THE EFFECT OF SEGMENT AVERAGING

ON THE QUALITY OF THE BURG SPECTRAL ESTIMATOR

by

Md. Anisur Rahman

Thesis submitted to the Faculty of the

Virginia Polytechnic Institute and State University

in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

in

Electrical Engineering

APPROVED:

A. A. (Louis) Beex, Chairman

I. M. Besieris

T. Pratt

July, 1984

Blacksburg, Virginia

ACKNOWLEDGEMENTS

I would like to express my sincere appreciation to Dr. A. A. (Louis) Beex for his thoughtful suggestions and well directed guidance that made this research work a success. I also wish to acknowledge that he had formulated this problem. His constant motivation during the course of this research endeavor gave me a great deal of encouragement.

I want to express my gratitude to the members of my graduate committee, Professor I. M. Besieris for his helpful advice and Dr. T. Pratt for checking the original draft of this thesis. I like to thank Dr. K-B. Yu for his suggestions regarding the improvement of this thesis.

Financial support from the Electrical Engineering Department of Virginia Polytechnic Institute and State University is thankfully acknowledged.

A special note of appreciation is due my parents for their love and encouraging support.

Md. A. Rahman, July 1984.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

Table

page

Chapter I

INTRODUCTION

Power spectral density (PSD) estimates of a given time series of data have been widely used in signal processing, optics, geophysics, bio-engineering, statistical mathematics, and economics. In the last two and a half decades numerous techniques and algorithms [1] have been devised for efficiently estimating the PSD. The PSD provides a frequency domain description of the second order statistics of a wide sense stationary process while the autocorrelation gives the corresponding time domain description. In fact, the PSD and the autocorrelation form a Fourier transform pair. A time series of infinite length is required to compute its PSD. Since an infinitely long data record is not available, we have to estimate the PSD.

There are numerous applications of PSD estimates [2, pp. 179]. Especially, in the field of signal processing, the power spectral density is an important tool in analyzing data, filter design, target tracking, and parameter estimation. Geophysicists use PSD estimates extensively to study the behavior of various geological phenomena, e.g. earthquakes, geomagnetic micropulsations, sunspot numbers, etc. Spectral estimates are useful to aeronautical design

engineers in suggesting how various parts of the aircraft structure need to be designed to minimize the risk of structural damage due to the buffeting of turbulant air. Statisticians estimate the power spectrum in model building, design of experiments, and frequency response studies of economic time series. Finally, we mention the use of PSD estimates as a diagonistic tool, when applied to the electroencephalogram (EEG) or electrocardiogram (ECG).

Current methods of spectrum estimation can be broadly grouped into two classes. The first one is the classical approach, which includes the periodogram method, autocorrelation methods and its variants (Bartlett, 1953, [3]; Grenander and Rosenblatt, 1957, [4]; Blackman and Tukey, 1958, [5]; Jenkins and Watts, 1968, [6]; Koopmans, 1974, [7]). The second one is based on parametric modelling. This includes the maximum entropy method or MEM (Burg, 1967, [8]), one step linear prediction (Parzen, 1969, [9]), and spectral estimation using ARMA models (Tretter and Steiglitz, 1967, [10]; Gutowski, Robinson and Treitel, 1978, [11]). Parametric spectral estimation exhibits a superior resolution property for short data records in comparison to the classical approach. But the classical approach is popular because the classical methods are fairly easy to implement and can be computed efficiently by using the Fast

Fourier Transform (FFT). There is a considerable body of literature on parametric spectral estimators ([12], [13]). Here we study only the MEM estimator.

The MEM estimator is also known as the Burg spectral estimator (BSE). In analogy with the Welch spectral estimation procedure [14], it was initially expected that the quality of the BSE could be improved by segmenting the available data record, applying the BSE to each segment, and a subsequent averaging of the relevant parameters or functions associated with each segment. The underlying idea was that the BSE was suitable for short data records, so that relatively little would be lost in terms of spectral resolution and hopefully much could be gained in terms of spectral estimation variance. We note that this approach should not work for large records as in that case the Burg estimator already performs close to the Cramer-Rao bound. The aim is specifically for spectral estimator quality improvement when applied to relatively short records. Motivated by this idea, the author investigated the effect of segment averaging on the quality of the BSE estimator.

The BSE is classified as an auto-regressive (AR) estimator. In Chapter 2 a tutorial review of the asymptotic properties of the AR spectral estimator (ARSPE) is given. Mainly the findings of Akaike [15], Kromer [16], Berk [17],

and Sakai [18] have been presented. The mean and variance for the modified Burg spectral estimator (MBSE) using several averaging techniques are derived in Chapter 3. The MBSE is a modified form of the BSE, where some averaging techniques are employed. The available N data points are segmented into M non-overlapping sections, so that each section has N/M data points. Three types of averaging technique are used with the MBSE. In the first method (AVA) the AR parameters, evaluated for each section, are averaged and the resulting average AR parameters yield an associated spectral density estimate. The second method (AVK) averages the reflection coefficients, the average of which leads to the corresponding spectral estimate. The final approach (AVP) evaluates the PSD estimate associated with each segment and then averages these directly. In this thesis, sometimes the reference to AVA, AVK or AVP will essentially mean the modified Burg spectral estimator using the corresponding averaging technique. In Chapter 3 the lower bound for the mean and variance of the reflection coefficients is also derived. Simulation results are presented and carefully studied in Chapter 4. The empirical and approximate theoretical variances of the respective estimation errors are compared against the theoretical Cramer-Rao lower bound (CRLB). The performances of the Welch

spectral estimator and the MBSE using averaging techniques are also compared. Finally, the conclusion of this investigation is given in Chapter 5. Some guidelines are given so that the results for the MBSE may be generalized to all other AR estimators.

Chapter II

ASYMPTOTIC PROPERTIES OF THE AR SPECTRAL ESTIMATOR (ARSPE)

2.1 INTRODUCTION

In time series analysis two approaches have dominated. One approach uses time domain analysis and the other uses the frequency domain. In the time domain approach some parametric model is postulated and this model is then fitted to the observed data by estimating its parameters. One of these parametric models is the AR or all-pole model. This model is widely used because the estimation of its parameters requires the solution of a system of linear equations only, which can be easily done. Moreover, it is efficient in representing the data because lower order models generally yield satisfactory results in fitting the data. Assuming the input of this model to be a zero mean white gaussian noise, the PSD estimate of the observed time series is given by the product of the variance of the white input noise and the magnitude squared of the model transfer function, evaluated on the unit circle in the z-domain. The AR model gives a good resolution of the power spectral density peaks.

Until now only the asymptotic properties of the ARSPE have been derived. The term 'asymptotic' implies that these properties provide a good prediction of the actual performance of the estimator when the amount of data is sufficiently large with respect to the order of the system. Initially, the error covariance matrices of the estimates of AR parameters using the Yule-Walker (YW) equations were derived [19, 20]. The derivation of the variance of the ARSPE estimates followed [15, 16, 17, 18, 21, 22].

The first classical paper on the asymptotic distribution of the estimates of the AR coefficients for a process of known order with independent innovations was presented by Mann and Wald (1943) [19]. Later, Anderson and Walker (1964) [20] gave the limit distribution of AR parameters from any linear stochastic process.

Akaike (1969) [15] has derived the variance of the ARSPE estimates in a limit distribution sense and his theorems (Theorem 1 and 2, [15]) are based on the assumption that the time series was produced by a finite autoregression. Then Kromer (1970) [16] presented significant results on the asymptotic behaviour of the ARSPE. Followed by this, Berk (1973) [17], Baggeroer (1976) [21], Huzzi (1977) [22] and Sakai (1978) [18] published papers on the asymptotic properties of the ARSPE estimates. In all of them, the

authors assumed some strict conditions on the data which are mentioned in the next section. Since their assumptions and approaches are different, the expressions of the variances of the PSD estimates are not the same. However, numerical evaluation of these expressions showed the same results.

The ARSPE class consists of a large number of spectral estimators, each of which estimates the AR coefficients in a different way. But asymptotically all these estimates converge to the same solution. So, by the words 'asymptotic properties of the ARSPE' we include the asymptotic statistical properties of all types of AR spectral estimators. In the following section we shall review some of the more significant work. We like to mention that the results shown in the next section will not be used in our investigation for we have dealt with a moderate number of data points. However, we feel it necessary to give the readers an overview of the results obtained for the asymptotic case.

2.2 VARIANCE OF THE ARSPE ESTIMATE

2.2.1 Akaike Derivation

Akaike [15] has considered a time series $\{x_t\}$ from a realization of a finite order autoregressive process defined by

$$x_t = - \sum_{m=1}^{p} a_m x_{t-m} + \varepsilon_t \qquad (2\text{-}1)$$

where $\{\varepsilon_t\}$ is a zero mean white noise with finite variance $E(\varepsilon^2_t)=\sigma^2$ and finite fourth order moment. The roots of the characteristic equation

$$1 + \sum_{m=1}^{p} a_m z^m = 0 \qquad (2\text{-}2)$$

are lying outside the unit circle in the z-plane. The total number of observations we will use is N. The estimates $\{\hat{a}_m, m=1,2,\ldots,p\}$ of the autoregressive coefficients are obtained by solving the approximate YW equations

$$\sum_{m=1}^{p} \hat{r}_{k-m} \hat{a}_m = - \hat{r}_k, \qquad k=1,2,\ldots,p \qquad (2\text{-}3)$$

where

$$\hat{r}_k = \sum_{t=1}^{N-|k|} x_{t+|k|} x_t / N$$

The power spectral density function $P(\omega)$ of the process $\{x_t\}$ is given by

$$P(\omega) = \frac{\sigma^2}{\left| 1 + \sum_{m=1}^{p} a_m \exp(-j\omega m) \right|^2} \qquad (2\text{-}4a)$$

An estimate $\hat{\sigma}^2$ of σ^2 is given by

$$\hat{\sigma}^2 = \hat{r}_0 + \sum_{m=1}^{p} \hat{a}_m \hat{r}_m \qquad (2\text{-}4b)$$

Akaike then defines

$$\Delta a_m = \hat{a}_m - a_m, \qquad \Delta\sigma^2 = \hat{\sigma}^2 - \sigma^2 \qquad (2\text{-}5a)$$

$$\underline{a} = [a_1, a_2, \ldots, a_p]^T \qquad \Delta P(\omega) = \hat{P}(\omega) - P(\omega) \qquad (2\text{-}5b)$$

$$A_p(\omega) = 1 + \sum_{m=1}^{p} a_m \exp(-j\omega m) \qquad (2\text{-}5c)$$

$$\Delta|A_p(\omega)|^2 = |\hat{A}_p(\omega)|^2 - |A_p(\omega)|^2 \qquad (2\text{-}5d)$$

If the process $\{x_t\}$ is strictly stationary and ergodic, $\hat{a}_m$ and $\hat{\sigma}^2$ converge to a_m and σ^2 with probability one as N, the number of data points, tends to infinity. Even though the estimate of the power spectral density is reciprocal of the quadratic function of estimates of the autoregressive coefficients, in the limit its variability is attributed mainly to the linear term. On the basis of a result of Anderson and Walker, Akaike has introduced a linear transformation of the data to get a set of mutually orthogonal variables $\sqrt{N}\Delta a_m$ $(m=1,2,\ldots,p)$. Under the assumption of strict stationarity and mutual independence of $\{\varepsilon_t\}$, he has proved that the distribution of $\sqrt{N}\Delta a_m$

converges, as N tends to infinity, to a p-dimensional gaussian distribution with zero mean vector and variance matrix $\sigma^2 R^{-1}$ where R is a pxp matrix with (i,j) element equal to $E(x_{t-i}x_{t-j})$. Akaike also obtained the result that $\sqrt{N}\{\sum_{t=1}^{N}(\varepsilon^2{}_t/N) - \sigma^2\}$ and $\sqrt{N}\Delta a_m$ tend to be independent in their limit distribution. He further showed that $\sqrt{N}\Delta P(\omega)/P(\omega)$ has a limit distribution with a variance composed of two components: one due to the relative variation of $\hat{\sigma}^2$ and the other due to that of $|\hat{A}_p(\omega)|^2$. His final result is given below

$$E_\infty\{N\frac{\Delta P(\omega_i)}{P(\omega_i)}\frac{\Delta P(\omega_j)}{P(\omega_j)}\} = \{\frac{E\varepsilon^4}{(E\varepsilon^2)^2} - 1\} + 4P(\omega_i)P(\omega_j)\sum_{k=0}^{p-1}\mathrm{Re}\{C_k(\omega_i)\}\mathrm{Re}\{C_k(\omega_j)\} \tag{2-6}$$

where

$$C_k(\omega) = \frac{A_p(\omega)}{\sigma_p}\frac{A_k(\omega)}{\sigma_k}\exp\{j\omega(k+1)\}$$

and $E_\infty\{.\}$ denotes the expectation in the limit distribution of the quantity within the brace. When i=j it gives the asymptotic variance of the ARSPE estimates. It should be noted that the foregoing result is also valid when the model order p is greater than the order of the process.

