorrelation function series expansion which allows one to define an infinite number of estimators including the pulse-pair. Estimator performance issues for both the Fourier and autocorrelation based estimators are also addressed.

Spectral Moments

The power spectral density can be interpreted as a function describing the distribution of power over a range of frequencies. This distribution is assured to take on values greater than or equal to zero over all frequencies due to the positivity constraint of the PSD. The estimated PSD can also be normalized to contain unit area. The normalized PSD for a discrete time random process is

$$\hat{P}_{xx}^{Norm}(\omega) = \frac{\hat{P}_{xx}(\omega)}{\int_{-\pi}^{\pi} \hat{P}_{xx}(\omega)\ d\omega} \tag{49}$$

which has the properties of a probability density function where

$$\hat{P}_{xx}^{Norm}(\omega) \geq 0 \quad for\ all\ \omega \tag{50}$$

and

$$\int_{-\pi}^{\pi} \hat{P}_{xx}^{Norm}(\omega) = 1\ . \tag{51}$$

One can now define moments of the power spectral density by

$$M_n = E\left[\omega^n\right] = \int_{-\pi}^{\pi} \omega^n \, \hat{P}_{xx}^{Norm}(\omega) \, d\omega \; . \tag{52}$$

The mean frequency is defined as

$$\mu = \int_{-\pi}^{\pi} \omega \, \hat{P}_{xx}^{Norm}(\omega) \, d\omega \; , \tag{53}$$

and the variance of the spectral random variable can also be computed using the relationship

$$\sigma^2 = E\left[\omega^2\right] - E^2\left[\omega\right] \; . \tag{54}$$

There are basically two methods for estimating the moments associated with a PSD. Fourier based techniques which require an estimate of the PSD and autocorrelation based techniques which take advantage of the Fourier transform pair relationship between the ACF and the PSD and the characteristic function defined in probability theory.

Fourier Based Spectral Moment Estimation

Given an estimate of the PSD obtained from traditional or modern spectral estimation techniques, the desired spectral moments can be estimated in the frequency domain using

$$\hat{M}_n = \sum_i \omega_i^n \, \hat{P}_{xx}^{Norm}\left(\omega_i\right) \; . \tag{55}$$

The spectral moment estimates reflect a composite estimate of the signal plus noise contained in the sampled data set. Except for the large signal-to-noise ratio case, noise suppression techniques are required to reduce the bias [32]. Noise suppression techniques include spectral filtering based on a priori defined thresholds to reduce the noise bias. Zrnic [40] has shown that Fourier methods are inferior to autocorrelation methods (which will be defined in the next sections) at low SNR's and for narrow spectral widths.

Pulse-Pair Mean and Variance Estimators

Due to the large number of samples collected by some measurement systems, a method for the estimating the spectral mean and variance was needed which reduced the processing time relative to that required by Fourier based spectral moment estimation techniques. The algorithm used in many present day systems is the pulse-pair algorithm developed by W. D. Rummler in 1968 while at Bell Telephone Labs [28]. The pulse-pair algorithm is based on the Fourier transform pair relationship between the autocorrelation function and the power spectral density and the characteristic function defined in probability theory. A derivation of the pulse-pair algorithm will be given at this time in order to show its relationship to similar estimators developed in later sections.

Assume that there exist a zero-mean complex stationary Gaussian process, $x(t)$, and an additive independent complex zero-mean Gaussian noise process, $n(t)$. Also, assume that the noise process is white. The total process $y(t) = x(t) + n(t)$ has an autocorrelation function defined by

$$r_{YY}(\tau) = E\left[y(t+\tau)\, y^*(t)\right] , \tag{56}$$

and since $x(t)$ and $n(t)$ are independent, the autocorrelation function for the total process can be written as the sum of the autocorrelation functions of the individual processes

$$r_{YY}(\tau) = r_{XX}(\tau) + r_{NN}(\tau) . \tag{57}$$

The ACF for the white noise process is defined as

$$r_{NN}(\tau) = \sigma^2 \delta(\tau) . \tag{58}$$

From probability theory, the characteristic function associated with a probability density function is defined as

$$\Phi(\tau) = \int_{-\infty}^{\infty} f(\omega) \exp(j\omega\tau)\, d\omega \tag{59}$$

where $f(\omega)$ is a density function. Since the noise is white, the derivation will proceed using only the signal process, and the noise will be compensated for at the end of the derivation, if needed. Let $P_{XX}(\omega)$ be the PSD associated with the signal process $x(t)$. Then the normalized PSD can be written as

$$\bar{P}_{XX}(\omega) = \frac{P_{XX}(\omega)}{\int_{-\infty}^{\infty} P_{XX}(\omega)\, d\omega} \tag{60}$$

where

$$\int_{-\infty}^{\infty} \bar{P}_{XX}(\omega)\, d\omega = 1\,. \tag{61}$$

This normalized PSD can now be used to define nth moment

$$M_n = \int_{-\infty}^{\infty} \omega^n \bar{P}_{XX}(\omega)\; d(\omega) \tag{62}$$

of the random variable, ω. As shown in Appendix A, the moments of a density function are related to the characteristic function through the following relationship

$$M_n = (-j)^n \frac{d^n \Phi(\tau)}{d\tau^n}\,. \tag{63}$$

Now, let the total energy in the signal process be defined as

$$E = \int_{-\infty}^{\infty} P_{XX}(\omega)\; d\omega\,. \tag{64}$$

Since $r_{XX}(\tau)$ and $P_{XX}(\omega)$ form a Fourier transform pair, $r_{XX}(\tau)$ can be written as

$$r_{XX}(\tau) = \frac{1}{2\pi} \int_{-\infty}^{\infty} P_{XX}(\omega) \exp(j\omega\tau)\; d\omega\,. \tag{65}$$

Normalizing both sides by $2\pi E$ yields

$$\frac{2\pi r_{XX}(\tau)}{E} = \int_{-\infty}^{\infty} \bar{P}_{XX}(\omega)\, exp(j\omega\tau)\; d\omega \tag{66}$$

where the right hand side can be interpreted as the characteristic function associated with the density function $\bar{P}_{XX}(\omega)$. Therefore, calculating the first moment using Equations 59, 63, and 66 yields

$$M_1 = -j\frac{2\pi}{E}\frac{dr_{XX}(\tau)}{d\tau} \tag{67}$$

which is the mean of the PSD.

Let the signal autocorrelation function be expressed as

$$r_{XX}(\tau) = A(\tau)\exp(j\Theta(\tau)) \tag{68}$$

where $A(\tau)$ is the magnitude function and is real and even, and $\Theta(\tau)$ is the phase function and is real and odd. Therefore, the mean can be written as

$$\begin{aligned} M_1 &= -j\frac{2\pi}{E}\frac{d}{d\tau}\left[A(\tau)\exp(j\Theta(\tau))\right]\Big|_{\tau=0} \\ &= -j\frac{2\pi}{E}\left[A'(\tau)\exp(j\Theta(\tau)) + A(\tau)\Theta'(\tau)\exp(j\Theta(\tau))\right]\Big|_{\tau=0} \end{aligned} \tag{69}$$

Now the even and odd properties of the autocorrelation magnitude and phase functions result in the following:

- $A'(\tau)$ is odd $\Rightarrow A'(0) = 0$.
- $\Theta(\tau)$ is odd $\Rightarrow \Theta(0) = 0$.

Therefore, the mean can be written as

$$M_1 = \frac{2\pi}{E}A(0)\,\Theta'(0) \tag{70}$$

and $A(0)$ is

$$A(0) = r_{XX}(0) = \frac{1}{2\pi}\int_{-\infty}^{\infty} P_{XX}(\omega)\exp(j\omega\tau)\ d\omega = \frac{E}{2\pi}\,. \tag{71}$$

Using Equations 70 and 71, the mean reduces to

$$M_1 = \Theta'(0)\ . \tag{72}$$

Now the derivative at zero can be approximated by following first order difference equation

$$\Theta'(\tau) \cong \frac{\Theta(\tau+T) - \Theta(0)}{T} = \frac{\Theta(T)}{T} \tag{73}$$

where T is the sampling period for a given measurement system . In terms of frequency in Hertz, the mean estimator can be expressed as

$$\hat{f} = \frac{1}{2\pi T}\arg(r_{YY}(T))\ . \tag{74}$$

Note that the mean estimate does not depend on the zeroth autocorrelation lag and therefore avoids any bias due to the noise power in $r_{YY}(0)$.

The variance estimate can also be obtained in a similar manner. Given that

$$\sigma^2 = E\{X^2\} - E^2\{X\} \,. \tag{75}$$

Using the characteristic function, the second moment about zero is defined in terms of the autocorrelation function as

$$\begin{aligned} M_2 &= (-j)^2 \frac{2\pi}{E} \frac{d^2}{d\tau^2} \left[A(\tau) \exp(j\Theta(\tau)) \right] \Big|_{\tau=0} \\ &= (-j)^2 \frac{2\pi}{E} \Big[A''(\tau) \exp(j\Theta(\tau)) + A'(\tau)\Theta'(\tau) \exp(j\Theta(\tau)) + \\ &\quad A'(\tau)\Theta'(\tau) \exp(j\Theta(\tau)) + A(\tau)\Theta''(\tau) \exp(j\Theta(\tau)) + \\ &\quad A(\tau)(\Theta'(\tau))^2 \exp(j\Theta(\tau)) \Big] \Big|_{\tau=0} \,. \end{aligned} \tag{76}$$

Using the even and odd properties of the magnitude and phase, Equation 76 reduces to

$$M_2 = \frac{-2\pi A''(0)}{E} + \left[\Theta'(0)\right]^2 \,. \tag{77}$$

Therefore, Equation 75 can be written as

$$\sigma^2 = \frac{-2\pi A''(0)}{E} \,. \tag{78}$$

A first order estimate of the second derivate of the phase can be written as

$$A''(\tau) \cong \frac{A(\tau + 2T) - 2A(\tau + T) + A(\tau)}{T^2} \,. \tag{79}$$

Papoulis [24] has shown that

$$A(0) - A(T) \cong \frac{A(0) - A(2T)}{4} \tag{80}$$

for small T. Therefore, Equation 79 can be approximated as

$$A''(0) \cong \frac{2A(T) - 2A(0)}{T^2} \tag{81}$$

for $\tau = 0$. Substituting this approximation of the second derivative of the correlation magnitude at $\tau = 0$ into Equation 77 yields

$$\begin{aligned} \sigma^2 &= \frac{-4\pi}{E}\frac{A(T) - A(0)}{(T)^2} \\ &= 2\left[1 - \frac{A(T)}{A(0)}\right] . \end{aligned} \tag{82}$$

The pulse-pair width estimator can then be expressed in terms of Hertz as

$$\hat{\sigma}_f = \frac{\sqrt{2}}{(2\pi T)}\left[1 - \frac{|r_{XX}(T)|}{r_{XX}(0)}\right]^{\frac{1}{2}} . \tag{83}$$

Accounting for the white noise power in the zeroth autocorrelation lag, Equation 83 can be written as

$$\hat{\sigma}_f = \frac{\sqrt{2}}{(2\pi T)}\left[1 - \frac{|r_{YY}(T)|}{r_{YY}(0) - r_{NN}(0)}\right]^{\frac{1}{2}} . \tag{84}$$

Equations 74 and 84 form the pulse-pair mean and width estimators which are a function of the sampling period and the zeroth and first autocorrelation lags.

Zrnic [37] shows that the pulse-pair mean estimator performs well in the case of narrowband symmetric spectra and is unbiased in the presence of white noise. The variance in the mean frequency estimator is a function of the spectral width. Zrnic states that the performance of the mean estimator is best when the width of the spectrum is less than 0.25π [37]. The variance or width estimator (square root of the variance) is not as well behaved. Figures 16 and 17 show the standard deviation and bias for the width estimator, respectively, as a function of SNR and spectral width assuming a Gaussian shaped power spectrum [37].

Estimators Based on Assumed Spectral Shape

Additional mean and width (standard deviation) estimators can be derived from assumed PSD shapes. As previously stated, in the case of Doppler power spectra associated with weather return, a Gaussian shaped PSD is often observed [9]. A Gaussian shaped PSD is characterized by

$$S(\omega) = \frac{K}{\sqrt{2\pi\sigma^2}} \exp\left(\frac{-\omega^2}{2\sigma^2}\right) \tag{85}$$

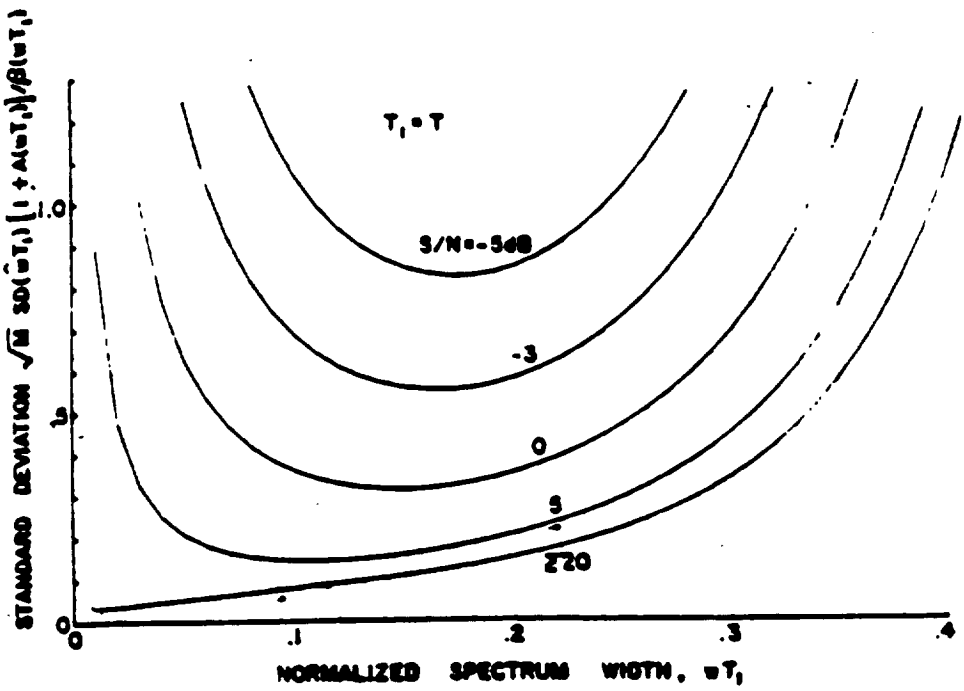


Figure 16. A plot of the normalized standard deviation associated with the pulse-pair width estimator. (Source: ©1977 IEEE, Zrnic, "Spectral Moment Estimates from Correlated Pulse Pairs", IEEE Trans. on AES, July 1977.)

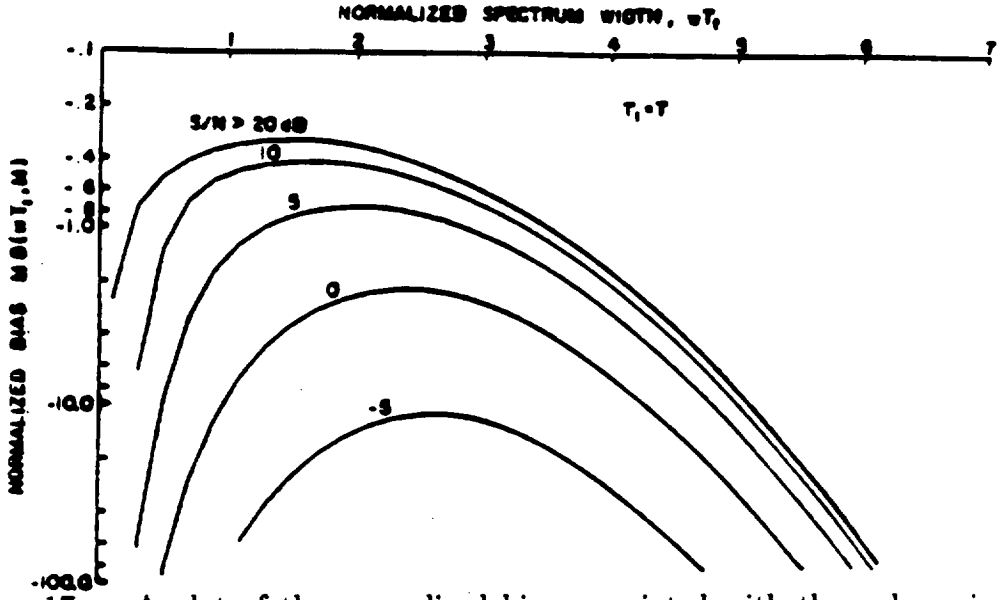


Figure 17. A plot of the normalized bias associated with the pulse-pair width estimator. (Source: ©1977 IEEE, Zrnic, "Spectral Moment Estimates from Correlated Pulse Pairs", IEEE Trans. on AES, July 1977.)

where K is an energy scale factor and σ is the standard deviation. The resulting ACF is also Gaussian shaped and is expressed as

$$r(\tau) = \frac{K}{2\pi} \exp\left(\frac{-\sigma^2\tau^2}{2}\right) . \tag{86}$$

Two new width estimators result from this expression for the ACF. The first estimator

$$\sigma = \sqrt{2 \ln\left[\frac{r(0)}{r(1)}\right]} \tag{87}$$

is a function of $r(0)$ which represents the total energy associated with the process including the noise power. If the noise is white, an estimate of the noise power can be subtracted from $r(0)$ to obtain a reduced bias estimate of the spectral width. For cases where the slope of the autocorrelation magnitude function near zero is small, a first order approximation of the natural logarithm in Equation 87 yields the pulse-pair width estimator given in Equation 84. A second estimator which circumvents the need to estimate the noise power in the white noise case is

$$\sigma = \sqrt{\frac{2}{3} \ln\left[\frac{r(1)}{r(2)}\right]} \tag{88}$$

Passarelli [25] has identified additional moment estimators which are based on an assumed PSD shapes. His work includes defining PSD models for the weather, clutter, and noise in a Doppler weather radar return and using these to obtain mean estimates.

Zrnic [38] has also investigated the width estimators given in Equations 87 and 88. Figure 18 shows the standard deviation in the the log based estimator given in Equation 87. The standard deviation of the estimator increases with decreasing SNR and for narrow and broad spectra. The bias associated with the estimator is shown in Figure 19. The bias is also a function of SNR, and as seen in Figure 19, the bias increases for narrow and broad spectra. Zrnic shows that for large SNR the standard deviation in the width estimate increases linearly with the true width below

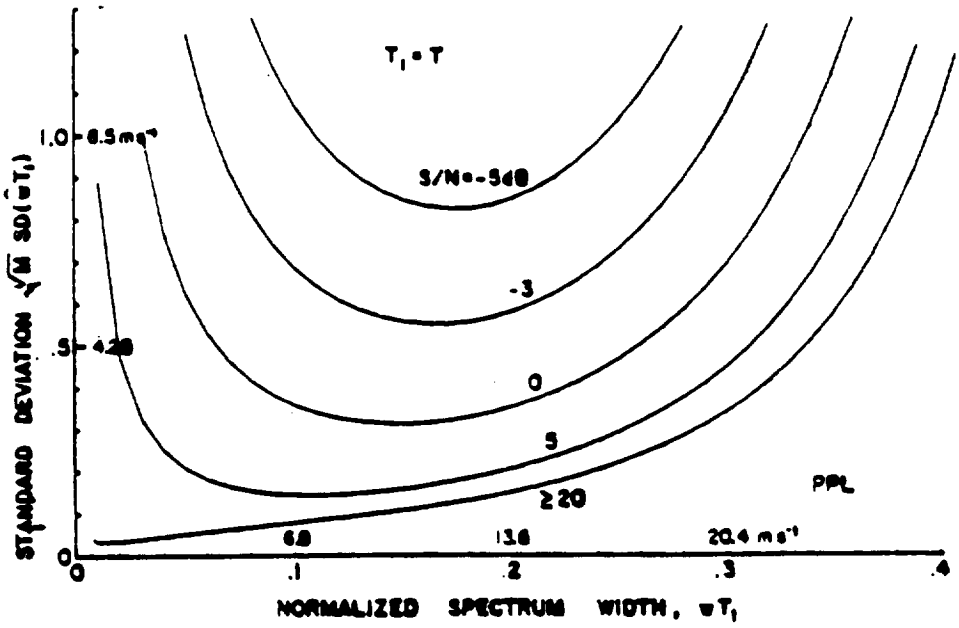


Figure 18. A plot of the normalized standard deviation associated with the width estimator derived assuming a Gaussian shape. (Source: ©1979 IEEE, Zrnic, "Spectrum Width Estimates for Weather Echos ", IEEE Trans. on AES, September 1979.)

a normalized width of 0.3. The standard deviation in the width estimator then increases exponentially for widths greater than 0.3. Figure 20 shows the standard deviation in the width estimator in Equation 88. Without the inclusion of the zeroth lag in the estimate, there is an increase in the standard deviation over that in Figure 18, especially at the larger widths.

General Autocorrelation Based Moment Estimators

Passarelli [26] has derived a general expression for autocorrelation based moment estimators in terms of a McLauren series expansion of the complex autocorrelation function. As in the pulse-pair derivation, the Fourier transform pair relationship between the autocorrelation function and the power spectral density is used to relate the two domains. The following discussion steps through Passarelli's derivation.

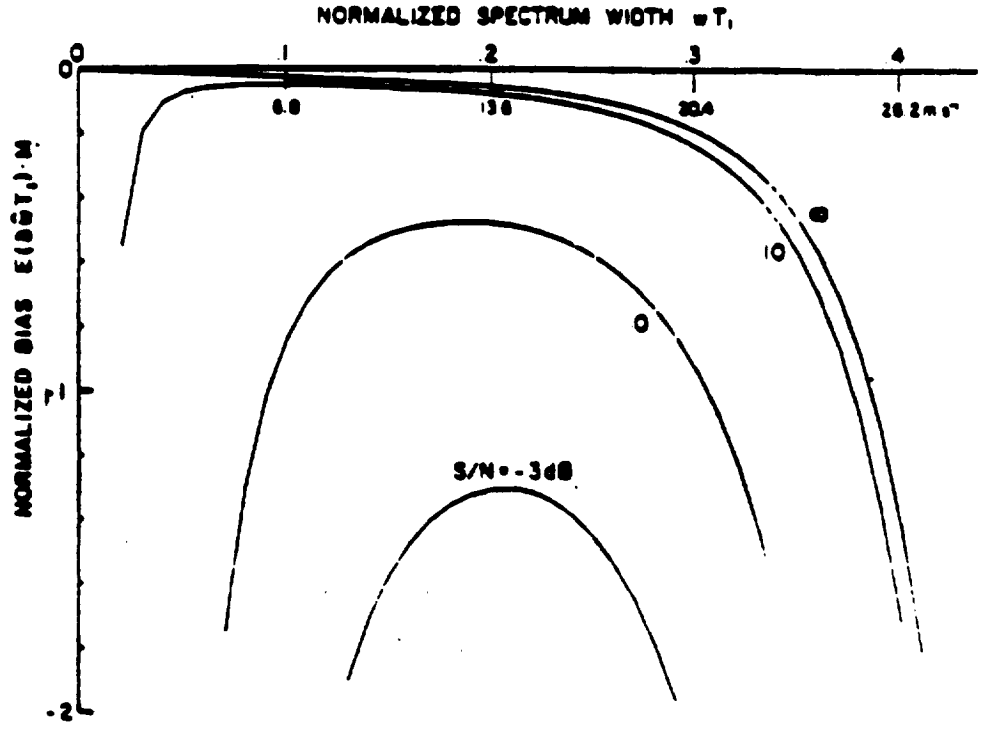


Figure 19. A plot of the normalized bias associated with the width estimator derived assuming a Gaussian shape. (Source: ©1979 IEEE, Zrnic, "Spectrum Width Estimates for Weather Echos ", IEEE Trans. on AES, September 1979.)

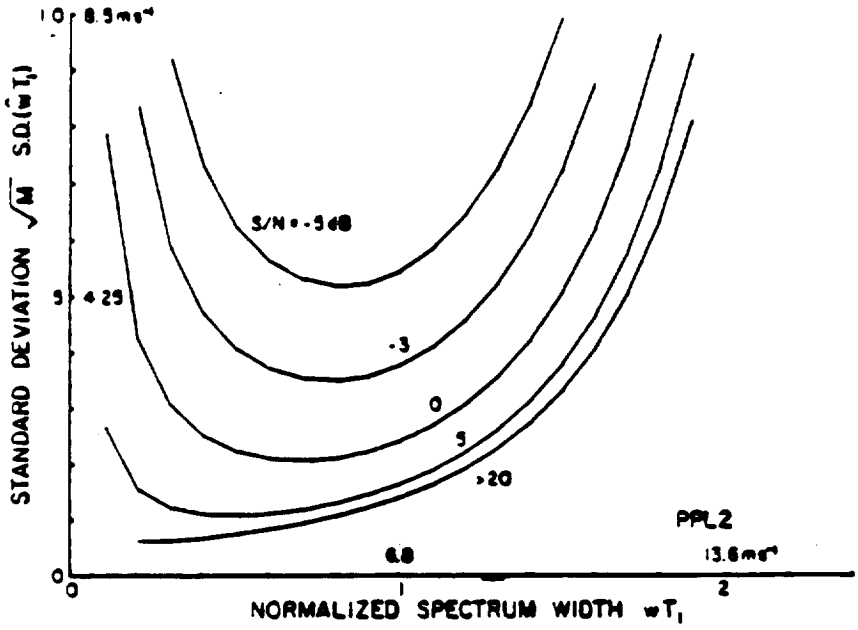


Figure 20. A plot of the normalized standard deviation associated with the width estimator derived assuming a Gaussian shape and using autocorrelation lags one and two. (Source: ©1979 IEEE, Zrnic, "Spectrum Width Estimates for Weather Echos ", IEEE Trans. on AES, September 1979.)

For a continuous time random process, the power spectral density is defined as

$$P_{XX}(\omega) = \int_{-\infty}^{\infty} r_{XX}(\tau)\, \exp(-j\omega\tau)\, d\tau \tag{89}$$

where the autocorrelation function for the complex WSS random process is defined as

$$r_{XX}(\tau) = E\left\{x(t+\tau)x^*(t)\right\} \ . \tag{90}$$

Through the Fourier transform inverse, the autocorrelation function is stated in terms of the PSD, as

$$r_{XX}(\tau) = \frac{1}{2\pi}\int_{-\infty}^{\infty} P_{XX}(\omega)\, \exp(j\omega\tau)\, d\omega \ . \tag{91}$$

For a complex WSS random process, the autocorrelation function can be written as

$$r_{XX}(\tau) = \frac{h(\tau)}{2\pi} exp(j\, g(\tau)) = \frac{A}{2\pi} + j\frac{B}{2\pi} \tag{92}$$

where $h(\tau)$ is referred to as the magnitude and $g(\tau)$ is referred to as the phase. The real and imaginary parts, A and B, of the autocorrelation function are defined in terms of the inverse Fourier transform where

$$A(\tau) \equiv Re[2\pi r_{XX}(\tau)] = \int_{-\infty}^{\infty} P_{XX}(\omega)\, \cos(\omega\tau)\, d\omega \tag{93}$$

and

$$B(\tau) \equiv Re[2\pi r_{XX}(\tau)] = \int_{-\infty}^{\infty} P_{XX}(\omega)\, \sin(\omega\tau)\, d\omega\ . \tag{94}$$

Note that both A and B are functions of τ, but henceforth, the τ dependence will be dropped for notational convenience. Therefore, the magnitude, $h(\tau)$, can be written as

$$h(\tau) = \left[A^2 + B^2\right]^{\frac{1}{2}} \tag{95}$$

and the phase, $g(\tau)$, can be written as

$$g(\tau) = \arctan\left[\frac{B}{A}\right]\ . \tag{96}$$

Note that $h(\tau)$ is an even function, and $g(\tau)$ is an odd function given that in Equations 93 and 94, A is an even function of τ, and B is odd function of τ. The autocorrelation function magnitude and phase can be expanded as a McLauren series about $\tau = 0$ where

$$h(\tau) = \frac{h(0)}{0!} + \frac{h'(0)\,\tau}{1!} + \frac{h''(0)\,\tau^2}{2!} + \frac{h'''(0)\,\tau^3}{3!} + \ldots \tag{97}$$

and

$$g(\tau) = \frac{g(0)}{0!} + \frac{g'(0)\,\tau}{1!} + \frac{g''(0)\,\tau^2}{2!} + \frac{g'''(0)\,\tau^3}{3!} + \ldots\ . \tag{98}$$

Now, since $h(\tau)$ is an even function, and $g(\tau)$ is an odd function, the polynomial expansions must also be even and odd, respectively, therefore

$$h(\tau) = \frac{h(0)}{0!} + \frac{h''(0)\,\tau^2}{2!} + \frac{h^{iv}(0)\,\tau^4}{4!} + \ldots \tag{99}$$

and

$$g(\tau) = \frac{g'(0)\,\tau}{1!} + \frac{g'''(0)\,\tau^3}{3!} + \frac{g^{v}(0)\,\tau^5}{5!} + \ldots\ . \tag{100}$$

Moments of PSD are defined in terms of a normalized PSD where

$$M_n = \frac{1}{P_0} \int_{-\infty}^{\infty} \omega^n P_{XX}(\omega)\, d\omega \tag{101}$$

and

$$P_0 = \int_{-\infty}^{\infty} P_{XX}(\omega)\, d\omega\ . \tag{102}$$

Moments of the PSD can also be defined in terms of the autocorrelation function through the characteristic function given in Equation 59 where

$$M_n = \frac{1}{r_{XX}(0)}(-j)^n \left.\frac{d^n r_{XX}(\tau)}{d\tau^n}\right|_{\tau=0}\ . \tag{103}$$

Applying Equation 103 to the definition of the autocorrelation function in Equation 92 yields,

$$\begin{aligned} M_1 &= \frac{1}{r_{XX}(0)}(-j) \left.\frac{dr_{XX}(\tau)}{d\tau}\right|_{\tau=0} \\ &= \frac{-j}{2\pi r_{XX}(0)} \left[\int_{-\infty}^{\infty} -\omega P_{XX}(\omega)\sin(\omega\tau)\, d\omega\Big|_{\tau=0} + j\int_{-\infty}^{\infty} \omega P_{XX}(\omega)\cos(\omega\tau)\, d\omega\Big|_{\tau=0}\right] \\ &= 0 + \left.\frac{B'}{P_0}\right|_{\tau=0} \end{aligned} \tag{104}$$

$$\begin{aligned} M_2 &= \frac{1}{r_{XX}(0)}(-j)^2 \left.\frac{d^2 r_{XX}(\tau)}{d\tau^2}\right|_{\tau=0} \\ &= \frac{(-j)^2}{2\pi r_{XX}(0)} \left[\int_{-\infty}^{\infty} -\omega^2 P_{XX}(\omega)\cos(\omega\tau)\, d\omega\Big|_{\tau=0} + j\int_{-\infty}^{\infty} -\omega^2 P_{XX}(\omega)\sin(\omega\tau)\, d\omega\Big|_{\tau=0}\right] \\ &= 0 - \left.\frac{A''}{P_0}\right|_{\tau=0} \end{aligned} \tag{105}$$

$$\begin{aligned} M_3 &= \frac{1}{r_{XX}(0)}(-j)^3 \left.\frac{d^3 r_{XX}(\tau)}{d\tau^3}\right|_{\tau=0} \\ &= \frac{(-j)^3}{2\pi r_{XX}(0)} \left[\int_{-\infty}^{\infty} \omega^3 P_{XX}(\omega)\sin(\omega\tau)\, d\omega\Big|_{\tau=0} + j\int_{-\infty}^{\infty} -\omega^3 P_{XX}(\omega)\cos(\omega\tau)\, d\omega\Big|_{\tau=0}\right] \\ &= 0 - \left.\frac{B'''}{P_0}\right|_{\tau=0} \end{aligned} \tag{106}$$

and

$$\begin{aligned} M_4 &= \frac{1}{r_{XX}(0)}(-j)^4 \left.\frac{d^4 r_{XX}(\tau)}{d\tau^4}\right|_{\tau=0} \\ &= \frac{(-j)^4}{2\pi r_{XX}(0)} \left[\int_{-\infty}^{\infty} \omega^4 P_{XX}(\omega)\cos(\omega\tau)\, d\omega\Big|_{\tau=0} + j\int_{-\infty}^{\infty} \omega^4 P_{XX}(\omega)\sin(\omega\tau)\, d\omega\Big|_{\tau=0}\right] \\ &= 0 + \left.\frac{A^{iv}}{P_0}\right|_{\tau=0} \end{aligned} \tag{107}$$

for the moments M_1 - M_4. After several iterations, a pattern develops relating the moments of the PSD to the derivatives of either the real or imaginary parts of the autocorrelation function. This pattern is summarized in the following set of equations

$$\left\{ \begin{array}{ll} A|_{\tau=0} = P_0 & B|_{\tau=0} = 0 \\ A'\big|_{\tau=0} = 0 & B'\big|_{\tau=0} = P_0 M_1 \\ A''\big|_{\tau=0} = -P_0 M_2 & B''\big|_{\tau=0} = 0 \\ A'''\big|_{\tau=0} = 0 & B'''\big|_{\tau=0} = -P_0 M_3 \\ A^{iv}|_{\tau=0} = P_0 M_4 & B'''\big|_{\tau=0} = 0 \\ A^{v}|_{\tau=0} = 0 & B^{v}|_{\tau=0} = P_0 M_5 \\ A^{vi}|_{\tau=0} = -P_0 M_6 & B^{v}|_{\tau=0} = 0 \end{array} \right\} . \tag{108}$$

A relationship between the coefficients of the McLauren series expansion and the central moments of the PSD can be defined. Writing the magnitude of the autocorrelation function in terms of A and B yields

$$h = (A^2 + B^2)^{\frac{1}{2}} \tag{109}$$

and its first derivative with respect to τ is

$$h' = \frac{1}{2}(A^2 + B^2)^{\frac{-1}{2}}(2AA' + 2BB') . \tag{110}$$

Now let

$$G = AA' + BB' . \tag{111}$$

Then h and its derivatives can be written in terms of h and G where

$$\left.\begin{aligned} h' &= \tfrac{G}{h} \\ h'' &= \tfrac{1}{h}(G' - (h')^2) \\ h''' &= \tfrac{1}{h}(G'' - 3h'h'') \\ h^{iv} &= \tfrac{1}{h}(G''' - 4h'h'''3(h'')^2) \end{aligned}\right\}. \tag{112}$$

The derivatives of G are expressed in terms of A and B and their derivatives where

$$\left.\begin{aligned} G &= AA' + BB' \\ G' &= AA'' + (A')^2 + BB'' + (B')^2 \\ G'' &= AA''' + 3A'A'' + BB''' + 3B'B'' \\ G''' &= AA^{iv} + 4A'A''' + 3(A'')^2 + BB^{iv} + 4B'B''' + 3(B'')^2 \end{aligned}\right\}. \tag{113}$$

Given the relationship between the moments of the PSD and the derivatives of the real and imaginary parts of the autocorrelation function in Equation 108, the derivatives of G can be expressed as

$$\left.\begin{aligned} G &= 0 \\ G' &= -(P_0)^2 M_2 + (P_0 M_1)^2 \\ G'' &= 0 \\ G''' &= (P_0)^2 M_4 + 3(P_0)^2 (M_2)^2 - 4(P_0)^2 M_1 M_3 \end{aligned}\right\}. \tag{114}$$

The derivatives of magnitude of the autocorrelation function can then be expressed as

$$\left.\begin{aligned} h' &= 0 \\ h'' &= -P_0(M_2 - (M_1)^2) = -P_0 \mathcal{M}_2 \\ h''' &= 0 \\ h^{iv} &= P_0(M_4 + 3(M_2)^2 - 4M_1 M_3 - 3(M_2 - (M_1)^2)^2) \\ &= P_0(M_4 - 4M_1 M_3 + 6M_2 M_1^2 - 3(M_1)^4) = P_0 \mathcal{M}_4 \end{aligned}\right\}. \tag{115}$$

Note that the relationships in Equation 115 have been expressed in terms of central moments about the mean, M_1, where

$$\mathcal{M}_n = \frac{1}{P_0} \int_{-\infty}^{\infty} (\omega - M_1)^n \, P_{XX}(\omega) \, d\omega \, . \tag{116}$$

The resulting McLauren series expansion for the magnitude of the autocorrelation function can then be expressed as

$$h(\tau) = P_0 \left[1 - \frac{\mathcal{M}_2 \tau^2}{2!} + \frac{\mathcal{M}_4 \tau^4}{4!} - \frac{(\mathcal{M}_6 - 10\mathcal{M}_3^2)\tau^6}{6!} + \ldots \right] . \tag{117}$$

Passarelli did not extend the expansion in terms of central moments beyond what is given here. He indicates that no general pattern was observed.