2.2.2 Kromer Derivation

Kromer [16] has determined the asymptotic distribution of the estimated spectral density as first the number N of observations and subsequently the order p of the autoregression goes to infinity. He has compared the asymptotic properties of the ARSPE estimates with that of conventional windowed periodogram estimates. His findings are summarized by the following results.

a. The ARSPE estimates are asymptotically unbiased.
b. The ARSPE estimates are asymptotically normal.
c. The variance of the estimated PSD is given by

$$\mathrm{Var}\{\hat{P}(\omega)\} = (2/\nu)P^2(\omega) \qquad (2\text{-}7)$$

where ν is the number of degrees of freedom and is related to the order of the AR process p by $\nu=N/p$. This result holds for large N and p and is valid for smooth spectra where $\delta P(\omega)/\delta\omega$ is not high.

2.2.3 Berk Derivation

Berk [17] has dealt with a stationary process $\{x_t\}$ with some regularity assumptions which are mentioned later. He has shown that the autoregression yields a consistent estimator of the spectral density of $\{x_t\}$ when the order p is asymptotically sufficient to overcome the bias. Assuming

p goes to infinity so that the bias from using a finite autoregression vanishes at a sufficient rate, the ARSPE estimates are furthermore asymptotically normally distributed, and uncorrelated at different fixed frequencies. Comparing with the spectral estimates based on a windowed periodogram he has found that the asymptotic variances for both estimators are the same.

Berk has defined the linear process $\{x_t\}$ by

$$x_t = \varepsilon_t + b_1\varepsilon_{t-1} + b_2\varepsilon_{t-2} + \cdots \tag{2-8a}$$

where b_1, b_2, ... are real numbers and $\{\varepsilon_t\}$ is a sequence of independent identically distributed (i.i.d.) zero mean random variables with variance $E(\varepsilon^2_t)=\sigma^2$. The polynomial $B(z)=1+b_1z+b_2z^2+\ldots$ is bounded away from zero for $|z|\leqslant 1$ so that the process is invertible. We have therefore

$$x_t + a_1x_{t-1} + a_2x_{t-2} + \cdots = \varepsilon_t \tag{2-8b}$$

where

$$A(z) = 1 + a_1z + a_2z^2 + \cdots = 1/B(z)$$

is bounded away from zero, $|z|\leqslant 1$. Berk uses a least squares predictor of order p for fitting the data with the AR model of (2-8b), or actually a truncation thereof. The estimated AR coefficients are determined from approximate YW equations

$$\hat{R}\underline{\hat{a}} = -\underline{\hat{r}} \tag{2-9}$$

where

$$\hat{R} = \sum_{j=p}^{N-1} \underline{X}_j \underline{X}_j^T/(N-p)$$

$$\hat{\underline{r}} = \sum_{j=p}^{N-1} \underline{X}_j x_{j+1}/(N-p)$$

$$\underline{X}_j = [x_j, x_{j-1}, \ldots, x_{j-p+1}]^T$$

The resulting estimated PSD is given by

$$\hat{P}(\omega) = \frac{\hat{\sigma}^2}{2\pi|\hat{A}_p(e^{j\omega})|^2} \qquad (2\text{-}10)$$

where

$$\hat{A}_p(z) = 1 + \sum_{i=1}^{p} \hat{a}_i z^i$$

$$\hat{\sigma}^2 = \hat{r}_0 + \sum_{m=1}^{p} \hat{a}_m \hat{r}_m$$

The corresponding theoretical quantities are $P(\omega)$, $A_p(z)$, σ^2, and R. If $0<P_1<P(\omega)<P_2$ and $\omega_1<\omega_2<\ldots<\omega_p$ are the eigenvalues of R, then [23]

$$0 < 2\pi P_1 \leq \omega_1 < \omega_2 < \ldots < \omega_p \leq 2\pi P_2 \qquad (2\text{-}11)$$

Using the latter and assuming the following regularity conditions

i) $A(\omega)$ is nonzero for $-\pi<\omega<\pi$,

ii) $E(\varepsilon^4_t) < \infty$,

iii) The choice of p in terms of N is such that $p^3/N \to 0$,

iv) The choice of p in terms of N is such that $\sqrt{n}(|a_{p+1}| + |a_{p+2}| + \ldots) \to 0$

Berk obtained the following results.

a) $\hat{P}(\omega)$ converges to $P(\omega)$ in probability under the above mentioned conditions.

b) The joint asymptotic distribution of

$$\sqrt{(N/p)}\{\hat{P}(0)-P(0)\},\ \sqrt{(N/p)}\{\hat{P}(\omega_1)-P(\omega_1)\},\ \ldots,$$

$$\sqrt{(N/p)}\{\hat{P}(\omega_p)-P(\omega_p)\},\ \sqrt{(N/p)}\{\hat{P}(\pi)-P(\pi)\},\quad 0<\omega_1<\ldots<\omega_p<\pi$$

is independent, normal, zero mean with variances

$$4P^2(0),\ 2P^2(\omega_1),\ \ldots,\ 2P^2(\omega_p),\ 4P^2(\pi)$$

2.2.4 Sakai Derivation

Sakai [18] has used the periodogram technique to investigate the properties of ARSPE estimates. He has shown numerically that the behavior of the variances is similar to the earlier results of Kromer and Berk for stationary

processes. He has considered a time series $\{x_t\}$ which consists of q sinusoids and additive zero mean stationary gaussian noise ε_t. So,

$$x_t = \sum_{i=1}^{q} (B_i e^{j\omega_i t} + B_i^* e^{-j\omega_i t}) + \varepsilon_t \tag{2-12}$$

where * denotes the complex conjugate and $\{B_i\}$ are complex constants. The AR parameters are then estimated from approximate YW equations, where the autocorrelation matrix is estimated from the data in the same manner as Akaike has done. The ARSPE estimate is given by

$$\hat{P}(\omega) = \frac{\hat{\sigma}^2}{2\pi\hat{A}(\omega)\hat{A}(-\omega)} \tag{2-13}$$

where

$$\hat{A}(\omega) = \sum_{i=0}^{p} \hat{a}_i e^{ji\omega}, \qquad \hat{a}_0 = 1$$

Sakai then proves that for a sufficiently large number of data points the ARSPE estimate $\hat{P}(\omega)$ can be expressed in terms of the periodogram $I_N(s)$ as follows.

$$\hat{P}(\omega) \simeq \int_{-\pi}^{\pi} G(\omega,s) I_N(s) ds \tag{2-14}$$

where

$$I_N(s) = \left| \sum_{t=1}^{N} x_t e^{-jts} \right|^2/(2\pi N)$$

$$G(\omega,s) = P(\omega)A(s)\left[\sum_{i=1}^{p} \{R^{-1}\underline{H}(\omega)\}e^{jis} + \sigma^{-2}A(-s) \right]$$

$$\triangleq P(\omega)g(\omega,s)$$

$$\underline{H}(\omega) = 2Re[\underline{\xi}(e^{j\omega})/A(\omega)]$$

$$\underline{\xi}(e^{j\omega}) = [e^{j\omega}, e^{j2\omega}, \ldots, e^{jp\omega}]^T$$

The symbol ≈ indicates that both sides are asymptotically equivalent. Thus, the ARSPE can be viewed as a smoothed periodogram with a data-dependent "spectral window" $G(\omega,s)$. Finally, Sakai expressed the asymptotic covariance between $P(\omega_i)$ and $P(\omega_j)$ by the following relation

$$E\left[\frac{\Delta P(\omega_i)}{P(\omega_i)} \frac{\Delta P(\omega_j)}{P(\omega_j)} \right]$$

$$\approx \int_{\pi}^{-\pi} \int_{\pi}^{-\pi} g(\omega_i,s_1)g(\omega_j,s_2)Cov\{I_N(s_1)I_N(s_2)\}ds_1ds_2 \qquad (2\text{-}15)$$

where

$$\Delta P(\omega_i) = \hat{P}(\omega_i) - P(\omega_i)$$

At present the above expression does not permit clear-cut analytic interpretations.

2.3 SUMMARY

Let us now summarize the significant findings of Akaike, Kromer, Berk and Sakai. All of them showed that asymptotically the ARSPE is consistent. Furthermore, from their results it is found that $\sqrt{(N/p)}\{\hat{P}(\omega)-P(\omega)\}/P(\omega)$ has a normal limiting distribution with zero mean and variance

$$N.E\left[\left\{\frac{\Delta P(\omega)}{P(\omega)}\right\}^2\right] \simeq \begin{cases} 2p & \text{for } \omega \neq 0, \pi \\ 4p & \text{for } \omega = 0, \pi \end{cases} \qquad (2\text{-}16)$$

The above expression (2-16) indicates that the asymptotic variance of the ARSPE estimates is approximately equivalent to that of the windowed periodogram with a suitably chosen truncation length [24].

Chapter III

STATISTICAL PROPERTIES OF THE MODIFIED BURG SPECTRAL ESTIMATOR (MBSE)

3.1 INTRODUCTION

In the past decade, there has been strong interest and much activity in developing high resolution power spectrum estimation techniques, particularly for short data records. Especially the Burg spectral estimation technique, also called maximum entropy spectral analysis (MESA), has received much attention in this regard. The MEM spectrum of a stationary process results from maximizing the entropy of that process. It has been found that the MEM of spectral estimation is equivalent to the least squares fitting of an AR model to the process [25]. It was also shown that if the maximum entropy spectra are calculated for m=1,2,...,p, where p is the order of the model, and the average of the reciprocals of these spectra is determined, then this average is equal to the reciprocal of the maximum likelihood spectrum [26, 27].

The BSE is based on the assumption that the received data sequence was generated by a white noise driven AR system. In this thesis, we are assuming that the stochastic process generating the N data points $\{x_t\}_{t=0}^{N-1}$ is zero mean,

gaussian, wide sense stationary and ergodic (A1)[1]. The Burg spectral estimate [28] is then given by

$$\hat{P}(\omega) = \frac{\hat{\sigma}^2}{\hat{A}(z)\hat{A}(z^{-1})}\Big|_{z=e^{j\omega}}, \qquad -\pi \leq \omega \leq \pi \tag{3-1}$$

where

$$\hat{A}(z) = 1 + \sum_{i=1}^{p} \hat{a}_{i,p} z^i$$

$$\hat{\sigma}^2 = \sum_{i=0}^{N-1} x^2{}_i / (N\hat{r}_0)$$

with $\hat{r}_0$ the autocorrelation at lag zero associated with the filter $1/\hat{A}(z)$. It is to be noted that the estimated gain factor $\hat{\sigma}^2$ is different from that proposed by Burg [28]. Burg estimates the gain factor from the sum of the squares of forward and backward errors which is minimized with respect to the reflection coefficient. The order p of the AR model (Fig. 1) is assumed known. The AR coefficients $\{\hat{a}_{i,p}\}_{i=1}^{p}$ are determined by using the Levinson-Durbin algorithm and the reflection coefficients $\{\hat{K}_m\}_{m=1}^{p}$ are estimated according to the Burg method [29].

---

[1] This assumption is denoted by A1.

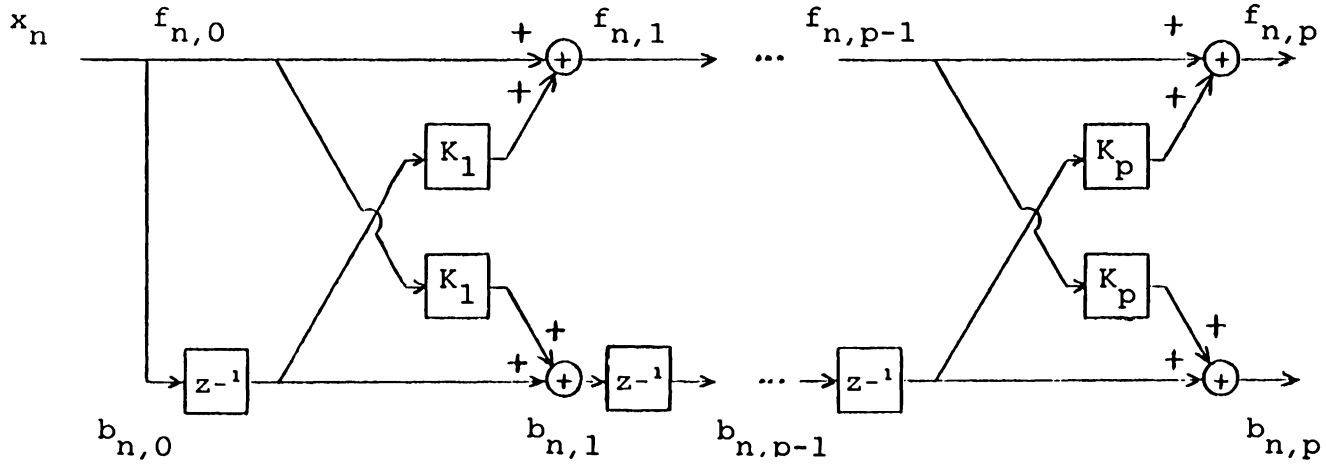


Figure 1: The Autoregressive Lattice Filter

$$\hat{K}_m = - \frac{2 \sum_{t=m}^{N-1} f_{t,m-1} b_{t-1,m-1}}{\sum_{t=m}^{N-1} (f^2_{t,m-1} + b^2_{t-1,m-1})} \qquad (3\text{-}2)$$

where

$$f_{t,m} = \sum_{i=o}^{m} \hat{a}_{i,m} x_{t-i}$$

$$b_{t,m} = \sum_{i=0}^{m} \hat{a}_{i,m} x_{t-m+i}$$

the forward respectively backward error at the mth stage. The reflection coefficient is defined as the partial correlation coefficient between the forward and backward prediction residuals. In the AR lattice filter approach reflection coefficient is analogous to the reflection coefficient of an acoustic tube model. The forward prediction residual or error is equal to $(x_n - \hat{x}_n)$ where x_n is generated by a one step linear predictor using the last p samples. Similarly the backward error is equal to $(x_n - \hat{x}_n)$ where x_n is generated by a one step linear predictor using the next p samples. It is worth mentioning that the Burg method for estimating reflection coefficients is equivalent to the harmonic mean method where the sum of forward and

backward errors is minimized with respect to the reflection coefficient. Here no assumptions are made regarding the data outside of the observation interval and all the available data are maximally utilized.

Let us now focus our attention on the statistical properties of the BSE. In the previous chapter we have discussed the significant results of the asymptotic behavior of the ARSPE. Since the BSE is purely AR, the asymptotic mean and variance can be derived from those results. We are, however, dealing with a moderate number of data points so that large sample results do not necessarily apply. By a moderate number of data points we mean that the length of the record is a few times the order of the estimated AR model. The Burg method is data adaptive and hence nonlinear. Still there is a one-to-one relation between the reflection coefficients and the corresponding AR coefficients, i.e. $\hat{\underline{a}}_p = f(\hat{\underline{K}}_p)$, where

$$\hat{\underline{a}}_p = [\hat{a}_{1,p}, \hat{a}_{2,p}, \ldots, \hat{a}_{p,p}]^T \tag{3-3a}$$

$$\hat{\underline{K}}_p = [\hat{K}_1, \hat{K}_2, \ldots, \hat{K}_p]^T \tag{3-3b}$$

Also there is a one-to-one relationship between the AR coefficients and the corresponding PSD estimate (3-1). If the statistics of $\hat{\underline{K}}_p$ were known then using a Jacobian

transformation twice, the mean and variance of the AR PSD estimate could be determined. However, the function $f(\hat{\underline{K}}_p)$ is nonlinear and the exact statistics of $\hat{\underline{K}}_p$ for a moderate number of data points are difficult to evaluate. Considering all these problems, we are led to derive approximate values for the mean and variance of the BSE estimate.

Here we have investigated the effect of segment averaging on the quality of the BSE. We have defined the modified Burg spectral estimator (MBSE) as the Burg spectral estimator where some averaging techniques are applied. Note that when number of sections is 1 the MBSE and the BSE are same. Three types of averaging technique are applied. The available data sequence is segmented into M nonoverlapping sections where each section has N/M data points. In the first method (AVA), the AR parameters are evaluated for each section, then averaged, and the resulting average AR parameters yield an associated spectral density estimate. The second method (AVK) averages the reflection coefficients, the average of which leads to the corresponding spectral density estimate. The final approach (AVP) evaluates the PSD estimate associated with each segment, and then averages these directly.

3.2 SEPARATION OF ERROR

Let us assume that the rational transfer function H(z), for the system generating the data, is known. The exact autocorrelation sequence $\{r_i\}_{i=0}^{p}$ for the data can then be determined [30]. The optimum pth order AR model which fits the data string with minimum mean square error, satisfies the YW equations,

$$R_{p,p}\tilde{\underline{a}}_p = -\underline{r}_p, \tag{3-4}$$

where

$$\underline{r}_p = [r_1, r_2, \ldots, r_p]^T$$

$$\tilde{\underline{a}}_p = [\tilde{a}_{1,p}, \tilde{a}_{2,p}, \ldots, \tilde{a}_{p,p}]^T$$

and $R_{p,p}$ is a symmetric Toeplitz matrix with first row $(r_0 \quad \underline{r}_{p-1}^T)$.

Let $P(\omega)$ be the actual PSD of the data, and let $\tilde{P}(\omega)$ be the PSD of the output of the filter $1/\tilde{A}(z)$ driven by white noise. Then $\Delta\tilde{P}(\omega)=P(\omega)-\tilde{P}(\omega)$ is the fixed deterministic error for modelling the given data as an AR(p) process. In the real world, r_i is not known. Let the MBSE give $\hat{\underline{a}}_p$ as an estimate of $\tilde{\underline{a}}_p$ with $\hat{P}(\omega)$ as the corresponding PSD estimate. Then the bias of the PSD estimate can be written as

$$B\{\hat{P}(\omega)\} = P(\omega) - E\{\hat{P}(\omega)\}$$

$$= \{P(\omega) - \tilde{P}(\omega)\} - [E\{\hat{P}(\omega)\} - \tilde{P}(\omega)]$$

$$= \Delta\tilde{P}(\omega) - E\{\Delta\hat{P}(\omega)\} \tag{3-5a}$$

For the mean square error we find

$$MSE\{\hat{P}(\omega)\} = E\{P(\omega) - \hat{P}(\omega)\}^2$$

$$= \Delta\tilde{P}^2(\omega) - 2\Delta\tilde{P}(\omega)E\{\Delta\hat{P}(\omega)\} + E\{\Delta\hat{P}^2(\omega)\} \tag{3-5b}$$

The mean and variance expressions are

$$E\{\hat{P}(\omega)\} = \tilde{P}(\omega) + E\{\Delta\hat{P}(\omega)\} \tag{3-5c}$$

$$Var\{\hat{P}(\omega)\} = MSE\{\hat{P}(\omega)\} - B^2\{\hat{P}(\omega)\} \tag{3-5d}$$

We have thus introduced an intermediate function $\tilde{P}(\omega)$ to find the mean and variance of the PSD estimate. It is evident that as N tends to infinity $\hat{P}(\omega)$ approaches $\tilde{P}(\omega)$. Furthermore, note that, unlike often done, we have not assumed that $E\{\hat{P}(\omega)\}$ equals $\tilde{P}(\omega)$ as N is not very large. If H(z) is known then $\Delta\tilde{P}(\omega)$ can be evaluated. Next we have to derive $E\{\Delta\hat{P}(\omega)\}$ and $E\{\Delta\hat{P}^2(\omega)\}$ in order to evaluate (3-5a) to (3-5d).

3.3 APPROXIMATION OF $E\{\Delta\hat{P}(\omega)\}$ AND $E\{\Delta\hat{P}^2(\omega)\}$

Two different methods are used to evaluate approximate values of $E\{\Delta\hat{P}(\omega)\}$ and $E\{\Delta\hat{P}^2(\omega)\}$. One of them is the approximation of Sakai and the other one uses the Taylor series approximation. Both of these methods yield good results, even for a moderate number of data points.

3.3.1 Approximation according to Sakai

Sakai et al. [31] have applied a periodogram technique to derive the ARSPE variance. That is not our purpose however, as we use his approximation technique to find an expression for $E\{\Delta\hat{P}(\omega)\}$ only. We will refer to this method as 'Sakai Approximation'. From (eqn. 31,[31]) (see Appendix A for the derivation of this equation)

$$E\{\Delta\hat{P}(\omega)\} \approx \tilde{P}(\omega)[\tilde{\sigma}^{-2}E(\Delta\hat{\sigma}^2) - \underline{H}^T(e^{j\omega})E(\Delta\underline{\hat{a}}_p)] \tag{3-6a}$$

where

$$\underline{H}(e^{j\omega}) = 2\mathrm{Re}\{\frac{\underline{\xi}(e^{j\omega})}{\tilde{A}(e^{j\omega})}\}$$

$$\underline{\xi}(e^{j\omega}) = [e^{j\omega},\ e^{j2\omega},\ \ldots,\ e^{jp\omega}]^T$$

$$\Delta\hat{\sigma}^2 = \hat{\sigma}^2 - \tilde{\sigma}^2, \qquad \Delta\underline{\hat{a}}_p = \underline{\hat{a}}_p - \underline{\tilde{a}}_p$$

and

$$E\{\Delta\hat{P}^2(\omega)\} \approx \tilde{P}^2(\omega)[\tilde{\sigma}^{-4}E\{(\Delta\hat{\sigma}^2)^2\} + \underline{H}^T(e^{j\omega})E(\Delta\underline{\hat{a}}_p\Delta\underline{\hat{a}}_p^T)\underline{H}(e^{j\omega}) - 2\tilde{\sigma}^{-2}\underline{H}^T(e^{j\omega})E(\Delta\hat{\sigma}^2\Delta\underline{\hat{a}}_p)] \tag{3-6b}$$

It is worth noting that the second term on the right hand side of equation (3-6b) has the same form as the CRLB for parametric spectral estimators. It has been found from simulation results that the first and last terms on the right hand side of equation (3-6b) are much less significant than the second term. Hence, the variance of the MBSE estimates using this equation is expected to be nearly the same as the variance computed from the CRLB. We assume that $\Delta\hat{\sigma}^2$ and $\Delta\underline{\hat{a}}_p$ are uncorrelated (A2)[2]. We have also considered

$$E(\Delta\hat{\sigma}^2) = \sum_{i=0}^{N-1} x^2{}_i \left(\frac{1}{\bar{r}_0} - \frac{1}{\tilde{r}_0} \right)/N \tag{3-7a}$$

$$E\{(\Delta\hat{\sigma}^2)^2\} = E^2(\Delta\hat{\sigma}^2) \tag{3-7b}$$

where $\tilde{r}_0$ and $\bar{r}_0$ are the autocorrelation at lag zero associated with the filters $1/\tilde{A}(z)$ and $1/\bar{A}(z)$ respectively, and

$$\bar{A}(z) = 1 + \sum_{i=1}^{p} E(\hat{a}_{i,p})z^i \tag{3-7c}$$

Now we have to derive $E(\Delta\underline{\hat{a}}_p)$ and $E(\Delta\underline{\hat{a}}_p\Delta\underline{\hat{a}}_p^T)$ in order to compute $E\{\Delta\hat{P}(\omega)\}$ and $E\{\Delta\hat{P}^2(\omega)\}$ according to (3-6a) and (3-6b). But $E(\Delta\underline{\hat{a}}_p)$ and $E(\Delta\underline{\hat{a}}_p\Delta\underline{\hat{a}}_p^T)$ depend on the method of

---

[2] This assumption is denoted by A2.

averaging. Later it will be shown that $E(\Delta\underline{\hat{a}}_p)$ and $E(\Delta\underline{\hat{a}}_p\Delta\underline{\hat{a}}_p^T)$ depend implicitly on the number of observations N.

3.3.2 Approximation using Taylor series expansion

Since we are estimating $\hat{A}(z)$ for $\tilde{A}(z)$ and both have the same order p, it is expected that the error $\{|1/\hat{A}(z)|^2-|1/\tilde{A}(z)|^2\}$ will be small for a moderate number of data points. So, in this method, using the Taylor series we have expanded $\hat{P}(\omega)$ around $\tilde{P}(\omega)$ and neglected all the terms higher than the first order. Taking the expected value, we get

$$E\{\Delta\hat{P}(\omega\} \simeq \left(\frac{\delta\tilde{P}}{\delta\underline{\tilde{a}}_p}\right)^T E(\Delta\underline{\hat{a}}_p) \tag{3-8a}$$

and

$$E\{\Delta\hat{P}^2(\omega\} \simeq \left(\frac{\delta\tilde{P}}{\delta\underline{\tilde{a}}_p}\right)^T E(\Delta\underline{\hat{a}}_p\Delta\underline{\hat{a}}_p^T)\left(\frac{\delta\tilde{P}}{\delta\underline{\tilde{a}}_p}\right) \tag{3-8b}$$

where

$$\left(\frac{\delta\tilde{P}}{\delta\underline{\tilde{a}}_p}\right) = \left[\frac{\delta\tilde{P}}{\delta\tilde{a}_{1,p}}, \frac{\delta\tilde{P}}{\delta\tilde{a}_{2,p}}, \ldots, \frac{\delta\tilde{P}}{\delta\tilde{a}_{p,p}}\right]^T$$

The quantities $E(\Delta\underline{\hat{a}}_p)$ and $E(\Delta\underline{\hat{a}}_p\Delta\underline{\hat{a}}_p^T)$ found in the Sakai method are also used in (3-8a) and (3-8b). It is to be noted that in the Taylor method we have considered the gain factor σ^2 to constitute a priori information, resulting from a

previous estimation step. If $\Delta\hat{\sigma}^2$ is zero both methods give identical results.