A similar approach can be taken in defining the McLauren series coefficients for the expansion of the phase. Let the phase associated with the autocorrelation function be written in terms of the real and imaginary parts of the autocorrelation where

$$g = arctan \left[\frac{B}{A} \right] \tag{118}$$

and its first derivative with respect to τ is

$$g' = \frac{AB' - BA'}{A^2 + B^2} \, . \tag{119}$$

Now, let

$$H = AB' - BA' \tag{120}$$

and

$$Q = A^2 + B^2 \, . \tag{121}$$

Then the higher order derivatives of the phase can be expressed in terms of H and Q as

$$\left.\begin{aligned} g &= arctan \left[\tfrac{B}{A} \right] \\ g' &= \tfrac{H}{Q} \\ g'' &= \tfrac{H' - Q' g'}{Q} \\ g''' &= \tfrac{H'' - 2Q' g'' - Q'' g'}{Q} \end{aligned}\right\} \, . \tag{122}$$

The higher order derivatives of H and Q, in Equation 122, can be expressed in terms of A and B as

$$\left.\begin{aligned} H &= AB' - BA' \\ H' &= AB'' - BA'' \\ H'' &= AB''' + A'B'' - B'A'' - BA''' \\ H''' &= AB^{iv} + 3A'B^{iv} - 2A^{iv}B' - BA^{iv} \end{aligned}\right\} \tag{123}$$

and

$$\left.\begin{aligned} Q &= A^2 + B^2 \\ Q' &= 2AA' + 2BB' \\ Q'' &= 2((A')^2 + AA'' + (B')^2 - BB'') \\ Q''' &= 2(AA''' + 3A'A'' + 3B'B'' + BB''') \end{aligned}\right\}. \tag{124}$$

Using the relationships between the derivatives of the real and imaginary parts of the autocorrelation function and the moments of the PSD defined in Equations 104 through 107, H and Q and its derivatives can be expressed as

$$\left.\begin{aligned} H &= P_0^2 M_1 \\ H' &= 0 \\ H'' &= -P_0^2 M_3 + P_0^2 M_1 M_2 \\ H''' &= 0 \end{aligned}\right\} \tag{125}$$

and

$$\left.\begin{aligned} Q &= P_0^2 \\ Q' &= 0 \\ Q'' &= -2P_0^2 M_2 + 2P_0^2 M_1^2 \\ Q''' &= 0 \end{aligned}\right\}. \tag{126}$$

Combining Equations 122, 125, and 126 yields

$$\left.\begin{aligned} g &= 0 \\ g' &= M_1 \\ g'' &= 0 \\ g''' &= -M_3 + 3M_1M_2 - 2M_1^3) = -\mathcal{M}_3 \end{aligned}\right\} . \tag{127}$$

The series expansion in terms of central moments can be written as

$$g(\tau) = M_1\tau - \frac{\mathcal{M}_3\tau^3}{3!} + \frac{(\mathcal{M}_5 - 10\mathcal{M}_2\mathcal{M}_3)\tau^5}{5!} + \ldots . \tag{128}$$

Again, Passarelli was not able to define the expansion in terms of central moments beyond what is given here.

The preceding derivation assumed a continuous autocorrelation function; however, in the case of a coherent pulsed Doppler radar, the return pulses are sampled resulting in a discrete-time random process. The resulting estimated autocorrelation function is therefore discrete, and the corresponding power spectrum is defined in terms of the discrete-time Fourier transform. The series expansion in Equations 117 and 128 can easily be written in terms of central moments of the PSD for a discrete-time random process.

Consider the nth central moment for a continuous random process where

$$\mathcal{M}_n = \frac{1}{P_0}\int_{-\infty}^{\infty}(\omega - \bar{\omega})^n P_{XX}(\omega)\, d\omega \tag{129}$$

and

$$P_0 = \int_{-\infty}^{\infty} P_{XX}(\omega)\, d\omega . \tag{130}$$

Also, the nth central moment for a discrete random process is defined as

$$\bar{\mathcal{M}}_n = \frac{1}{\bar{P}_{XX}(0)}\int_{-\pi}^{\pi}(\theta - \bar{\theta})^n P_{XX}(\theta)\, d\theta \tag{131}$$

and

$$\bar{P}_0 = \int_{-\pi}^{\pi} \bar{P}_{XX}(\theta)\, d\theta . \tag{132}$$

Now, the power spectrum for a discrete-time random process with sampling period T is related to the power spectrum of a continuous time random process by

$$\bar{P}_{XX}(\theta) = \frac{P_{XX}(\omega)}{T} \tag{133}$$

provided the Nyquist condition is met [22]. Therefore, the total power, $\bar{P}_0$, in the discrete-time random process can be written in terms of the power spectrum of a continuous time random process by

$$\bar{P}_0 = \int_{-\infty}^{\infty} \frac{P_{XX}(\omega)}{T}\, d\omega\, T = P_0\,. \tag{134}$$

The nth central moment for the discrete-time process can be expressed in terms of $P_{XX}(\omega)$ as

$$\begin{aligned} \bar{\mathcal{M}}_n &= \frac{1}{P_0}\int_{-\infty}^{\infty} T^n\,(\omega-\bar{\omega})^n\,\frac{P_{XX}(\omega)}{T}\,d\omega\,T \\ &= T^n\,\frac{1}{P_0}\int_{-\infty}^{\infty}(\omega-\bar{\omega})^n\,P_{XX}(\omega)\,d\omega \\ &= \mathcal{M}_n T^n\,. \end{aligned} \tag{135}$$

Now, sampling the expansions given for the magnitude and phase of the autocorrelation functions in Equations 117 and 128 at $\tau = kT$ where T is the sampling period yields

$$h(kT) = P_0\left[1 - \frac{\mathcal{M}_2(kT)^2}{2!} + \frac{\mathcal{M}_4(kT)^4}{4!} - \frac{(\mathcal{M}_6 - 10\mathcal{M}_3^2)(kT)^6}{6!} + \ldots\right] \tag{136}$$

and

$$g(kT) = M_1(kT) - \frac{\mathcal{M}_3(kT)^3}{3!} + \frac{(\mathcal{M}_5 - 10\mathcal{M}_2\mathcal{M}_3)(kT)^5}{5!} + \ldots\,. \tag{137}$$

Using the relationship found in Equation 135, Equations 136 and 137 can be expressed as

$$h(kT) = P_0\left[1 - \frac{\bar{\mathcal{M}}_2(k)^2}{2!} + \frac{\bar{\mathcal{M}}_4(k)^4}{4!} - \frac{(\bar{\mathcal{M}}_6 - 10\bar{\mathcal{M}}_3^2)(k)^6}{6!} + \ldots\right] \tag{138}$$

and

$$g(kT) = \bar{M}_1(k) - \frac{\bar{\mathcal{M}}_3(k)^3}{3!} + \frac{(\bar{\mathcal{M}}_5 - 10\bar{\mathcal{M}}_2\bar{\mathcal{M}}_3)(k)^5}{5!} + \ldots . \tag{139}$$

The case of additive white noise can be added to the model yielding

$$h(kT) = N\delta(0) + P_0\left[1 - \frac{\bar{\mathcal{M}}_2(k)^2}{2!} + \frac{\bar{\mathcal{M}}_4(k)^4}{4!} - \frac{(\bar{\mathcal{M}}_6 - 10\bar{\mathcal{M}}_3^2)(k)^6}{6!} + \ldots\right] \tag{140}$$

where N is the noise power.

Passarelli proposed that a closed system of equations be formed by truncating the series expansion found in Equations 140 and 137. The closed system of equations are then solved to yield estimators for different moments such as the mean, $\bar{M}_1$, the variance, $\bar{\mathcal{M}}_2$, and the skewness, $\bar{\mathcal{M}}_3$. Passarelli states that two factors will contribute to the performance of the estimators:

1. The rate of convergence of the series.
2. The accuracy of the estimates of the autocorrelation function.

Passarelli also states that for broad or skewed spectra more terms will be needed in the expansion to model the autocorrelation function. For narrow or symmetric spectra, Passarelli states that only a few terms will be required for convergence and that the use of additional autocorrelation lags will only add to the variance of the moment estimator.

For purely symmetric spectra, the odd central moments are zero. Passarelli shows that for symmetric spectra the expansion reduces to

$$h(kT) = P_0 \sum_{n=0}^{\infty} (-1)^n \frac{\bar{\mathcal{M}}_{2n} k^{2n}}{2n!} \tag{141}$$

and

$$g(kT) = \bar{M}_1 k \, . \tag{142}$$

Poly-Pulse-Pair Mean Estimators

Lee and Strauch [35] proposed a poly-pulse-pair approach for improving the mean estimate of weak narrow symmetric spectra. The poly-pulse-pair approach

turns out to be an average of the mean estimators derived from Equation 142. The poly-pulse-pair mean estimator can be written as [26]

$$\hat{\bar{M}}_1 = \frac{1}{N}\sum_{k=1}^{N}\frac{g(kT)}{k} . \tag{143}$$

Passarelli's Variance Estimators

Passarelli investigates his proposed approach for generating moment estimators by assessing the performance of several estimators for the variance, $\bar{\mathcal{M}}_2$. Passarelli derives four variance estimators using a closed system consisting two equations and two unknowns or three equations and three unknowns. The first variance estimator is derived using

$$h(0) = r_{XX}(0) + N \tag{144}$$

and

$$h(1) = r_{XX}(0)[1 + \frac{\bar{\mathcal{M}}_2}{2}] \tag{145}$$

yielding

$$\bar{\mathcal{M}}_2 = 2\left[1 - \frac{h(1)}{h(0) - N}\right] \tag{146}$$

which is the pulse-pair variance estimator. Passarelli goes on to define three other estimators using lags (1,2), (0,1,2), and (1,2,3) and a closed system of equations. For notational purposes these estimators will be referred to by the autocorrelation lags used to derive them. The three additional variance estimators are

$$[0\,1\,2] \quad \bar{\mathcal{M}}_2 = \frac{5}{2} - \frac{1}{h(0)}\left[\frac{8h(1)}{3} - \frac{h(2)}{6}\right] \tag{147}$$

$$[1\,2] \quad \bar{\mathcal{M}}_2 = 2\frac{h(1) - h(2)}{4h(1) - h(2)} \tag{148}$$

and

$$[1\,2\,3] \quad \bar{\mathcal{M}}_2 = \bar{\mathcal{M}}_2[1\,2]\left[1 + \frac{\bar{\mathcal{M}}_4[1\,2\,3]}{4!}\frac{16h(1) - h(2)}{h(1) - h(2)}\right] \tag{149}$$

where

$$\bar{\mathcal{M}}_4[1\,2\,3] = 4!\frac{(h(1) - h(2))(9h(1) - h(3)) - (h(1) - h(3))(4h(1) - h(2))}{(81h(1) - h(3))(4h(1) - h(2)) - (16h(1) - h(2))(9h(1)h(3))} . \tag{150}$$

Note that the estimators [0 1] and [0 1 2] require an estimate of the noise power where [1 2] and [1 2 3] do not.

Passarelli evaluated the variance estimators assuming a Gaussian shaped power spectra which commonly models weather returns. The test data was generated using the algorithm described by Sirmans and Bumgarner [32] for weather returns. Passarelli also assumed an additive white noise source. Passarelli used the four variance estimators listed in Equations 146, 147, 148, and 149 to estimate the variance of a Gaussian shaped spectrum with a standard deviation, σ, between 0.05π and 0.5π. Passarelli used forty realizations for each standard deviation (spectral width) and 128 point autocorrelations. Passarelli measured the performance of the estimators in terms of the normalized error in the variance estimate (normalized by the true variance). Figure 21 shows the normalized error as a function of spectral width and the particular estimator employed for a SNR of 10 dB.

Based on Figure 21, Passarelli points out that each estimator is optimum (minimum error) over a different range of spectral widths. In the case of a narrowband process (widths less that 0.1π), the [1 2] estimator is shown to produce the minimum error. Passarelli states that for the narrowband case, the use of the fewest number of fundamental terms (e.g. N, h(0), h(1), h(2), etc.) in computing $\bar{\mathcal{M}}_2$ will yield the optimum estimator since the series exhibits fast convergence in the narrowband case and any additional fundamental terms only add uncertainty to the estimate. For the wideband case, Passarelli shows that low order contiguous estimates of $h(k)$ are needed since the series requires more terms for convergence and the low order terms contain most of the information resulting from a broad spectrum. In other words, the broad spectra results in an autocorrelation function which falls off quickly based on the uncertainty principle of the Fourier transform.

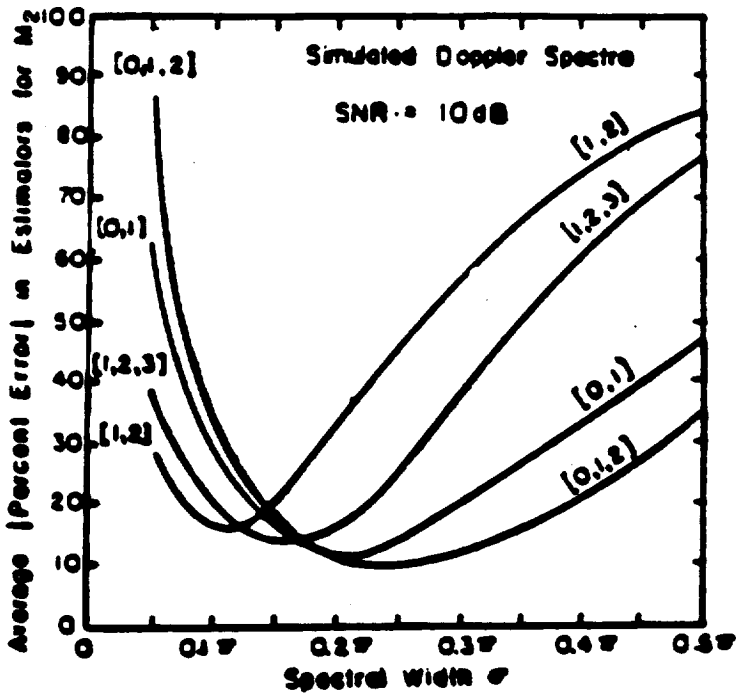


Figure 21. A plot of the normalized error associated with four variance estimators derived from a closed system equations at a SNR of 10 dB. (Source: ©1983 AMS, Passarelli and Siggia, "The Autocorrelation Function and Doppler Spectral Moments: Geometric and Asymptotic Interpretations", IEEE Trans. on AES, July 1977.)

Conclusions

As noted by Passarelli, of all the autocorrelation based variance estimators presented in this chapter an "optimum" estimator resulting in a minimum bias and standard deviation in the estimate over the entire range of spectral widths does not exists. In addition, the range of spectral widths in the narrowband region over which a particular variance estimator derived from Passarelli's expansion is optimum is less than 0.05π. There exist the need to develop variance estimators that exhibit a small bias and standard deviation over a larger range of spectral widths. In other words, a more robust variance estimator is sought.

CHAPTER IV

THE OVERDETERMINED SYSTEM

Introduction

Autocorrelation based spectral moment estimators provide a means for reducing processing requirements over that of Fourier based methods and also provide better performance under various spectral constraints and low signal-to-noise ratios [9]. However, performance of autocorrelation based spectral moment estimators is driven by the quality of the autocorrelation function estimates. The bias and standard deviation associated with the autocorrelation function estimates is a function of the estimator, the number of samples used in the estimate, and the signal-to-noise ratio. In fielded systems, the number of samples available for use in estimating the autocorrelation function and the SNR is often limited by the physical environment and system requirements.

In practice, estimates of the autocorrelation function at only the first few lags are used in autocorrelation based spectral moment estimators. As seen in Chapter III, the pulse-pair mean and width estimators and width estimators based on an assumed Gaussian shaped power spectrum use estimates of the autocorrelation function at lags zero through two. For variance estimators defined by Passarelli's closed systems [26], the use of additional autocorrelation function estimates at the higher order lags only tends to increase the bias in the variance estimate over the narrowband spectral width region.

In this work, Passarelli's series expansion is extended to develop an overdetermined system. This framework allows an assessment of the value of using additional autocorrelation lag estimates for any spectral moment of interest and for any power spectrum shape and corresponding autocorrelation function. The overdetermined

system approach is supported by similar techniques found in ARMA spectral estimation to extract additional information from the available autocorrelation lags [4, 34, 5, 3].

In order to investigate the application of an overdetermined system to increase the performance of a given autocorrelation based spectral moment estimator and to identify under what conditions the overdetermined system might be applied, the variance spectral moment estimator is chosen for evaluation for the case where the power spectrum is Gaussian shaped. The Gaussian shaped spectrum is an interesting case and finds application in many fields including Doppler processing of meteorological returns. In order to investigate the contribution of the overdetermined system to estimator performance, the overdetermined variance estimators will be applied to simulated Doppler weather radar returns under various signal-to-noise conditions and at different spectral widths [15]. Since both the SNR and the number of samples used in the estimate of the autocorrelation function effects the quality of the autocorrelation estimates, the number of samples will be chosen to be a fixed, and the SNR will be varied to evaluate performance. In this work, the overdetermined variance estimators will be shown to improve estimator performance by extending the range of spectral widths and SNR's over which an estimator can perform. In addition, the overdetermined system based on Passarelli's series expansion is shown to yield the least squares solution. The resulting pseudo-inverse is also shown to be pre-computable and can be stored for use in a real-time environment.

It needs to be stressed that the overdetermined system is not limited to variance estimators, but improvements in performance for other moments such as the mean, skewness and kurtosis might be obtained under this framework. However, the scope of this work does not include the investigation of other spectral moments. This work is meant to define a framework upon which additional analysis might be performed. This approach is not limited to the Gaussian shaped spectrum case; however, the

Gaussian shaped spectra is seen in radar meteorological returns and is applicable in other fields as well.

Overdetermined Systems

This section casts Passarelli's truncated autocorrelation expansion as an overdetermined set of equations to be solved for different moments of interest. A solution of this overdetermined system using the pseudo-inverse is shown to be the least squares solution.

Defining the Overdetermined System

A truncated version of Passarelli's expansion in Equations 140 and 139 can be written in terms of a closed system of equations in matrix form as

$$\begin{bmatrix} 1 & 0 & \cdots & 0 \\ 1 & -\frac{(1)^2}{2!} & \cdots & (-1)^N \frac{(1)^{(2N)}}{(2N)!} \\ \vdots & \vdots & \vdots & \vdots \\ 1 & -\frac{(N-1)^2}{2!} & \cdots & (-1)^N \frac{(N-1)^{(2N)}}{(2N)!} \\ 1 & -\frac{(N)^2}{2!} & \cdots & (-1)^N \frac{(N)^{(2N)}}{(2N)!} \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ \vdots \\ a_N \end{bmatrix} = \begin{bmatrix} h(0) \\ h(1) \\ \vdots \\ h(N-1) \\ h(N) \end{bmatrix} \tag{151}$$

and

$$\begin{bmatrix} 1 & -\frac{(1)^3}{3!} & \cdots & (-1)^{(M-1)} \frac{(1)^{(2M-1)}}{(2M-1)!} \\ 2 & -\frac{(2)^3}{3!} & \cdots & (-1)^{(M-1)} \frac{(2)^{(2M-1)}}{(2M-1)!} \\ \vdots & \vdots & \vdots & \vdots \\ M-1 & -\frac{(M-1)^3}{3!} & \cdots & (-1)^{(M-1)} \frac{(M-1)^{(2M-1)}}{(2M-1)!} \\ M & -\frac{(M)^3}{3!} & \cdots & (-1)^{(M-1)} \frac{(M)^{(2M-1)}}{(2M-1)!} \end{bmatrix} \begin{bmatrix} b_0 \\ b_1 \\ \vdots \\ b_{M-1} \end{bmatrix} = \begin{bmatrix} g(1) \\ g(2) \\ \vdots \\ g(M-1) \\ g(M) \end{bmatrix} \tag{152}$$

where $a_0 = P_0$, $a_1 = P_0\mathcal{M}_2$, $b_0 = \mu$, $b_1 = \mathcal{M}_3$, and so forth. This closed set of equations can then be cast as an overdetermined system by considering the case of N_{OD} equations for the magnitude function, $h(k)$, based on a system with (N+1) unknowns (terms in the series expansion) where $N_{OD} > (N+1)$ and M_{OD} equations for the phase function, $g(k)$, based on a system with M unknowns where $M_{OD} > M$.

The overdetermined systems can be expressed in matrix form as

$$\begin{bmatrix} 1 & 0 & \cdots & 0 \\ 1 & -\frac{(1)^2}{2!} & \cdots & (-1)^N \frac{(1)^{(2N)}}{(2N)!} \\ \vdots & \vdots & \vdots & \vdots \\ 1 & -\frac{(N)^2}{2!} & \cdots & (-1)^N \frac{(N)^{(2N)}}{(2N)!} \\ 1 & -\frac{(N+1)^2}{2!} & \cdots & (-1)^N \frac{(N+1)^{(2N)}}{(2N)!} \\ 1 & -\frac{(N+2)^2}{2!} & \cdots & (-1)^N \frac{(N+2)^{(2N)}}{(2N)!} \\ \vdots & \vdots & \vdots & \vdots \\ 1 & -\frac{(N_{OD})^2}{2!} & \cdots & (-1)^N \frac{(N_{OD})^{(2N)}}{(2N)!} \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ \vdots \\ a_N \end{bmatrix} = \begin{bmatrix} h(0) \\ h(1) \\ \vdots \\ h(N) \\ h(N+1) \\ h(N+2) \\ \vdots \\ h(N_{OD}) \end{bmatrix} \tag{153}$$

and

$$\begin{bmatrix} 1 & -\frac{(1)^3}{3!} & \cdots & (-1)^{(M-1)} \frac{(1)^{(2M-1)}}{(2M-1)!} \\ 2 & -\frac{(2)^3}{3!} & \cdots & (-1)^{(M-1)} \frac{(2)^{(2M-1)}}{(2M-1)!} \\ \vdots & \vdots & \vdots & \vdots \\ M & -\frac{(M)^3}{3!} & \cdots & (-1)^{(M-1)} \frac{(M)^{(2M-1)}}{(2M-1)!} \\ M+1 & -\frac{(M+1)^3}{3!} & \cdots & (-1)^{(M-1)} \frac{(M+1)^{(2M-1)}}{(2M-1)!} \\ M+2 & -\frac{(M+2)^3}{3!} & \cdots & (-1)^{(M-1)} \frac{(M+2)^{(2M-1)}}{(2M-1)!} \\ \vdots & \vdots & \vdots & \vdots \\ M_{OD} & -\frac{(M_{OD})^3}{3!} & \cdots & (-1)^{(M-1)} \frac{(M_{OD})^{(2M-1)}}{(2M-1)!} \end{bmatrix} \begin{bmatrix} b_0 \\ b_1 \\ \vdots \\ b_{M-1} \end{bmatrix} = \begin{bmatrix} g(1) \\ g(2) \\ \vdots \\ g(M) \\ g(M+1) \\ g(M+2) \\ \vdots \\ g(M_{OD}) \end{bmatrix} . \tag{154}$$

Solution of the Overdetermined System

In general, an overdetermined system is defined in matrix form as

$$\mathcal{A}\underline{x} = \underline{b} \tag{155}$$

where $\mathcal{A}$ is a matrix of dimension ((n x m) : n > m), $\underline{x}$ is a vector of dimension (m x 1) and $\underline{b}$ is a vector of dimension (n x 1). An approach to solving the overdetermined system is to multiply both sides by the transpose of the $\mathcal{A}$ matrix, $\mathcal{A}^T$, to form the system

$$\mathcal{A}^T\mathcal{A}\underline{x} = \mathcal{A}^T\underline{b} . \tag{156}$$

Now, if the square matrix $\mathcal{A}^T\mathcal{A}$ has full rank, then the system in Equation 156 has only one solution. Also, it can be shown that if $\mathcal{A}$ has full column rank [7, 8], then

$\mathcal{A}^T\mathcal{A}$ has full rank. However, if the matrix $\mathcal{A}$ does not have full column rank, then there are an infinite number of solutions.

For the case where $\mathcal{A}$ has full column rank, the solution can be expressed as

$$\underline{x} = \left[\mathcal{A}^T\mathcal{A}\right]^{-1}\mathcal{A}^T\underline{b} \tag{157}$$

and is often termed the least squares solution. In Equation 155, the vector $\underline{b}$ may not lie in the space spanned by the columns of $\mathcal{A}$. Note that since the rank of $\mathcal{A}$ is not equal to n, the columns of $\mathcal{A}$ do not span $\Re^n$ space in which $\underline{b}$ lies. If one is trying to find the solution $\underline{x}$ that minimizes the norm of the error vector (i.e. the least squares solution) defined by

$$\| \underline{e} \| = (\underline{b} - \mathcal{A}\underline{x})^T (\underline{b} - \mathcal{A}\underline{x}) \tag{158}$$

where the error is defined as

$$\underline{e} = (\underline{b} - \mathcal{A}\underline{x}) \tag{159}$$

then taking the partial derivative of the norm of the error vector with respect to $\underline{x}$, setting it equal to zero, and solving for $\underline{x}$ yields

$$\underline{x} = \left[\mathcal{A}^T\mathcal{A}\right]^{-1}\mathcal{A}^T\underline{b} \tag{160}$$

which is equivalent to the answer given in Equation 157. In general, this form is often termed the pseudo-inverse.

Least Squares Solution

The overdetermined systems defined in Equations 153 and 154 can be shown to have full column rank and therefore yield the least squares solution. To prove the property of full column rank consider the following square matrix

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & x_1^2 & x_1^4 & x_1^6 \\ 1 & x_2^2 & x_2^4 & x_2^6 \\ 1 & x_3^2 & x_3^4 & x_3^6 \end{bmatrix}. \tag{161}$$

This is a matrix with a similar form as that given in Equation 151 but with some of the normalizations removed from each column. This matrix is similar to the Vandermonde matrix used to fit a polynomial to sampled data [8]. We will seek to prove that this square matrix has full rank and then extend the results to the overdetermined system in Equation 153. To prove that the matrix has full rank, it will be shown that the determinant of the matrix is not zero provided that $x_i \neq x_j$ for i $\neq$ j and $x_i \neq 0$ for all i.

The first step is to perform the following column operations $-x_1^2 * C_{n-1} + C_n$, $-x_1^2 * C_{n-2} + C_{n-1}$, ..., $-x_1^2 * C_2 + C_3$ which yield

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & x_1^2 & 0 & 0 \\ 1 & x_2^2 & -x_1^2x_2^2 + x_2^4 & -x_1^2x_2^4 + x_2^6 \\ 1 & x_3^2 & -x_1^2x_3^2 + x_3^4 & -x_1^2x_3^4 + x_3^6 \end{bmatrix} \tag{162}$$

and grouping terms yields

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & x_1^2 & 0 & 0 \\ 1 & x_2^2 & x_2^2(x_2^2 - x_1^2) & x_2^4(x_2^2 - x_1^2) \\ 1 & x_3^2 & x_3^2(x_3^2 - x_1^2) & x_3^4(x_3^2 - x_1^2) \end{bmatrix} . \tag{163}$$

Next, perform the following column operations $-x_2^2 * C_{n-1} + C_n$, $-x_2^2 * C_{n-2} + C_{n-1}$, ..., $-x_2^2 * C_3 + C_4$ which yield

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & x_1^2 & 0 & 0 \\ 1 & x_2^2 & x_2^2(-x_1^2 + x_2^2) & 0 \\ 1 & x_3^2 & x_3^2(-x_1^2 + x_3^2) & -x_2^2(x_3^2(-x_1^2 + x_3^2)) + x_3^4(-x_1^2 + x_3^2) \end{bmatrix} . \tag{164}$$

and grouping terms yields

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & x_1^2 & 0 & 0 \\ 1 & x_2^2 & x_2^2(x_2^2 - x_1^2) & 0 \\ 1 & x_3^2 & x_3^2(x_3^2 - x_1^2) & x_3^2(x_3^2 - x_1^2)(x_3^2 - x_2^2) \end{bmatrix} . \tag{165}$$

Since Equation 165 is lower triangular, the resulting determinant is the product of the diagonal elements which yields

$$det = x_1^2x_2^2x_3^2(x_2^2 - x_1^2)(x_3^2 - x_2^2)(x_3^2 - x_1^2) . \tag{166}$$

Equation 166 states that the matrix has full rank, or the determinant is nonzero provided that $x_i \neq x_j$ for i ≠ j and $x_i \neq 0$ for all i.

This approach can be applied to any square matrix (NxN) of the form

$$\begin{bmatrix} 1 & 0 & \dots & 0 \\ 1 & x_1^2 & \dots & x_1^{2N} \\ \vdots & \vdots & \vdots & \vdots \\ 1 & x_{N-1}^2 & \dots & x_{N-1}^{2N} \end{bmatrix} \tag{167}$$

yielding the following expression for the determinant

$$det = \prod_k x_k^2 \prod_{i,j} (x_i^2 - x_j^2) \text{ for i > j and i} \neq \text{j}, \ \ i,j,k = 1,\dots,N \ . \tag{168}$$

This approach can also be extended to the following matrix form

$$\begin{bmatrix} x_1 & x_1^3 & \dots & x_1^{2N-1} \\ x_2 & x_2^3 & \dots & x_2^{2N-1} \\ \vdots & \vdots & \vdots & \vdots \\ x_N & x_N^3 & \dots & x_N^{2N-1} \end{bmatrix} \tag{169}$$

where the resulting determinant can be expressed as

$$det = \prod_k x_k \prod_{i,j} (x_i^2 - x_j^2) \text{ for i > j and i} \neq \text{j}, \ \ i,j,k = 1,\dots,N \ . \tag{170}$$

Now, since the square matrices in Equations 167 and 169 have been shown to have full rank provided the elements x_i are not zero or equal, then one can add additional rows to these matrices while maintaining full *column* rank and yielding an overdetermined system defined by

$$\begin{bmatrix} 1 & 0 & \dots & 0 \\ 1 & x_1^2 & \dots & x_1^{2N} \\ \vdots & \vdots & \vdots & \vdots \\ 1 & x_{N-1}^2 & \dots & x_{N-1}^{2N} \\ 1 & x_N^2 & \dots & x_N^{2N} \\ \vdots & \vdots & \vdots & \vdots \\ 1 & x_{N_OD}^2 & \dots & x_{N_OD}^{2N} \end{bmatrix} \tag{171}$$

and

$$\begin{bmatrix} x_1 & x_1^3 & \dots & x_1^{2N-1} \\ x_2 & x_2^3 & \dots & x_2^{2N-1} \\ \vdots & \vdots & \vdots & \vdots \\ x_N & x_N^3 & \dots & x_N^{2N-1} \\ x_{N+1} & x_{N+1}^3 & \dots & x_{N+1}^{2N-1} \\ \vdots & \vdots & \vdots & \vdots \\ x_{N_O D} & x_{N_O D}^3 & \dots & x_{N_O D}^{2N-1} \end{bmatrix} . \tag{172}$$

Now the columns of Equations 171 and 172 can be multiplied by nonzero scale factors without changing the rank of the matrix. With the appropriately chosen scale factors, Equations 171 and 172 can be transformed into Equations 153 and 154 without changing the rank of the matrices. Therefore, it has been shown that the overdetermined systems in Equations 153 and 154 have full column rank, and therefore, the solution to the overdetermined system will be the least squares solution.

Overdetermined Variance Estimators

As previously stated, autocorrelation based spectral moment estimators are a function of the estimated autocorrelation function which is obtained from samples of a random process. The quality of the autocorrelation function estimates therefore directly affects the performance of the moment estimators. The overdetermined systems in Equations 153 and 154 will be shown to improve autocorrelation based spectral moment estimator performance. To focus in on how this improvement in performance might be achieved, the overdetermined system will be applied to several of the variance estimators obtained from a truncation of Passarelli's series expansion for the case of a Gaussian shaped spectrum.

The variance estimators investigated by Passarelli [26] were obtained from closed systems formed from a truncation of the series expansion. In investigating variance estimator performance, Passarelli assumed a Gaussian shaped spectrum due to its applicability to meteorological returns. The corresponding magnitude function associated with a Gaussian shaped autocorrelation function can be defined as

$$R(k) = \frac{P_0}{2\pi} \exp\left(\frac{-\sigma^2 k^2}{2}\right) \tag{173}$$

where P_0 is the total power in the spectrum and σ is the spectral width. The four variance estimators defined by Passarelli in Chapter III can be applied to samples of this function to yield performance bounds in the presence of complete knowledge of the autocorrelation function. Figure 22 is a plot the normalized error in the variance estimate as a function of normalized spectral width when applied to samples of the autocorrelation function defined in Equation 173. The normalized error is defined as the absolute value of the difference between the estimated variance and the true value which is then normalized by the true value.

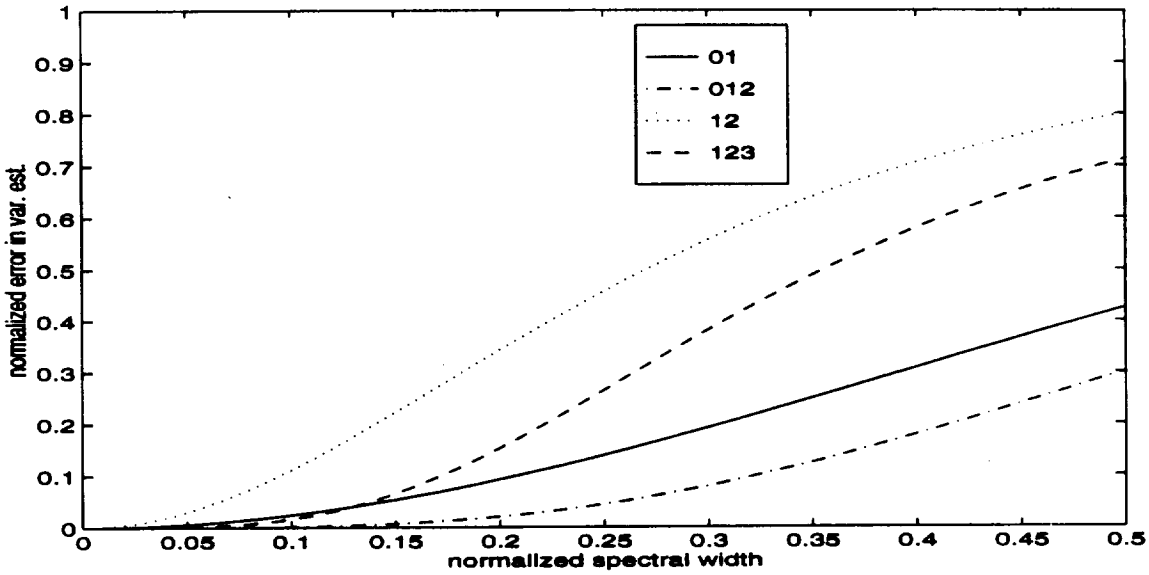


Figure 22. The normalized error in the variance estimate using estimators [01], [12], [012], and [123] and assuming a Gaussian shaped autocorrelation function.

As seen in Figure 22, the four variance estimators yield their minimum error at the smaller spectral widths. The [0 1 2] estimator, which includes the zeroth lag and contains the largest number of terms in the series expansion, outperforms the other estimators over the entire range of spectral widths. This performance is due to the fact that more terms in the series expansion allow for a better representation of the Gaussian shaped autocorrelation function. At the larger spectral widths, the higher

order derivatives of the autocorrelation function are much larger than at the smaller spectral widths, and therefore, the McLauren series expansion converges more slowly and requires more terms in the expansion as noted by Passarelli [26]. This explains the degradation in performance as the spectral width increases.