The gain factor is, however, not a known parameter, and it needs to be estimated. Then

$$E\{\Delta\hat{P}(\omega)\} \simeq \left(\frac{\delta\tilde{P}}{\delta\tilde{\sigma}^2}\right)E(\Delta\hat{\sigma}^2) + \left(\frac{\delta\tilde{P}}{\delta\underline{\tilde{a}}_p}\right)^T E(\Delta\underline{\hat{a}}_p) \tag{3-9a}$$

and

$$E\{\Delta\hat{P}^2(\omega)\} \simeq \left(\frac{\delta\tilde{P}}{\delta\tilde{\sigma}^2}\right)^2 E\{(\Delta\hat{\sigma}^2)^2\} + \left(\frac{\delta\tilde{P}}{\delta\underline{\tilde{a}}_p}\right)^T E(\Delta\underline{\hat{a}}_p\Delta\underline{\hat{a}}_p^T)\left(\frac{\delta\tilde{P}}{\delta\underline{\tilde{a}}_p}\right) + 2\left(\frac{\delta\tilde{P}}{\delta\tilde{\sigma}^2}\right)\left(\frac{\delta\tilde{P}}{\delta\underline{\tilde{a}}_p}\right)^T E(\Delta\underline{\hat{a}}_p\Delta\hat{\sigma}^2) \tag{3-9b}$$

Since

$$\frac{\delta\tilde{P}}{\delta\tilde{\sigma}^2} = \tilde{\sigma}^{-2}\tilde{P}(\omega)$$

$$\frac{\delta\tilde{P}}{\delta\underline{\tilde{a}}_p} = -2\tilde{P}(\omega)\mathrm{Re}\left\{\frac{\underline{\xi}(e^{j\omega})}{\tilde{A}(e^{j\omega})}\right\} = -\tilde{P}(\omega)\underline{H}(e^{j\omega})$$

we find that (3-9a) and (3-9b) are identical to the expressions in (3-6a) and (3-6b) respectively. Thus Sakai approximation and Taylor approximation will give the same result. Both methods behave exactly the same since only the first order linear terms have been considered. However, we can introduce the second order term in the expression for

$E\{\Delta\hat{P}(\omega)\}$ because up to the second order statistics of $\{\Delta\hat{\underline{a}}_p\}$ can be determined.

$$E\{\Delta\hat{P}(\omega)\} \simeq \left(\frac{\delta\tilde{P}}{\delta\tilde{\sigma}^2}\right)E(\Delta\hat{\sigma}^2) + \left(\frac{\delta\tilde{P}}{\delta\tilde{\underline{a}}_p}\right)^T E(\Delta\hat{\underline{a}}_p) + \frac{1}{2}E(\Delta\hat{\underline{a}}_p^T W \Delta\hat{\underline{a}}_p) \tag{3-10}$$

where

$$W = \left(\frac{\delta}{\delta\tilde{\underline{a}}_p}\right)\left(\frac{\delta}{\delta\tilde{\underline{a}}_p}\right)^T\tilde{P}(\omega)$$

$$E(\Delta\hat{\underline{a}}_p^T W \Delta\hat{\underline{a}}_p) = \sum_{i=1}^{p}\sum_{j=1}^{p} W_{i,j}E(\Delta\hat{a}_{i,p}\Delta\hat{a}_{j,p})$$

But from the simulation results, it has been found that the third term on the right hand side of (3-10) overestimates the error $\{\Delta\hat{P}(\omega)\}$ near the location of poles which are close to the unit circle in the z-domain (see Appendix B for details). In our simulation results we have used (3-8a) and (3-8b) for the Taylor approximation to analyze the effect of neglecting the variance of $\hat{\sigma}^2$.

3.4 AR COEFFICIENT AVERAGING (AVA)

In this averaging method, the predictor coefficients are computed from each of the M sections where each section has N/M data points. The predictor coefficients are then averaged to obtain the estimated AR filter $1/\hat{A}(z)$. The following relations can be derived easily.

$$\hat{\underline{a}}_p = \sum_{m=1}^{M} \hat{\underline{a}}_{m,p}/M \tag{3-11}$$

$$E(\Delta\hat{\underline{a}}_p) = E[(\sum_{m=1}^{M} \hat{\underline{a}}_{m,p}/M) - \tilde{\underline{a}}_p] = E(\hat{\underline{a}}_{m,p}) - \tilde{\underline{a}}_p \tag{3-12a}$$

$$E(\hat{\underline{a}}_{m,p}) = E\begin{bmatrix} (I + \hat{K}_{m,p}J)\hat{\underline{a}}_{m,p-1} \\ \hat{K}_{m,p} \end{bmatrix} \tag{3-12b}$$

where J is the reverse operator matrix which is defined as

$$J \triangleq \begin{bmatrix} & & & 1 \\ & & \cdot & \\ & 0 & & 0 \\ & \cdot & & \\ 1 & & & \end{bmatrix} \tag{3-12c}$$

Assuming that $\hat{K}_{m,p}$ and any one element of $\hat{\underline{a}}_{m,p-1}$ are weakly correlated (A3)[3], it is found from (3-12b) that

$$E(\hat{\underline{a}}_{m,p}) \simeq \begin{bmatrix} \{I + E(\hat{K}_{m,p}J)\}E(\hat{\underline{a}}_{m,p-1}) \\ E(\hat{K}_{m,p}) \end{bmatrix} \tag{3-12d}$$

---

[3] This assumption is denoted by A3.

We note that assumption A3 was experimentally found to hold, even for a moderate number of data points. Therefore, $E(\hat{\underline{a}}_{m,p})$ can be calculated recursively from (3-12d) and (3-19a); the derivation of (3-19a) will be shown later.

To approximate $E(\Delta\hat{\underline{a}}_p \Delta\hat{\underline{a}}_p^T)$, we first find

$$E(\Delta\hat{\underline{a}}_p \Delta\hat{\underline{a}}_p^T) = E(\hat{\underline{a}}_p - \tilde{\underline{a}}_p)(\hat{\underline{a}}_p - \tilde{\underline{a}}_p)^T$$

$$= E(\hat{\underline{a}}_p \hat{\underline{a}}_p^T) + \tilde{\underline{a}}_p \tilde{\underline{a}}_p^T - \tilde{\underline{a}}_p E(\hat{\underline{a}}_p^T) - E(\hat{\underline{a}}_p)\tilde{\underline{a}}_p^T \qquad (3\text{-}13a)$$

and

$$E(\hat{\underline{a}}_p \hat{\underline{a}}_p^T) = E\Big[\sum_{m=1}^{M} \sum_{k=1}^{M} \{\hat{\underline{a}}_{m,p} \hat{\underline{a}}_{k,p}^T\}/M^2 \Big]$$

$$= \sum_{m=1}^{M} E(\hat{\underline{a}}_{m,p} \hat{\underline{a}}_{m,p}^T)/M^2 + \sum_{m=1}^{M} \sum_{\substack{k=1 \\ k\neq m}}^{M} E(\hat{\underline{a}}_{m,p} \hat{\underline{a}}_{k,p}^T)/M^2$$

$$= E(\hat{\underline{a}}_{m,p} \hat{\underline{a}}_{m,p}^T)/M + (M-1)E(\hat{\underline{a}}_{m,p})E(\hat{\underline{a}}_{k,p}^T)/M \qquad (3\text{-}13b)$$

by assuming $\hat{\underline{a}}_{m,p}$ and $\hat{\underline{a}}_{k,p}$ $(m\neq k)$ are uncorrelated, i.e. any two segments are uncorrelated (A4)[4]. For $E(\hat{\underline{a}}_{m,p} \hat{\underline{a}}_{m,p}^T)$ we have

$$E(\hat{\underline{a}}_{m,p} \hat{\underline{a}}_{m,p}^T) = E\left[\begin{matrix} (I+\hat{K}_{m,p}J)\hat{\underline{a}}_{m,p-1} \\ \hat{K}_{m,p} \end{matrix} \right] \left[\hat{\underline{a}}_{m,p-1}^T (I+\hat{K}_{m,p}J) \quad \hat{K}_{m,p} \right] \qquad (3\text{-}13c)$$

---

[4] This assumption is denoted by A4.

Now using assumption A3 it can be shown that

$$E(\hat{\underline{a}}_{m,p}\hat{\underline{a}}^{T}_{m,p}) = T_{2x2} \qquad (3\text{-}13d)$$

where

$$T_{1,1} = \Gamma + E(\hat{K}_{m,p})(J\Gamma+\Gamma J) + E(\hat{K}^{2}_{m,p})J\Gamma J$$

$$T_{1,2} = E(\hat{K}_{m,p})E(\hat{\underline{a}}_{m,p-1}) + E(\hat{K}^{2}_{m,p})JE(\hat{\underline{a}}_{m,p-1})$$

$$T_{2,1} = [T_{1,2}]^{T} \qquad T_{2,2} = E(\hat{K}^{2}_{m,p})$$

$$\Gamma = E(\hat{\underline{a}}_{m,p-1}\hat{\underline{a}}^{T}_{m,p-1})$$

So, $E(\hat{\underline{a}}_{m,p}\hat{\underline{a}}^{T}_{m,p})$ can be calculated recursively from (3-12d), (3-13d), (3-19a) and (3-19b). The derivation of (3-19a) and (3-19b) will be shown later. Thus using (3-12a) to (3-13d) we can determine $E(\Delta\hat{\underline{a}}_{p})$ and $E(\Delta\hat{\underline{a}}_{p}\Delta\hat{\underline{a}}^{T}_{p})$ for the AVA method. Finally, the approximate mean and variance of the MBSE estimates using this method are evaluated from equations (3-5a) to (3-5d).

3.5 REFLECTION COEFFICIENT AVERAGING (AVK)

Here the reflection coefficients computed from each of M sections are averaged. The AR parameters associated with the averaged reflection coefficients are then computed using the Levinson-Durbin algorithm. For the approximate value of $E(\Delta\hat{\underline{a}}_p)$, applying A3, we get

$$\hat{K}_p = \sum_{m=1}^{M} \hat{K}_{m,p}/M \tag{3-14}$$

$$E(\Delta\hat{\underline{a}}_p) = E(\hat{\underline{a}}_p) - \tilde{\underline{a}}_p \tag{3-15a}$$

$$E(\hat{\underline{a}}_p) \simeq \begin{bmatrix} \{I + E(\hat{K}_{m,p})J\}E(\hat{\underline{a}}_{p-1}) \\ E(\hat{K}_{m,p}) \end{bmatrix} \tag{3-15b}$$

This has the same form as (3-12d) in the AVA method. For this method $E(\Delta\hat{\underline{a}}_p \Delta\hat{\underline{a}}_p^T)$ is the same as (3-13a). Now using A3 and A4 we find

$$E(\hat{\underline{a}}_p\hat{\underline{a}}_p^T) = E\left[\begin{array}{c} \{I + \sum_{m=1}^{M}(\hat{K}_{m,p})J/M\}\hat{\underline{a}}_{p-1} \\ \\ \sum_{m=1}^{M}(\hat{K}_{m,p})/M \end{array}\right]$$

$$\times \left[\hat{\underline{a}}_{p-1}^T(I + \sum_{m=1}^{M}\hat{K}_{m,p}J/M) \qquad \sum_{m=1}^{M}\hat{K}_{m,p}/M\right]$$

$$= \left[\begin{array}{ll} \Gamma+E(\hat{K}_{m,p})(J\Gamma+\Gamma J)+\Omega J\Gamma J & E(\hat{K}_{m,p})E(\hat{\underline{a}}_{p-1}) \\ & + \Omega JE(\hat{\underline{a}}_{p-1}) \\ E(\hat{K}_{m,p})E(\hat{\underline{a}}_{p-1}^T)+\Omega E(\hat{\underline{a}}_{p-1}^T)J & \Omega \end{array}\right] \tag{3-15c}$$

where

$$\Gamma = E(\hat{\underline{a}}_{p-1}\hat{\underline{a}}_{p-1}^T)$$

$$\Omega = E(\hat{K}^2_{m,p})/M + (M-1)E^2(\hat{K}_{m,p})/M$$

Hence, as in the AVA method both $E(\hat{\underline{a}}_p)$ and $E(\hat{\underline{a}}_p\hat{\underline{a}}_p^T)$ can be calculated using a recursive technique.

3.6 POWER SPECTRAL DENSITY AVERAGING (AVP)

First, the PSD function associated with each segment is evaluated. This averaging technique gives the average of these estimates. This method is analogous to the Welch spectral estimation method for reducing the variance of periodogram estimators. The approach to determine the approximate mean and variance of the MBSE estimates is somewhat different from those used in the AVA and AVK methods. Here we have

$$\hat{P}(\omega) = \sum_{m=1}^{M} \hat{P}_m(\omega)/M \tag{3-16}$$

Under assumption A4 we get

$$E\{\Delta\hat{P}(\omega)\} = E\{\Delta\hat{P}_m(\omega)\}, \qquad \Delta\hat{P}_m(\omega) = \hat{P}_m(\omega) - \tilde{P}(\omega) \tag{3-17a}$$

$$E\{\Delta\hat{P}^2(\omega)\} = E\{\Delta\hat{P}^2_m(\omega)\}/M + (M-1)E^2\{\Delta\hat{P}_m(\omega)\}/M \tag{3-17b}$$

Equations (3-6a) and (3-6b) are used to evaluate $E\{\Delta\hat{P}_m(\omega)\}$ and $E\{\Delta\hat{P}^2_m(\omega)\}$. $E(\Delta\hat{\underline{a}}_p)$ and $E(\Delta\hat{\underline{a}}_p\Delta\hat{\underline{a}}_p^T)$ are computed from (3-12a) to (3-13d) where M is always set to one. Finally, from (3-5a) to (3-5d) the mean and variance of the PSD estimate can be determined.

3.7 DERIVATION OF $E(\hat{K}_m)$ AND $E(\hat{K}^2_m)$

The magnitude of the reflection coefficients using the Burg method are always less than one, which corresponds to a stable model. Since it is difficult to get the exact value, an approximate value of the above quantities will be derived. For convenience, we have dropped the subscript indicating the segment number from the notation for the reflection coefficients. Let us rewrite (3-2) as y/z where

$$N \leftarrow N/M$$

$$y = -2 \sum_{t=m}^{N-1} f_{t,m-1} b_{t-1,m-1} \tag{3-18a}$$

$$z = \sum_{t=m}^{N-1} (f^2_{t,m-1} + b^2_{t-1,m-1}) \tag{3-18b}$$

Here N is replaced by N/M because for any mth section there are only N/M data points and $\hat{K}_m$ is determined using this number of data points. We assume that the probability densities p(y,z) are concentrated near their center of gravity (η_y, η_z), and $\hat{K}_m$ is smooth in the vicinity of this point (A5)[5]. Then taking the expected value of the terms in the Taylor series expansion of y/z around (η_y, η_z) and retaining up to second order moments, the following results

---

[5] This assumption is denoted by A5.

are obtained (see Appendix C).

$$E(\hat{K}_m) \simeq \frac{\eta_y}{\eta_z} - \frac{\mu_{y,z}}{\eta_z^2} + \frac{\eta_y \mu_{0z}}{\eta_z^3} \tag{3-19a}$$

and

$$E(\hat{K}^2_m) \simeq \frac{\eta_y^2 + \mu_{y0}}{\eta_z^2} + 3\frac{\eta_y^2 \mu_{0z}}{\eta_z^4} - 4\frac{\eta_y \mu_{yz}}{\eta_z} \tag{3-19b}$$

where

$$\eta_y = E(y), \qquad \eta_z = E(z),$$

$$\mu_{y0} = Var(y), \qquad \mu_{0z} = Var(z), \qquad \mu_{yz} = Cov(yz)$$

Using the gaussian fourth order moment rule, it can be shown (see Appendix C) that

$$\eta_y = -2(N-m)\gamma_1/N \tag{3-20a}$$

$$\eta_z = 2(N-m)\gamma_2/N \tag{3-20b}$$

$$\mu_{y0} = 4\{\kappa_y + \sum_{t=1}^{N-m-1} 2(N-m-t)(\gamma_3^2 + \gamma_4\gamma_5)\}/N^2 \tag{3-20c}$$

$$\mu_{0z} = \{\kappa_z + \sum_{t=1}^{N-m-1} 4(N-m-t)(2\gamma_3^2 + \gamma_4^2 + \gamma_5^2)\}/N^2 \tag{3-20d}$$

$$\mu_{yz} = -2\{\kappa_{yz} + \sum_{t=1}^{N-m-1} 4(N-m-t)\gamma_3(\gamma_4 + \gamma_5)\}/N^2 \tag{3-20e}$$

where

$$\gamma_1 = \sum_{i=0}^{m-1} \sum_{j=0}^{m-1} E(\hat{a}_{i,m-1})E(\hat{a}_{j,m-1})r_{m-i-j}$$

$$\gamma_2 = \sum_{i=0}^{m-1} \sum_{j-0}^{m-1} E(\hat{a}_{i,m-1})E(\hat{a}_{j,m-1})r_{i-j}$$

$$\gamma_3 = \sum_{i=0}^{m-1} \sum_{j=0}^{m-1} E(\hat{a}_{i,m-1})E(\hat{a}_{j,m-1})r_{t+i-j}$$

$$\gamma_4 = \sum_{i=0}^{m-1} \sum_{j=0}^{m-1} E(\hat{a}_{i,m-1})E(\hat{a}_{j,m-1})r_{t+m-i-j}$$

$$\gamma_5 = \sum_{i=0}^{m-1} \sum_{j=0}^{m-1} E(\hat{a}_{i,m-1})E(\hat{a}_{j,m-1})r_{t-m+i+j}$$

$$\kappa_y = (N-m)(\gamma_1^2 + \gamma_2^2)$$

$$\kappa_z = 4(N-m)(\gamma_1^2 + \gamma_2^2)$$

$$\kappa_{yz} = 4(N-m)\gamma_1\gamma_2$$

In the above derivations we have treated the predictor coefficients ($\hat{a}_{i,m}$) to behave as constant parameters $E(\hat{a}_{i,m})$. It is to be noted that $E(\hat{K}_m)$ and $E(\hat{K}^2_m)$ depend implicitly on the number of data points, N/M, in each section. We also note that the the variance of any reflection coefficient decreases to zero as N/M approaches infinity.

3.8 LOWER BOUND FOR THE MEAN AND VARIANCE OF $\hat{K}_m$

In this section we derive the lower bound for the mean and variance of the reflection coefficient $\hat{K}_m$ for any stage m. From (3-2) we get

$$\frac{1}{N}\sum_{t=m}^{N-1}\hat{K}_m\{f^2_{t,m-1} + b^2_{t-1,m-1}\} = -\frac{2}{N}\sum_{t=m}^{N-1} f_{t,m-1}b_{t-1,m-1} \tag{3-21a}$$

$$\sum_{t=m}^{N-1}\{\hat{K}_m - E(\hat{K}_m)\}\{f^2_{t,m-1} + b^2_{t-1,m-1}\}$$

$$= -\sum_{t=m}^{N-1}[2f_{t,m-1}b_{t-1,m-1} + E(\hat{K}_m)\{f^2_{t,m-1} + b^2_{t-1,m-1}\}] \tag{3-21b}$$

Taking expected values on both sides yields

$$\sum_{t=m}^{N-1} E[\{\hat{K}_m - E(\hat{K}_m)\}\{f^2_{t,m-1} + b^2_{t-1,m-1}\}]$$

$$= -\sum_{t=m}^{N-1}[2C_{t,m-1} + E(\hat{K}_m)(F_{t,m-1} + B_{t-1,m-1})] \tag{3-22}$$

where

$$F_{t,m-1} = E(f^2_{t,m-1}), \qquad B_{t-1,m-1} = E(b^2_{t-1,m-1})$$

$$C_{t,m-1} = E(f_{t,m-1}b_{t-1,m-1})$$

Now we can write

$$\sum_{t=m}^{N-1} |E[\{\hat{K}_m - E(\hat{K}_m)\}f^2_{t,m-1}]| + \sum_{t=m}^{N-1} |E[\{\hat{K}_m - E(\hat{K}_m)\}b^2_{t-1,m-1}]|$$

$$\geq |(N-m)\{2C_{m-1} + E(\hat{K}_m)(F_{m-1} + B_{m-1})\}| \tag{3-23}$$

It is to be noted that the forward error $f_{t,m-1}$ and the backward error $b_{t-1,m-1}$ are wide sense stationary, so that their mean values are independent of t. Using the Cauchy-Schwartz inequality, we find

$$|E[\{\hat{K}_m - E(\hat{K}_m)\}f^2_{t,m-1}]|^2 \leq E[\{\hat{K}_m - E(\hat{K}_m)\}^2]E(f^4_{t,m-1}) \tag{3-24a}$$

and

$$|E[\{\hat{K}_m - E(\hat{K}_m)\}b^2_{t-1,m-1}]|^2 \leq E[\{\hat{K}_m - E(\hat{K}_m)\}^2]E(b^4_{t-1,m-1}) \tag{3-24b}$$

Hence we find

$$(N-m)\sqrt{Var(\hat{K}_m)}[\sqrt{E(f^4_{t,m-1})} + \sqrt{E(b^4_{t-1,m-1}}]$$

$$\geq (N-m)\{2C_{m-1} + E(\hat{K}_m)(F_{m-1} + B_{m-1})\} \tag{3-25}$$

For the stationary case F_{m-1} equals B_{m-1} and $E(f^4_{t,m-1})$ equals $E(b^4_{t-1,m-1})$, so that a theoretical lower bound for the variance of $\hat{K}_m$ is given by

$$Var(\hat{K}_m) \geq \frac{[C_{m-1} + E(\hat{K}_m)F_{m-1}]^2}{E(f^4_{t,m-1})} \tag{3-26}$$

From the derivation of (3-26) it is difficult to show whether this inequality gives a tight lower bound or not. Also this bound does not exhibit any relationship between the number of observations and the variance of the reflection coefficient at stage m. However, from the simulated results we find that this bound holds for the reflection coefficient at the first stage. For the first stage, the error sequences are the same as the data series. But for subsequent stages the forward and backward residuals become nonlinear functions of the data and the exact statistics of these residuals are difficult to derive from the known statistics of the data. We can not, as a result, readily evaluate the lower bound for the variance of the reflection coefficients of the higher order AR models.

Taking the expected value on both sides of (3-21a) yields

$$\sum_{t=m}^{N-1} E\{\hat{K}_m(f^2_{t,m-1} + b^2_{t-1,m-1})\} = -2(N-m)C_{m-1} \qquad (3\text{-}27)$$

We know that $\mu_{11} \leq \mu_{02}\mu_{20}$ [32], so we can write

$$E\{\hat{K}_m f^2_{t,m-1}\} \leq \sqrt{\{Var(\hat{K}_m)Var(f^2_{t,m-1})\}} + E(\hat{K}_m)E(f^2_{t,m-1}) \qquad (3\text{-}28a)$$

As a result

$$E(\hat{K}_m)(F_{m-1} + B_{m-1}) + \sqrt{\{Var(\hat{K}_m)Var(f^2_{t,m-1})\}} + \sqrt{\{Var(\hat{K}_m)Var(b^2_{t-1,m-1})\}} \geq -2C_{m-1} \tag{3-28b}$$

This yields

$$E(\hat{K}_m) \geq -\frac{C_{m-1} + \{Var(\hat{K}_m)Var(f^2_{t,m-1})\}^2}{F_{m-1}} \tag{3-28c}$$

because for stationarity F_{m-1} equals B_{m-1} and $Var(f^2_{t,m-1})$ equals $Var(b^2_{t-1,m-1})$. Substituting the lower limit of $Var(\hat{K}_m)$ in (3-28c), the lower bound for the mean of the estimate of $\hat{K}_m$ is given by

$$E(\hat{K}_m) \geq -\frac{C_{m-1}}{F_{m-1}} \tag{3-29}$$

It is interesting to note that the total number of observations N is not explicitly present in the expression for the lower bound of $Var(\hat{K}_m)$. However, when the mean value of $\hat{K}_m$ approaches its lower bound then the variance of $\hat{K}_m$ goes to zero. This theoretical lower bound can be achieved when the number of observations approaches infinity. In this regard it is worth mentioning that the asymptotic covariance matrix for the reflection coefficients has been derived elsewhere [33].

3.9 THE CRLB FOR THE VARIANCE OF THE MBSE ESTIMATE

A useful method for evaluating the performance of an estimator is to study the variance of the estimation error. The Cramer-Rao lower bound (CRLB) gives the theoretical lower bound on the estimation error variance of any unbiased estimator [34, 35]. The CRLB is generally used as a tool for measuring the efficiency of any estimator. It has been shown [36] that the achievable accuracy from the maximum likelihood estimator (MLE) in terms of covariance is based on the Cramer-Rao inequality, i.e.

$$\mathrm{Cov}[\hat{\underline{\beta}}] \geq [-E\{\frac{\delta^2 \ln L}{\delta \underline{b} \delta \underline{b}^T}\}]^{-1} \triangleq F^{-1} \tag{3-30}$$

where $\hat{\underline{\beta}}$ is the vector of estimated parameters for $\underline{b}$, F is the so-called Fisher information matrix, and L is the likelihood function for N observations. Almost all the common methods (e.g. autocorrelation, autocovariance, forward-backward, Burg method, etc.) for estimating the reflection coefficients are ML estimates when the data length becomes asymptotically large [33]. Therefore, we see that this lower bound can be achieved asymptotically by any unbiased estimator.

Here we are interested in the CRLB for the variance of the MBSE estimates. Detailed derivations of the CRLB for

the variance of a parametric PSD estimate $\hat{S}(\underline{g},\omega)$ of a stationary zero mean process are given in [37]. From (eqn. 3-23,[37]) we get

$$\mathrm{Var}\{\hat{S}(\underline{g},\omega)\} \geq \{\nabla_{\underline{g}} S(\underline{g},\omega)\}^T F^{-1} \{\nabla_{\underline{g}} S(\underline{g},\omega)\} \tag{3-31}$$

where

$$\hat{S}(\underline{g},\omega) = \hat{\sigma}^2 \frac{\hat{B}(z)\hat{B}(z^{-1})}{\hat{A}(z)\hat{A}(z^{-1})}\Big|_{z=e^{j\omega}}$$

$$\underline{g} = [a_1, \ldots, a_n, b_1, \ldots, b_m, \sigma]^T$$

$$\nabla_{\underline{g}} = [\frac{\delta}{\delta a_1} S(\underline{g},\omega), \ldots, \frac{\delta}{\delta a_n} S(\underline{g},\omega), \frac{\delta}{\delta b_1} S(\underline{g},\omega), \ldots, \frac{\delta}{\delta b_m} S(\underline{g},\omega), \frac{\delta}{\delta \sigma} S(\underline{g},\omega)]^T$$

A theorem due to Whittle (Theorem 9, [38]) can be used to compute F numerically.

$$F = \frac{N}{4\pi} \int_{-\pi}^{\pi} \underline{V}(\underline{g},\omega)\underline{V}^T(\underline{g},\omega)d\omega \tag{3-32}$$

where

$$\underline{V}(\underline{g},\omega) = \frac{1}{S(\underline{g},\omega)} \nabla_{\underline{g}}\{S(\underline{g},\omega)\}$$