However, one never obtains an unperturbed estimate of the autocorrelation function. The bias and variance in the autocorrelation lag estimate results in a perturbation in the true shape of the autocorrelation function. In an effort to represent the effects of perturbations in the autocorrelation function estimate, a random term is added to the true Gaussian autocorrelation function to yield

$$R(k) = \frac{P_0}{2\pi} \exp(\frac{-\sigma^2 k^2}{2}) + \alpha\eta(k) \tag{174}$$

where η is uniformly distributed between $\frac{-1}{2}$ and $\frac{1}{2}$, and the perturbation factor α is used to scale the standard deviation of the random term. The standard deviation of $\eta(k)$ is therefore $\alpha\sqrt{\frac{1}{12}}$. This form of the autocorrelation function is used for illustrative purposes only. In the next section, analysis will be performed using estimates of the autocorrelation function obtained from simulated Doppler weather radar returns embedded in white noise.

Figure 23, shows a comparison of the unperturbed and perturbed Gaussian autocorrelation functions for normalized spectral widths of 0.01 and 0.2, where π is normalized to one. The perturbation factor, α, equals 0.01 in this example. Again, the perturbation factor is a relative number used for illustrative purposes only. As seen in the figure, a change in the slope caused by a bounded size perturbation (between $\frac{-\alpha}{2}$ to $\frac{\alpha}{2}$) is larger for the the autocorrelation function associated with the smaller spectral width where the unperturbed slope is small.

Figure 24 is plot of the normalized error in the four variance estimators using a perturbation factor of 0.005 and averaged over 30 iterations. As seen in this plot, the four estimators exhibit a large error in the estimate at the smaller spectral widths. However, at the larger spectral widths only a slight change in the normalized error is

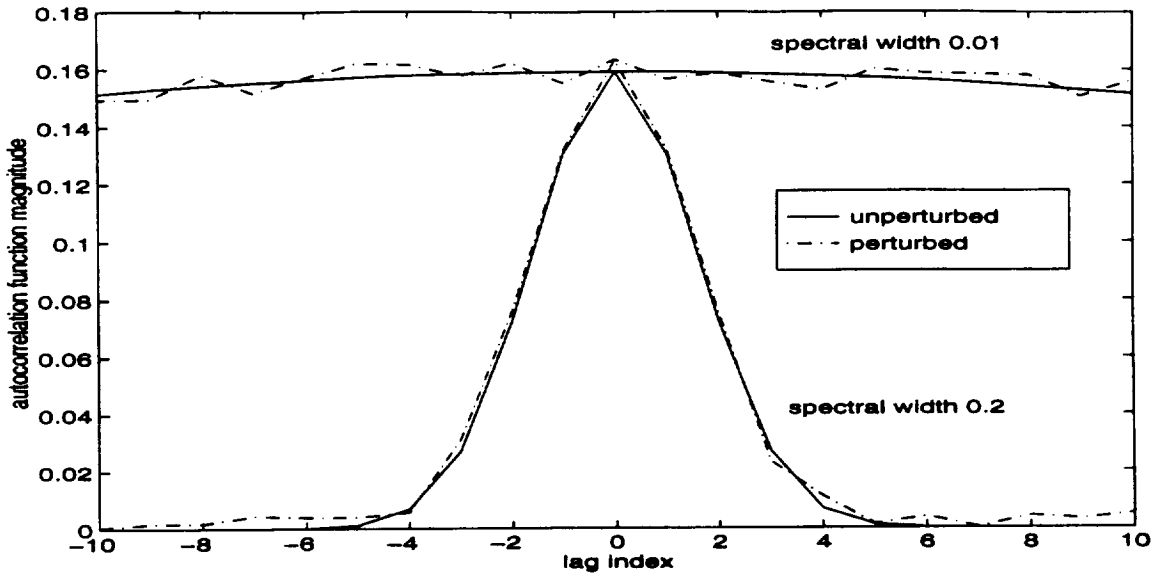


Figure 23. A plot of the unperturbed and perturbed Gaussian shaped autocorrelation functions with normalized spectral widths of 0.01 and 0.2.

observed over that seen in Figure 22. This is due to the fact that the perturbation has less of an effect on the slope of the autocorrelation function over the lags of interest.

The following analysis focuses on the narrow band spectral width region. There are several reasons for focusing in on variance estimators operating in this region. Note that in Figure 22 the closed systems using only a few terms in the series expansion yield their best estimates of the variance at the smaller spectral widths. Therefore, from a closed system point of view, the number of autocorrelation estimates needed to obtain acceptable performance is small. "Acceptable" performance is a loose term that is dependent upon the application and how well one needs to estimate the variance. Assuming that a low order closed system will yield some degree of acceptable performance, then if an overdetermined system is to be applied, the minimum number of equations needed is equal to the number required by the closed system. From an implementation standpoint, one would like to use additional lags but at the same time place a limit on the number in order to keep

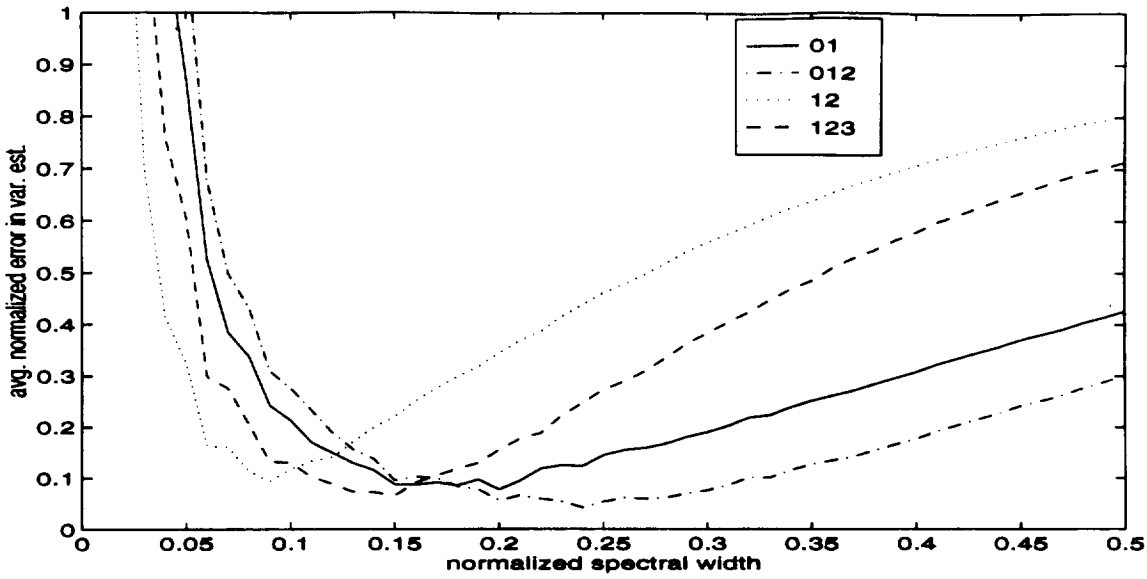


Figure 24. The normalized error in the variance estimate using estimators [01], [12], [012], and [123] and assuming a Gaussian shaped autocorrelation function with a perturbation factor of 0.005 .

computational requirements to a minimum. In contrast, the larger spectral width regions require more terms in the series expansion for convergence and therefore a larger closed system estimator.

Additional autocorrelation function estimates at the higher order lags are applied assuming that lags away from the center of the autocorrelation function also contain information which can be used by the estimator. The rate at which the autocorrelation function falls off is an indicator of which lags may contain useful information, and based on the Fourier uncertainty principle, narrowband spectra exhibit a slow roll-off of the autocorrelation function which will be exploited in this case to extract additional information. Finally, the spectral width of a process may effect the performance of other moment estimators. The pulse-pair mean estimator is known to exhibit a bias due to large spectral widths [40]. In this chapter, variance estimators are applied to Gaussian shaped power spectral densities with

spectral widths (standard deviations) less than or equal to a normalized width of 0.2 where π is normalized to 1.

Figure 25 is a plot of the normalized error as a function of the number of equations (or lags) used in the overdetermined [0 1] variance estimator assuming that a perfect estimate of the Gaussian autocorrelation function is available. The overdetermined system is applied over the region of normalized spectral widths from 0.01 to 0.2 in steps of 0.01 as seen in the figure legends. As a reference point, the smallest spectral width is represented by the bottom curve in each of the figures. Each curve represents a bound on the obtainable performance of the overdetermined [0 1] variance estimator. Figure 25 shows that the closed system, the first point on each curve, yields the smallest error in the variance estimate when the autocorrelation estimate is unperturbed. The rate at which the normalized error increases with respect to the number of equations in the overdetermined system is proportional to the spectral width. The [0 1] estimator is a good approximation to the Gaussian autocorrelation function at the lower spectral widths. Therefore, the overdetermined system is able to apply a least squares fit of the model to a range of autocorrelation lags while maintaining less than a 20% normalized error in the variance estimate. A 20% normalized error in the variance estimate corresponds to approximately a 10% normalized error in the width estimate (the square root of the variance estimate). In the absence of an unbiased estimator, the degree of acceptable bias over which meaningful information may be obtained is very application specific.

Figure 26 is a plot of the normalized error in the [0 1 2] variance estimator applied as an overdetermined system. As noted in Figure 22, the [0 1 2] estimator is a better estimator of the variance at the smaller spectral widths than the [0 1] estimator. In Figure 26, the normalized error is reduced over the entire range of spectral widths as compared to that observed in Figure 25 for the [0 1] overdetermined variance estimator. Figures 27 and 28 are the performance bounds for [1 2]

and [1 2 3] overdetermined variance estimators using the Gaussian shaped autocorrelation function. The performance can be characterized in terms of an initial bias due to the closed system and the slope of the performance curve. The more terms used in the series expansion, the smaller the slope of the performance curves. One might assume that the overdetermined system employing a large number of terms in the series expansion would always yield the best estatimator. In the following sections, this will be shown not to be the case for all spectral widths.

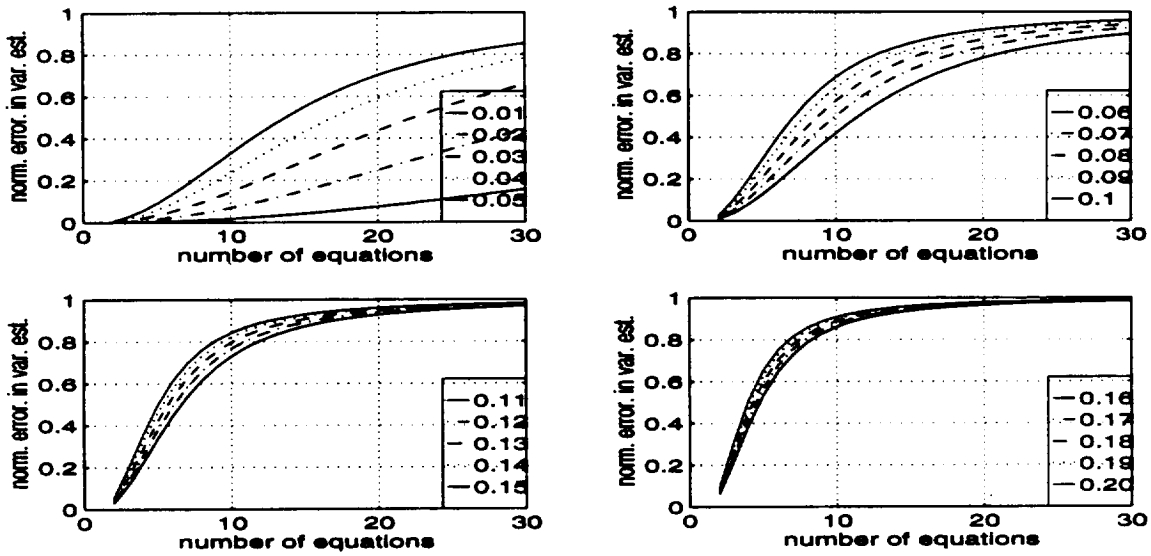


Figure 25. The normalized error in the variance estimate using the [0 1] estimate in an overdetermined system for a Gaussian shaped autocorrelation function.

Figure 29 is a plot of the [0 1] overdetermined variance estimator applied to the exact autocorrelation function in Equation 173 and the perturbed autocorrelation function with a perturbation factor of 0.01. For normalized spectral widths less than 0.15, the perturbation of the autocorrelation function has a significant impact on the performance of the closed system (the initial point on each of the curves). However, as more equations are used in the overdetermined system the averaged normalized error is reduced. The least squares fit minimizes the effects of the perturbations

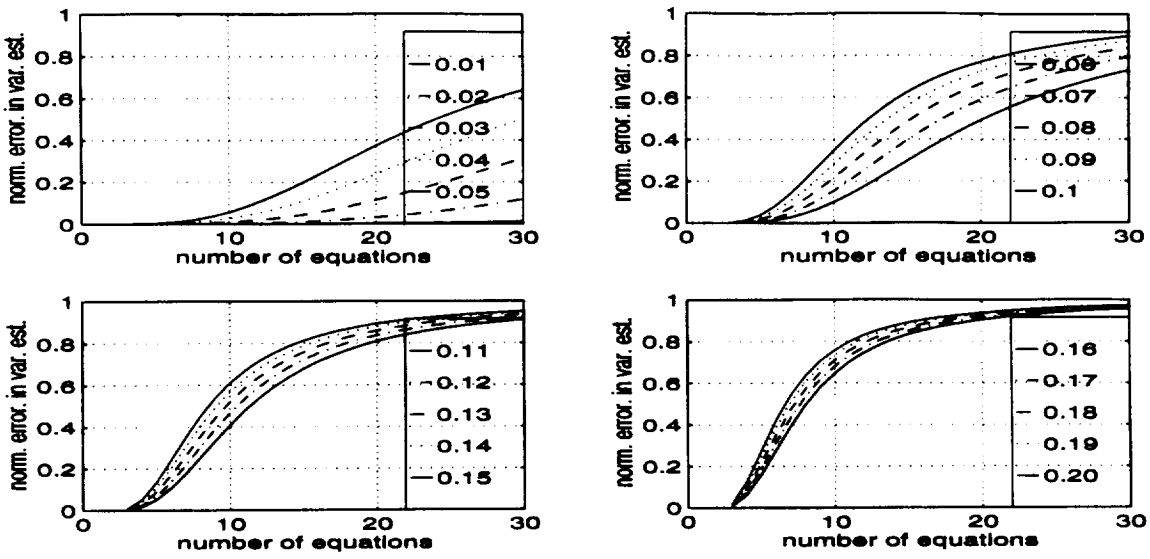


Figure 26. The normalized error in the variance estimate using the [0 1 2] estimate in an overdetermined system for a Gaussian shaped autocorrelation function.

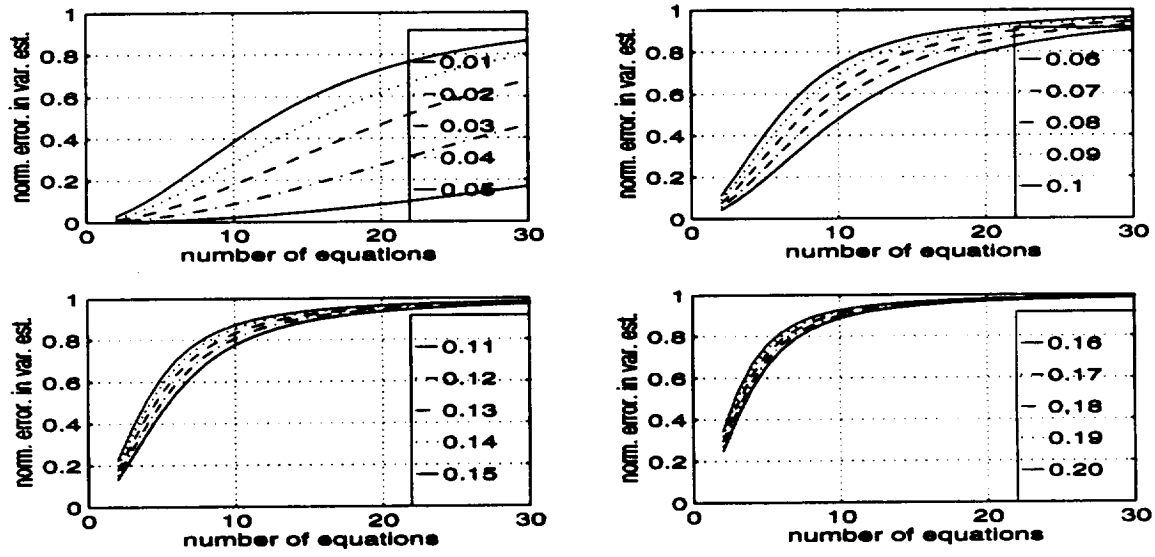


Figure 27. The normalized error in the variance estimate using the [1 2] estimate in an overdetermined system for a Gaussian shaped autocorrelation function.

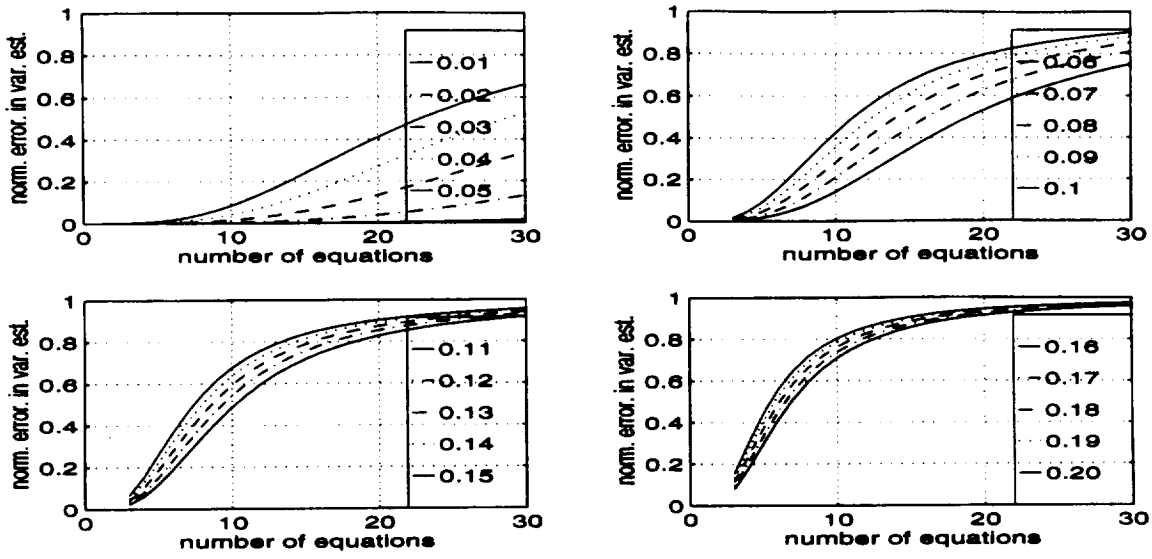


Figure 28. The normalized error in the variance estimate using the [1 2 3] estimate in an overdetermined system for an Gaussian shaped autocorrelation function.

as more autocorrelation lags are used in the overdetermined system. There is a point where the performance of the estimator reaches that achievable for the overdetermined variance estimator applied to the true Gaussian shaped autocorrelation function. The number of equations required to best fit the observed data to the estimator is a function of the spectral widths. However, in the spectral width region from 0.05 to 0.15 only a few additional autocorrelation lags are required to reach the minimum obtainable error. Also, as seen in Figure 29, the number of equations required for the error to reach a minimum tends to group near an optimum number for a given region of spectral widths. The grouping can be explained by the separation and slope of the performance curves for the [0 1] overdetermined variance estimator in Figure 25.

In Figure 30, the overdetermined [0 1 2] variance estimator is applied to the perturbed autocorrelation lag estimates. The number of equations needed for the unperturbed and perturbed curves to intersect has increased over that given in Figure 29. However, the [0 1 2] estimator is a better estimator at the smaller spectral

widths, and therefore, the obtainable normalized error may be lower when applying the overdetermined system. This represents a tradeoff between the size of the closed system and the number of additional lags needed in the overdetermined system to obtain a certain level of performance. Additional lags imply additional computations required by the signal processor. This will be investiaged further using simulated Doppler weather radar returns.

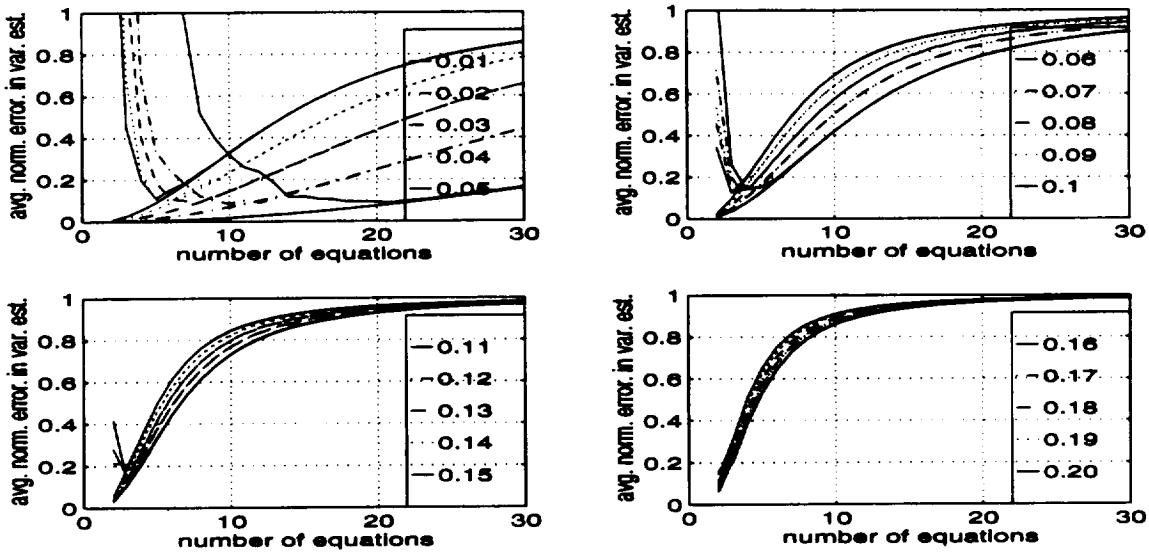


Figure 29. The normalized error in the variance estimate using the [0 1] estimate in an overdetermined system for an Gaussian shaped autocorrelation function with a perturbation of 0.01 .

Doppler Weather Radar Example

The variance estimator is often used in meteorological processing to measure the turbulence associated with an event, and much of the literature has focused on the bias and standard deviation observed in autocorrelation based variance estimators. As seen in Chapter III, autocorrelation based variance estimators are known to exhibit a bias and standard deviation that increases for narrow and wideband spectra as a function of SNR and the number of samples used to estimate the auto-

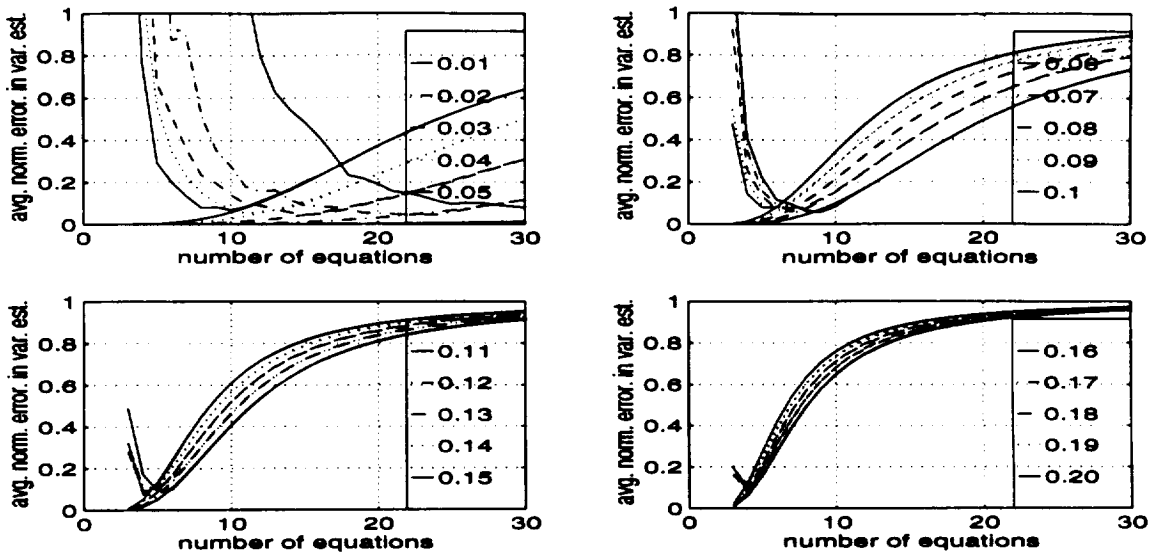


Figure 30. The normalized error in the variance estimate using the [0 1 2] estimate in an overdetermined system for an Gaussian shaped autocorrelation function with a perturbation of 0.01 .

correlation function. This section focuses on the overdetermined system applied to variance estimators which are used to measure turbulence associated with Doppler weather radar returns. The perturbations in the autocorrelation function estimates as noted in the previous section will be introduced into the autocorrelation function through a simulation of Doppler weather radar returns embedded in white noise. The performance of the overdetermined variance estimators will serve to support the case for the possible inclusion of additional autocorrelation lags in estimating other moments in order to improve performance.

As in the previous section, the analysis will be limited to the narrow band case and an assumed Gaussian shaped power spectrum. The narrowband case is important because Zrnic has suggested that weather radar systems should be designed such that the observed normalized spectral width is constrained to be less than 0.25 in order to reduce the bias in the moment estimates [37]. This can be achieved through the proper selection of the pulse repetition rate (i.e. sampling rate). As

previously noted, Passarelli's expansion is not based on any assumed form of the power spectrum or corresponding autocorrelation function. However, the data used in this analysis assumes a Gaussian shaped spectrum as is often observed in meteorological returns [9] and as is often used throughout the literature for evaluating estimator performance. The data is simulated using the approach given by Zrnic and Bumgarner [39, 32] (Appendix B).

As with any autocorrelation based or Fourier based moment estimator, the estimator can only approach it's optimum performance in the absence of noise and clutter (returns from objects not associated with the target of interest). Before applying any autocorrelation based moment estimator, it is necessary to reduce the noise through correlation subtraction or filtering in the frequency domain. For this analysis, the noise is assumed to be white, and an a priori average noise power level will be subtracted from the zeroth autocorrelation lag in order to remove the estimated noise power. In radar systems, the average noise power may be estimated during periods when the radar is not transmitting or receiving.

In this analysis, the random process is assumed to be stationary or at least locally stationary and ergodic in the autocorrelation over the time of observation. A pulsed Doppler radar collects tens to hundreds of range samples from spatial cells in a small time interval over which the process is assumed to be stationary. Therefore, the autocorrelation lag estimates in this analysis are based on 100 complex data samples using the asymptotically unbiased autocorrelation lag estimator defined in Appendix A. As a Monte Carlo measure of performance, the results presented in this dissertation are based on estimates averaged over 40 independent iterations.

A key result of Passarelli's analysis is that a single variance estimator is not optimum over the entire range of spectral widths. Passarelli formed a closed system of equations for four variance estimators using the following sets of lags (0,1), (0,1,2), (1,2), and (1,2,3). The bias exhibited in Chapter III for these autocorrelation based variance estimators warrants the application of additional lags in an

attempt to improve estimator performance in the narrow spectrum width region. This analysis will also lead to a better understanding of the information contained in the added lag values of the autocorrelation function and to how they may used by an overdetermined system to increase estimator performance.

Reduction in Bias

The performance of the variance estimators evaluated in this section is measured in terms of the bias (or normalized error) and the standard deviation in the estimate. Figure 31 shows the averaged normalized error in the variance estimate as a function of the number of equations used in the overdetermined system. This system is defined by the truncated autocorrelation series expansion in Equation 153 using only two terms, [0 1], (this is the traditional pulse-pair variance estimator). In this figure, the signal-to-noise ratio (SNR) is 5 dB. The initial estimate in all the plots (at N equations) represents Passarelli's results for a closed system of size (N x N). In each sub-figure, the plot contains five curves each representing the estimates for a given normalized spectral width ranging from 0.01 to 0.2 as given in the figure legends.

As seen in Figure 31, the number of equations required to obtain the minimum normalized error increases as the spectral width decreases. This can be explained by the following discussion based on the rate of series convergence and the Fourier uncertainty principle. The series coefficients of the autocorrelation function expansion are defined as the derivatives of the magnitude of the autocorrelation function evaluated at zero. Therefore, the information is contained in each lag of the autocorrelation function to some degree as noted by the series expansion of $h(k)$ in terms of the coefficients. The autocorrelation function falls off at a rate that is inversely proportional to the spectral width as explained by the Fourier uncertainty principle, and so for the narrowband case, these overdetermined systems are attempting to estimate the value of a small number (the derivative or slope of the autocorrelation

function at zero) from the corrupted autocorrelation lag estimates. The additional equations required at the smaller spectral widths are needed by the overdetermined system to extract the information from the perturbed autocorrelation lags whose slope (information content) may be significantly changed between two samples from that given in the true autocorrelation function. At the higher spectral widths the slope of the autocorrelation function is much steeper and the perturbations in the lag estimates have less of an effect on the slope of the autocorrelation function.

From Figure 31, it is evident that the variance estimates for the spectral widths less than or equal to 0.15 benefit from the use of additional lags. The knee or minimum in the curve is representative of the overdetermined variance estimator having extracted all the information it can from the available data. It is observed that the number of additional equations needed to obtain the minimum normalized error tends to group over a range of spectral widths. In the region of normalized spectral widths from 0.03 to 0.05, the number of equations needed to obtain the minimum normalized error is seven to eleven. In the range of normalized spectral widths from 0.06 to 1.0, the number is three to six, and in the range from 0.11 to 0.15, the number of equations is three to four. In comparing Figures 5.13 and 5.7, the tendency for the minimum errors to group around a common number of equations over a region of spectral widths is due to the slope and spread of the performance curves for the overdetermined system. If the range of observable spectral widths could be known a priori, then an average number of equations needed in that region to reach the minimum error could be applied to improve estimator performance. The overdetermined system is able to extend the [0 1] estimator over a larger region of spectral widths than would have been possible using just the closed system estimator (the initial point on each curve).

One approach in a real system might be to use the overdetermined system to get a first cut estimate of the region in which one is operating and then apply a more optimal estimator for that region which may be derived from either a closed or

overdetermined system. Passarelli noted that to use the closed system variance estimators in an optimal fashion, one would need to either estimate the spectral width region in which one is operating or have a priori knowledge of the expected spectral widths. An overdetermined system can allow operation over a larger region of spectral widths by reducing the observed bias and the observed standard deviation as will be seen in the following section. Also, for implementation purposes, it should be noted that the matrix in Equation 153 does not depend on the estimated autocorrelation lags, and therefore, the pseudo inverse, $[\mathcal{A}^T\mathcal{A}]^{-1}\mathcal{A}^T$, can be computed off-line and stored for use in a real-time environment.

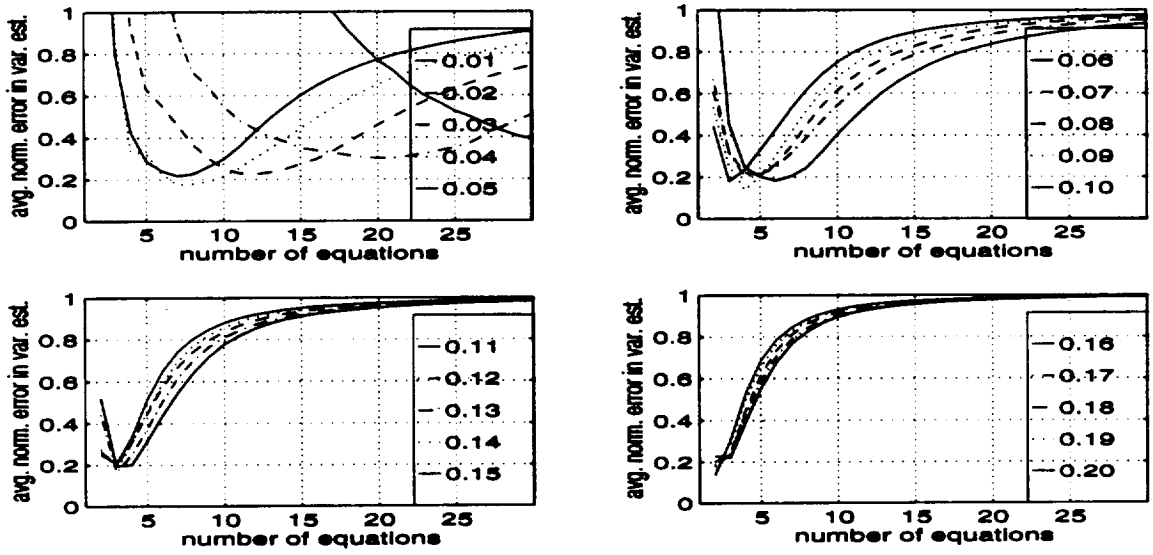


Figure 31. The averaged normalized error in the variance estimate for the 5 dB SNR case using two terms in the overdetermined system.

Passarelli's [26] work shows that estimators derived from different size closed systems are optimal over different spectral width regions. The effect of more terms in the series expansion can be also be evaluated for the overdetermined system. Figure 32 is a plot of the averaged normalized error in the variance estimate for the

case where three terms in the expansion, [0 1 2], are used to form the overdetermined system with a SNR of 5 dB. One can see that for spectral widths between 0.11 and 0.2, the use of additional terms in conjunction with an overdetermined system results in a reduction in estimate error over that in Figure 31. The additional terms, however, require a higher order overdetermined system to reach this minimum error. This is due to the fact that the least squares solution must satisfy the minimum error constraint using more terms in the series expansion. However, more terms do not produce minimum errors that are measurably better in the region of spectral widths from 0.01 to 0.1, and more equations are required for the same performance. For the spectral widths in the region 0.11 to 0.2, the addition of more terms provides a better model of the autocorrelation function and when used in conjunction with the overdetermined system yields less estimate error in this spectral width region.

These two examples (Figures 31 and 32) have shown that the best model for the autocorrelation function in a particular spectral width region obtained from Passarelli's expansion can be used in conjunction with additional autocorrelation lags in an overdetermined system to improve estimator performance. However, the "best" model is not the same for all spectral widths as noted by Passarelli. The [0 1 2] estimator is a better estimator at the larger spectral widths where the slope of the performance curve (in Figure 26) permits more equations to be used in trying to obtain the minimum error or the most information extraction.

Reducing the Effects of a Low SNR

The overdetermined system can be used to reduce the effects of operating at a lower SNR when estimating the autocorrelation lags from a fixed number of samples. Figure 33 is a plot of the average normalized error in the variance estimate for a SNR of 0 dB using an overdetermined system containing two terms in the series expansion, [0 1]. As previously stated, an average white noise power estimate has been subtracted from the zeroth autocorrelation lag, but the minimum obtainable error

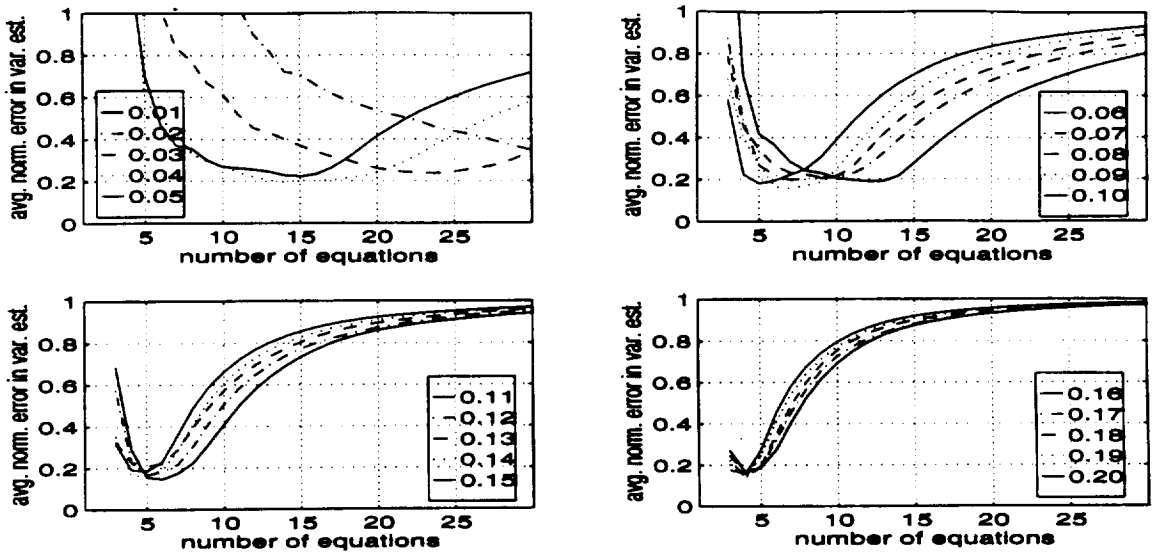


Figure 32. The averaged normalized error in the variance estimate for the 5 dB SNR case using three terms in the overdetermined system.

has increased for each of the spectral widths due to an increase in the variance of the estimated autocorrelation function. This increase in the variance of the estimated autocorrelation function could be reduced by using additional samples in computing the autocorrelation function, but in pulsed Doppler radar systems confined by the pulse repetition rate and the observation time, this may not be an option. Even though the estimate error has increased in all spectral width regions, the minimum error still occurs near the same number of equations as that in Figure 31. In the 0.15 to 0.2 spectral width region, additional equations (lags) are needed to reduce the estimate error as compared to Figure 31 where the minimum was reached for the closed system case (NxN). Even though the minimum errors reached in Figure 31 are not obtained here, the overdetermined systems are still able to significantly decrease the error over the (2 x 2) closed system (the initial point on each curve). In terms of the performance bound in Figure 25, the noise produces sufficient degradation in the autocorrelation function estimates such that the overdetermined system is not able to compensate and extract as much information.

Figure 34 is a plot of the average normalized error for the 0 dB SNR case using three terms, [0 1 2], in the series expansion. The additional terms tend to improve performance in the spectral width region from 0.11 to 0.2 over that in Figure 33. The improved performance is attributable to the three terms of the series expansion used in the overdetermined system which provide a better model of the autocorrelation function in this region. In comparing Figure 33 and 34 in the spectral width region from 0.01 to 0.1, the noise has narrowed the performance gap between the two estimators, [0 1] and [0 1 2]. However, the number of equations needed to reach a minimum is larger in the [0 1 2] case than in the [0 1] case since the overdetermined system is fitting the data to a different model containing additional terms in the series expansion.

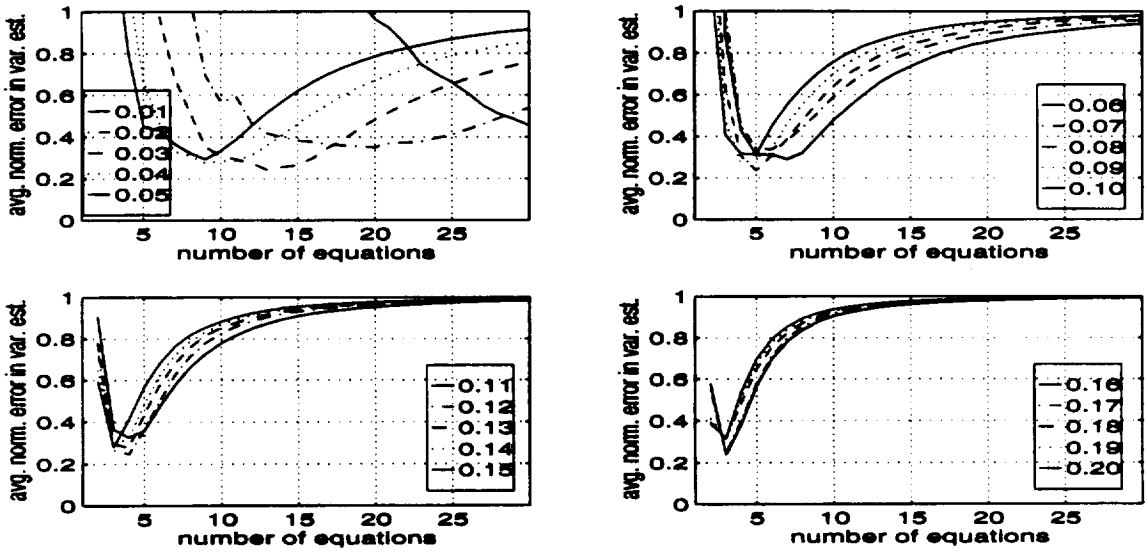


Figure 33. The averaged normalized error in the variance estimate for the 0 dB SNR case using two terms in the overdetermined system.

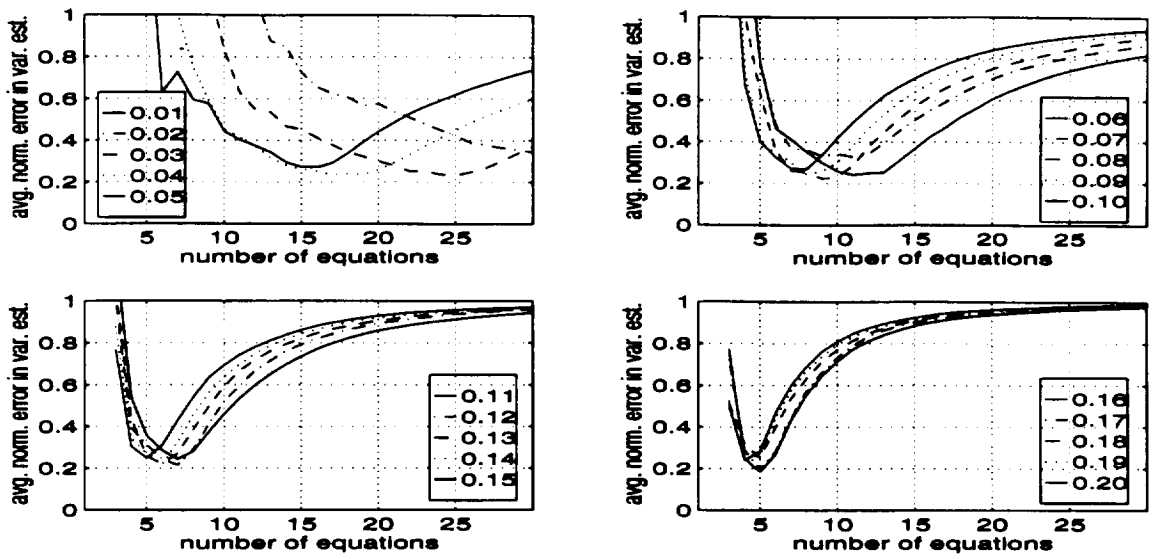


Figure 34. The averaged normalized error in the variance estimate for the 0 dB SNR case using three terms in the overdetermined system.

Gaussian Fit

As a comparison to previous work in fitting an assumed Gaussian autocorrelation function to measured data in order to estimate the variance [1, 27], a least squares fit of a Gaussian shaped autocorrelation function is applied to the data. Figure 35 is a plot of the average normalized error in the variance estimate as a function of the number of lags used in the least squares fit. As one can see from Figure 35, the least squares fit is needed over the entire range of spectral widths to reduce estimator error. The achievable minimum errors in the spectral width region of 0.01 to 0.1 are higher than the those presented in Figures 31 and 32. This can possibly be explained by the fact that the Gaussian model represents an infinite series expansion which may not yield the best estimator (model) for the very narrow widths when applied in an overdetermined system. In the spectral width region from 0.11 to 0.2, the least squares fit yields comparable results to that in Figures 31 and 32. Theoretically, the performance curve for the overdetermined system using perfect knowledge of the autocorrelation function would be flat with zero error over the entire range of

equations applied to the system. However, the overdetermined system must fit what is equivalent to an infinite series expansion to the data. There are effectively more terms that must be estimated. Another important note is the lack of a clustering or grouping of minimum errors around a common number of equations in the 0.01 to 0.1 spectral width region. This would be important in terms of an implementation scheme. Another point that should be made here is that one may not always have a closed form of the autocorrelation function to which to fit the data. Passarelli's expansion is not limited to the Gaussian case and therefore moment estimators based on the overdetermined system is a more robust approach.

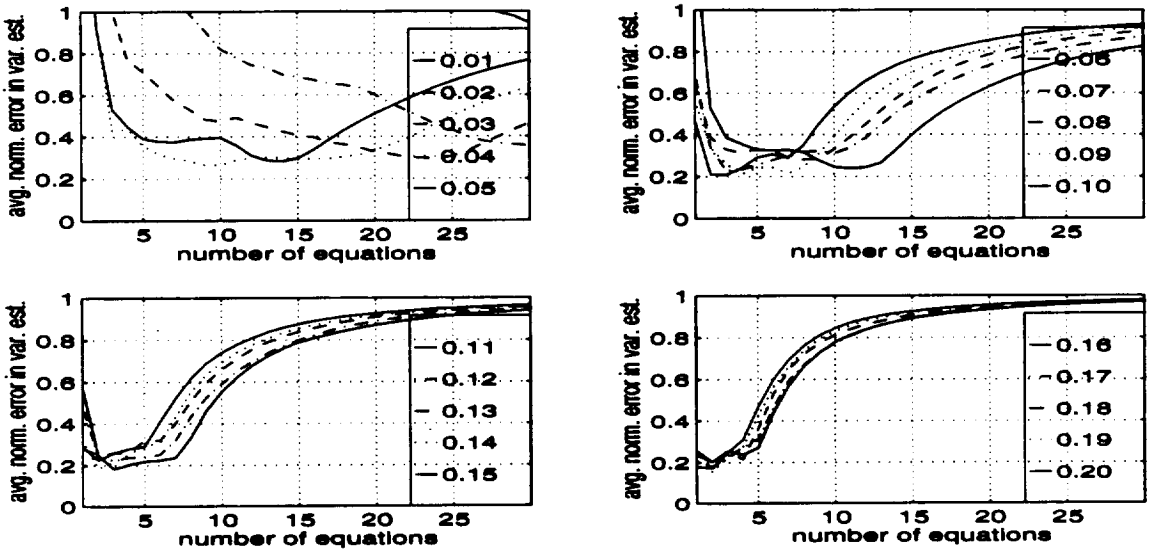


Figure 35. The averaged normalized error in the variance estimate for the 5 dB SNR Gaussian case.

Reduction in Standard Deviation

As seen in the previous section, the use of additional equations in an overdetermined system can reduce the estimator bias (normalized error). In this section, the overdetermined system will be shown to reduce the standard deviation in the

estimate. Figure 36 shows the standard deviation (std) in the variance estimate for the 5 dB case using only two terms in the series expansion. The standard deviation is computed for the variance estimate in terms of the normalized frequency (0 to 1). As seen in Figure 36, the standard deviation is reduced as more equations are used in the estimate. Figure 37 is plot of the standard deviation in the estimate for the 5 dB case using three terms in the series expansion. As Passarelli had expected, the larger closed system with more terms in the series expansion results in a larger standard deviation in the estimate. But note that the use of additional equations in the overdetermined system results in a decrease in the standard deviation over that initially observed for the (N x N) closed system. Figures 38 and 39 are plots of the standard deviation for the 0 dB SNR case with two and three terms in the autocorrelation expansion, respectively. As expected, the decrease in SNR results in an increase in the standard deviation which can be reduced by the overdetermined system. The reduction in the standard deviation is most critical at the smaller spectral widths where the standard deviation may be on the order of the quantity being estimated.

The Zeroth Lag

In each of the overdetermined systems presented thus far, the zeroth lag was included. However, Passarelli has shown that for certain spectral widths and a low SNR, the omission of the zeroth lag in the closed system may be of benefit. Remember that in the white noise case, the zeroth lag contains the noise power, and an a priori estimate of the noise power must be subtracted from the zeroth lag which may introduce additional errors in the estimate. Figure 40 is a plot of the normalized error in the variance estimate for the 5 dB SNR case using two terms in the series expansion starting with lag one. The minimum normalized error in the spectral width regions from 0.01 to 0.15 is similar to that observed in Figure 31 which included the zeroth lag in the overdetermined system. Therefore, the [1 2] estimator in the

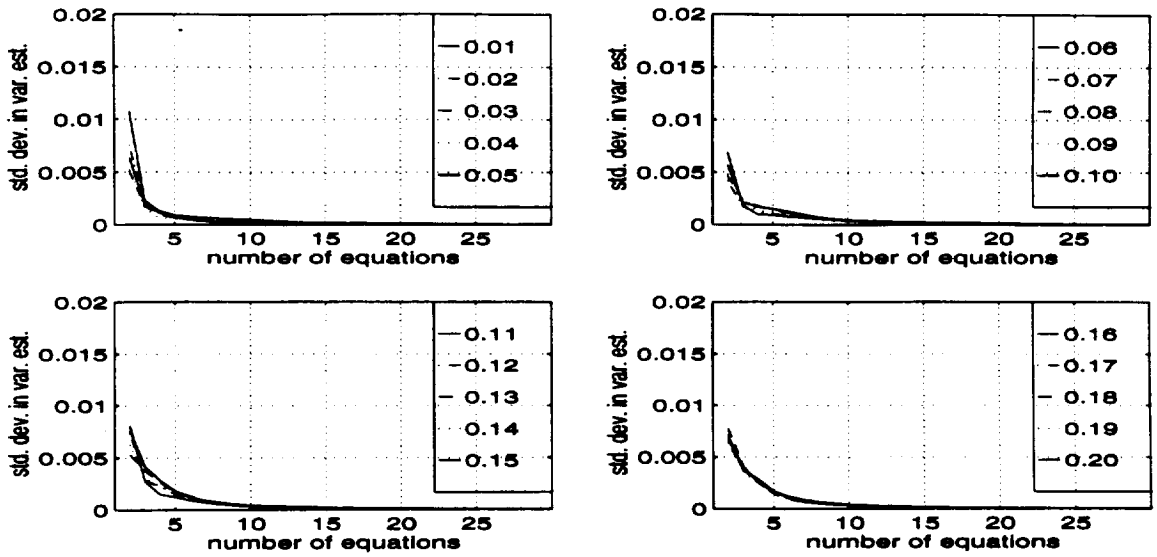


Figure 36. Standard deviation in the variance estimate for the 5 dB SNR case using two terms in the overdetermined system.

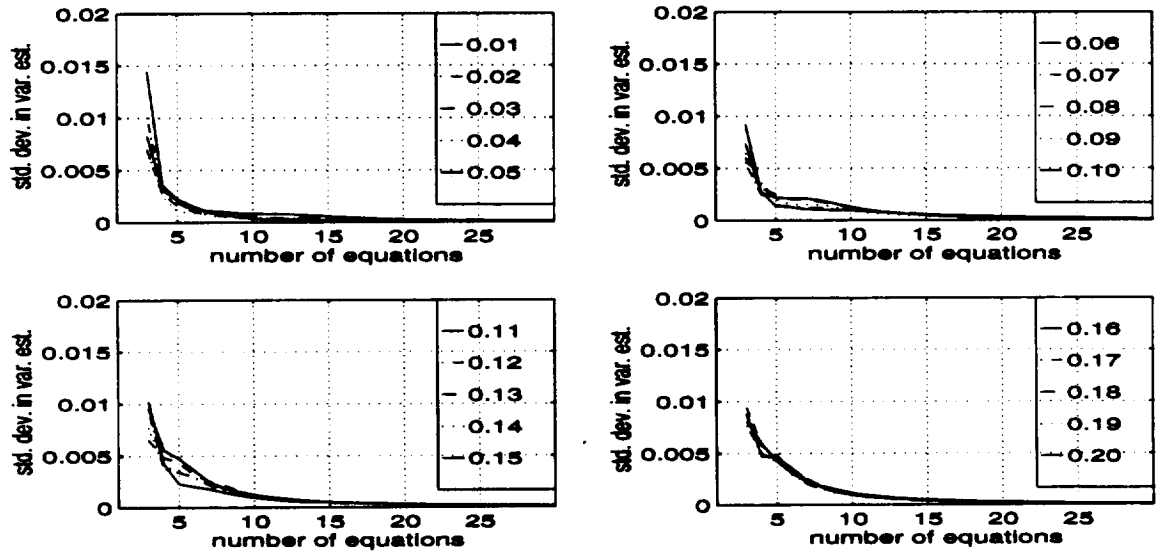


Figure 37. Standard deviation in the variance estimate for the 5 dB SNR case using three terms in the overdetermined system.

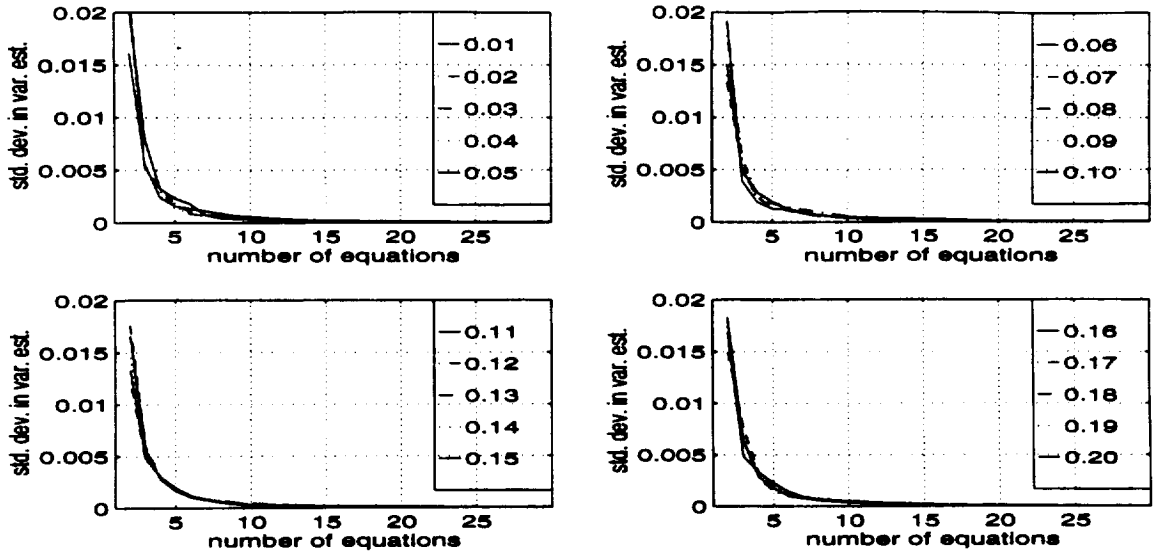


Figure 38. Standard deviation in the variance estimate for the 0 dB SNR case using two terms in the overdetermined system.

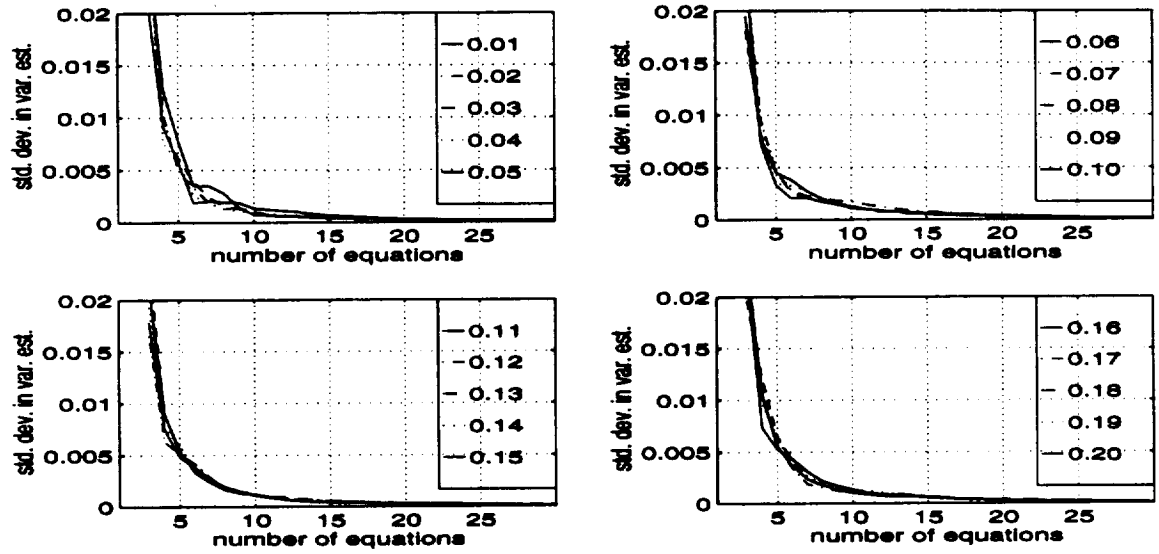


Figure 39. Standard deviation in the variance estimate for the 0 dB SNR case using three terms in the overdetermined system.

overdetermined system yields a result similar to that observed for the [0 1] estimator. This should be expected when comparing the performance bounds seen in Figures 25 and 27 over this region. However, the [0 1] and [1 2] estimator's performance did not track in Figure 21 when the closed system approach was applied over this region. The overdetermined system has improved the performance of the two estimators and caused the two estimators to yield comparable performance over this region. However, in the spectral width region 0.16 to 0.2, the minimum normalized error has increased over the case in Figure 31 containing the zeroth lag. Again, in comparing the performance bounds in Figures 25 and 27, the overdetermined system for the [1 2] estimator can not reach the performance of the [0 1] estimator.

Figure 41 is a plot of the normalized error variance estimate for the 5 dB SNR case using three terms in the series expansion starting with lag one. In this case, the normalized error has been reduced over that observed in Figure 40 in the spectral width region from 0.16 to 0.2 and is comparable to that observed in Figure 32. This analysis reveals that the zeroth lag is important in the region 0.16 to 0.2 when only a few terms are used in the expansion. The zeroth lag contains a major portion of the information in these "medium" width cases. The use of additional terms, however, tends to compensate for the omission of the zeroth lag. In the spectral width region 0.01 to 0.15, the zeroth lag is less important due to the fact that the autocorrelation falls off much slower, and the information is spread over a larger number of autocorrelation lags. In the region 0.01 to 0.1, the performance of the [1 2 3] estimator in an overdetermined system is similar to the [0 1 2] estimator.

Closed System Performance

One might ask the question, "Why choose an overdetermined system instead of a larger closed system?". Figure 42 is a plot of the normalized error in the variance estimate using closed systems of size (NxN) and a SNR of 5 dB. The number of equations as defined by the horizontal axis represents the dimension of the closed

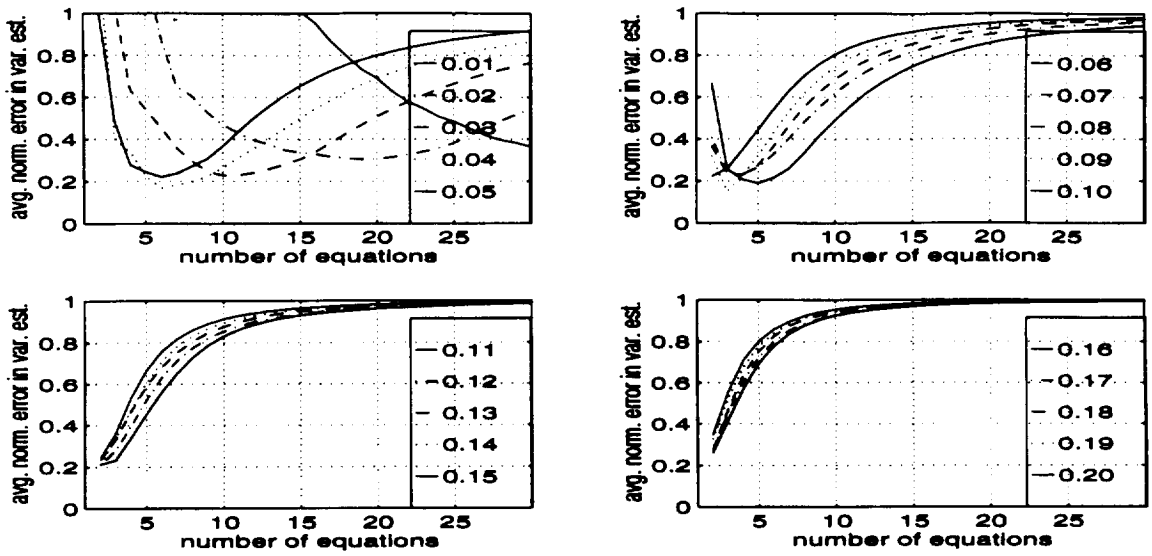


Figure 40. The averaged normalized error in the variance estimate for the 5 dB SNR case using the [1 2] estimator in an overdetermined system.

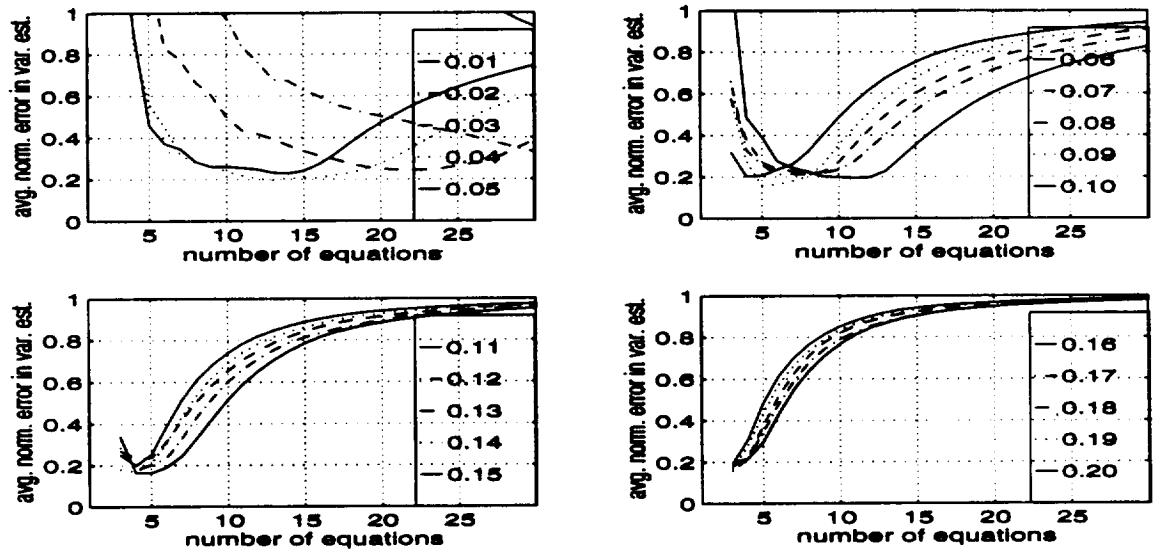


Figure 41. The averaged normalized error in the variance estimate for the 5 dB SNR case using the [1 2 3] estimator in an overdetermined system.

system. As one can observe, the larger closed systems only result in an increase in the normalized error. For the narrowband case, it has been observed that, in general, it is better to use an estimator based on a few terms in Passarelli's expansion in an overdetermined system, than to include additional lags based on more terms in a larger closed system.

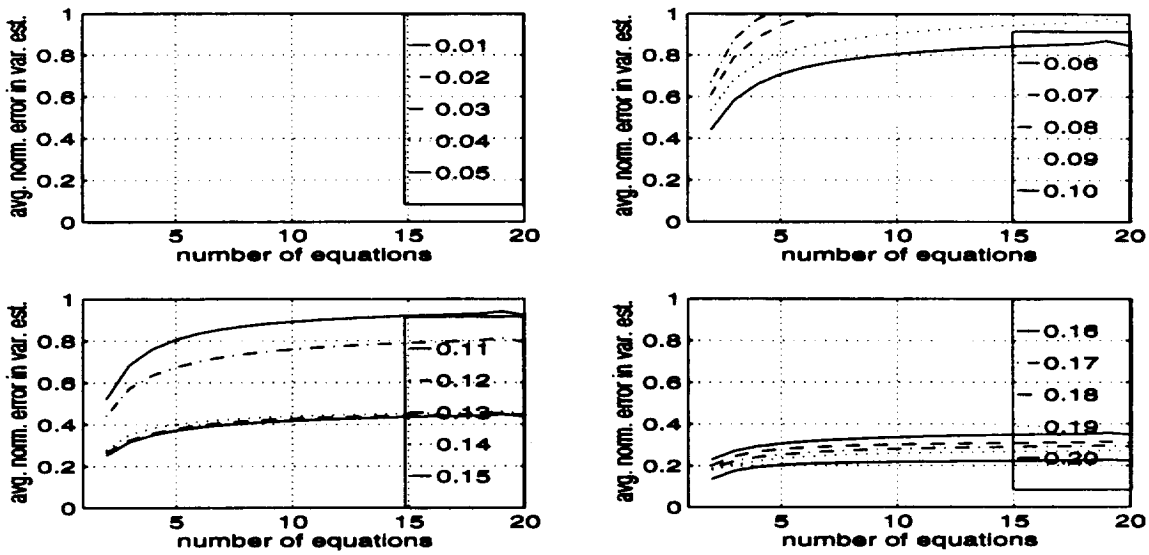


Figure 42. The averaged normalized error in the closed system for the 10 dB SNR case.

CHAPTER V

CONCLUSIONS

Motivation

This research was motivated by the need to improve the performance of autocorrelation based spectral moment estimators. In general, the performance of autocorrelation based spectral moment estimators is limited by the quality of the autocorrelation function estimate. The current suite of autocorrelation based spectral moment estimators is derived using an assumed model of the autocorrelation function, the characteristic function defined in probability theory, and the Fourier transform relationship between the autocorrelation function and the power spectrum. Autocorrelation based spectral moment estimators are important because for certain moments such as the mean and variance, they are shown to exhibit better performance at narrow spectral widths and low signal-to-noise ratios as compared to Fourier based methods. Also, autocorrelation based spectral moment estimators provide a means for reducing the computational load over that of the Fourier based techniques which require an application of the discrete Fourier transform and discrete centroiding techniques.

Passarelli has defined a series expansion of the general complex autocorrelation function which relates central moments of the power spectrum to coefficients in the series expansion. A truncation of Passarelli's series expansion results in a closed system of equations that can be solved for the moment or moments of interest. A comparison of this closed system of equations to the Yule-Walker equations defined in autoregressive spectral estimation is made. Autoregressive spectral estimation is based on the spectral factorization property which states that the power spectrum associated with a random process can be represented as the output of a linear system driven by white noise. The autoregressive model represents the case where the

linear system is assumed to be modeled as an all pole filter. The Yule-Walker equations are obtained from the AR model by computing the second order statistics (the autocorrelation) of the system. The Yule-Walker equations result in a set of linear equations relating the coefficients of the linear system to lags of the autocorrelation function. In an effort to improve the spectral estimate for a given model order, the overdetermined Yule-Walker equations [4] have been defined. The overdetermined Yule-Walker equations are an attempt to extract additional information from the autocorrelation function estimate at higher order lags.

This research focuses on adapting techniques developed in the field of modern power spectral estimation for use in the field of autocorrelation based spectral moment estimation to improve estimator performance. In particular, an overdetermined system in terms of a truncation of Passarelli's series expansion of the complex autocorrelation function can be defined for improving estimator performance similar to that in modern spectral estimation for the Yule Walker equations. Some initial attempts have been made at using estimates of the autocorrelation function at higher order lags to improve autocorrelation based spectral moment estimator performance, but, other than the poly-pulse-pair mean estimator which is used to reduce estimator variance, no rigorous work has been published on the relationship between use of additional lags and estimator performance in an overdetermined system.

Summary of Results and Contributions

This dissertation has defined a framework in which to incorporate additional estimates of the autocorrelation function at higher order lags into autocorrelation based spectral moment estimators in order to improve estimator performance. In particular, this work defines the structure in terms of an overdetermined system applied to a truncation of Passarelli's series expansion for the general complex autocorrelation function.

This work starts by relating the overdetermined system techniques found in the field of modern spectral estimation to the field of autocorrelation based spectral moment estimation in order to improve spectral moment estimator performance. Chapter II is a review of modern spectral estimation techniques. In Chapter II, the overdetermined Yule-Walker equations are defined, and the use of additional autocorrelation lags to improve the power spectrum estimate is presented. In Chapter III, the current suite of autocorrelation based moment estimators is defined including Passarelli's series expansion. The performance of autocorrelation based spectral moment estimators presented in Chapter III is a function of the quality of the autocorrelation function estimate.

An overdetermined system, defined in terms of Passarelli's series expansion, has no solution, but solutions yielding minimum norms can be defined. It is shown here that the overdetermined system defined by a truncated version of Passarelli's expansion yields a matrix having full column rank. This implies that the system of equations can be solved using pseudo-inverse techniques which yields the least squares solution.

The overdetermined system defined in this work is shown to improve autocorrelation based spectral moment estimator performance. There are theoretically an infinite number of moments that could be assessed under this overdetermined framework. However, in fielded systems the spectral moments of primary interest are the zeroth moment (the power), the first moment (the mean), and the second central moment (the variance). The spectral variance estimator is chosen in this analysis because it is particularly vulnerable to bias and large standard deviations over a range of spectral widths in the presence of low SNR's and where a limited number of observations are available. The variance estimate or its square root (the width estimate) is typically used in meteorological processing as a measure of turbulence. The premise here is to show that an overdetermined system can be used to improve the performance of autocorrelation based spectral estimators over estimators derived

from the traditional closed system approach. However, it is not intended here to exhaustively assess all the possibilities.

Even though various power spectral shapes might be considered, an exhaustive assessment is not presented. In this work, the Gaussian shaped power spectrum is chosen for assessing overdetermined variance estimators ability to improve estimator performance. The Gaussian shaped spectrum is found to model Doppler weather radar returns [9].

For the overdetermined variance estimator, it is shown that the narrowband case offers the most opportunity for additional information extraction since the information is spread of over a larger number of autocorrelation lags. This is explained by the Fourier uncertainty principle. The narrowband case is also important as noted by Zrnic in reducing the bias in moment estimates. Based on his analysis, Zrnic has proposed that any fielded system should employ autocorrelation based spectral moment estimators only when the normalized spectral widths are constrained to be less than 0.25π.

The overdetermined variance estimators evaluated in this work are defined by a a truncation of Passarelli's series expansion using different numbers of terms in the expansion. The performance of the estimators is assessed in terms of the overdetermined systems ability to reduce the bias and standard deviation in the estimate over that of the estimator defined by the closed system. This work defines performance bounds for the normalized bias in the variance estimate when overdetermined variance estimators are applied to the case of Gaussian shaped spectra. The slope and separation of the performance curves, as a function of both the number of equations used in the overdetermined system and the spectral width, are relevant in characterizing the behavior of the estimator. In order to assess the performance of the overdetermined variance estimators, simulated Doppler weather radar returns were generated at different spectral widths and signal-to-noise ratios.

Especially at the narrow and wide spectral widths, the bias in the closed system variance estimators is shown by Passarelli to be a function of quality of the estimate of the autocorrelation function. Here it is shown that the overdetermined systems could be used to reduce the bias in the variance estimate by incorporating additional estimates of the autocorrelation function at higher order lags. The performance of the overdetermined variance estimators applied to the simulated weather radar returns is shown to be bound by the performance curves derived from perfect knowledge of the autocorrelation function. The slope of the performance curves is an indicator of how quickly the truncated series expansion degrades in modeling the true autocorrelation function. Therefore, the slope is an indicator of how quickly the overdetermined system must extract information before the model fails to represent the autocorrelation function, and the bias increases beyond some unacceptable level. The separation between performance curves in a given spectral width region is an indicator of the variation in the number of equations needed in the overdetermined system to reduce the bias in the estimate. This is relevant in defining a single overdetermined variance estimator to operate over an extended region of spectral widths.

It has been shown that the overdetermined variance estimators could be used to reduce the bias in the narrow band spectral width region. The number of terms to use in the series expansion and the number of equations needed in the overdetermined system is a function of the spectral widths and degradation of the autocorrelation estimate. However, it is shown that a single variance estimator can be made more robust by applying additional lags in an overdetermined system. As noted by Passarelli and as observed in this work, the use of larger closed systems incorporating additional estimates of the autocorrelation function at higher order lags only contributes to increase the bias in the estimate. The overdetermined system actually uses the higher order lags by applying a least squares fit to the data in a manner that allows one to extract additional information. It was also shown that the standard

deviation in the variance estimates could be reduced by incorporating additional autocorrelation function estimates at the higher order lags.

Using the estimator bias and standard deviation as a measure of performance, this work showed that an overdetermined system using only a few terms in the series expansion could be used to increase the region of spectral widths over which a given variance estimator might perform. In an application, the overdetermined variance estimator might be applied over the region of narrowband spectral widths in order to obtain an estimate of the region in which one is operating. An overdetermined or closed system variance estimator might then be applied in a more defined region to improve estimator performance.