But this method can give numerical problems for spectra with sharp peaks unless special care is taken in the numerical integration algorithm. Moreover, numerical integration may

not give the exact integral value. Friedlander [39, 40] has presented an elegant method for the exact computation of the Fisher information matrix. From his results we find that the CRLB for $\mathrm{Var}\{\hat{S}(\underline{g},\omega)\}$ can be expressed as an explicit function of the model parameters based on certain properties of stationary covariance matrices. The Fisher information matrix for the autoregressive moving average (ARMA) parameters is given by

$$F = N \begin{bmatrix} R_{xx} & -R_{xz} & \underline{0} \\ -R_{zx} & R_{zz} & \underline{0} \\ \underline{0}^T & \underline{0}^T & 2/\sigma^2 \end{bmatrix} \qquad (3\text{-}33a)$$

where

$$R_{xx} = [A_1 A_1^T - A_2 A_2^T]^{-1}$$

$$R_{zz} = [B_1 B_1^T - B_2 B_2^T]^{-1}$$

$$R_{xz} = [A_1 B_1^T - B_2 A_2^T]^{-1}$$

$$A_1 = \begin{bmatrix} 1 & & & & \\ a_1 & \cdot & & & \\ \cdot & \cdot & \cdot & & \\ \cdot & \cdot & \cdot & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ a_{q-1} & \cdot & \cdot & \cdot & a_1 \quad 1 \end{bmatrix}$$

$$A_2 = \begin{bmatrix} a_q & & & & \\ a_{q-1} & \cdot & & & \\ \cdot & \cdot & \cdot & & \\ & & & & 0 \\ \cdot & \cdot & \cdot & \cdot & \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ a_1 & \cdot & \cdot & \cdot & a_{q-1} \quad 1 \end{bmatrix}$$

$$q = \max\{ m, n \}$$

and B_1, B_2 are similarly defined as A_1, A_2 respectively. Hence, for a pure AR process we get

$$F^{-1} = \frac{1}{N} \begin{bmatrix} A_1 A_1^T - A_2 A_2^T & \underline{0} \\ \underline{0}^T & \sigma^2/2 \end{bmatrix} \qquad (3\text{-}33b)$$

and for a pure moving average (MA) process

$$F^{-1} = \frac{1}{N} \begin{bmatrix} B_1 B_1^T - B_2 B_2^T & \underline{0} \\ \underline{0}^T & \sigma^2/2 \end{bmatrix} \qquad (3\text{-}33c)$$

Thus we find that F^{-1} decreases as $1/N$ and hence asymptotically the variance of the PSD estimates decreases with the inverse of the number of observations, and no unbiased estimator exists whose error variance dies out faster than $1/N$.

Although the MBSE is not an unbiased estimator for short data records and the CRLB provides a tight bound only asymptotically, we still compute the CRLB for the variance of the MBSE estimates in order to provide a standard against which to measure our results. It can be shown that all three averaging methods become equivalent when the data length is asymptotically large and the covariance of the estimated parameters then depends on the total number of data points. For example, in the AVA method, the covariance of the estimated parameters for any section is reduced by M/N. But when the AR parameters from all the segments are averaged, the covariance matrix is again reduced by 1/M i.e. in effect the covariance of the averaged AR parameters is reduced by 1/N.

Here we have computed the CRLB in two ways. The CRLB is evaluated for the system generating the data ('Cramer data') and also for the AR model $1/\tilde{A}(z)$ which is to be estimated ('Cramer model'). The 'Cramer data' bound should give the lowest bound achievable by the most efficient estimator for the given particular process. The 'Cramer model' bound should give the lowest bound achievable by any ARSPE for the given model order. We like to emphasize again that these bounds hold for unbiased estimators when the data length is asymptotically large.

3.10 SUMMARY

In this chapter we have formulated two approximation techniques to derive the theoretical mean and variance of the modified Burg spectral estimator (MBSE). We have defined an MBSE as the Burg spectral estimator (BSE) with some averaging technique applied. Three different types of segment averaging are considered. The first one (AVA) averages the autoregressive (AR) coefficients computed from each section. The resulting averaged AR parameters then yield the power spectral density (PSD) estimate. The second one (AVK) averages the reflection coefficients evaluated from each segment; these averaged reflection coefficients are used to compute the corresponding spectral density estimate. The final approach (AVP) evaluates the PSD estimate associated with each segment, and then averages these directly. Two approximation methods, namely the Sakai approximation and the Taylor approximation, are formulated to evaluate the statistical properties of the modified Burg spectral estimator. The Taylor approximation is different from the other one mainly because it neglects the bias and variance of the gain factor of the power spectral density estimate. Both approximation methods use a recursion technique to evaluate the approximate mean and variance of the MBSE estimates. We have also derived the lower bound for

the mean and variance of the reflection coefficient estimates. But we have failed to present any significant analysis of this theoretical bound for the variance of the reflection coefficient computed from the Burg method. Finally, we have discussed the theoretical Cramer-Rao lower bound for the variances of parametric spectral estimators with special reference to the modified Burg spectral estimator.

Chapter IV

SIMULATION RESULTS

In this chapter the effect of segment averaging on the quality of the BSE is investigated experimentally. When any segment averaging technique is applied on the BSE, the resulting estimator is defined as a modified Burg spectral estimator (MBSE). We compute the PSD estimate from a given time series generated by computer. We are concerned specifically with the statistical properties of these PSD estimates. Three different types of time series are considered. They are generated by MA, AR, and ARMA filters respectively, driven by white noise. AVA, AVK and AVP methods are then applied to each type of process for estimating the PSD using the MBSE. The analytical values of the mean and variance of the MBSE estimates are compared against the corresponding sample mean and sample variance. The CRLB bounds are also computed for analyzing the results. The Welch method, which uses a segment averaging technique, has also been used to evaluate PSD estimates and bias, variance and mean square error are compared with those obtained from the MBSE procedure. The numerical methods presented here are by no means exhaustive but they do serve to illustrate some interesting properties as pointed out in

subsequent sections. All observations made are pertinent to the few data types considered, but as they represent high pass, low pass, and band pass processes it is expected that our results would be applicable for a large class of processes. In Appendix D all computer programs (in FORTRAN) used in this simulation are included.

4.1 EXPERIMENTAL PROCEDURE

We have simulated three different classes of time series by driving AR, MA, and ARMA filters respectively, with zero mean white Gaussian noise. Thus we have obtained data representing MA, AR, and ARMA processes. These data sequences result in PSD estimates for the processes. Fifty statistically independent realizations for each type of time series were generated in order to compute the sample mean and the sample variance of the PSD estimates. The first 128 data points of each realization were discarded to allow the transient, that arises in the generation of the process, to decay. The variance of the noise is adjusted in such a way that the average power of the filter output is unity. Alternatively, the data points are normalized so that the sample autocorrelation at lag zero is unity. This normalization should correspond to the average power of unity since theoretically the following relation holds.

$$r_0 = \frac{1}{2\pi} \int_{-\pi}^{\pi} P(\omega) d\omega \qquad (4\text{-}1)$$

For convenience, we have also normalized the radian frequency ω by dividing it by π.

The sequences so generated were analyzed using four different segment averaging techniques: AVA, AVK, AVP and the Welch procedure. The N data points are divided into M nonoverlapping sections, so that each section has N/M data points. For each section, the AR parameters, the reflection coefficients, the MBSE estimates, and the Fourier transforms of the weighted data as in the Welch procedure are computed. According to the type of averaging technique the appropriate parameters are then averaged. In this experiment we have taken 128 data points which can be considered a moderate amount. By 'a moderate amount' we mean that the data length is a few times greater than the order of the estimated model or the data length is small enough so that asymptotic results do not apply. The performance of the above methods is investigated for 1, 2, 4, 8 and 16 sections.

The order p of the AR model which is used to fit the data is assumed known. For AVA the PSD estimates are given by

$$\hat{P}(\omega) = \frac{\hat{\sigma}^2}{|1 + \hat{\underline{a}}_p^T \underline{\xi}(e^{j\omega})|^2} \qquad (4\text{-}2a)$$

where the $\hat{\underline{a}}_p$ is the average of the AR parameter vector $\hat{\underline{a}}_{m,p}$ evaluated for each section m, and $\hat{\sigma}^2$ is the estimated gain factor. For AVK the PSD estimates are given by

$$\hat{P}(\omega) = \frac{\hat{\sigma}^2}{|1 + \hat{\underline{\hat{a}}}_p^T \underline{\xi}(e^{j\omega})|^2} \tag{4-2b}$$

where $\hat{\underline{\hat{a}}}_p$ is the AR coefficient vector associated with the averaged reflection coefficients. For AVP the PSD estimates are given by

$$\hat{P}(\omega) = \frac{1}{M} \sum_{m=1}^{M} \frac{\hat{\sigma}^2_m}{|1 + (\hat{\underline{a}}_{m,p})^T \underline{\xi}(e^{j\omega})|^2} \tag{4-2c}$$

It is noted that for the AVP method $\hat{\sigma}^2_m$, m=1,2,...,M, is computed from the mth section consisting of N/M samples whereas in the AVA and AVK methods we compute the gain factor from N samples. For the Welch method the PSD estimates at equally spaced frequencies are given by

$$B^W_{xx}(2\pi Mk/N) = \frac{1}{M} \sum_{i=1}^{M} W^i_{N/M}(2\pi Mk/N), \qquad k=0,1,\ldots,\frac{N}{M}-1 \tag{4-2d}$$

where

$$W^i_{N/M}(2\pi Mk/N) = \frac{M}{NU} \left| \sum_{t=0}^{(N/M)-1} x^i_t w_t \exp(-j2\pi Mkt/N) \right|^2$$

$$U = \frac{M}{N} \sum_{t=0}^{(N/M)-1} w^2_t$$

We have only used the rectangular window and hence the weighting function w_t is unity. An appropriate FFT algorithm is used to compute $W^i_{N/M}(2\pi Mk/N)$ so that the number of computations is reduced. It is easy to implement an FFT on a data sequence of length 2 to the power M. There are different ways to estimate the gain factor for spectral matching. For example, σ^2 can be estimated from the squares of the forward and the backward residual sequences which is minimized with respect to the reflection coefficient; it can also be estimated from (2-4b). However, we have found that the following estimate gives satisfactory results in the sense that both the estimated and the actual PSD have the same average power level.

$$\hat{\sigma}^2 = \sum_{t=0}^{N-1} x_t^2/(N\hat{r}_0) \qquad (4\text{-}3)$$

The CRLB for the variance of the MBSE estimates is computed for measuring the efficiency of this estimator. The CRLB is evaluated in two ways. It is computed for the system generating the data ('Cramer data') and also for the AR model $1/\tilde{A}(z)$ which is implicitly to be estimated ('Cramer model'). Though the CRLB gives a tight bound for a large

amount of data for unbiased estimators, it is still found to provide a useful reference for the MBSE which operates on a moderate amount of data.

4.2 MODIFIED BURG SPECTRAL ESTIMATOR (MBSE)

4.2.1 MA Data

A third order low pass MA filter is used to generate the data sequence of an MA process. The zeros of the MA filter are: a real one at -0.60538 and a pair of complex conjugate zeros at $-0.072085 \pm j0.63833$ in the z-plane. The true autocorrelation sequence of this MA process is evaluated by means of the algorithm presented by Dugre et al. [30]. The autocorrelation sequence is then normalized and shown in Fig. 2a. It is worth to mention that the exact autocorrelation at lag (q+1) or more is always zero for an MA(q) process.

We have assumed a fifth order AR model for fitting the data of the MA process. Now AVA, AVK and AVP methods are applied to this data sequence for estimating the PSD of the process. The actual PSD of the MA process has a relatively small dynamic range, i.e. $\delta P(\omega)/\delta\omega$ is not too high. Let us first analyze the results of the AVA method. From Fig. 3a it is evident that the bias increases with an increase in the number of segments, particularly in the frequency band

where most of the power lies. However, it is observed that this increment is small as long as the number of sections, M, does not exceed 4. For sixteen sections, the AVA method failed to give a good estimate of the actual PSD. The number of data points per section is only 8 when M is 16; and this number of data points is very small with respect to the order of the AR(5) model. To estimate the AR parameters for each section we first have to estimate the autocorrelation sequence from lag zero to lag five from the 8 data points. For short data records the variance of $\hat{r}_i$ is very high. As a result the variance of the AR parameters is also high. We get therefore, a poor estimate of the PSD when M is 16. In fact, it has been observed that when the number of segments exceeds M_o, the transition level for 128 data points, the performance of the AVA method starts to deteriorate rapidly; this will be shown later. We define M_o sections as a transition point or level because if the number of sections exceeds M_o the efficiency of the estimator starts to degrade. The sample mean of the PSD estimates is then compared against the theoretical mean evaluated from the Sakai approximation and the Taylor approximation. It is shown in Figs. 3b, 3c and 3d that the expected mean computed from the Sakai approximation follows the sample mean more closely than that computed from the Taylor approximation.

Hence, it can be stated that the Sakai approximation gives better results in comparison to the Taylor approximation. This observation is expected because in the Taylor approximation $\Delta\hat{\sigma}^2$ is assumed zero, which is not true. Again, when M is 16, both approximation methods are substantially different and both fail to give the expected mean value of the PSD estimates (see Fig. 3d). The data from the immediately adjacent sections might well be correlated when the segment length is of the order of the process correlation and thus it could violate assumption A4 on which these approximation methods are based.