This work has shown that an overdetermined system in terms of Passarelli's series expansion can be used to define spectral moment estimators which are able to use the available autocorrelation function estimates in a manner that extracts more of the available information. The overdetermined system is not limited to a particular moment or shape of the autocorrelation function. Potential areas for continued work using this framework are discussed in the following section.

Future Work

Future work in this area would involve the application of the overdetermined system to other central moments of interest and to other power spectrum shapes. One example would be the application of the overdetermined system to the mean estimator. As noted previously, the poly-pulse-pair mean estimator was shown to exhibit a reduction in the variance of the estimate by averaging over successive lags of the phase of the autocorrelation function. However, the overdetermined system using one term in the series expansion, yields the

$$\hat{M}_1 = \frac{\sum_{k=1}^{N_{OD}} kg(k)}{\sum_{k=1}^{N_{OD}} k^2} \tag{175}$$

estimator which is a different estimator than poly-pulse-pair estimator in Equation 143 and yields the least squares estimate and not just a weighted average. The

overdetermined system could also be applied to the estimation of spectral skewness which is limited due to the fact that one is trying to measure the third derivative of the phase function at zero from discrete estimates of the autocorrelation function. However, since this information is embedded in the autocorrelation function at each lag as noted by Passarelli's series expansion, the overdetermined system may allow one to extract the desired information by incorporating the higher order lags. In addition, the investigation of adaptive techniques for determining the optimum number of equations might be applied. Techniques found in robust linear regression analysis may lend some insight into developing adaptive techniques.

APPENDICES

Appendix A

Probability Theory and Random Processes

Introduction

This chapter includes a review of probability theory and random processes. The autocorrelation function and its Fourier transform, the power spectrum, are defined. Also, basic probability measures such as moments and characteristic functions which will be used to relate the autocorrelation and spectral domains are defined. In addition, estimator performance issues are addressed.

Probability Theory

Probability Assignment

Probability theory is the branch of mathematics used to described random events in some meaningful way. A random event is the outcome, α, of some experiment whose possible outcomes can be defined, but the knowledge of which outcome will result on a given experiment is unknown until the experiment has been performed. Probability theory attempts to characterize the likelihood of an event or events at the outcome of a given experiment. A common example is the roll of a dice where the possible set of outcomes on a single roll of the dice is the set $\{1, 2, 3, 4, 5, 6\}$ and the likelihood of a particular result is 1/6. The collection of all possible outcomes is called a probability space, Ω. A grouping of possible experimental outcomes is termed an event, A, or a subset of the space.

The partitioning and grouping of subsets in the probability space can be described in terms of set theory. Set theory provides for the partitioning of the probability space into subsets where the event A is a subset of the probability space denoted as $A \subset \Omega$. A special subset of all probability spaces is the null set or empty set denoted $\{\phi\}$. Subsets of the probability space can be related through the union and intersection operators which are both commutative and associative. The

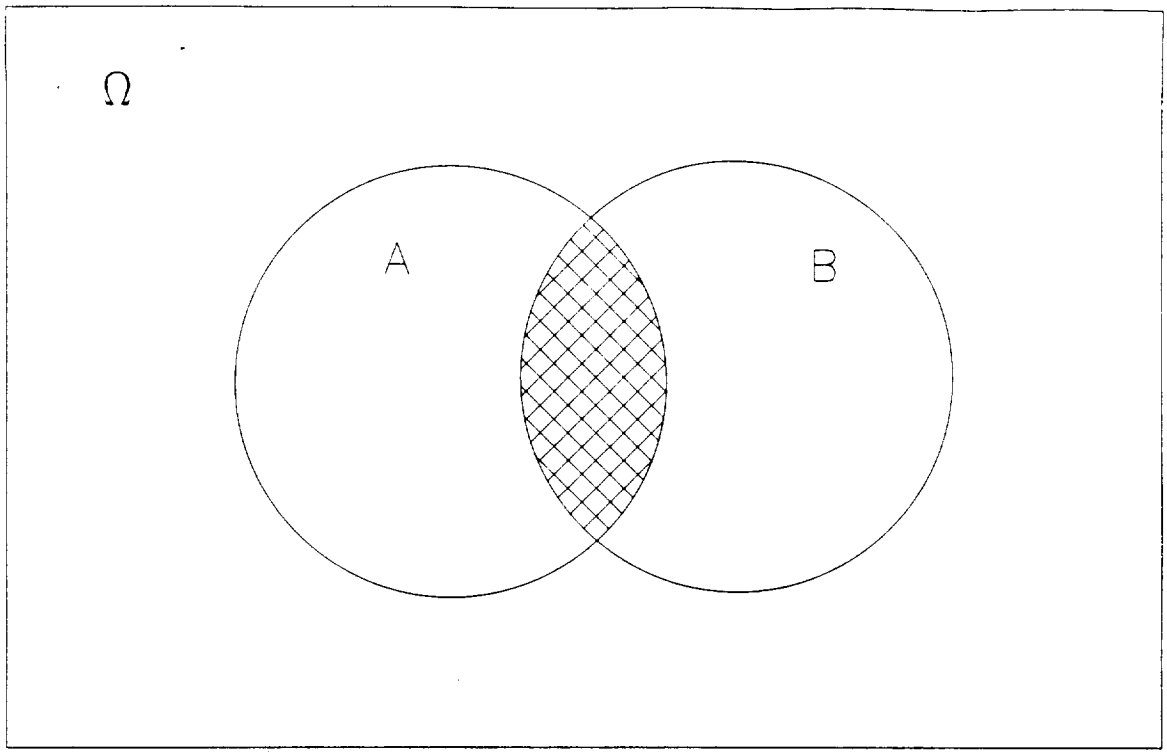


Figure A-1. Venn diagram representing the intersection of two sets.

union operator, ∪, combines two subsets into a single set. For example, given the two subsets A and B of a probability space, then

$$C = A \cup B \tag{A-1}$$

forms a new subset, C, containing the elements of A and B. The intersection operator, ∩, selects only the elements common to two sets and forms a new set from the overlap. This is best visualized through the use of a Venn diagram as shown in Figure A-1 where the hatched area is the new set containing the intersection of the sets A and B. Two sets which have as their intersection the null set

$$A \cap B = \{\phi\} \tag{A-2}$$

are said to be mutually exclusive or having no common elements.

In a probability space, each event, A, in the probability space, Ω, is assigned a real number between 0 and 1 as an indicator of its likelihood of occurrence. This

probability assignment is denoted P(A) where

$$P(A) \geq 0 \tag{A-3}$$

$$P(\Omega) = 1 \tag{A-4}$$

and

$$\text{if } (A \cap B) = \{\phi\} \text{ then } P(A \cup B) = P(A) + P(B) \ . \tag{A-5}$$

In the general case, the probability of the union of two events can be written as

$$P(A \cup B) = P(A) + P(B) - P(A \cap B) \ . \tag{A-6}$$

Besides the probability assignment given to unions and intersections, probabilities can be defined in terms of an event A conditioned on an event B. The conditional probability, $P(A \mid B)$ is defined as

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)} \ . \tag{A-7}$$

Two important theorems result from the conditional probability. The first theorem is the total probability theorem which states that if the probability space, Ω, can be partitioned into a set of mutually exclusive sets $\{A_1, A_2, \ldots, A_n\}$, then any event B in Ω can be written as

$$P(B) = P(B \mid A_1)P(A_1) + P(B \mid A_2)P(A_2) + \ldots + P(B \mid A_n)P(A_n) \tag{A-8}$$

using Equation A-7. The second important theorem is Bayes theorem which states that

$$P(A \mid B) = \frac{P(B \mid A)\, P(A)}{P(B)} \ . \tag{A-9}$$

The probability P(A) is often termed the a priori probability which indicates that this value is known before the experiment, and $P(A \mid B)$ is termed the a posteriori probability which indicates that the value is only known after the experiment. In the case where two events are independent of each other, then

$$P(A \cap B) = P(A)P(B) \ . \tag{A-10}$$

The independence of two events results in Equation A-7 being written as

$$P(A \mid B) = \frac{P(A)P(B)}{P(B)} = P(A) \, . \tag{A-11}$$

Random Variables

A random variable is a mapping, $X(\alpha)$, of all possible outcomes, α, of an experiment which are contained in Ω to the real number line. If the mapping is from a countable number of outcomes, then the random variable will take on only discrete values along the real number line. Events or subsets of the probability space can be defined as intervals on the real number line where

$$\{X \leq x\} \tag{A-12}$$

is an event or subset of the samples containing the possible outcomes mapped to the real line in the region less than some value x. The random variable is any function satisfying the following two conditions

1. $\{X \leq x\}$ is an event for all x.
2. $P(X = \infty) = 0$ and $P(X = -\infty) = 0$.

For future reference, a complex random variable is defined as

$$Z = X + jY \tag{A-13}$$

where $j = \sqrt{-1}$ and X and Y are real random variables.

The probability assignments given to the events associated with Ω are now defined as a mapping from the real line (defining the events) to the output of a cumulative distribution function (CDF), $F_X(x)$, where

$$F_X(x) = P\left(\{X \leq x\}\right) \, . \tag{A-14}$$

A statement listing all the properties of the CDF would be of considerable length; however, a few of the more general ones are worth listing. Properties of a CDF include:

1. $F_X(+\infty) = 1$ and $F(-\infty) = 0$.
2. $F_X(x)$ is a non-decreasing function.
3. $P(x_1 < X \leq x_2) = F_X(x_2) - F_X(x_1)$.
4. $P(X = x) = F_X(x) - F_X(x^-)$.

The derivate of the cumulative distribution function

$$f_X(x) = \frac{dF_X(x)}{dx} \tag{A-15}$$

is termed a probability density function (PDF) for the random variable X and has the following properties:

1. $f_X(x) \geq 0$
2. $\int_{-\infty}^{\infty} f_X(x)\, dx = 1$
3. $F_X(x_2) - F_X(x_1) = \int_{x_1}^{x_2} f_X(x)\, dx$

A PDF may be a continuous or discrete function depending on the type of random variable. There are many probability density functions defined in the literature which model random events. One of the most commonly applied PDF is the Gaussian PDF. The Gaussian PDF is continuous function

$$f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(\frac{(x-\mu)^2}{2\sigma^2}\right) \text{ for } -\infty \leq x \leq \infty \tag{A-16}$$

and is defined by two parameters, μ and σ.

Expectation

Often, one would like to characterize a random variable by the value it is most likely to take on. This value is often termed the "expected" value or mean, μ, of the random variable. The expected value of a random variable is computed as the first moment of its probability density function, where

$$\mu = E\{X\} = \int_{-\infty}^{\infty} x\, f_X(x)\, dx\,. \tag{A-17}$$

Higher order moments can also be defined as

$$M_n = E\{X^n\} = \int_{-\infty}^{\infty} x^n\, f_X(x)\, dx\,. \tag{A-18}$$

Another important moment used to characterize a random variable is the second moment about the mean which is often termed the variance, σ^2, of the random variable. The variance is defined as

$$\sigma^2 = E\{(X-\mu)^2\} = E\{x^2\} - \mu^2 \, . \tag{A-19}$$

Characteristic Function

The Fourier transform of a probability density function is defined as the characteristic function. The characteristic function offers some useful properties and completely characterizes the random variable without any loss of information. This is due to the Fourier transform pair relationship. The characteristic function is defined as

$$\Phi(\omega) = \int_{-\infty}^{\infty} f_X(x) \, \exp(j\omega x) \, dx \tag{A-20}$$

and the inverse relationship as

$$f_X(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \Phi(\omega) \, \exp(-j\omega x) \, dx \, . \tag{A-21}$$

The characteristic function offers a method for calculating moments using derivatives. Taking the first derivate of Equation A-20 with respect to ω yields

$$\frac{d\,\Phi(\omega)}{d\omega} = \int_{-\infty}^{\infty} jx \, f_X(x) \, \exp(j\omega x) \, dx \tag{A-22}$$

and evaluating at $\omega = 0$ yields

$$\frac{d\,\Phi(\omega)}{d\omega} \, |_{\omega=0} = \int_{-\infty}^{\infty} jx \, f_X(x) \, dx = j\mu \tag{A-23}$$

which equates the mean of the random variable to the derivative of the characteristic function evaluated at zero. In general, higher order moments are related to the characteristic equation through

$$M_n = (-j)^n \frac{d^n \Phi(\omega)}{d\omega^n} |_{\omega=0} \, . \tag{A-24}$$

Joint Random Variables

In many cases, an experiment will result in an outcome consisting of two or more random variables. An example is the blind selection of objects from a bag containing plastic objects varying in shape (e.g. square, round, triangular, etc.) and also varying in color (e.g. red, blue, green, etc.). The experiment outcome will consists of two random variables X and Y representing the shape and color of the objects, respectively. Probability theory defines the joint occurrence of the two random variables in terms of joint PDF's and CDF's. The probability associated with two random variables is defined as the probability that the condition on X and the condition on Y will be jointly satisfied or in other words

$$P\{X \leq x \text{ and } Y \leq y\} = P\{X \leq x\,,\, Y \leq y\}\,. \tag{A-25}$$

This joint probability is defined in terms of a joint CDF as

$$F_{XY}(x,y) = P\{X \leq x\,,\, Y \leq y\} \tag{A-26}$$

and the corresponding joint PDF is

$$f_{XY}(x,y) = \frac{dF_{XY}(x,y)}{dx\,dy}\,. \tag{A-27}$$

The joint CDF and PDF can be extended to N random variables $\{X_1, X_2, \ldots, X_N\}$ where the joint CDF is defined as

$$F_{(X_1,X_2,\ldots,X_N)}(x_1,x_2,\ldots,x_n) = P(X_1 \leq x_1,\, X_2 \leq x_2,\, \ldots,\, X_N \leq x_N) \tag{A-28}$$

and the joint PDF is defined as

$$f_{(X_1,X_2,\ldots,X_N)}(x_1,x_2,\ldots,x_n) = \frac{dF_{X_1,X_2,\ldots,X_N}(x_1,x_2,\ldots,x_n)}{dx_1\,dx_2\,\ldots dx_n}\,. \tag{A-29}$$

Moments are also defined for joint random variables. For the two dimensional case, the expected value of two random variables, X and Y, is defined as

$$r_{XY} = E\{XY\} = \int\int xy\, f_{XY}(x,y)\,dx\,dy \tag{A-30}$$

which is termed the correlation between X and Y. This can be easily be extended to N random variables. Another important moment in the two random variable case is the covariance which is defined as

$$c_{XY} = E\{(X - \mu_x)(Y - \mu_y)\} . \tag{A-31}$$

An important property of random variables is that of statistical independence. Statistical independence implies that the probability associated with one random variable does not depend on the probability associated with another random variable or

$$P\{X \leq x, Y \leq y\} = P\{X \leq x\} P\{Y \leq y\} \tag{A-32}$$

or in other words the two events occur independent of one another. Statistical independence is expressed in terms of the CDF and PDF as

$$F_{XY}(x, y) = F_X(x) F_Y(y) \tag{A-33}$$

and

$$f_{XY}(x, y) = f_X(x) f_Y(y) \tag{A-34}$$

respectively. Another property associated with joint random variables is the notion of correlation as defined in Equations A-30 and A-31. Two random variables are said to be uncorrelated if the covariance defined in Equation A-31 is

$$c_{XY} = 0 . \tag{A-35}$$

It can be shown that statistically independence implies uncorrelated; however, the converse is not necessarily true.

Functions of Random Variables

Functions of One Random Variable

As is often the case, the random variable under observation is a function of another random variable with known PDF. The problem becomes one of defining the PDF for the observed random variable. The mapping between the two random variables is defined as

$$y = g(x) \tag{A-36}$$

where X has a known PDF, $f_X(x)$, and known CDF, $F_X(x)$. The function $g()$ is also restricted to be monotonically increasing or decreasing. The probability that $\{Y \leq y\}$ is related to X through the inverse transformation where

$$P(Y \leq y) = P(X \leq g^{-1}(y)) \tag{A-37}$$

or

$$\begin{aligned} P(Y \leq y) &= F_Y(y) && \text{(A-38)} \\ &= P(X \leq g^{-1}(y)) && \text{(A-39)} \\ &= F_X(g^{-1}(y)) \,. && \text{(A-40)} \end{aligned}$$

The probability density function for Y can be obtained through the definition of the PDF and the chain rule

$$\begin{aligned} f_Y(y) &= \frac{dF_Y(y)}{dy} && \text{(A-41)} \\ &= \frac{dF_X(x)}{dx}\Big|_{g^{-1}(y)} \left|\frac{dx}{dy}\right| && \text{(A-42)} \\ &= f_X(g^{-1}(y)) \left|\frac{dx}{dy}\right| \,. && \text{(A-43)} \end{aligned}$$

Finding $g(x)$

Often one is interested in transforming a random variable with known PDF to another random variable with known PDF. An example is the generation of realizations of a random variable on a computer. Uniformly distributed samples are

easily obtained from simple algorithms on a computer. However, one is usually interested in obtaining samples which have some other distribution such as Gaussian, Rayleigh, exponential, etc.. With two known PDF's, the question becomes what transformation to use. This can be obtained by equating

$$P(Y \leq y) = P(X \leq x) \tag{A-44}$$

or

$$\int_{-\infty}^{x} f_X(\alpha)\, d\alpha = \int_{-\infty}^{y} f_Y(\alpha)\, d\alpha \tag{A-45}$$

and solving for

$$y = F_Y^{-1}(F_X(x)) = g(x)\,. \tag{A-46}$$

Functions of two random variables

Functions of two random variables are also common in physical systems. In general, the mapping is defined as

$$z = g(x, y)\,. \tag{A-47}$$

The CDF for Z can be obtained from the joint PDF of X and Y where

$$F_Z(z) = \int\int_{D_z \in \{x,y\}} f_{XY}(x, y)\, dx\, dy \tag{A-48}$$

and D_z is the set of all points $\{x, y\}$ for which $\{Z \leq z\}$. A very common situation is the case where the signal, X, is a random variable with additive noise, Y. The resultant random variable is $Z = X + Y$. If the signal and noise are assumed to be statistically independent, then their joint PDF is

$$f_{XY}(x, y) = f_X(x)\, f_Y(y)\,. \tag{A-49}$$

The CDF of Z is then obtained by

$$F_Z(z) = \int_{-\infty}^{\infty} \int_{-\infty}^{z-x} f_Y(y)\, dy f_X(x)\, dx\,. \tag{A-50}$$

Computing the integral on dy yields

$$F_Z(z) = \int_{-\infty}^{\infty} [F_Y(z-x) - F_Y(-\infty)]\, f_X(x)\, dx\,. \tag{A-51}$$

Now, differentiating $F_Z(z)$ to obtain the PDF yields

$$f_Z(z) = \int_{-\infty}^{\infty} f_Y(z-x)\, f_X(x)\, dx\, dz\,. \tag{A-52}$$

This states that the resultant PDF is a convolution of the individual PDF's if the random variables are statistically independent.

The Central Limit Theorem

The PDF for a sum of N statistically independent and identically distributed (i.i.d.) random variables can be shown to approach a Gaussian distribution in the limit that $N \to \infty$. More precisely, the Central Limit Theorem [17] states that if the x_i are i.i.d. and $S_N = \sum_{i=1}^{N} x_i$, then S_N is distributed such that

$$\lim_{N\to\infty} P(\frac{S_N - N\mu}{\sqrt{N}\sigma} \leq s) = \int_{-\infty}^{s} \frac{1}{2\pi} \exp(\frac{-\alpha^2}{2})\, d\alpha\,. \tag{A-53}$$

which is a Gaussian distribution with mean zero and variance one.

Estimation Theory

Parameter Estimation

There are a large number of density functions which can be used to model random events found in the physical world. Each density function has an associated parameter or parameters which define the density. The allowable ranges for the parameters associated with a particular density function form a family of density functions. In order to characterize a set of observations $\{x_1, x_2, \ldots, x_n\}$ obtained by sampling a random event with a known form of the density function (i.e., Gaussian, uniform, Rayleigh, etc.), an estimate of the density function parameters must be obtained from the observed sample data. The density function, $f_X(x)$, and its parameter, θ, or parameters is often expressed as

$$f_X(x;\theta) = f_X(x)\,. \tag{A-54}$$

The parameter estimate is a function of the observed sample data which consists of independent and identically distributed random variables $\{X_1, X_2, \ldots, X_N\}$. This parameter is therefore a function of random variables

$$\hat{\theta} = g(X_1, X_2, \ldots, X_N) \tag{A-55}$$

and is termed a statistic.

Estimator Bias

Since $\hat{\theta}$ is a function of random variables, it is also a random variable. To what degree this random variable is able to estimate the parameter is best measured in terms of properties of random variables. One of the most important properties of an estimator is that its expected value equal the true value of the parameter being estimated. The degree to which the expected value of the estimator approaches the true value is measured in termed of a bias

$$Bias = E\{\hat{\theta}\} - \theta \tag{A-56}$$

where θ is the true value, and an estimator is said to be unbiased if

$$E\{\hat{\theta}\} = \theta\,. \tag{A-57}$$

An estimator may also be defined as asymptotically unbiased when

$$\lim_{N\to\infty} E\{\hat{\theta}_N\} = \theta \tag{A-58}$$

where N is the number of samples used in computing the estimate.

Estimator Variance

In comparing unbiased estimators, one wants to consider the variation in the estimator about its mean or expected value. A "good" estimator is one which has a small variance. However, in comparing two unbiased estimators, one estimator

can be declared more efficient than another by comparing the variance of the two estimators. For example, if $\hat{\theta}_1$ and $\hat{\theta}_2$ are two unbiased estimators and

$$Var(\hat{\theta}_1) < Var(\hat{\theta}_2)$$

then $\hat{\theta}_1$ is more efficient than $\hat{\theta}_2$. The Cramer-Rao lower bound (CRB) puts a limit on the minimum variance obtainable for an estimator. The Cramer-Rao bound is

$$Var(\hat{\theta}) \geq \left[N \int_{-\infty}^{\infty} \left(\frac{\partial \ln f_X(x;\theta)}{\partial \theta} \right)^2 f_X(x;\theta)\, dx \right]^{-1} \tag{A-59}$$

$$= \left[N\, E\left\{ \left(\frac{\partial \ln f_X(x;\theta)}{\partial \theta} \right)^2 \right\} \right]^{-1} . \tag{A-60}$$

A derivation of the Cramer-Rao lower bound can be found in Larsen [17]. In the case of equality in Equation A-59, the estimator is termed an efficient estimator.

An seen in Equation A-59, the variance associated with the CRB decreases as a function of the number of samples used in the estimate. The variance of an estimator therefore should converge to zero as N tends to infinity

$$\lim_{N\to\infty} Var\{\hat{\theta}_N\} = \lim_{N\to\infty} E\left\{ \left| \hat{\theta}_N - E\left\{\hat{\theta}_N\right\} \right|^2 \right\} = 0 . \tag{A-61}$$

For use in defining an ergodic process in later sections, mean square consistency will be defined here. An estimator is said to be mean square consistent if the mean square error tends to zero in the limit

$$\lim_{N\to\infty} E\left\{ \left| \hat{\theta}_N - \theta \right|^2 \right\} = 0 . \tag{A-62}$$

Methods for Obtaining Estimators

There are three commonly used methods for developing parameter estimators. The three methods are: Bayesian estimators, maximum likelihood estimators, and estimators based on the method of moments. A brief discussion of each will follow in order to fully develop this section on parameter estimation.

Bayesian estimation assumes that the parameter to be estimated is no longer a constant but a random variable with a known density function, $f_\theta(\theta)$. This allows for the use of a priori information. The parameter estimate is then the expected value of the parameter given the observed values where

$$\hat{\theta} = E\{\theta|x_1, x_2, \ldots, x_N\} = \int_{-\infty}^{\infty} \theta \frac{f_{X_1,X_2,\ldots,X_N}(x_1, x_2, \ldots, x_N|\theta)}{f_{X_1,X_2,\ldots,X_N}(x_1, x_2, \ldots, x_N)} f_\theta(\theta)\, d\theta \,. \quad \text{(A-63)}$$

A maximum likelihood estimator is based on finding the value of the parameter that maximizes the likelihood function which is defined as

$$f_{X_1,X_2,\ldots,X_N}(x_1, x_2, \ldots, x_N; \theta) = f_{X_1}(x_1; \theta) f_{X_2}(x_2; \theta) \ldots f_{X_N}(x_N, \theta) \,. \quad \text{(A-64)}$$

Since the logarithm is a monotonic function, the value of θ that maximizes the likelihood function also maximizes the log of the likelihood function. Therefore, the log-likelihood function is often used in practice where the maximum is obtained by

$$\frac{\partial \ln[f_{X_1,X_2,\ldots,X_N}(x_1, x_2, \ldots, x_N; \theta)]}{\partial \theta} = 0 \quad \text{(A-65)}$$

and solving for θ. It can be shown that if the maximum likelihood estimator can be found, then maximum likelihood estimator variance will equal the CRB [14].

The method of moments is based on equating the theoretical moments to the sample moments and solving the system of equations for the unknown parameters. The sample moments are defined as

$$m_k = \frac{1}{N} \sum_{i=1}^{N} x_i^k \quad \text{(A-66)}$$

where k indicates the k-th sample moment. The system of equations to be solved is expressed as

$$\frac{1}{N} \sum_{i=1}^{N} x_i^k = \int_{-\infty}^{\infty} x^k f_X(x; \theta_1, \theta_2, \ldots, \theta_M))\, dx \quad \text{for } k = 1 \ldots M \,. \quad \text{(A-67)}$$

Random Processes

Random Processes Defined

In order to characterize the outcome of an experiment over time, random processes have been defined which describe each outcome as a new random variable indexed by time. Given the direction of this dissertation, only discrete-time random processes will be discussed; however, the concepts given here can easily be extended to continuous-time random processes. A discrete-time random process is denoted by a set of random variables, X, which are indexed by the time sample index, n, or $\{X(n) : n \in \text{Integers}\}$. Each random variable, $X(n)$, has a corresponding CDF

$$F_{X(n)} = P\{X(n) \leq x\} \tag{A-68}$$

and corresponding PDF

$$f_{X(n)} = \frac{dF_{X(n)}(x)}{dx} . \tag{A-69}$$

In order to characterize the total process, the joint CDF is required

$$F_{X(1),X(2),\ldots,X(k)}(x_1, x_2, \ldots, x_k) = P\left(X\left(1\right) \leq x_1,\, X\left(2\right) \leq x_2, \ldots, X\left(k\right) \leq x_k\right) \tag{A-70}$$

and the corresponding joint PDF

$$f_{X(1),X(2),\ldots,X(k)}(x_1, x_2, \ldots, x_k) = \frac{dF_{X(1),X(2),\ldots,X(k)}(x_1, x_2, \ldots, x_k)}{dx_1\, dx_2 \ldots dx_k} . \tag{A-71}$$

The possible set of output waveforms of a random process form a family or ensemble of waveforms.

Expectations

Since each time indexed output represents a random variable, the moments defined in Section A can be applied to yield

$$M_m(n) = E\left\{X^m\left(n\right)\right\} . \tag{A-72}$$

With Equation A-72, one is able to calculate the mean and variance for each random variable associated with the random process. Besides calculating moments for individual random variables comprising the process, joint moments between random variables may also be defined. Two commonly used moments defined for joint random processes are the autocorrelation defined as

$$r_{XX}(k,l) = E\{X(k)X^*(l)\} \tag{A-73}$$

and the autocovariance defined as

$$c_{XX}(k,l) = E\{[X(k) - \mu(k)][X(l) - \mu(l)]^*\} \tag{A-74}$$

which should be noted are functions of the time-indices. When comparing two random processes, $X(n)$ and $Y(n)$, similar moments can be defined where

$$r_{XY}(k,l) = E\{X(k)Y^*(l)\} \tag{A-75}$$

is termed the cross-correlation and

$$c_{XY}(k,l) = E\{[X(k) - \mu(k)][Y(l) - \mu(l)]^*\} \tag{A-76}$$

is termed the cross-covariance. Two random processes are said to be uncorrelated if

$$c_{XY}(k,l) = 0 \text{ for all k and l}. \tag{A-77}$$

This property shows that if two random processes, $X(n)$ and $Y(n)$, are uncorrelated, then the sum of the two random process

$$Z(n) = X(n) + Y(n) \tag{A-78}$$

yields an autocorrelation function that is the sum of the individual correlation functions

$$r_{ZZ}(k,l) = r_{XX}(k,l) + r_{YY}(k,l)\,. \tag{A-79}$$

This is a useful property when dealing with additive noise that is uncorrelated with the random process under observation.

Stationarity

Random processes can be divided into classes based on the form of the underlying joint densities. For the case of discrete-time random processes, if the PDF's of the random variables comprising the random process all have the same PDF

$$f_{X(n)}(x_n) = f_X(x) \tag{A-80}$$

then the process is termed a first-order stationary process. In this case, all first order statistics, (i.e., mean, variance, etc.), will be equivalent for each time-indexed random variable. For the case of a joint PDF consisting of two random variables, if the joint PDF is independent of absolute time

$$f_{X(n+k),X(n)}(x_n, x_k) = f_{X(m+k),X(m)}(x_m, x_k) \text{ for all m,n, and k} \tag{A-81}$$

and just depends on the separation in time between the two random variables, then the process is termed a second-order stationary process. Therefore, the second order statistics are independent of absolute time and just depend on the difference between the time-indexes. An example is the autocorrelation function where

$$r_{XX}(k, l) = r_{XX}(k - l, 0) \tag{A-82}$$

or for easier notation

$$r_{XX}(k, l) = r_{XX}(k - l)\ . \tag{A-83}$$

A random process having constant mean and an autocorrelation function that depends only on the distance between random variables is termed a wide sense stationary (WSS) random process. This is an important type of random process because the autocorrelation function is commonly used in process analysis. Other forms of stationarity exist, but will not be discussed here.

Ergodicity

Up to this point, the moments and correlations have been defined in terms of ensemble averages. However, seldom does one have more than one realization of a random process, and the ensemble averages are not known a priori. There are, however, random processes which exhibit a relationship between the ensemble averages and the sample averages that is of interest. Ergodic random processes are stationary processes for which the sample averages converge to the ensemble averages in the mean square sense. A random process may be ergodic in one moment and not another. Therefore, different types of ergodicity have been defined. Two commonly used types of ergodicity are ergodic random processes in the mean (mean ergodic) and ergodic random processes in the autocorrelation (autocorrelation ergodic). A WSS random process is mean ergodic if

$$\lim_{N\to\infty} E\left\{|\mu_N - \mu|^2\right\} = 0 \tag{A-84}$$

where

$$\mu_N = \frac{1}{N}\sum_{i=1}^{N} x_i \tag{A-85}$$

and μ is the ensemble mean. A WSS random process is autocorrelation ergodic if

$$\lim_{N\to\infty} E\left\{|r_N(k) - r(k)|^2\right\} = 0 \tag{A-86}$$

where

$$r_N(k) = \frac{1}{N}\sum_{i=1}^{N} x_i\, x_{i+k}^* \tag{A-87}$$

and $r(k)$ is the ensemble autocorrelation.

Gaussian Random Processes

A very common random process is the Gaussian random process whose joint PDF is defined as

$$f_{\underline{X}}(\underline{x}) = \frac{1}{(2\pi)^2|C|^{1/2}} \exp\left(\frac{-1}{2}(\underline{x} - \underline{m}_x)^T C^{-1} (\underline{x} - \underline{m}_x)\right) \tag{A-88}$$

where the mean vector, $\underline{m_x}$, is defined as

$$\underline{m_x} = [E\{X_1\}, E\{X_2\}, \ldots, E\{X_N\}]^T \tag{A-89}$$

and the covariance matrix, C, has elements c_{kl} defined by

$$c_{kl} = E\{(x_k - m_k)(x_l - m_l)\} = c(k,l) \; . \tag{A-90}$$

In the case of a WSS Gaussian random process, the mean vector reduces to

$$\underline{m_x} = E\{X\}\,[1, 1, 1, \ldots]^T \tag{A-91}$$

and the covariance matrix elements are defined as

$$c_{kl} = c(k-l) \; . \tag{A-92}$$

Autocorrelation Matrices

The covariance and autocorrelation matrices are used in many areas of process analysis and are worthy of a discussion, especially in the case of a WSS random process. The autocorrelation matrix for a WSS random process is defined as

$$E\{\underline{x}\,\underline{x}^H\} = \begin{bmatrix} r_{xx}(0) & r_{xx}^*(1) & \ldots & r_{xx}^*(N) \\ r_{xx}(1) & r_{xx}(0) & \ldots & r_{xx}^*(N-1) \\ \vdots & \vdots & \ldots & \vdots \\ r_{xx}(N) & r_{xx}(N-1) & \ldots & r_{xx}(0) \end{bmatrix} \tag{A-93}$$

where H represents conjugate transpose. The covariance matrix is related to the autocorrelation matrix by

$$C = R - \underline{m_x}\,\underline{m_x}^H \; . \tag{A-94}$$

There are several useful properties of the autocorrelation matrix when the process is assumed to be wide sense stationary. It is easily shown that the autocorrelation matrix is conjugate symmetric (Hermitian) since

$$r_{xx}(k) = E\{x_{n+k}\,x_n^*\} \tag{A-95}$$

and

$$r_{xx}(-k) = E\{x_n\, x^*_{n+k}\} \tag{A-96}$$

then

$$r_{xx}(k) = r^*_{xx}(-k)\ . \tag{A-97}$$

Also, since the values along the diagonal are all equal, the autocorrelation matrix is also defined to be Toeplitz. Another useful property of the autocorrelation matrix is that it is nonnegative definite (or positive semi-definite). A nonnegative definite matrix is one for which

$$\underline{b}^H R \underline{b} \geq 0 \tag{A-98}$$

for any vector $\underline{b}$. A simple proof is

$$\underline{b}^H R \underline{b} = \underline{b} E\{\underline{x}\,\underline{x}^H\}\underline{b} \tag{A-99}$$

$$= E\{\underline{b}\,\underline{x}\,\underline{x}^H\,\underline{b}\} \tag{A-100}$$

$$= E\{\left|\underline{b}^H\,\underline{x}\right|^2\} \tag{A-101}$$

and since $|\underline{b}^H\,\underline{x}|^2$ is greater than or equal to zero for any $\underline{b}$ then $\underline{b}^H R \underline{b} \geq 0$.

Power Spectrum

The Power Spectral Density

For a wide sense stationary random process, the second-order moments, (i.e., the autocorrelation function), have been defined as a function of time (or delay). Since the autocorrelation function is a deterministic quantity, one can apply Fourier analysis to the function in order to gain insight into the spectral content of the random process. The power spectrum or power spectral density (PSD) for a discrete-time random process is defined as

$$P_X(\omega) = \sum_{k=-\infty}^{\infty} r_{XX}(k)\, \exp(-j\omega k) \tag{A-102}$$

and the corresponding Fourier inverse is

$$r_{XX}(k) = \frac{1}{2\pi}\int_{-\pi}^{\pi} P_X(\omega)\,\exp(j\omega k)\,d\omega\,. \tag{A-103}$$

In terms of system analysis, the z-transform of the autocorrelation function is defined as

$$P_X(z) = \sum_{k=-\infty}^{\infty} r_{XX}(k)\,z^{-k}\,. \tag{A-104}$$

Note that Equation A-102 is the z-transform of the autocorrelation function evaluated along the unit circle.

Linear System Analysis

In many cases, a random process is observed at the output of a linear system. One can characterize the output in terms of a known or assumed input and the system response. Let x(n) represent a realization of a WSS random process as input to a linear system with impulse response h(n). The output of the system, y(n), is defined through the convolution sum where

$$y(n) = \sum_{k=-\infty}^{\infty} x(k)h(n-k)\,. \tag{A-105}$$

Taking the expectation of both sides yields

$$E\{y(n)\} = \sum_{k=-\infty}^{\infty} E\{x(k)\}h(n-k)\,. \tag{A-106}$$

Therefore, the expected value of the output is the mean of process weighted by the total energy of the system impulse response

$$E\{y(n)\} = \mu_X \sum_{k=-\infty}^{\infty} h(n-k)\,. \tag{A-107}$$

The expected value of the output does little to describe the information content of the random process over time. The time dependence is observed in the autocorrelation function. The second order statistics associated with the output are expressed as

$$E\{y(n+k)\,y^*(n)\} = \sum_{l=-\infty}^{\infty} E\{y(n+k)x^*(l)\}h^*(n-l)\,. \tag{A-108}$$

Equation A-108 can be expressed in terms of a correlation function where

$$r_{YY}(k) = \sum_{k=-\infty}^{\infty} r_{YX}(n+k-l)h^*(n-l) \; . \tag{A-109}$$

Now, let $m = l - n$ and change the index of summation such that

$$r_{YY}(k) = \sum_{m=-\infty}^{\infty} r_{YX}(k-m)h^*(-m) \tag{A-110}$$

or

$$r_{YY}(k) = r_{XY}(k) * h^*(-k) \; . \tag{A-111}$$

Now, computing $r_{YX}(k)$ yields

$$E\{y(n+k)\, x^*(n)\} = \sum_{k=-\infty}^{\infty} E\{x(n+k-l)x^*(n)\}h(l) \tag{A-112}$$

and therefore

$$\begin{aligned} r_{YX}(k) &= \sum_{k=-\infty}^{\infty} r_{XX}(k-l)h(l) && \text{(A-113)} \\ &= r_{XX}(k) * h(k) \; . && \text{(A-114)} \end{aligned}$$

Therefore, the second order input-output relationship of a random process passing through a linear system is expressed as

$$r_{YY}(k) = r_{XX}(k) * h(k) * h^*(-k) \; . \tag{A-115}$$

Note, that the z-transform of $h^*(-k)$ is expressed in terms of the z-transform of $h(k)$, $H(z)$, such that

$$Z\{h^*(-k)\} = H^*(\frac{1}{z^*}) \; . \tag{A-116}$$

A simple proof follows. Let the z-transforms of $h^*(-k)$ be defined as

$$H_1(z) = \sum_k h^*(-k)\, z^{-k} \; . \tag{A-117}$$

Now, conjugating both sides yields

$$H_1^*(z) = \sum_k h(-k)\, z^{*(-k)} \; . \tag{A-118}$$

Let $m = -k$, and summing over m

$$H_1^*(z) = \sum_m h(m)\, z^{*-(-m)} \tag{A-119}$$

$$= \sum_m h(m) \frac{1}{z^*}^{(-m)} . \tag{A-120}$$

Conjugating both sides yields

$$H_1(z) = \left[\sum_m h(m) \frac{1}{z^*}^{(-m)}\right]^* = H^*(\frac{1}{z^*}) . \tag{A-121}$$

Using this relationship and Equation A-115, the second-order input-output z-transform relationship is written as

$$P_{yy}(z) = P_{xx}(z)H(z)H^*(\frac{1}{z^*}) . \tag{A-122}$$

When the transfer function is evaluated along the unit circle, $z = \exp(j\omega)$, Equation A-122 becomes

$$P_{yy}(\omega) = P_{xx}(\omega)\,|H(\omega)|^2 . \tag{A-123}$$

This equation defines the power spectral density of the output as the product of the power spectral density of the input times the magnitude squared of the transfer function frequency response.

Positivity

A power spectrum describes the distribution of power over frequency and since power is defined to be a positive quantity, the power spectrum must be positive for all ω. Now the power spectrum is defined as the discrete-time Fourier transform of the autocorrelation function. Therefore, using the discrete-time Fourier inverse relationship, the average power in a realization of a random process, $y(n)$, can be written as

$$r_{YY}(0) = E\{|y(n)|^2\} = \frac{1}{2\pi}\int_{-\pi}^{\pi} P_{YY}(\omega)\, d\omega . \tag{A-124}$$

The magnitude squared of $y(n)$ constrains $r_{YY} \geq 0$. Now, assume that the realization $y(n)$ is the output of a linear time-invariant system with input $x(n)$. Then, $r_{YY}(0)$ can be written as

$$r_{YY}(0) = E\{|y(n)|^2\} = \frac{1}{2\pi}\int_{-\pi}^{\pi} |H(\omega)|^2 \, P_{XX}(\omega)\, d\omega \,. \tag{A-125}$$

The filter can be defined as an ideal bandpass filter where

$$H(\omega) = 1 \text{ for } \omega_1 \leq \omega \leq \omega_2 \tag{A-126}$$

$$= 0 \text{ for } |\omega_1| \text{ and } |\omega_2| \leq 0 \,. \tag{A-127}$$

Therefore, Equation A-125 can be written as

$$r_{YY}(0) = E\{|y(n)|^2\} = \frac{1}{2\pi}\int_{\omega_1}^{\omega_2} P_{XX}(\omega)\, d\omega \,. \tag{A-128}$$

Since $r_{YY}(0)$ is defined to be

$$r_{YY}(0) \geq 0 \tag{A-129}$$

$P_{XX}(\omega)$ must be

$$P_{XX}(\omega) \geq 0 \tag{A-130}$$

for $\omega_1 \leq \omega \leq \omega_2$ and $|H(\omega)|^2 = 1$. Now, $P_{XX}(\omega)$ and $|H(\omega)|^2$ are constrained to be positive or equal to zero for $\omega_1 \leq \omega \leq \omega_2$. Therefore, given Equation A-123, $P_{YY}(\omega)$ is constrained such that

$$P_{YY}(\omega) \geq 0 \tag{A-131}$$

for $\omega_1 \leq \omega \leq \omega_2$. Now the bandwidth of the filter is allowed to take on any value between $-\pi$ and π, therefore the condition

$$P_{YY}(\omega) \geq 0 \tag{A-132}$$

must hold for all ω. This says that the power spectral density is a positive semi-definite function provided it is the discrete-time Fourier transform of a true autocorrelation function.

Spectral Factorization

Spectral factorization is a property of the power spectrum which allows any power spectrum to be represented as the output of a causal and stable filter driven by white noise. The power spectrum was defined as

$$P_{XX}(\omega) = \sum_{k=-\infty}^{\infty} r_{XX}(k) \exp(-j\omega k) \tag{A-133}$$

which is a real-valued positive function that is periodic in 2π. The z-transform can be obtained by replacing $\exp(j\omega)$ with z

$$P_{XX}(z) = \sum_{k=-\infty}^{\infty} r_{XX}(k)\, z^{-k} \tag{A-134}$$

where $P_{XX}(z)$ is analytic in the annulus $\rho \leq |z| \leq \frac{1}{\rho}$ which includes the unit circle. Taking the logarithm of $P_{XX}(z)$, $\ln[P_{XX}(z)]$, yields another analytic [6] function in the annulus. This function can be expressed in terms of a Laurent series expansion [30] about zero where

$$\ln[P_{XX}(z)] = \sum_{k=-\infty}^{\infty} a_k\, z^{-k}\,. \tag{A-135}$$

The coefficients of the Laurent series, a_k, can be obtained by evaluating the series at $z = \exp(j\omega)$

$$\ln[P_{XX}(\omega)] = \sum_{k=-\infty}^{\infty} a_k \exp(-j\omega k) \tag{A-136}$$

and observing that this is the Fourier series representation of the periodic function $\ln[P_{XX}(z)]$. The coefficients are defined by

$$a_k = \frac{1}{2\pi} \int_{-\pi}^{\pi} \ln[P_{XX}(\omega)] \exp(j\omega k)\, d\omega\,. \tag{A-137}$$

Since the power spectrum is real, the coefficients are conjugate symmetric, $a_k = a^*_{-k}$ and a_0 is defined as

$$a_0 = \frac{1}{2\pi} \int_{-\pi}^{\pi} \ln[P_{XX}(\omega)]\, d\omega\,. \tag{A-138}$$

The power spectrum can now be written in terms of the expansion as

$$P_{XX}(z) = \exp(a_0)\ \exp(\sum_{k=1}^{\infty} a_k\, z^{-k})\ \exp(\sum_{k=-\infty}^{-1} a_k\, z^{-k})\,. \tag{A-139}$$

Now, let

$$H(z) = \exp(\sum_{k=1}^{\infty} a_k \, z^{-k}) \tag{A-140}$$

which is analytic in the region $|z| > \rho$. $H(z)$ may now be expanded in a power series such that

$$H(z) = 1 + h(1)z^{-1} + h(2)z^{-2} + \ldots \tag{A-141}$$

where

$$h(0) = lim_{z \to \infty} H(z) = 1 \tag{A-142}$$

given the definition of $H(z)$. Since the region of convergence includes the unit circle, $H(z)$ is stable filter. Also, $H(z)$ is causal given that $c(k) = 0$ for $k \leq 0$. Using the conjugate symmetry of the Laurent series coefficients, Equation A-139 can be written as

$$P_{XX}(z) = \sigma^2 \, H(z) \, H^*(\frac{1}{z^*}) \tag{A-143}$$

where $\sigma^2 = \exp(a_0)$. Evaluating $P_{XX}(z)$ on the unit circle yields

$$P_{XX}(\omega) = \sigma^2 \, |H(\omega)|^2 \tag{A-144}$$

which is the frequency response obtained from the output of linear system driven by white noise.

Power Spectrum Estimation

Autocorrelation Sequence Estimation

The definition of a power spectral density for a random process is conditioned on the assumption of WSS. However, in physical systems, the random process under observation is usually only locally WSS stationary. Locally WSS being defined as only slight variations in $r_{xx}[k]$ with respect to the time index n in Equation A-81 over a finite observation of the random process. Such physical systems include human speech, atmospheric returns from radar, and oceanographic returns from sonar. In addition to the WSS requirement, the autocorrelation lags of the process

are not known a priori and must be estimated from a finite number, N, of samples of a realization. In order to estimate the autocorrelation lags, the assumption of ergodicity must be exercised. With only N observations, one can at best obtain estimates of the autocorrelation function for lags between $-N \leq k \leq N$. Assuming an autocorrelation ergodic process and an estimate of the autocorrelation function, $\hat{r}_{XX}(k)$, the PSD estimate is defined as

$$\hat{P}_{xx}(f) = \sum_{k=-N}^{N} \hat{r}_{xx}[k] \exp(-j2\pi fk) \qquad \frac{-1}{2} \leq f \leq \frac{1}{2} . \tag{A-145}$$

There are two commonly used estimators for the autocorrelation function. The first one is an unbiased estimator of autocorrelation function defined as

$$\hat{r}_{XX}(k) = \frac{1}{N-k} \sum_{n=0}^{N-k} x(n+k)x^*(n) \text{ for } -(N-1) \leq k \leq N-1 . \tag{A-146}$$

Taking the expectation of both sides yields

$$E\{\hat{r}_{XX}(k)\} = \frac{N-k}{N-k} r_{XX}(k) = r_{XX}(k) \tag{A-147}$$

which is a statement of unbiasness. The variance of the estimator is approximately [29]

$$Var\{\hat{r}_{XX}(k)\} \approx \frac{N}{(N-k)^2} \sum_{l=-\infty}^{\infty} \left(r_{XX}^2(l) + r_{XX}(l+k)r_{XX}(l-k) \right) \tag{A-148}$$

for a real Gaussian random process and $N \gg k$. Another commonly used autocorrelation estimator is

$$\hat{r}_{XX}(k) = \frac{1}{N} \sum_{n=0}^{N-k} x(n+k)x^*(n) \text{ for } -(N-1) \leq k \leq N-1 \tag{A-149}$$

which is a biased estimator where

$$E\{\hat{r}_{XX}(k)\} = \frac{N-k}{N} r_{XX}(k) . \tag{A-150}$$

However, the autocorrelation estimator in Equation A-149 is asymptotically unbiased where

$$\lim_{N\to\infty} E\{\hat{r}_{XX}(k)\} = \lim_{N\to\infty} \frac{N-k}{N} r_{XX}(k) = r_{XX}(k) . \tag{A-151}$$

The variance of the estimator in Equation A-149 is approximately [29]

$$Var\{\hat{r}_{XX}(k)\} \approx \frac{1}{N} \sum_{l=-\infty}^{\infty} \left(r_{XX}^2(l) + r_{XX}(l+k) r_{XX}(l-k) \right) . \qquad \text{(A-152)}$$

In most physical environments and especially at the higher order lags, Marple [29] states that the sum of the variance and the squared bias is larger for the unbiased estimator in Equation A-146 than for the biased estimator in Equation A-149.

Another problem is that the estimator in Equation A-146 may yield invalid autocorrelation sequences. For a WSS random, it can be shown that

$$r_{XX}(0) \geq |r_{XX}(k)| \text{ for all } k . \qquad \text{(A-153)}$$

A simple proof follows from the positivity constraint and the definition of the power spectrum. From the definition, let

$$r_{XX}(k) = \frac{1}{2\pi} \int_{-\pi}^{\pi} P_X(\omega) \exp(j\omega k) \, d\omega . \qquad \text{(A-154)}$$

Using the Schwartz inequality and the positivity constraint, it follows that

$$|r_{XX}(k)| \leq \frac{1}{2\pi} \int_{-\pi}^{\pi} |P_X(\omega)| \, |\exp(j\omega k)| \, d\omega \qquad \text{(A-155)}$$

$$= \frac{1}{2\pi} \int_{-\pi}^{\pi} P_X(\omega) \, d\omega . \qquad \text{(A-156)}$$

Noting that Equation A-156 is the definition for $r_{XX}(0)$, then

$$|r_{XX}(k)| \leq r_{XX}(0) \text{ for all } k . \qquad \text{(A-157)}$$

If the condition in Equation A-153 is not met, the autocorrelation sequence is invalid. It can be shown that this condition will always be met by the biased estimator given in Equation A-149. Let $\underline{x}$ represent a vector containing N samples of a realization from an ergodic random process where

$$\underline{x} = [x_1, x_2, \ldots, x_N]^T . \qquad \text{(A-158)}$$

Also, define a shift operator S^k which shifts a vector k samples up or down and fills the empty elements after the shift with zeros. For example $S^1\{\underline{x}\}$ yields

$$S^1\{\underline{x}\} = [0, x_1, x_2, \ldots, x_{N-1}]^T . \tag{A-159}$$

The biased autocorrelation sequence estimate in Equation A-149 can be written in terms of the shift operator as

$$\hat{r}_{XX}(k) = \frac{1}{N} S^k\{\underline{x}(n)\}^T \underline{x}^*(n) \text{ for } -(N-1) \leq k \leq N-1 . \tag{A-160}$$

Let $\underline{y}_k = S^k\{\underline{x}\}$ for any k, then the biased autocorrelation estimate can be written as

$$\hat{r}_{XX}(k) = \frac{1}{N} \underline{y}_k^T \underline{x}^*(n) \text{ for } -(N-1) \leq k \leq N-1 . \tag{A-161}$$

Applying the Schwartz inequality yields

$$||\hat{r}_{XX}(k)|| \leq \frac{1}{N} ||\underline{y}_k^T|| \, ||\underline{x}^*(n)|| \text{ for } -(N-1) \leq k \leq N-1 . \tag{A-162}$$

The norm operator $||\underline{x}||$ is defined as

$$||\underline{x}|| = \underline{x}^T \underline{x} \tag{A-163}$$

Now, the norm of $\underline{y}_k$ is less than or equal to the norm of $\underline{x}$ since $\underline{y}_k$ contains a subset of the elements of $\underline{x}$ and all other elements have been set to zero.

Therefore, for the case when $k = 0$,

$$\hat{r}_{XX}(0) = \frac{1}{N} \underline{x}^T \underline{x} \tag{A-164}$$

and for $k \neq 0$

$$\hat{r}_{XX}(k) = \frac{1}{N} \underline{y}_k^T \underline{x} . \tag{A-165}$$

Now, since the $||\underline{x}|| \geq ||\underline{y}||$ then

$$\hat{r}_{XX}(0) \geq \hat{r}_{XX}(k) \text{ for all } k . \tag{A-166}$$

This condition does not hold for all estimates of the autocorrelation function using the unbiased estimator in Equation A-146. This can be shown by a simple example. Let $\underline{x} = [1.1\ 1.05\ 1.07]^T$, then the unbiased autocorrelation estimate is

$$\hat{r}_{XX}(0) = 1.15 \tag{A-167}$$

$$\hat{r}_{XX}(1) = 1.14 \tag{A-168}$$

$$\hat{r}_{XX}(2) = 1.18\,. \tag{A-169}$$

Note that lag two in the autocorrelation estimate is greater than the zeroth lag. Therefore, the autocorrelation estimate is not a valid one based on the property given in Equation A-153. Since the unbiased autocorrelation estimate may lead to an invalid autocorrelation sequence, and since the biased estimator is asymptotically unbiased and its variance tends to zero in the limit, this dissertation will hence forth use the term "autocorrelation estimator (or estimate)" to refer to the biased estimator of the autocorrelation sequence. If the unbiased estimator is used, it will be stated explicitly.

Power Spectrum Estimation Techniques

Spectral estimation techniques can be divided into traditional and modern techniques. The traditional spectral estimation techniques focus on the use of the discrete Fourier transform (DFT) and an estimate of the ACF. These techniques include the periodogram method and the Blackman-Tukey method for spectral estimation. The Blackman-Tukey method applies window functions to the estimated autocorrelation function to reduce the variance in the spectral estimates. It can be shown that the Blackman-Tukey method reduces to the periodogram method when using a rectangular window. The modern spectral estimation techniques are based on an assumed model for generating the random process. These techniques include autoregressive moving average modeling (ARMA) which seeks to model the random process as the output of a linear system driven by white noise and Prony's method

which seeks to apply a deterministic exponential model to the data. Chapter 3 develops in some detail the ARMA modeling process.

Periodogram Method

The periodogram estimate of the power spectral density is computed as the magnitude squared Fourier transform of a finite length realization of the random process. The periodogram estimate is

$$\hat{P}_{XX}(\omega) = \frac{1}{N}\left|\sum_{n=0}^{N-1} x(n)\exp(-j\omega n)\right|^2 . \tag{A-170}$$

This is an unbiased estimator of the power spectral density. This can be shown by changing the summation index such that

$$\hat{P}_{XX}(\omega) = \frac{1}{2M+1}\left|\sum_{n=-M}^{M} x(n)\exp(-j\omega n)\right|^2 . \tag{A-171}$$

Now, let $M \to \infty$ and taking the expectation of both sides yields,

$$\lim_{M\to\infty} E\{\hat{P}_{XX}(\omega)\} = \frac{1}{2M+1}\sum_{n=-M}^{M} x(n)\exp(-j\omega n)\sum_{m=-M}^{M} x^*(m)\exp(j\omega m) \tag{A-172}$$

$$= \frac{1}{2M+1}\sum_{n=-M}^{M}\sum_{m=-M}^{M} r_{XX}(n-m)\exp(-j\omega(n-m)) . \tag{A-173}$$

Letting $k = n - m$ yields

$$\lim_{M\to\infty} E\{\hat{P}_{XX}(\omega)\} = \frac{1}{2M+1}\sum_{k=-2M}^{2M}(2M+1-|k|)r_{XX}(k)\exp(-j\omega(k)) \tag{A-174}$$

$$= \frac{1}{2M+1}\sum_{k=-2M}^{2M}(2M+1-|k|)r_{XX}(k)\exp(-j\omega(k)) \tag{A-175}$$

and in the limit,

$$\lim_{M\to\infty} E\{\hat{P}_{XX}(\omega)\} = P_{XX}(\omega) \tag{A-176}$$

which states that the periodogram is an unbiased estimator. However, it can be shown that the variance of the estimator does not approach zero as the number of samples increases. Kay [14] has shown that the variance of the periodogram is approximately,

$$Var[\hat{P}_{XX}(\omega)] \approx P_{XX}^2(\omega) . \tag{A-177}$$

However, an averaging of periodograms generated from M non-overlapping, independent, and identically distributed finite realizations of the random process can be used to reduce the variance in the estimate. The averaged periodogram can be expressed as

$$\hat{P}_{XX}(\omega)_{avg} = \frac{1}{M} \sum_{m=1}^{M} \hat{P}_{XX}^{m}(\omega) \,. \tag{A-178}$$

The variance of the average periodogram estimator is reduce by a factor of M over the periodogram estimator given in Equation A-170. Since several realizations of the random process are seldom available in practice, a single realization is partitioned into M non-overlapping sequences of length N. The variance of the periodogram estimate is, however, no longer reduced by M, but by a factor slightly less than M [14].

Blackman-Tukey Method

The Blackman-Tukey method for spectral estimation is an attempt to reduce the variance of the estimate through data windowing. The Blackman-Tukey power spectral density estimator is defined as

$$\hat{P}_{XX}^{BT}(\omega) = \sum_{k=-(N-1)}^{N-1} \hat{r}_{XX}(k)\, w(k) \exp(-j\omega k) \tag{A-179}$$

where $w(k)$ is a time-domain weighting function. The weighting function is applied to reduce the variation in the latter lags of the estimated autocorrelation sequence. Since the latter lags are estimated using fewer and fewer samples, the weighting has the effect of reducing the variance of Blackman-Tukey estimator. Kay [14] has shown the variance to be approximately

$$Var[\hat{P}_{XX}^{BT}(\omega)] \approx \frac{P_{XX}^{2}(\omega)}{N} \sum_{k=-N}^{N} w^{2}(k) \,. \tag{A-180}$$

However, an additional bias is imposed due to the corresponding convolution operation occurring in the frequency domain due to the windowing operation. Further discussion of Fourier based estimators can be found in [14].

Appendix B

Generation of Gaussian Shaped Spectra

Zrnic [39] has developed a method for generating in-phase and quadrature data based on a specified shape for the power spectrum. This technique has been used in weather echo analysis where the Doppler return is often Gaussian shaped [9]. This appendix contains the simulation algorithm proposed by Zrnic.

The complex return from weather plus white noise can be modeled as

$$I(m) = s(m)cos[\phi(m)] + n(m)cos[\psi(m)] \tag{B-1}$$

and

$$Q(m) = s(m)sin[\phi(m)] + n(m)sin[\psi(m)] \tag{B-2}$$

where I(m) and Q(m) are the in-phase and quadrature returns, respectively. The signal and noise return envelopes s(m) and n(m) are Rayleigh distributed and the signal and noise phase terms $\phi(m)$ and $\psi(m)$ are uniformly distributed between 0 and 2π. This signal can be written in terms of its power per frequency bin through the discrete Fourier transform where

$$I(m) + jQ(m) = \frac{1}{N}\sum_{k=0}^{N-1} P_k^{1/2}\exp(j\theta_k)\exp(-j\frac{2\pi}{N}mk)\ . \tag{B-3}$$

The power P_k per frequency bin can be shown [9] to be exponentially distributed and the phase θ_k is uniformly distributed between 0 and 2π. The density function for P_k can be written as

$$F(P_k) = \frac{1}{S_k + W}\exp\left(\frac{-P_k}{S_k + W}\right) \tag{B-4}$$

where S_k is the signal power per frequency bin and W is the white noise power per frequency bin. Samples from a uniform density function are easily generated on most computers. A transformation from a uniformly distributed random variable to an exponentially distributed random variable as given in Equation B-4 is needed

in order to generate samples of P_k. The transformation is obtained by equating probabilities over the domains of the two density functions. This is expressed as

$$\int_0^{P_k} \frac{1}{S_k + W} \exp\left(\frac{-P_k}{S_k + W}\right) dP_k = \int_0^{X_k} dX_k \, . \tag{B-5}$$

Computing the integrals and solving for P_k yields

$$P_k = -(S_k + N) \ln(1 - X_k) \, . \tag{B-6}$$

The shape of the signal spectrum, S_k, is arbitrary and for the purposes of this dissertation is chosen to be Gaussian shaped. The signal and noise powers can be combined to form the signal-to-noise ratio (SNR)

$$SNR = \frac{\sum_{k=0}^{N-1} S_k}{NW} \, . \tag{B-7}$$

Appendix C

The Averaged Normalized Error

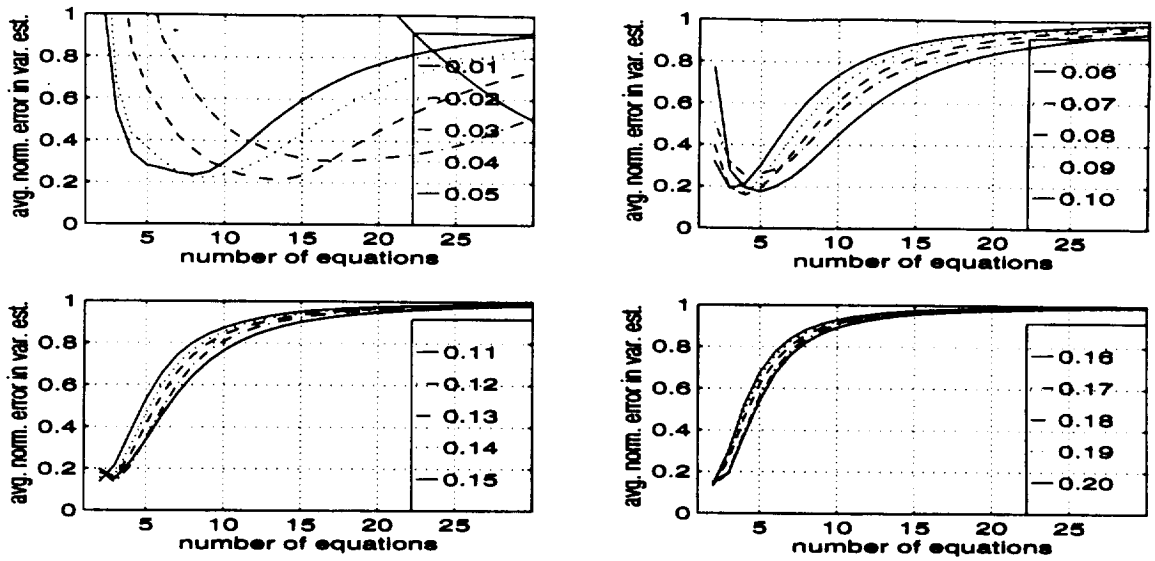


Figure C-1. The averaged normalized error in the variance estimate for the 10 dB SNR case using the [0 1] estimator in an overdetermined system.

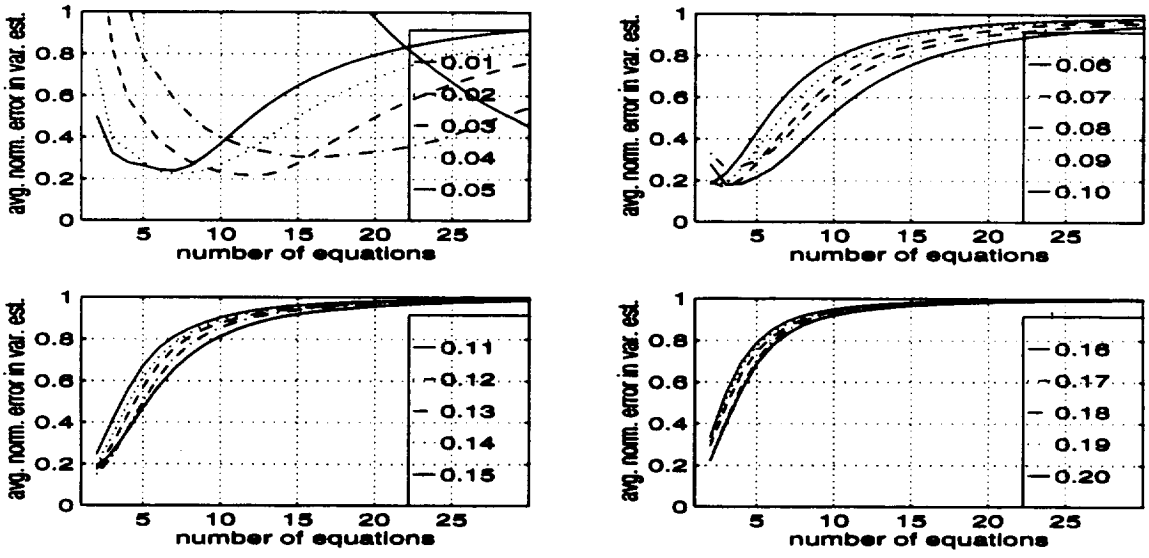


Figure C-2. The averaged normalized error in the variance estimate for the 10 dB SNR case using the [1 2] estimator in an overdetermined system.

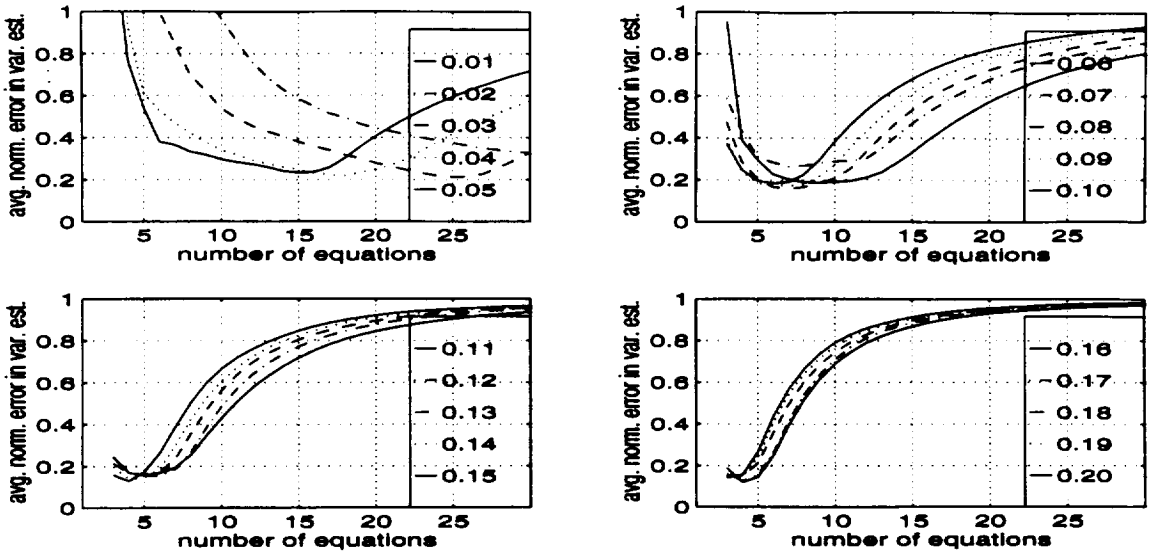


Figure C-3. The averaged normalized error in the variance estimate for the 10 dB SNR case using the [0 1 2] estimator in an overdetermined system.

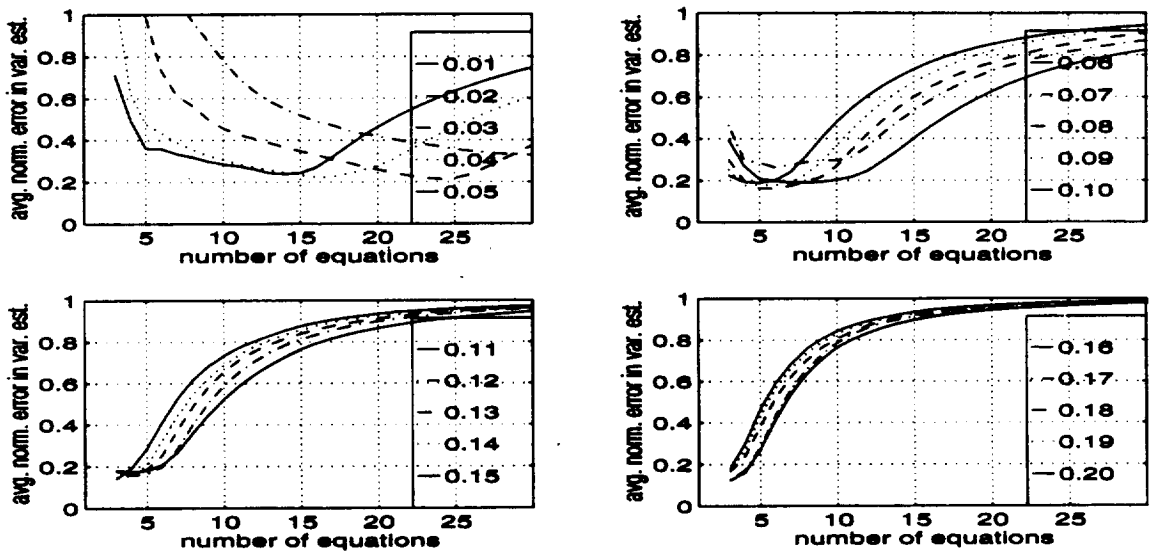


Figure C-4. The averaged normalized error in the variance estimate for the 10 dB SNR case using the [1 2 3] estimator in an overdetermined system.

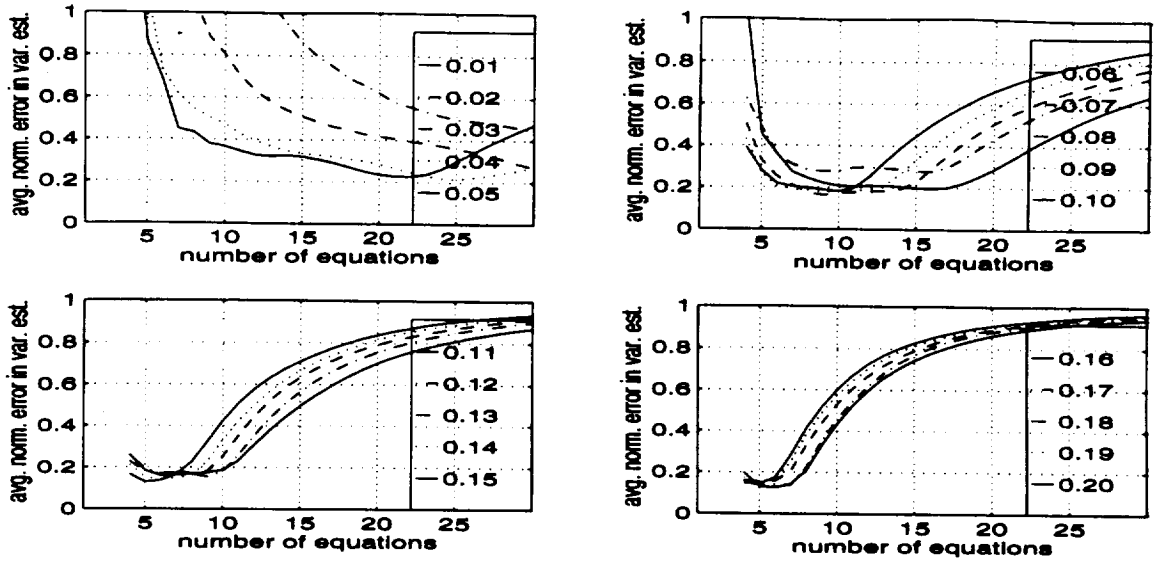


Figure C-5. The averaged normalized error in the variance estimate for the 10 dB SNR case using the [0 1 2 3] estimator in an overdetermined system.

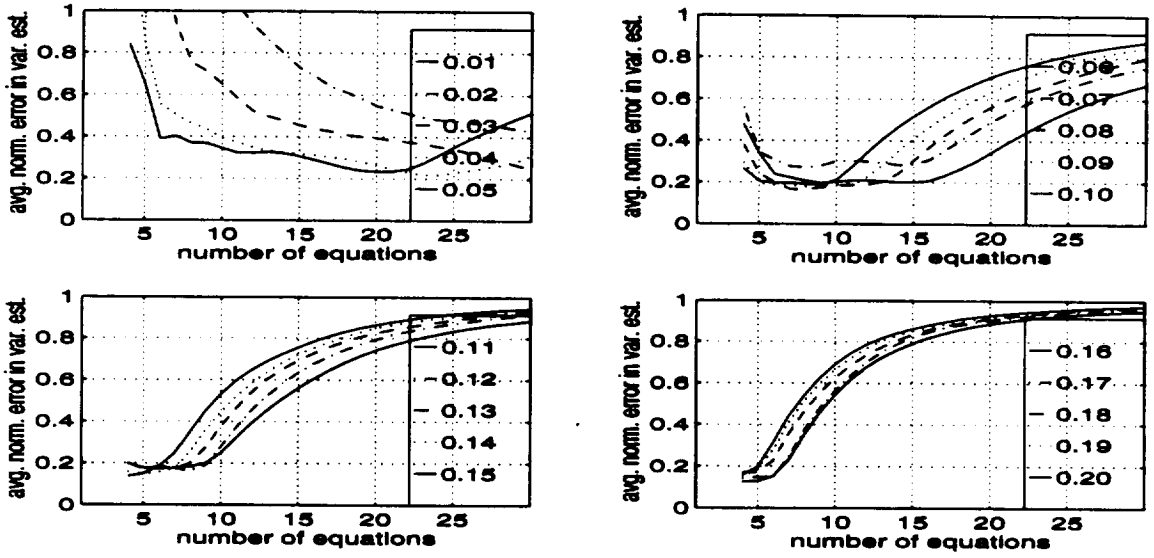


Figure C-6. The averaged normalized error in the variance estimate for the 10 dB SNR case using the [1 2 3 4] estimator in an overdetermined system.