From Fig. 4a we find that the variance of the PSD estimates does not change appreciably for 1, 2 and 4 sections. However only at the zero frequency, which is the main peak frequency of the actual PSD, it is observed that the variance decreases with an increase in the number of sections, but this decrement is small. So, it can be inferred that the variance of the PSD estimates does not steadily decrease with an increase in the number of sections. It appears that up to a certain number of sections, say M_o, the segmentation has little effect on the variance; as if the variance depends on the total number of data points, a behavior similar to the CRLB. The variance of the PSD estimates for 16 sections is high in comparison to

the variance for 1, 2 and 4 sections. The reason is that when M is 16, the estimated AR parameters for each section vary too much from their mean values. The theoretical variance of the PSD estimates computed from the Sakai approximation and the Taylor approximation are shown in Figs. 4b, 4c and 4d. They are compared against the corresponding sample variance. It has been observed that both approximation methods give almost the same variance. So we can see that the gain factor has a negligible effect on the variance. The 'Cramer data' bound and the 'Cramer model' bound are also shown in Figs. 4b, 4c and 4d. It is noted from these figures that in most cases the 'Cramer data' bound is lower than the sample variance of the PSD estimates, even though the MBSE is not an unbiased estimator for short data records. The 'Cramer model' lower bound has a shape similar to the sample variance but the average levels are different. The sample variance of the MBSE estimates is found to be close to the theoretical lower bound when the number of segments is less than or equal to 4, i.e. the MBSE is still efficient for a moderate number of data points. The 'Cramer model' bound is supposed to give the theoretical lower bound for the variance of the MBSE. But we have found that generally the 'Cramer model' bound is not lower than the sample variance in the frequency band where most of the

power lies. The underlying reason is that the CRLB is based on the unbiasedness property of the parametric estimator. It is mentioned in [37] that the exact bound of the variance of a biased parametric estimator depends on the gradient of the bias with respect to the true parameters. We quote [37]

> *"...no matter how small the bias is, as long as its gradient is not zero, parametric estimates with variances lower than the bound can be expected."*

An interesting relation is found between the 'Cramer model' bound and the variances evaluated from the approximation methods. The 'Cramer model' bound is found to follow closely the variances given by the approximation methods and the 'Cramer model' bound always remains lower than those variances.

The mean square error (MSE) of the PSD estimates is shown in Fig. 5a. It is noted from this figure that the MSE increases by a small amount when the number of segments is increased up to 4. So, a little will be lost if we divide the 128 data points in up to 4 segments. It is evident that when M is 16, the MSE will be pretty high for the reasons mentioned earlier. In Figs. 5b to 5d the theoretical MSEs of the MBSE estimates, computed from the Sakai approximation and the Taylor approximation, are shown. The behavior of

these MSEs is found to be similar to that of the sample MSE. The theoretical lower bound of the MSE of the estimates can be computed from the 'Cramer model' bound and is defined as follows

$$\text{MSE('Cramer model')} = \Delta\tilde{P}^2(\omega) + \text{Var('Cramer model')} \tag{4-4}$$

Again we find that the sample MSE of the PSD estimates is lower than this theoretical lower bound for the reasons mentioned earlier.

Let us now focus our attention on the results obtained from the AVK method. In Figs 6a to 6c the PSD estimates for different number of sections are shown. Simultaneously, the expected mean computed from the approximation methods and the true PSD of the MA process are also shown. In Figs. 7a to 7c the sample variance and the corresponding analytical variances of the PSD estimates are shown. Finally, the sample MSE of the estimates and the corresponding MSE evaluated from the approximation methods are plotted in Figs. 8a to 8c. The theoretical lower bound is also computed for the purpose of comparison. After close scrutiny of these figures we find that the AVA and AVK methods give almost the same results. However, it is noticed that the performance of the AVK method is comparatively better than that of the AVA method.

The AVP method is the worst among these three averaging techniques because it gives the largest variance of the PSD estimates. Especially when the number of sections is 16, the AVP method completely fails to estimate the PSD of the process (see Figs. 9c, 10c and 11c). It gives a very high variance at the normalized radian frequencies zero and one. Even for 4 sections the results are not very satisfactory (see Figs. 9b, 10b and 11b). The sample mean, sample variance, and sample MSE of the estimates are shown in Figs. 9a, 10a, and 11a for 1, 2, 4 and 16 sections. Since in this method the PSD estimates computed from each section are directly averaged, it is expected that the averaged PSD estimate will be more erratic. As the poles of the estimated AR model for each section cannot concentrate around their mean position for MA data, we observe that the sample variance is higher than that of the AVA and AVK methods. Unlike both AVA and AVK methods, it is observed that at the zero frequency the variance increases with the number of sections. We like to mention that for only one section, the AVA, AVK and AVP methods give identical results since no averaging is performed and in this case we are simply using the BSE.

4.2.2 AR Data

A fourth order band pass AR filter is used to generate the data sequence of an AR process. The poles of the AR filter are: two real ones at ±0.8 and a pair of complex conjugate poles at +0.5±j0.75 in the z-plane. The exact normalized autocorrelation sequence of the AR process is computed and shown in Fig. 2b. This autocorrelation sequence oscillates between positive and negative values and it goes nearly to zero for a lag higher than 32.

We have assumed the order of the AR model to be the same as that of the data generator. Now AVA, AVK and AVP methods are applied to this data sequence for estimating the PSD of the process. Let us study the results of the AVA method. From Fig. 12a it is evident that the bias increases with an increase in the number of segments, particularly in the frequency band where most of the power lies. However, it is observed that this increment is small as long as the number of sections does not exceed 4. For 16 sections the AVA method fails to give a good estimate of the actual PSD, because the data length per section is short with respect to the length of the process correlation. As a result, the variances of the AR parameters increase, which is reflected in the corresponding PSD estimates. It is evident that somewhere between 4 and 16 sections, there is a transition

point M_o; and the AVA method starts to lose its efficiency when M exceeds M_o. The sample mean of the PSD estimates is compared against the theoretical mean evaluated from the Sakai approximation and the Taylor approximation. It is shown in Figs. 12b and 12c that the mean computed from the Sakai approximation follows the sample mean more closely than that computed from the Taylor approximation. Hence, it can be stated that the Sakai approximation gives better results in comparison to the Taylor approximation. We made the same statement for MA data in the case of the AVA method. Another important characteristic of the approximation methods can be observed from Figs. 12b and 12c. This characteristic is that the main peak frequencies given by the approximation methods do not match exactly with the main peak frequency ω_p of the estimated PSD, when M is greater than one and the difference between the main peak frequencies increases with M. Again for 16 segments, both approximation methods fail to give the expected mean value of the PSD estimates (see Fig. 12d). Here the actual PSD has a large dynamic range. So, the first derivatives of $P(\omega)$ with respect to the AR parameters are very high near the main peak frequency ω_p of 0.31 for the actual PSD (see Appendix B). As a result it is observed that both the approximation methods change sharply near the main peak

frequency and the Taylor approximation method even gives negative values for the expected mean of the PSD estimates around that frequency. For 16 sections the bias of the estimated AR parameters is quite large so that the high values of the first derivatives are not offset. However, for 1, 2, and 4 sections the high values of $\underline{H}(e^{j\omega})$ are offset by small values of the bias. The second reason for the failure of the approximation methods using 16 sections, is the short data records per segment and hence successive sections might well be correlated. We mentioned earlier that the true autocorrelation of the AR process has decayed to near zero at lag 32 or more, so adjacent sections are actually correlated when M is 16. Thus assumption A4 certainly does not hold for 16 segments in the AR data case.

From Fig. 13a we find that the variances of the PSD estimates are almost the same for 1, 2, and 4 sections. However, at the peak frequencies 0, 0.31, and 1, the variance decreases with an increase in the number of sections but this decrement is small. As in the MA data case for the AVA method, we can state that the variance of the PSD estimates does not steadily decrease with an increase in the number of sections. It appears that up to a certain number of sections, the segmentation has only a small positive effect on the variance; as if the variance

depends largely on the total number of data points, a behavior similar to the CRLB. The variance of the PSD estimates for 16 sections is high in comparison to the variances for 1, 2, and 4. The reason is again that for such a short data record per section, the variances of the AR parameters are exceedingly high. The theoretical variance of the PSD estimates computed from the Sakai approximation and the Taylor approximation are shown in Figs. 13b, 13c and 13d. They are compared against the corresponding sample variance. It has been observed that both approximation methods give almost the same variance. So we see that the gain factor has a negligible effect on the variance. The 'Cramer data' bounds are also shown in Figs. 13b, 13c and 13d. Here we have not shown the 'Cramer model' bound because the system generating AR data and the AR model are the same. The 'Cramer data ' and the 'Cramer model' bounds will be identical therefore. Similar results as in the MA data case are observed with regard to the 'Cramer data' bound. Unlike in the MA data case for the AVA method, it is found that the sample variance is always less than the 'Cramer data' bound around the main peak frequency because the gradient of the bias is not negligible in that frequency band.

The MSE of the PSD estimates is shown in Fig. 14a. It is noted from these figures that the MSE is increased by a

small amount when the number of segments is increased up to 4 segments. So, a little will be lost if we divide the 128 data points in up to 4 segments and thereby we can operate on a smaller number of data points, but doing it more often. It is evident that for 16 sections the MSE will be fairly high. In Figs. 14b to 14d the theoretical MSEs of the MBSE estimates, computed from the Sakai approximation and the Taylor approximation, are shown. The behavior of these MSEs is found to be similar to that of the sample MSE. The theoretical lower bound of the MSE of the estimates can be computed from the 'Cramer model' bound. Again we find that the sample MSE of the PSD estimates is lower than this theoretical lower bound. In this case, the plausible explanation is that the gradient of the bias of the estimator may not be negligible whereas in the 'Cramer model' bound we assumed this gradient to be zero.

Let us now examine the results obtained from the AVK method. In Figs. 15a to 15c the PSD estimates for different number of sections are shown. Simultaneously, the expected mean computed from the approximation methods and the true PSD of the AR process are also shown. In Figs. 16a to 16c the sample variance and the corresponding analytical variances of the PSD estimates are shown. Finally, the sample MSE of the estimates and the corresponding MSE

evaluated from the approximation methods are plotted in Figs. 17a to 17c. The theoretical lower bound is also computed for the purpose of comparison. After an extensive study of these figures we find that the AVA and the AVK methods give almost the same results. However, as in the MA data case it is noticed that the performance of the AVK method is better than that of the AVA method to a certain degree, especially with respect to the variances of the PSD estimates.

The AVP method is the worst among these three averaging techniques because it gives the largest variance of the PSD estimates. Particularly when M is 16 the AVP method completely fails to estimate the PSD of the process (see Figs. 18c, 19c and 20c). It gives a very high variance at frequencies 0, 1, and .31; the latter is the main peak frequency ω_p. For 4 sections, the results are shown in Figs. 18b, 19b and 20b. Unlike in the MA data case for the AVP method, when M is 4, the results are quite satisfactory. The sample mean, sample variance, and sample MSE of the estimates are plotted in Figs. 18a, 19a, and 20a for 1, 2, 4 and 16 sections. Here, since both the generating system and the AR model are the same, the poles of the estimated AR model for each section are concentrated around the actual poles, at least for up to 4 sections. We see that the PSD

estimates are smoother than for the MA data case for the AVP method.

4.2.3 ARMA Data

The coloring filter used to generate the data for an ARMA(3,2) process has three poles: a real one at +0.8 and a pair of complex conjugate poles at -0.5±j0.5 in the z-plane. A pair of zeros of the filter is at +0.5±j0.5. So, this is a high pass filter. The exact normalized autocorrelation sequence of the ARMA process is computed and shown in Fig. 2c. This autocorrelation sequence oscillates between positive and negative values and it goes nearly to zero at lag 14. Here the actual PSD has a relatively large dynamic range.

We have assumed the fifth order AR model to fit the data. AVA, AVK and AVP methods are then applied to this data sequence for estimating the PSD of the process. Let us investigate the results of the AVA method. From Fig. 21a it is found that the bias increases with an increase in the number of segments, particularly in the frequency band where most of the power lies. However, it is observed that this increment is very insignificant as long as the number of segments does not exceed 4. Even for 16 sections the AVA method does not give a poor estimate of the actual PSD. The

reason is that the autocorrelation sequence dies out for a relatively small amount of lag and this process also has predominantly AR characteristics. We like to state that although the autocorrelation is zero for lags in excess of 3 in the MA data case, 16 segments do not give a good estimate because a finite order AR model cannot fit the data exactly. However, the PSD estimate for 16 segments is distinctly different from the corresponding estimates for 1, 2, and 4 sections. The peak frequency of the PSD estimates is quite far away from the actual peak frequency ω_p which is at 0.76 and it is expected that as the order of the AR model is increased, the estimate will resemble the actual PSD more. The sample mean of the PSD estimate is compared against the theoretical mean evaluated from the Sakai approximation and the Taylor approximation. It is shown in Figs. 21b and 21c that the mean computed from the Sakai approximation follows the sample mean more closely than that from the Taylor approximation. Hence, it can be stated that the Sakai approximation gives better results in comparison to the Taylor approximation. We made the same statement for the MA data case for the AVA method. Like the AR data case for the AVA method, it is noted that the main peak frequencies evaluated from the approximation methods do not correspond to the main peak frequency of the estimated PSD when M is

greater than one, and the difference between the main peak frequencies increases with M. Again for 16 sections both approximation methods fail to give the expected mean value of the PSD estimates (see Fig. 21d).