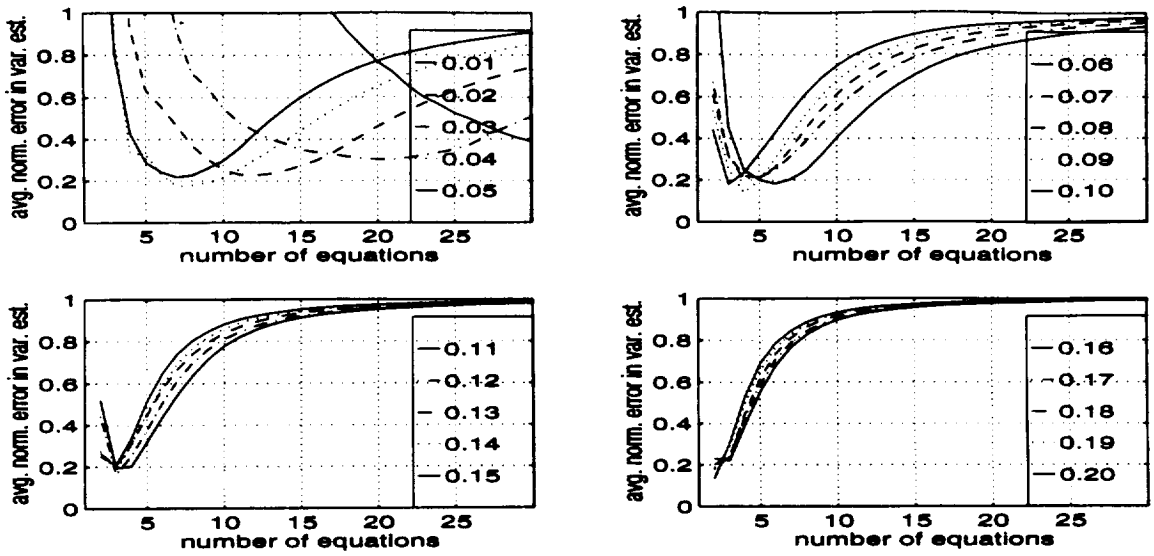


Figure C-7. The averaged normalized error in the variance estimate for the 5 dB SNR case using the [0 1] estimator in an overdetermined system.

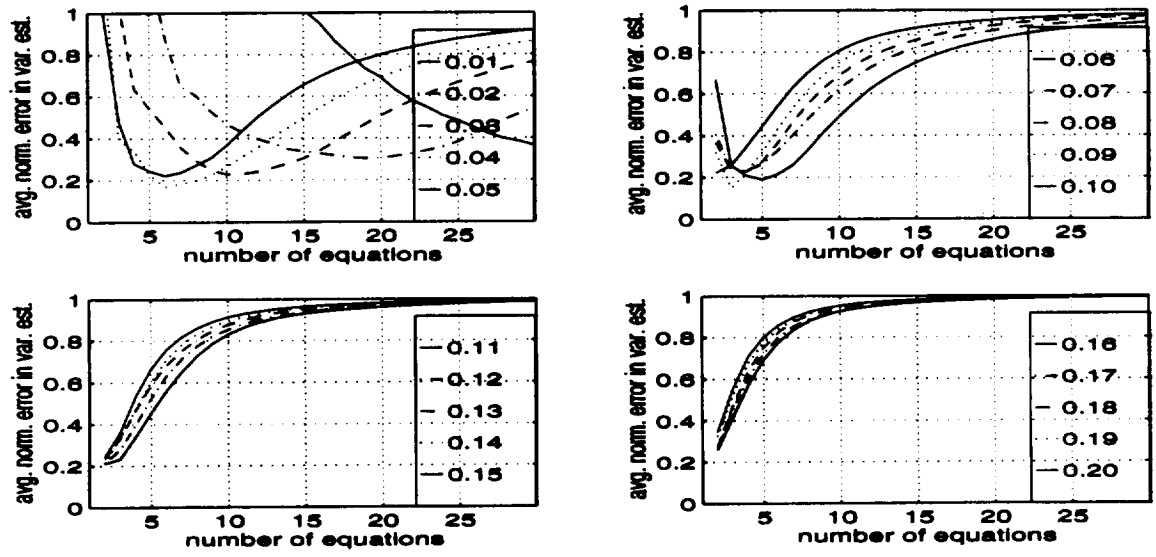


Figure C-8. The averaged normalized error in the variance estimate for the 5 dB SNR case using the [1 2] estimator in an overdetermined system.

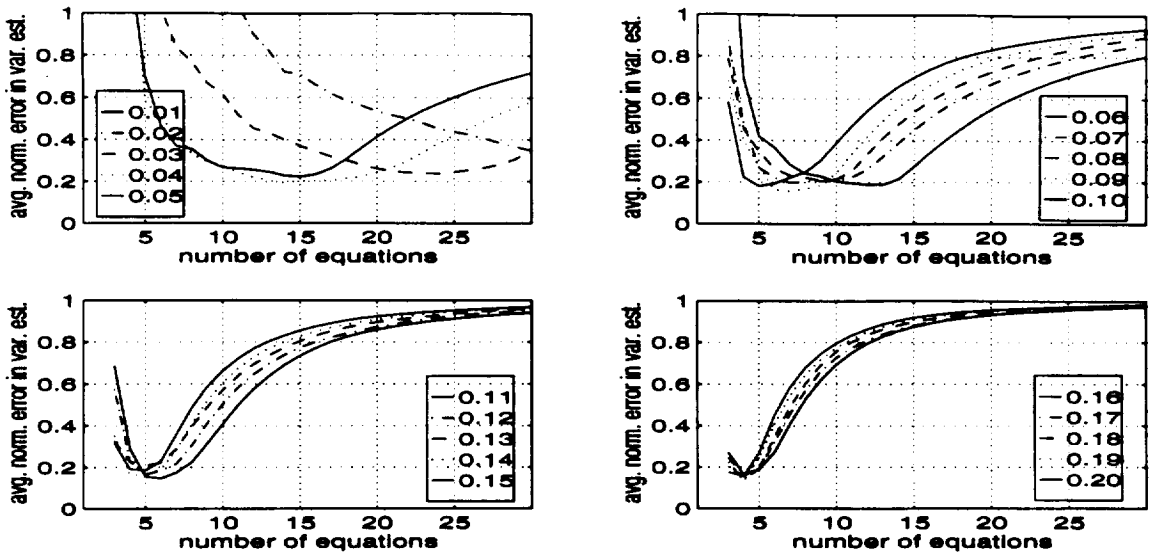


Figure C-9. The averaged normalized error in the variance estimate for the 5 dB SNR case using the [0 1 2] estimator in an overdetermined system.

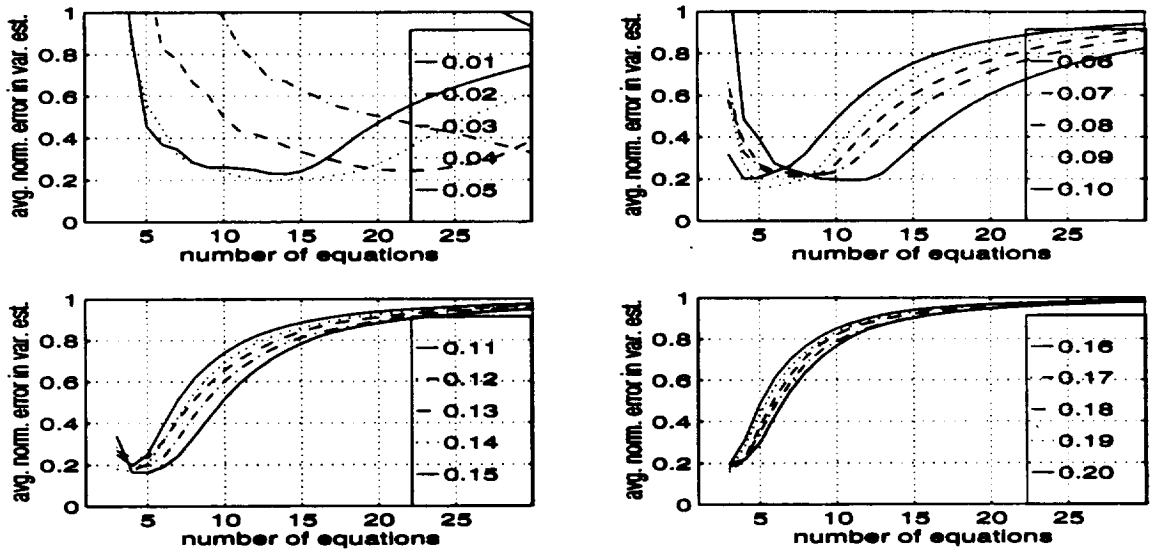


Figure C-10. The averaged normalized error in the variance estimate for the 5 dB SNR case using the [1 2 3] estimator in an overdetermined system.

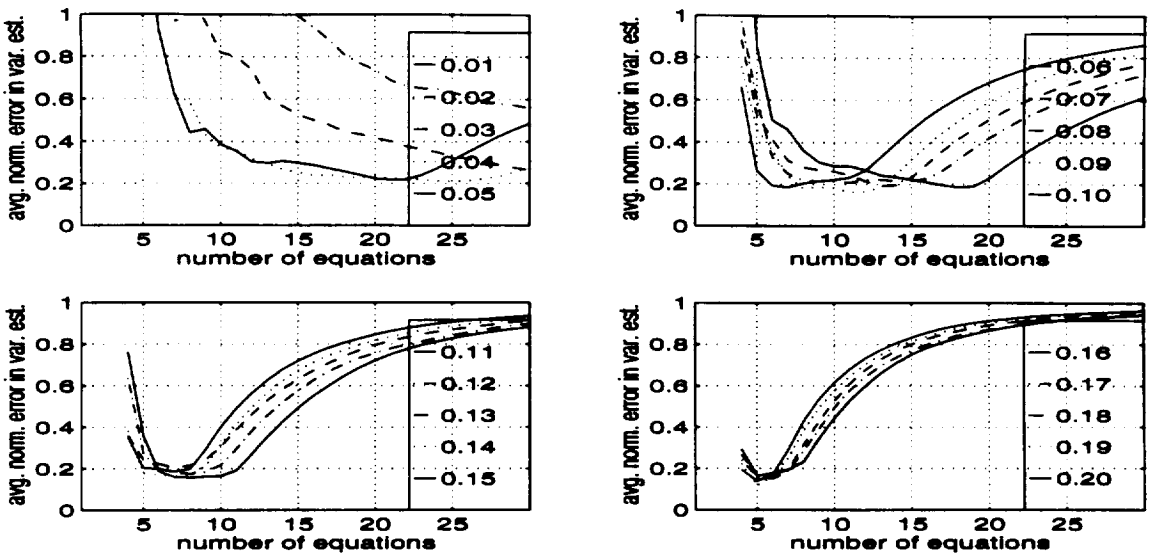


Figure C-11. The averaged normalized error in the variance estimate for the 5 dB SNR case using the [0 1 2 3] estimator in an overdetermined system.

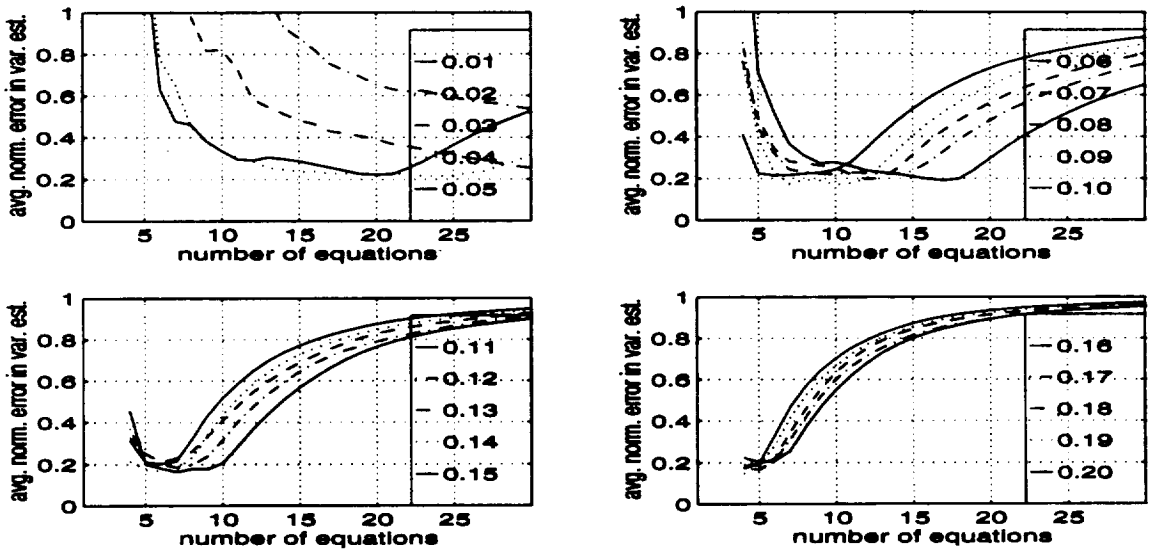


Figure C-12. The averaged normalized error in the variance estimate for the 5 dB SNR case using the [1 2 3 4] estimator in an overdetermined system.

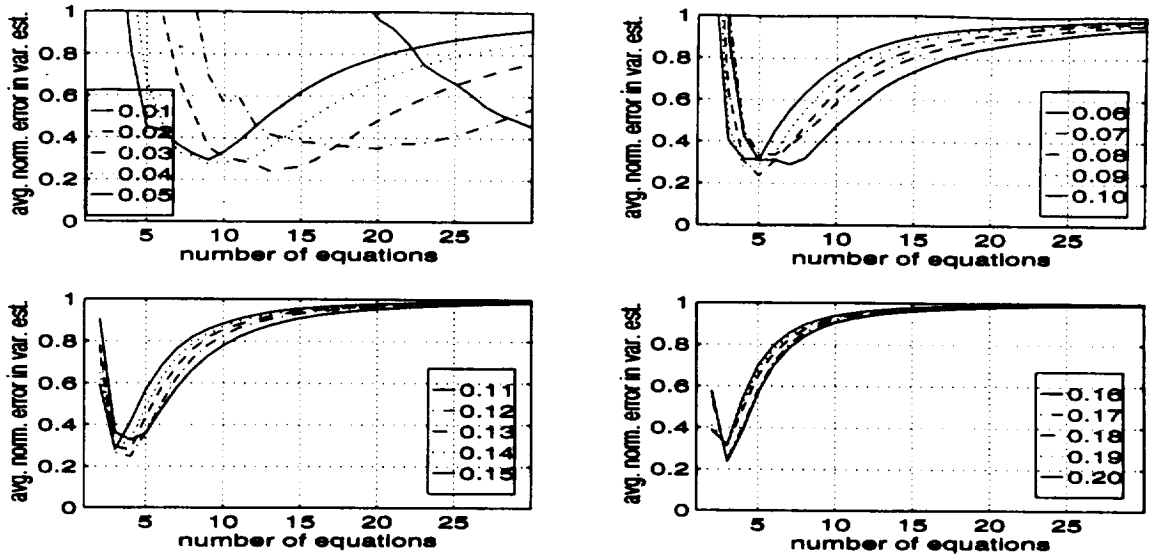


Figure C-13. The averaged normalized error in the variance estimate for the 0 dB SNR case using the [0 1] estimator in an overdetermined system.

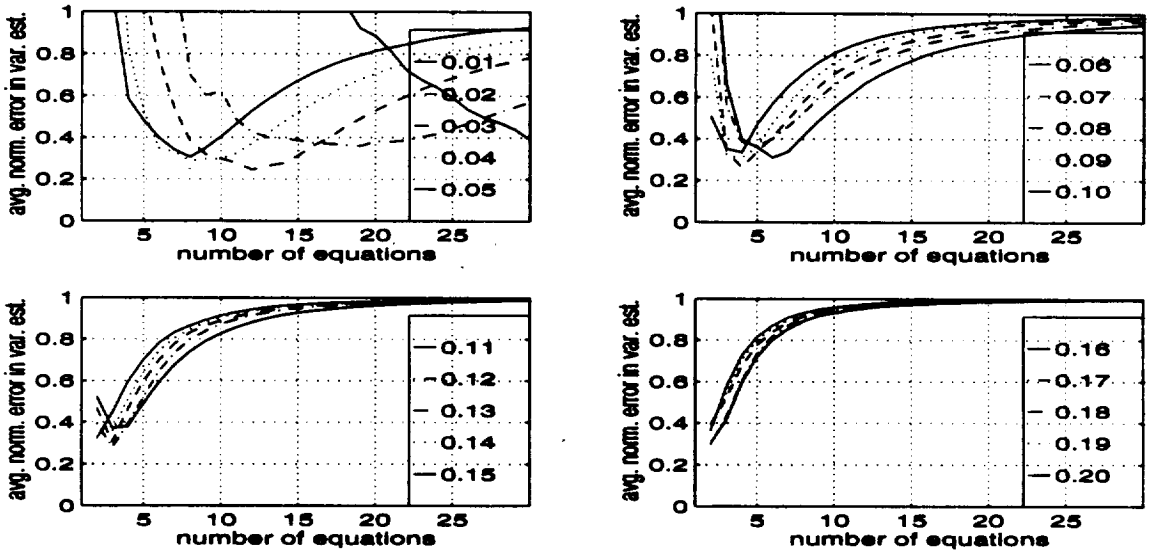


Figure C-14. The averaged normalized error in the variance estimate for the 0 dB SNR case using the [1 2] estimator in an overdetermined system.

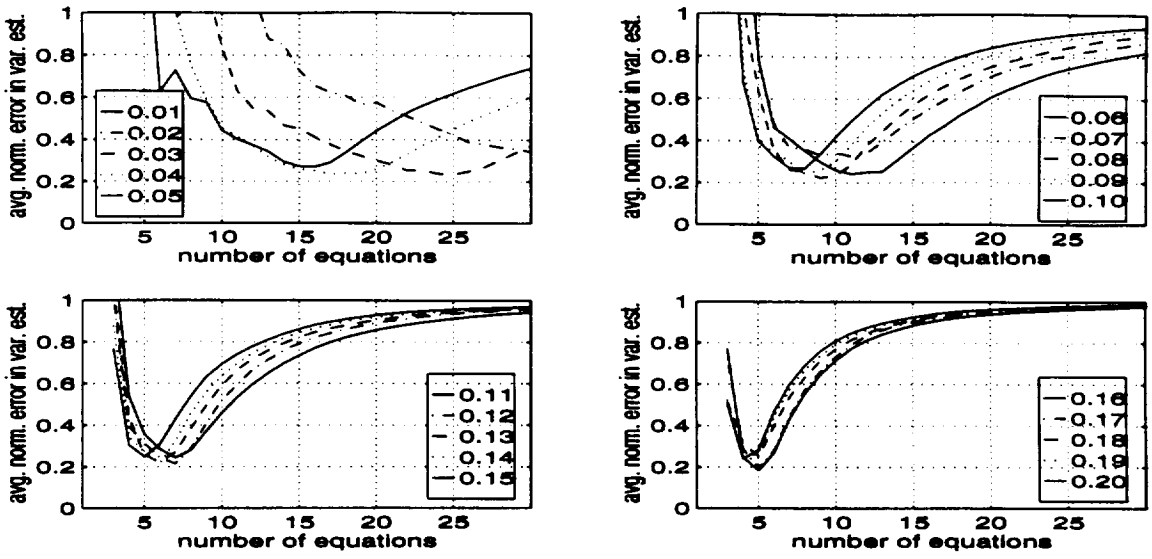


Figure C-15. The averaged normalized error in the variance estimate for the 0 dB SNR case using the [0 1 2] estimator in an overdetermined system.

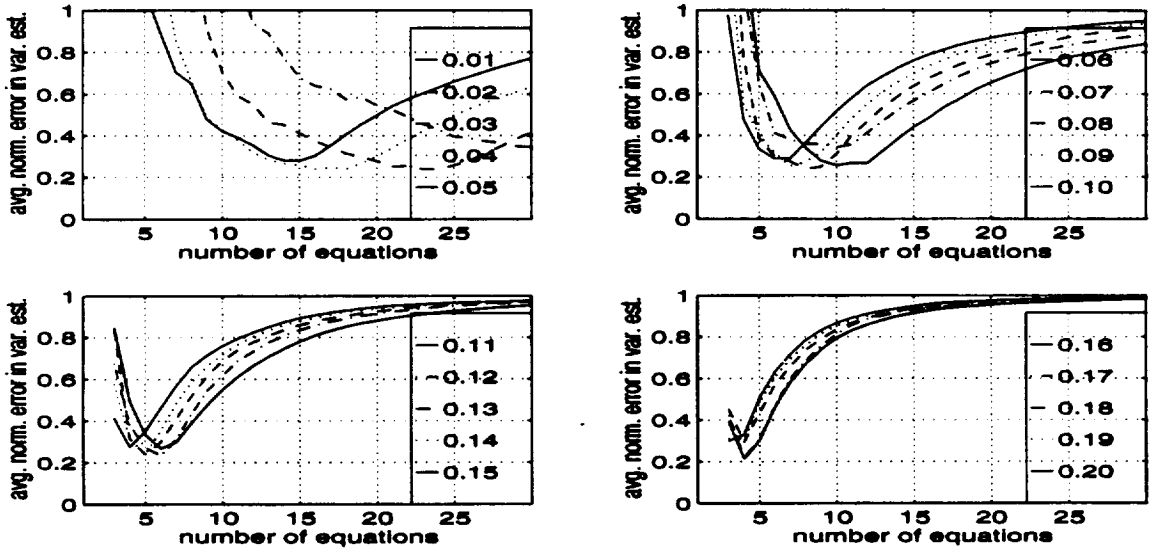


Figure C-16. The averaged normalized error in the variance estimate for the 0 dB SNR case using the [1 2 3] estimator in an overdetermined system.

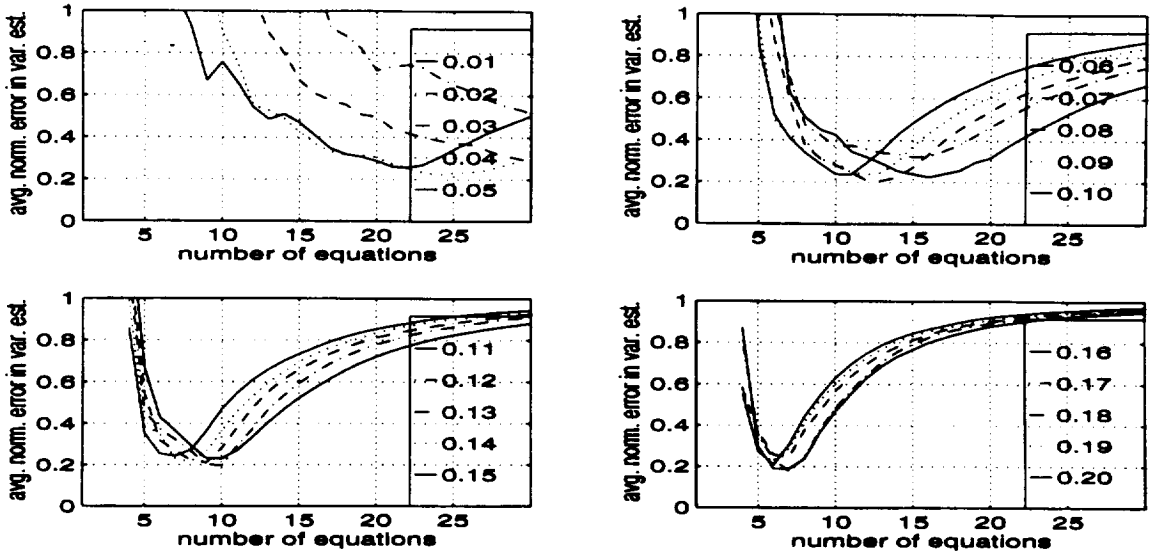


Figure C-17. The averaged normalized error in the variance estimate for the 0 dB SNR case using the [0 1 2 3] estimator in an overdetermined system.

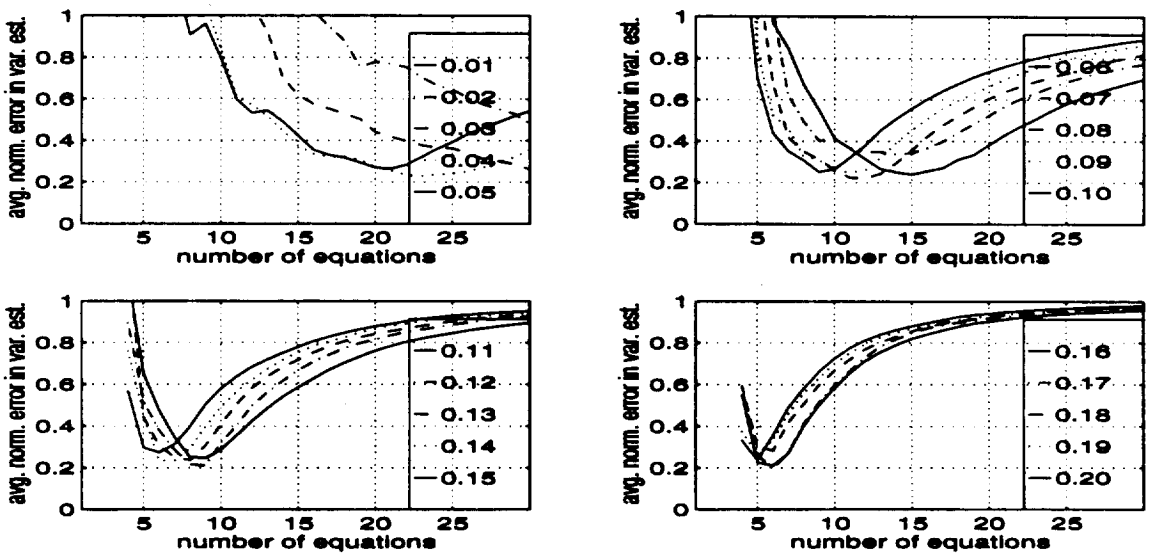


Figure C-18. The averaged normalized error in the variance estimate for the 0 dB SNR case using the [1 2 3 4] estimator in an overdetermined system.

Appendix D

The Standard Deviation in the Variance Estimate

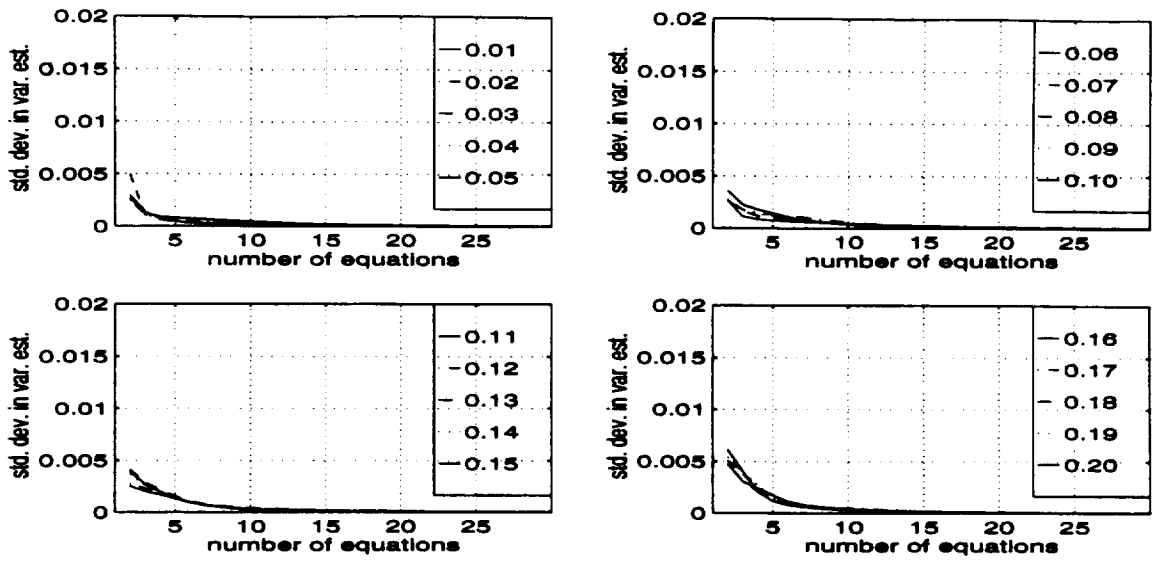


Figure D-1. The standard deviation in the variance estimate for the 10 dB SNR case using the [0 1] estimator in an overdetermined system.

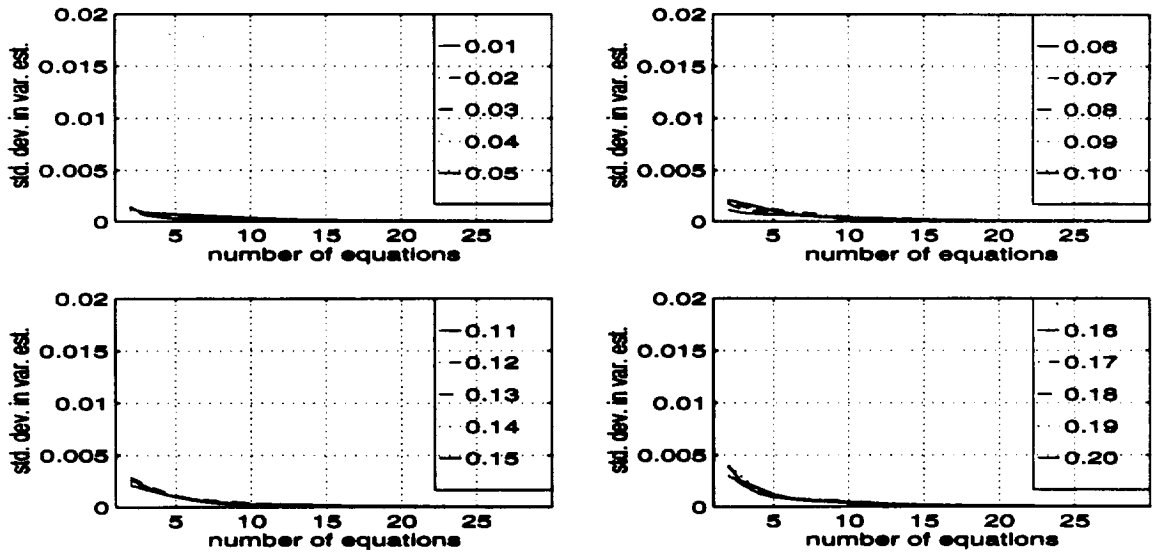


Figure D-2. The standard deviation in the variance estimate for the 10 dB SNR case using the [1 2] estimator in an overdetermined system.

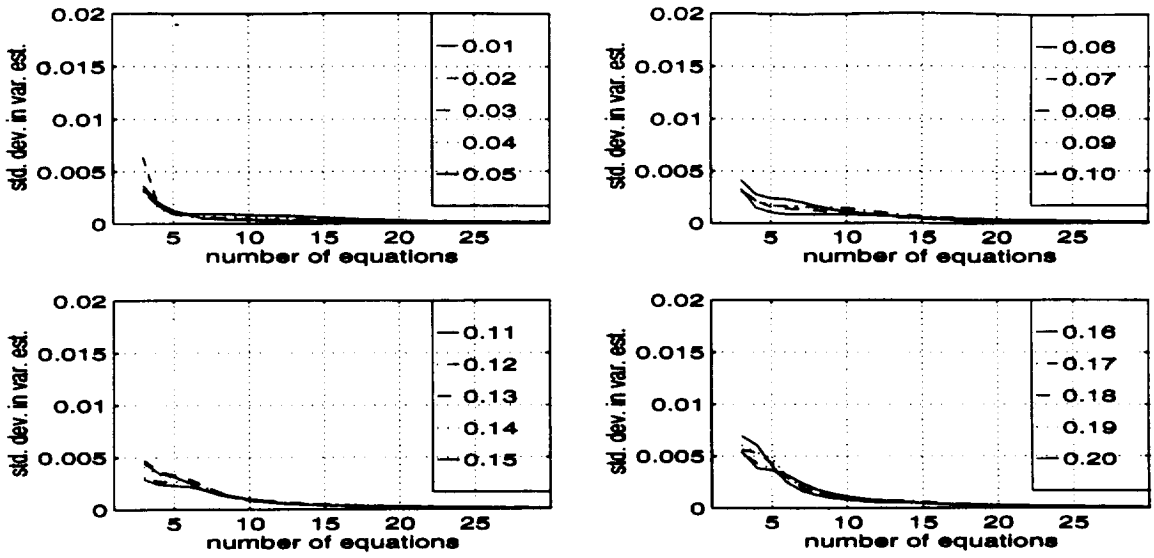


Figure D-3. The standard deviation in the variance estimate for the 10 dB SNR case using the [0 1 2] estimator in an overdetermined system.

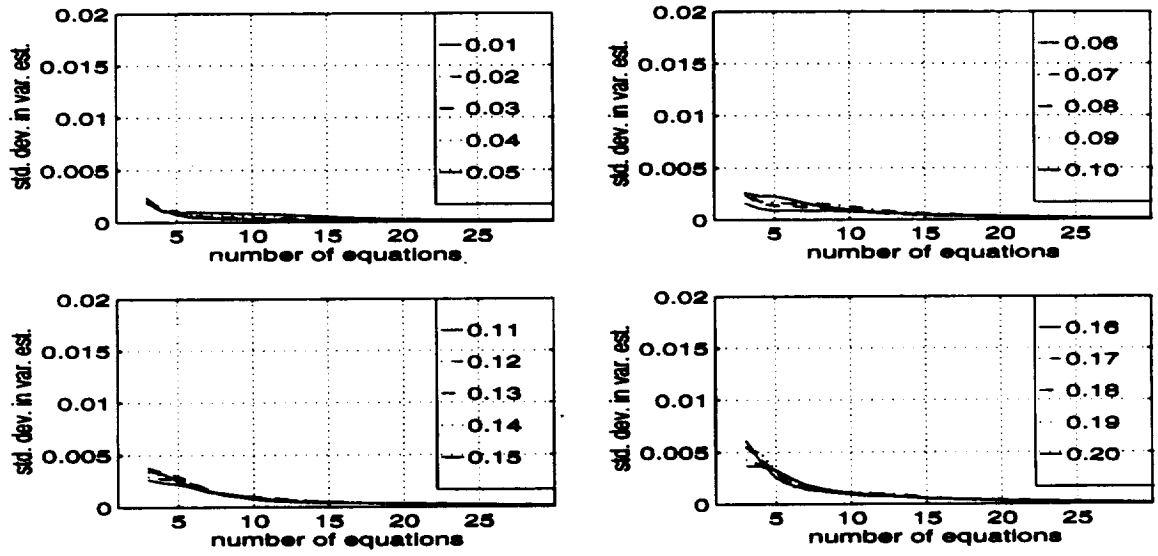


Figure D-4. The standard deviation in the variance estimate for the 10 dB SNR case using the [1 2 3] estimator in an overdetermined system.

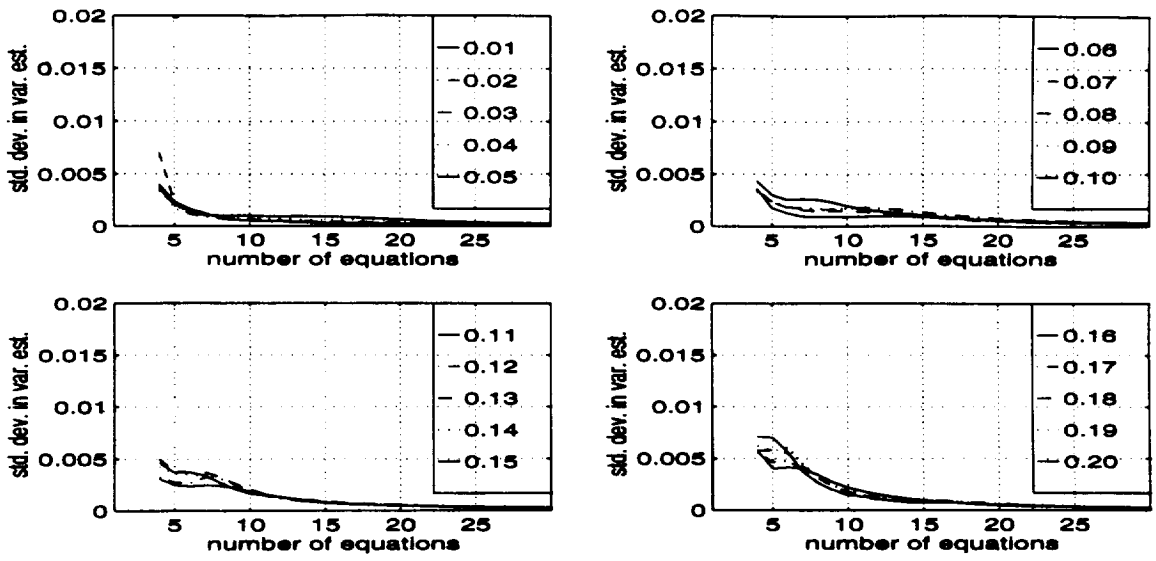


Figure D-5. The standard deviation in the variance estimate for the 10 dB SNR case using the [0 1 2 3] estimator in an overdetermined system.

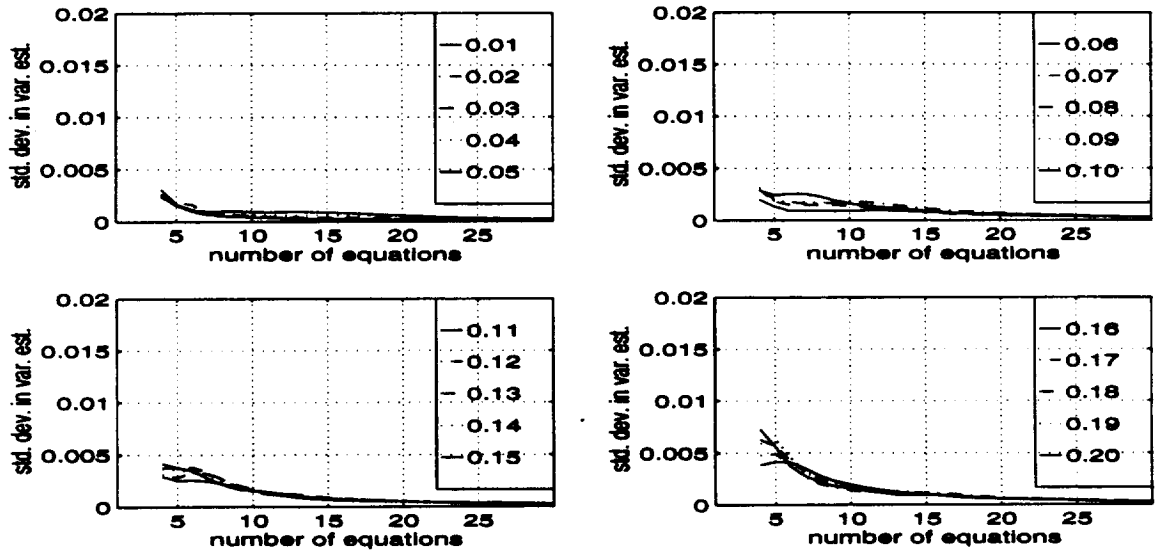


Figure D-6. The standard deviation in the variance estimate for the 10 dB SNR case using the [1 2 3 4] estimator in an overdetermined system.

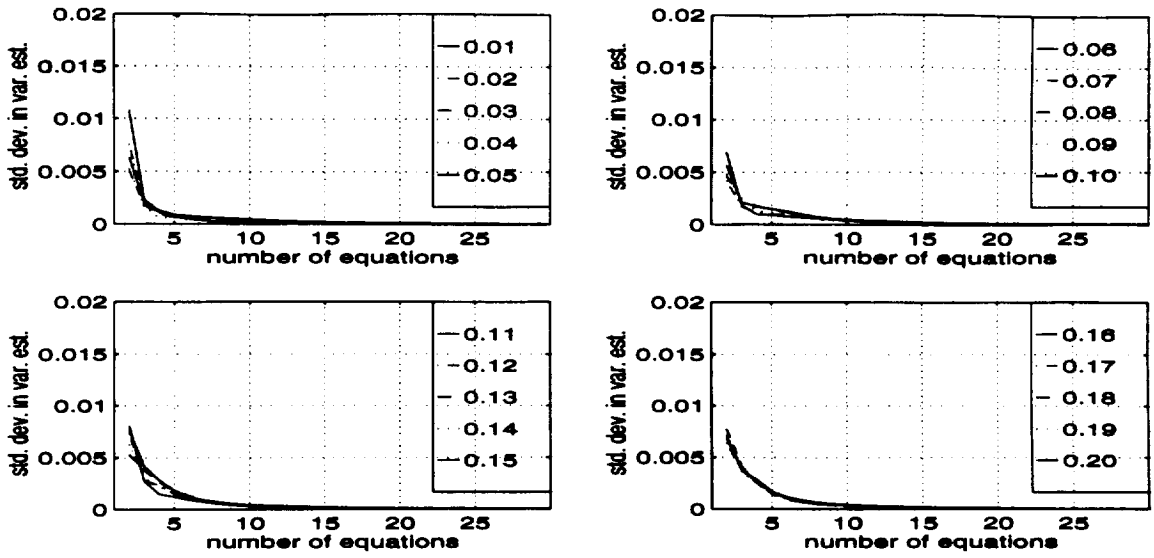


Figure D-7. The standard deviation in the variance estimate for the 5 dB SNR case using the [0 1] estimator in an overdetermined system.

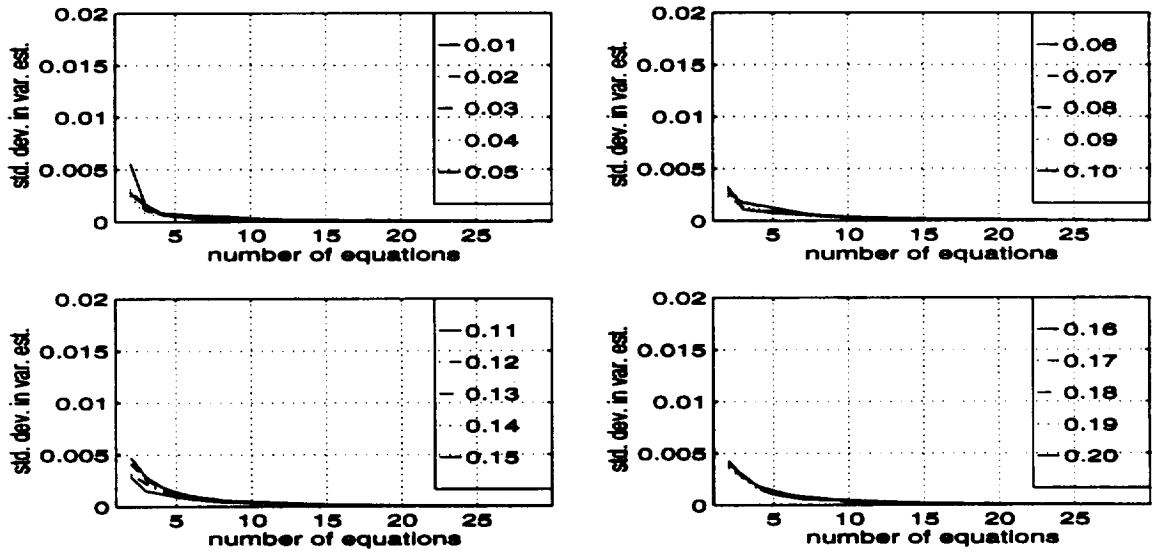


Figure D-8. The standard deviation in the variance estimate for the 5 dB SNR case using the [1 2] estimator in an overdetermined system.

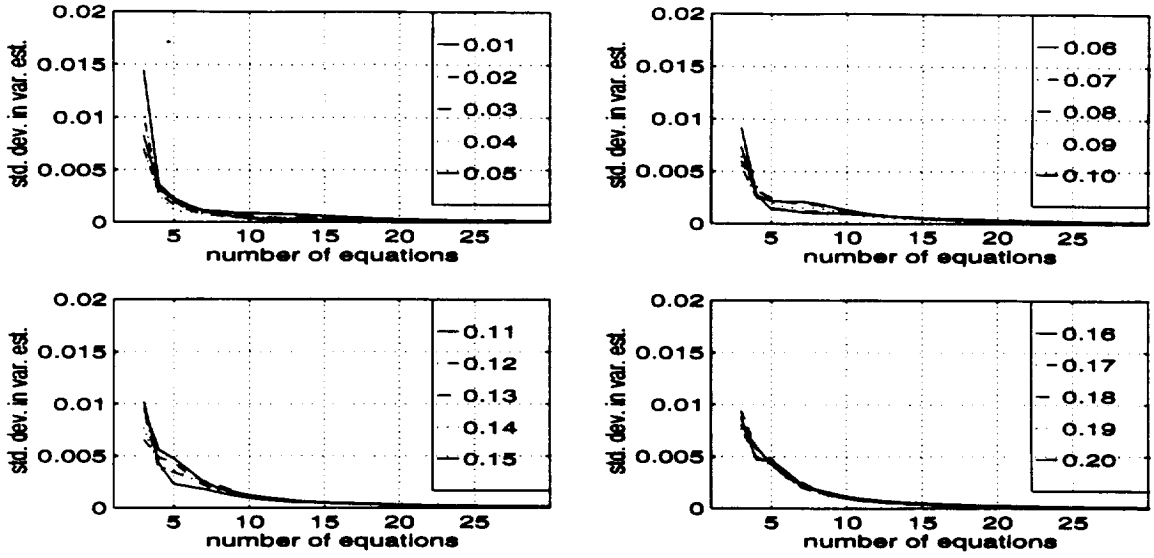


Figure D-9. The standard deviation in the variance estimate for the 5 dB SNR case using the [0 1 2] estimator in an overdetermined system.

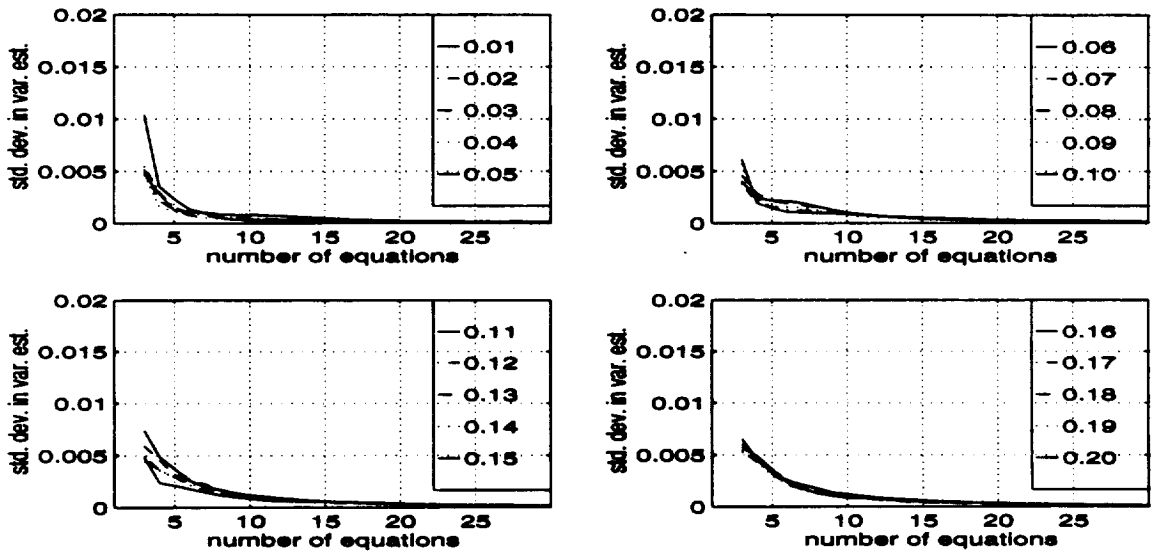


Figure D-10. The standard deviation in the variance estimate for the 5 dB SNR case using the [1 2 3] estimator in an overdetermined system.

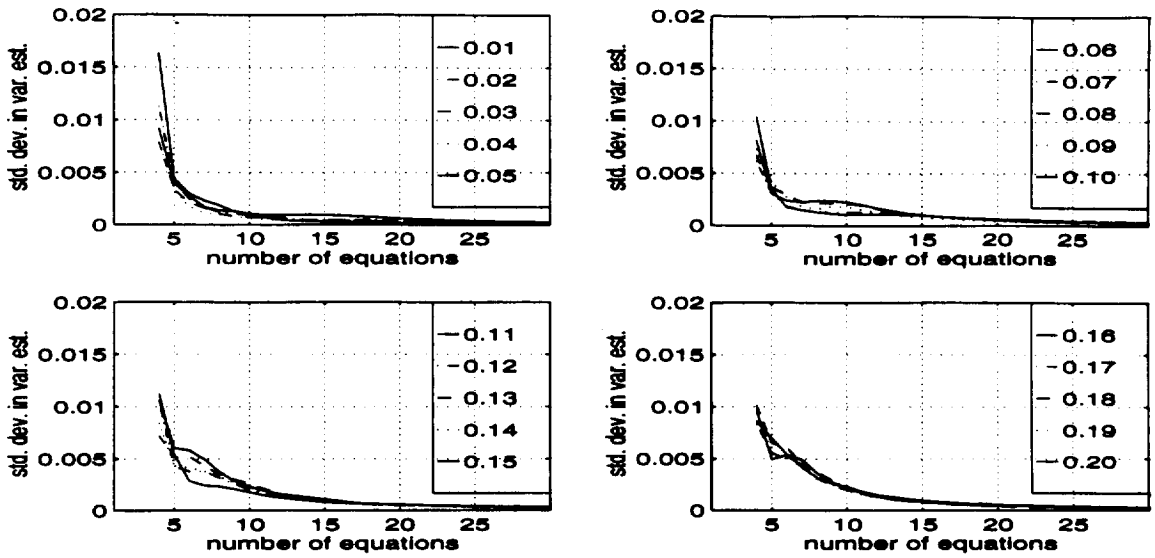


Figure D-11. The standard deviation in the variance estimate for the 5 dB SNR case using the [0 1 2 3] estimator in an overdetermined system.

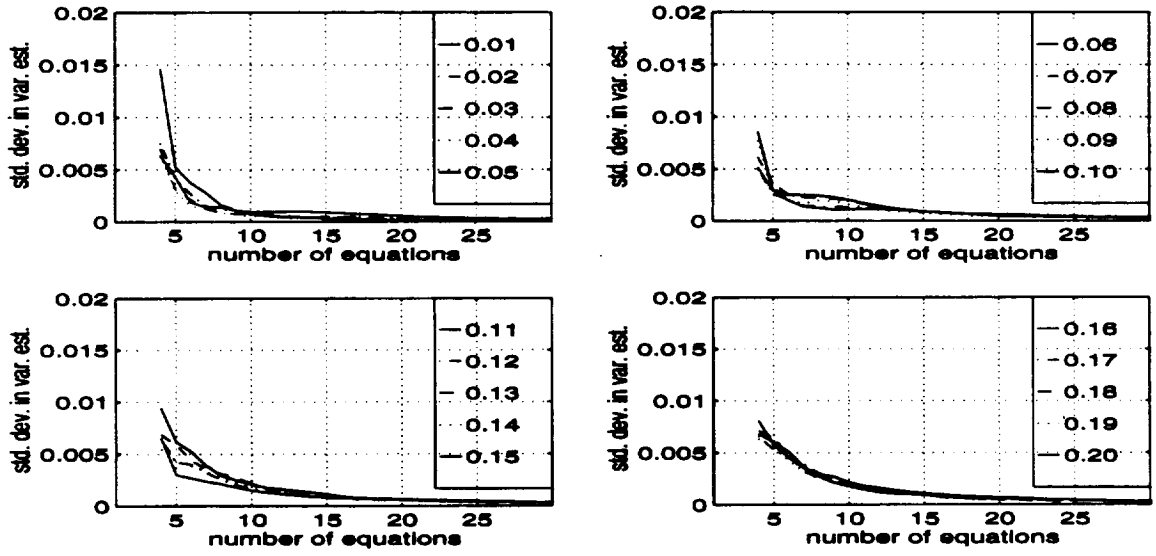


Figure D-12. The standard deviation in the variance estimate for the 5 dB SNR case using the [1 2 3 4] estimator in an overdetermined system.

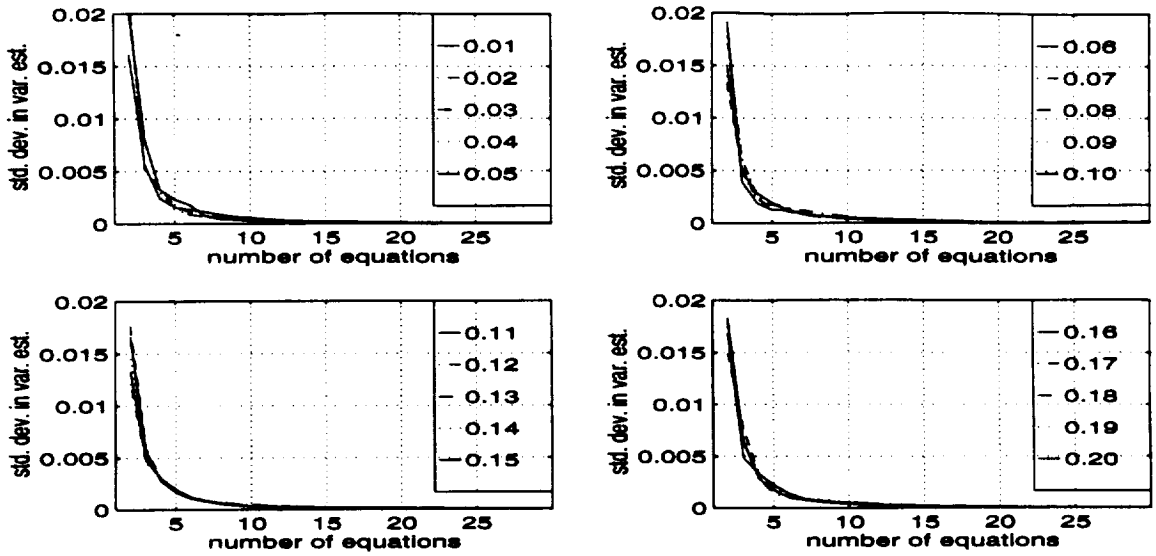


Figure D-13. The standard deviation in the variance estimate for the 0 dB SNR case using the [0 1] estimator in an overdetermined system.

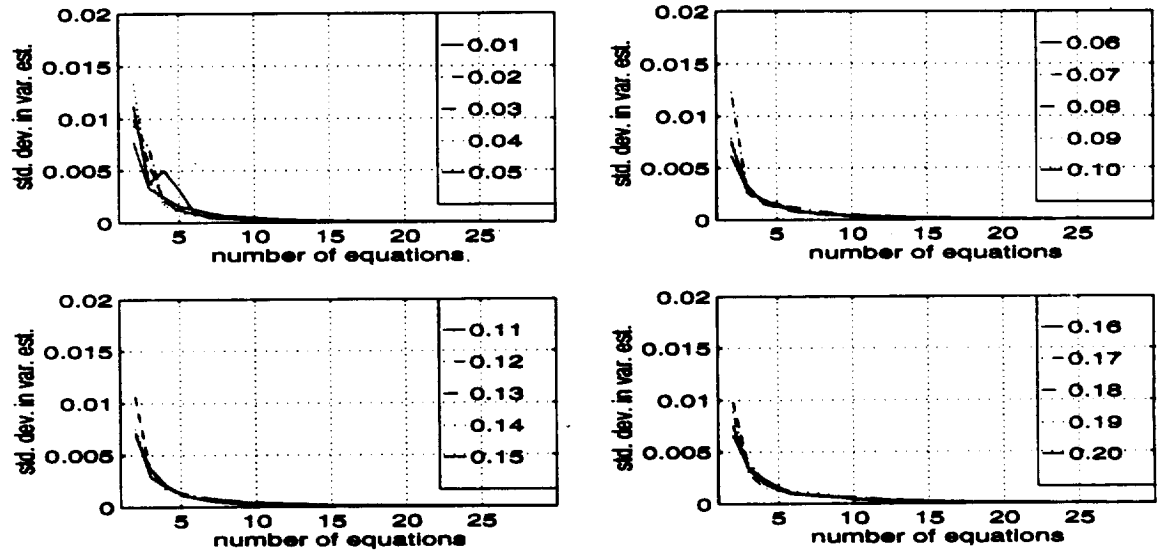


Figure D-14. The standard deviation in the variance estimate for the 0 dB SNR case using the [1 2] estimator in an overdetermined system.

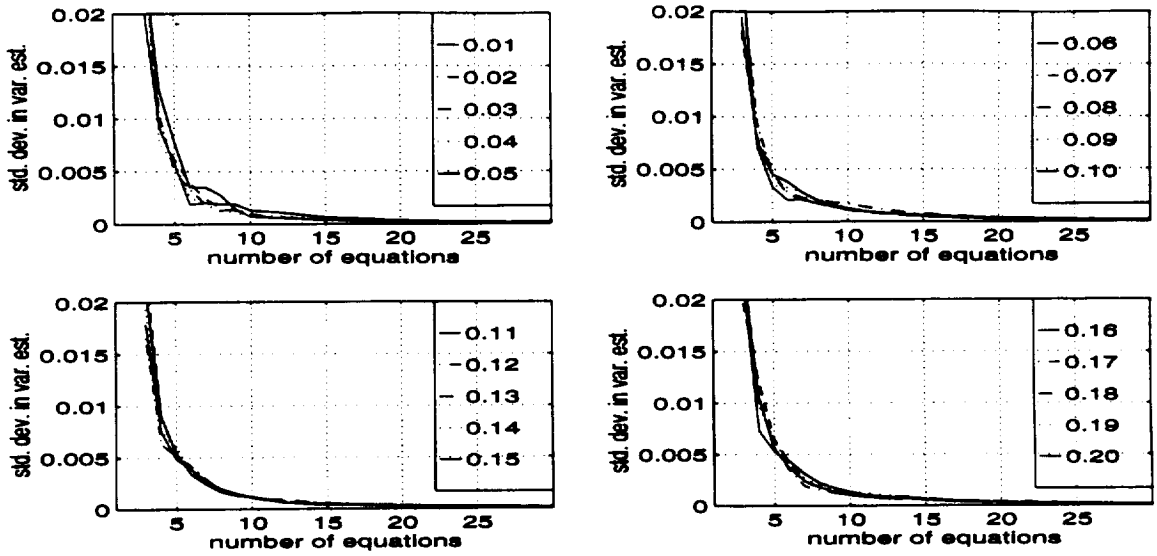


Figure D-15. The standard deviation in the variance estimate for the 0 dB SNR case using the [0 1 2] estimator in an overdetermined system.

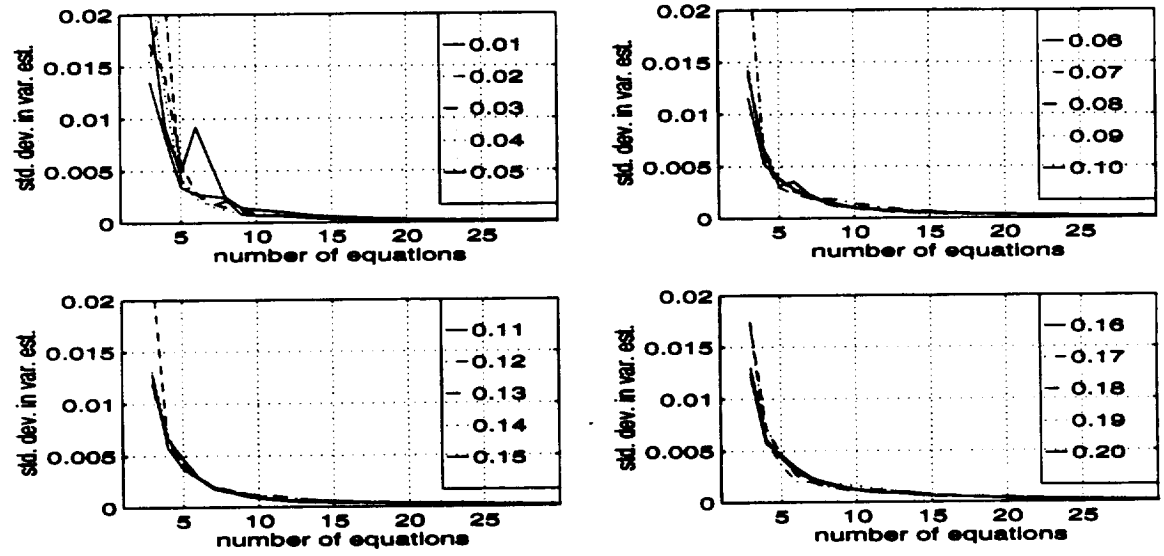


Figure D-16. The standard deviation in the variance estimate for the 0 dB SNR case using the [1 2 3] estimator in an overdetermined system.

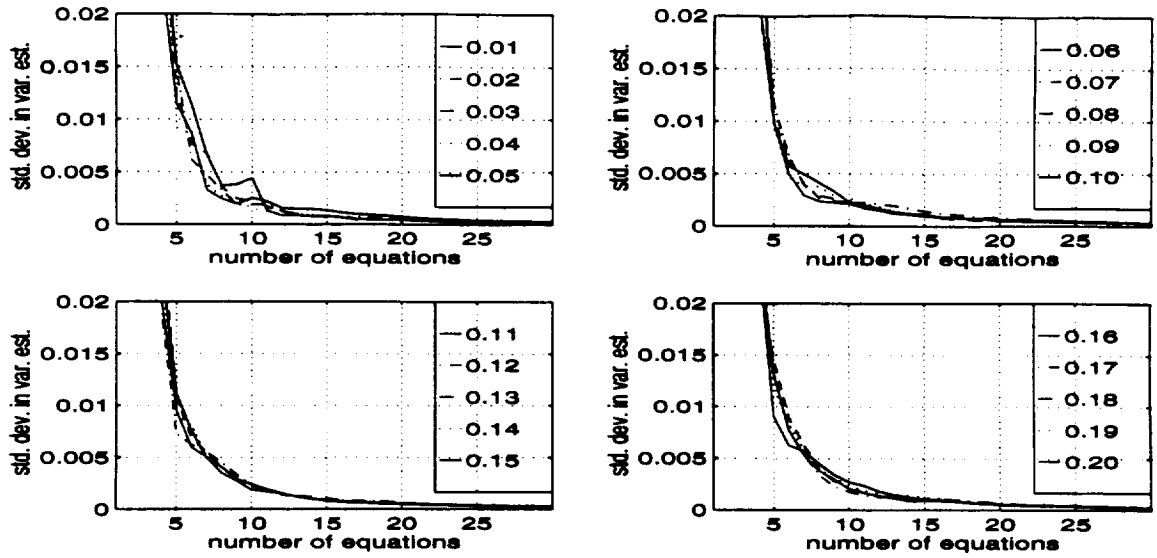


Figure D-17. The standard deviation in the variance estimate for the 0 dB SNR case using the [0 1 2 3] estimator in an overdetermined system.

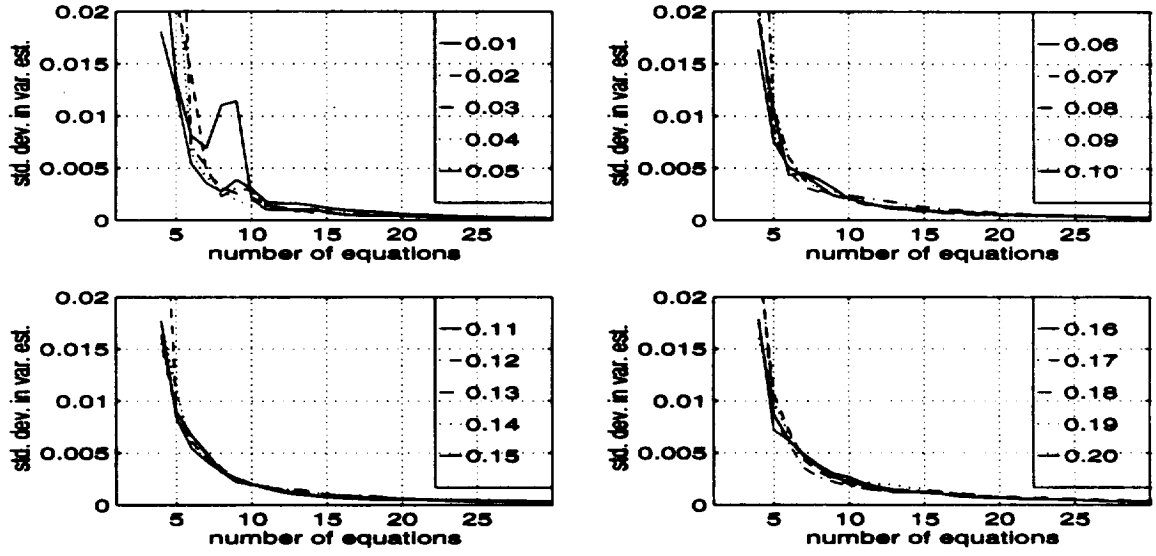


Figure D-18. The standard deviation in the variance estimate for the 0 dB SNR case using the [1 2 3 4] estimator in an overdetermined system.

REFERENCES

1. J. Bilbro, G. Fichtl, D. Fitzjarrald, M. Krause, and R. Lee. Airborne doppler lidar wind field measurements. *Bulletin American Meteorological Society*, 65:348–359, April 1984.

2. Ronald N. Bracewell. *The Fourier Transform and Its Applications.* McGraw-Hill, New York, second, revised edition, 1986.

3. Stephen P. Bruzzone and M. Kaveh. Information tradeoffs in using the sample autocorrelation function in arma parameter estimation. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, ASSP-32(4):701–715, August 1984.

4. James A. Cadzow. Spectral estimation: An overdetermined rational model equation approach. *Proceedings of the IEEE*, 70(9):907–939, September 1982.

5. Y. T. Chan and Randy P. Langford. Spectral estimation via the high-order yule-walder equations. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 30(5):689–698, October 1982.

6. Ruel V. Churchill and James Ward Brown. *Complex Variables and Applications.* McGraw-Hill, New York, fourth edition, 1984.

7. Charles G. Cullen. *Matrices and Linear Transformations.* Dover Publications, Inc., New York, second edition, 1972.

8. Biswa Nath Datta. *Numerical Linear Algebra and Applications.* Brooks/Cole Publishing Company, Pacific Grove, CA 93950, 1995.

9. Richard J. Doviak and Dusan S. Zrnic. *Doppler Radar and Weather Observations.* Academic Press, Inc., 1984. page 83.

10. Monson H. Hayes. *Statistical Digital Signal Processing and Modeling.* John Wiley and Sons, Inc., New York, 1996.

11. Simon Haykin, Brian W. Currie, and Stanislav B. Kesler. Maximum-entropy spectral analysis of radar clutter. *Proceedings of the IEEE*, 70(9):953–962, September 1982.

12. Steven M. Kay. The effects of noise on the autoregressive spectral estimator. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, ASSP-27(5):478–485, October 1979.

13. Steven M. Kay. Noise compensation for autoregressive spectral estimates. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, ASSP-28(3):292–303, June 1980.

14. Steven M. Kay. *Modern Spectral Estimation, Theory and Applications.* Prentice-Hall, Englewood Cliffs, New Jersey 07632, 1988.

15. B. M. Keel and E. G. Baxa Jr. An overdetermined approach to autocorrelation based spectral moment estimators for use in doppler weather radar. In *Proceedings of the 1997 IEEE National Radar Conference*, Syracuse, NY, May 1997. IEEE Aerospace and Electronics System Society. To Be Presented.

16. A. N. Kolmogorov. Interpolation und extrpolation von stationaren zufalligen folgen. *Bulletin of Academy of Science USSR Ser. Math*, 5:3–14, 1941.

17. Richard J. Larsen and Morris L. Marx. *An Introduction to Mathematical Statistics and Its Applications.* Prentice-Hall, Englewood Cliffs, New Jersey 07632, second edition, 1986.

18. Jae S. Lim and Alan V. Oppenmiem. Enhancement and bandwidth compression of noisy speech. *Proceedings of the IEEE*, 67(12):1586–1604, December 1979.

19. Pravas R. Mahapatra and Dusan S. Zrnic. Practical algorithms for mean velocity estimation in pulse doppler weather radars using a small number of samples. *IEEE Transactions on Geoscience and Remote Sensing*, GE-21(4):491–501, October 1983.

20. Kenneth S. Miller and Marvin W. Rochwarger. A covariance approach to spectral estimation. *IEEE Transactions on Information Theory*, IT-18(5):588–596, September 1972.

21. Leslie M. Novak and Nathan E. Lindgren. Maximum likelihood estimation of spectral parameters using burst waveforms. In *16th ASILOMAR*, pages 450–456, 1982.

22. Alan V. Oppenheim and Ronald W. Schafer. *Discrete-Time Signal Processing.* Prentice-Hall, Englewood Cliffs, New Jersey 07632, 1989.

23. Athanasios Papoulis. *The Fourier Integral and Its Applications.* McGraw-Hill, New York, 1962.

24. Athanasios Papoulis. *Probability, Random Variables, and Stochastic Processes.* McGraw-Hill, New York, 1965.

25. Richard E. Passarelli Jr. Parametric esimation of doppler spectral moments: An alternative ground clutter rejection technique. *Journal of Climate and Applied Meteorlogy*, pages 850–857, May 1983.

26. Richard E. Passarelli Jr. and Alan D. Siggia. The autocorrelation function and spectral moments: Geometric and asymptotic interpretations. *Journal of Climate and Applied Meteorology*, 22(22):1776–1787, October 1983.

27. P. K. Rastogi and R. F. Woodman. Mesospheric studies using the jicamarca incoherent-scatter radar. *Journal of Atmospheric and Terrestrial Physics*, 36:1217–1231, 1974.

28. W. D. Rummler. Introduction of a new estimator for velocity spectral parameters. Technical Memo MM-68-4121, Bell Telephone Labs, 1968.

29. Jr. S. Lawrence Marple. *Digital Spectral Analysis and Applications.* Prentice-Hall, Englewood Cliffs, New Jersey 07632, 1987.

30. E. B. Saff and A. D. Snider. *Fundamentals of Complex Analysis for Mathematics, Science, and Engineering.* Prentice-Hall, Englewood Cliffs, New Jersey 07632, second edition, 1993.

31. Dale Sirmans and Bill Bumgarner. Estimation of spectral density mean and variance by covariance argument techniques. *16th Radar Meteorological Conference*, pages 6–13.

32. Dale Sirmans and Bill Bumgarner. Numerical comparison of five mean frequency estimators. *Journal of Applied Meteorolgy*, 14:991–1001, September 1975.

33. R. C. Srivastava and A. R. Jameson. Time-domain computation of mean and variance of doppler spectra. *Journal of Applied Meteorology*, 18:189–194, February 1979.

34. Petre Stocia, Soderstrom Torsten, and Funan Ti. Asymptotic properties of the high-order yule-walker estimates of sinusoidal frequencies. *IEEE Transactions ASSP*, 37:1721–1734, November 1989.

35. R. G. Strauch, R. W. Lee, and et. al. Improved doppler velocity estimates by the poly-pulse-pair mehtod. In *18th Conference on Radar Meteorology*, pages 376–380, Atlanta, GA, March 1978. American Meteorology Society.

36. H. Wold. *A Study in the Analysis of Stationary Time Series.* Almqvist and Wiksell, Stockholm, 1954.

37. D. S. Zrnic. Spectral moment estimators from correlated pulse pairs. *IEEE Transactions on Aerospace and Electronic Systems*, AES-13(4):344–354, July 1977.

38. D. S. Zrnic. Spectrum width estimators for weather echos. *IEEE Transactions on Aerospace and Electronic Systems*, AES-15(5):613–619, September 1979.

39. Dusan S. Zrnic. Simulation of weatherlike doppler spectra and signals. *Journal of Applied Meteorolgy*, 14:619–620, June 1975.

40. Dusan S. Zrnic. Estimation of spectral moments for weahter echos. *IEEE Transactions on Geoscience Electronics*, GE-17(4):113–128, October 1979.