From Fig. 22a we have found that the variances of the PSD estimates are almost the same for 1, 2 and 4 sections. Unlike the MA and AR data case for the AVA method, the variance does not decrease at the main peak frequency with an increase in the number of sections. Hence it is shown that the variance of the PSD estimates does not steadily decrease with an increase in the number of sections. The variance of the PSD estimates for 16 sections is high in comparison to the variances for 1, 2 and 4 sections. Using only 8 data points per segment it is not possible to estimate the AR parameters with low variances when the order of the model is 5. The theoretical variance of the PSD estimates computed from the Sakai approximation and the Taylor approximation are shown in Figs. 22b, 22c and 22d. They are compared against the corresponding sample variance. It is observed that both approximation methods give almost the same variance. We see that the gain factor has a negligible effect on the variance. The 'Cramer data' and the 'Cramer model' bounds are also shown in Figs. 22b, 22c and 22d. As seen in the MA data case, we observe here the same

relations between the theoretical lower bounds and the sample variance. The 'Cramer data' bound is lower than the sample variance whereas the 'Cramer model' bound is higher, particularly in the frequency region where most of the power lies. Moreover, the shape of the 'Cramer model' bound is similar to the sample variance but the average levels are different.

The MSEs of the PSD estimates are shown in Fig. 23a. Note from these figures that the MSE increases by an insignificant amount when the number of segments is increased up to 4 sections. So, a little will be lost if we segment the data in up to 4 sections, and thereby we can operate on a smaller number of data points per section, but doing it M times, instead of operating once on the full length of data. It is evident that for 16 segments the MSE will be fairly high. In Figs. 23b to 23d the theoretical MSEs of the MBSE estimates, computed from the Sakai approximation and the Taylor approximation, are shown. The behavior of these MSEs is found to be similar to the sample MSE. The theoretical lower bound of the MSE of the estimates can be computed from the 'Cramer model' bound. Again we find that the sample MSE of the PSD estimates is lower than this theoretical lower bound because the unbiasedness assumption for the estimator is not valid.

Let us now analyze the results obtained from the AVK method. In Figs. 24a to 24c the PSD estimates for different number of sections are shown. Simultaneously, the expected mean computed from the approximation methods and the true PSD of the MA process are also shown. In Figs. 25a to 25c the sample variance and the corresponding analytical variances of the PSD estimates are shown. Finally, the sample MSE of the estimates and the corresponding MSE evaluated from the approximation methods are plotted in Figs. 26a to 26c. The theoretical lower bounds are also computed for the purpose of comparison. After an extensive study of these figures we find that the AVA and the AVK methods give almost the same results. However, as in the MA data case it is noted that the performance of the AVK method is better than that of the AVA method to a certain extent.

The AVP method is the worst among these three averaging techniques because it gives the largest variance of the PSD estimates. Particularly for 16 segments the AVP method completely fails to estimate the PSD of the process (see Figs. 27c, 28c and 29c). It gives a very high variance at frequencies 0, 1, and .76; the latter is again the main peak frequency. For 4 sections the results are shown in Figs 27b, 28b and 29b. Unlike in the MA data case for AVP, the results for 4 sections are quite satisfactory. The sample

mean, sample variance, and sample MSE of the estimates are plotted in Figs. 27a, 28a, and 29a for 1, 2, 4 and 16 sections. We note that the PSD estimates follow an erratic behavior when M is higher than 4.

We refer collectively to Figs. 30 to 38 where the mean, variance and mean square error of the PSD estimates for 8 sections are shown for all the three data cases using the three averaging techniques. Comparing with the corresponding estimates for 1, 2, 4 and 16 sections we find that in the MA data, AR data, and ARMA data cases, the transition point is reached for AVA, AVK and AVP methods when the number of sections is 8. In fact, it has been found that when the number of segments is more than eight for 128 data points, the performance of the three methods starts to degrade rapidly. Particularly in the ARMA data case when the number of sections exceeds 8 the estimated peak frequency starts to drift away from the peak frequencies for 1, 2, and 4 sections. We find therefore M_o is actually 8 for our three data cases using the three averaging methods.

Let us now find the number of operations (NOPS) required to compute the PSD estimates of the modified Burg spectral estimators. We consider each basic mathematical operation, namely multiplication, division, addition, and subtraction, as a single operation. In the following table (Table 1) the

TABLE 1

NUMBER OF COMPUTATIONS REQUIRED TO EVALUATE PSD ESTIMATES

Method	Number of operations
AVA	$2N(p+1) + M(\frac{11p^3}{6} + 5p^2 + \frac{43p}{6} - 8) + \frac{2p^3}{3} + \frac{5p^2}{2} + \frac{35p}{6} + 5$
AVK	$2N(p+1) + M(\frac{11p^3}{6} + 5p^2 + \frac{43p}{6} - 8) + \frac{2p^3}{3} + \frac{9p^2}{2} + \frac{23p}{6} + 5$
AVP	$2N(p+1) + M(\frac{15p^3}{6} + \frac{15p^2}{2} + 11p + 126)$

N = number of samples

p = order of AR model

M = number of sections

total number of operations for computing the PSD estimates using the AVA, AVK, and AVP methods are given. The plots of the number of operations versus number of sections for these methods are shown in Fig. 39. It is evident from this figure that the total NOPS is maximum for the AVP method whereas for the AVA method it is minimum. The difference between the NOPS for the AVA and AVK methods is very small for low order AR models. We note that for any averaging method the NOPS increases with the number of sections. So, the NOPS for modified Burg spectral estimator is higher than the corresponding NOPS for Burg spectral estimator.

4.3 WELCH PROCEDURE

The Welch procedure has been applied to estimate the PSD of the MA, AR, and ARMA processes which were generated. This method uses the segment averaging technique. So, it is useful to make a comparative study between this method and the MBSE. This method is a modification of the Bartlett procedure. In this case the window w_t is applied to the data segments directly before computation of the periodogram. We have considered only the rectangular window and nonoverlapping sections. Theoretically it has been proved that as the number of segments increases, the bias of the PSD estimates increases but the variance decreases. Welch

[41] shows that if the segments of x_t are nonoverlapping, then

$$\mathrm{Var}[B^{W}_{xx}(\omega)] \simeq I^{2}_{N/M}(\omega)/M \tag{4-5}$$

where $I_{N/M}(\omega)$ is the periodogram of the process. This method smoothes the periodogram. As a result resolution decreases.

In Figs. 30 to 38 we have shown the sample mean, sample variance and sample MSE of the PSD estimates computed from the Welch procedure. Simultaneously these are compared against the AVA, AVK and AVP methods. Here we have considered only 4 and 8 sections. After studying these figures, our findings indicate the following results.

a) The bias increases and the variance decreases with an increase in the number of sections when the Welch method is used for MA, AR, and ARMA data. At the same time, the MSE increases and resolution reduces. The variance approaches the theoretical lower bound ('Cramer data') as the number of sections is increased.

b) The AVA and AVK methods give comparatively better estimates in terms of variance than the Welch procedure for most of the cases.

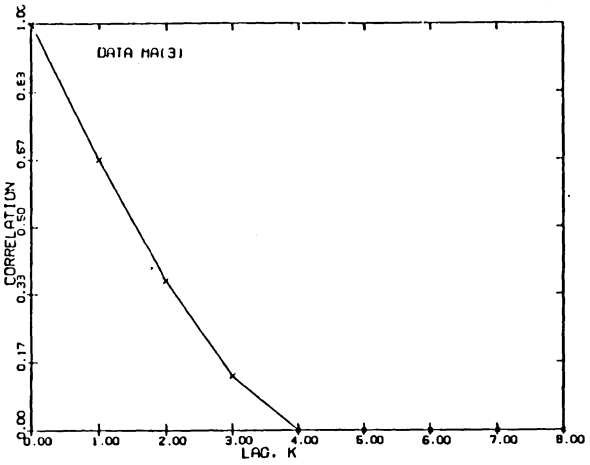
c) Though both the AVP method and the Welch procedure use the same averaging technique, but different spectral estimators, the AVP method gives relatively poor estimates at normalized frequencies 0, ω_p, and 1.

4.4 SUMMARY

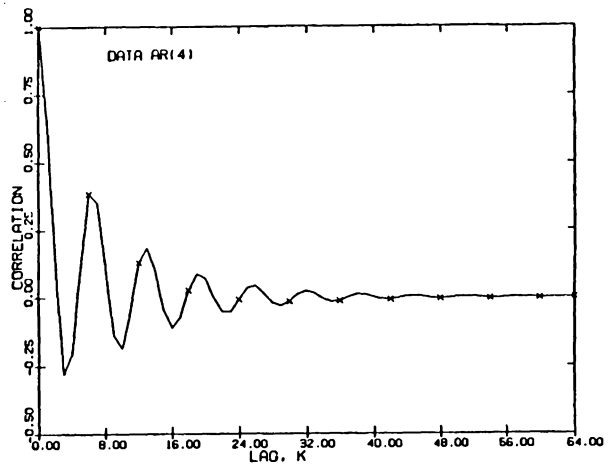
Let us now summarize the significant simulation results which are mentioned in the previous sections. The statistical properties of the AVA and AVK methods are almost the same for all three cases (see Fig. 40). Intensive studies reveal that the AVK method is slightly better than the AVA method with respect to the variance of the PSD estimates. The AVP method gives the worst results among the three averaging techniques. In most of the cases the variances at the peak frequencies of the actual PSD decrease as the number of sections is increased up to a certain point, say M_o, for both the AVA and AVK methods. However this reduction in the variances is not significant. From the experimental results we find that M_o is 8 when the total number of data is 128. The MSEs of the PSD estimates remains almost the same for these two methods up to M_o sections. When the number of sections exceeds M_o, both methods start to degrade. The MBSE estimates are found to follow the theoretical bound closely even for a moderate

number of data points provided the number of sections is less than or equal to M_o. In comparison to the Taylor approximation method, the Sakai approximation method gives better results in terms of mean and mean square error of the estimates. That is, the expected mean and the mean square error given by this method follow more closely the corresponding sample mean and sample mean square error. But both approximation methods give the same value for the variances of the power spectral density estimates. The Taylor approximation is different from the other one mainly because it neglects the bias and variance of the gain factor of the power spectral density estimate. Hence, these results imply that the variance of the gain factor has very little effect on the variance of the power spectral density estimates, while the bias of the gain factor has a considerable effect on the expected mean of the estimates. The 'Cramer data' bound is found to be lower than the sample variance in most of the cases and the 'Cramer model' bound has a shape similar to that of the sample variance, but the average levels are different. Finally, it is found that the approximation methods give good predictions and the segment averaging techniques do not reduce the variance of the PSD estimates appreciably.

a.



b.



c.

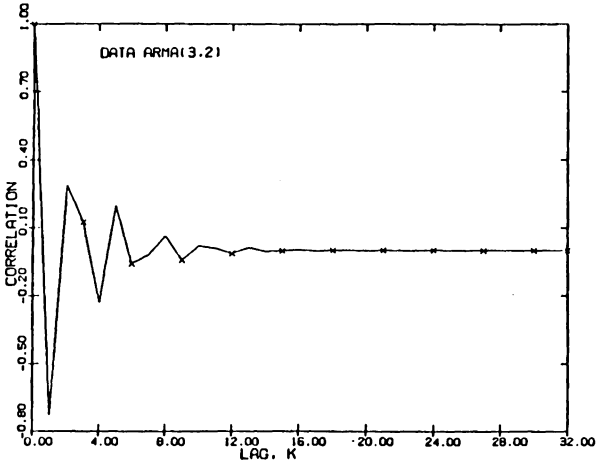
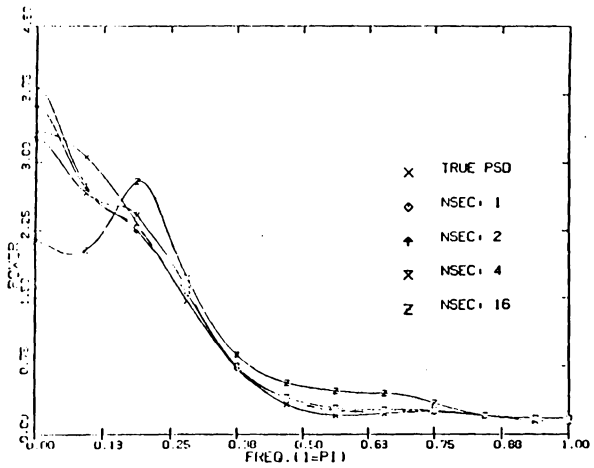
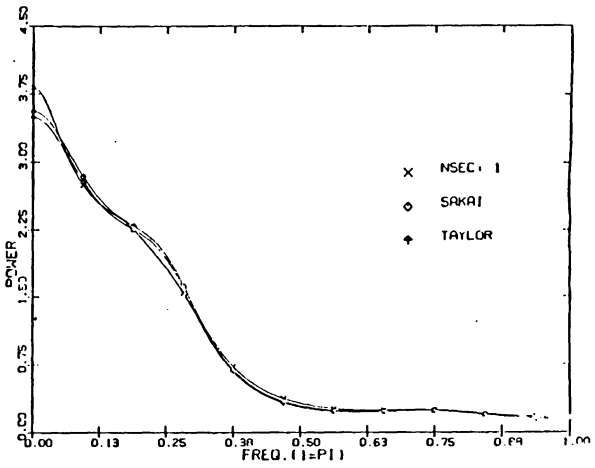


Fig. 2 Autocorrelation versus lag

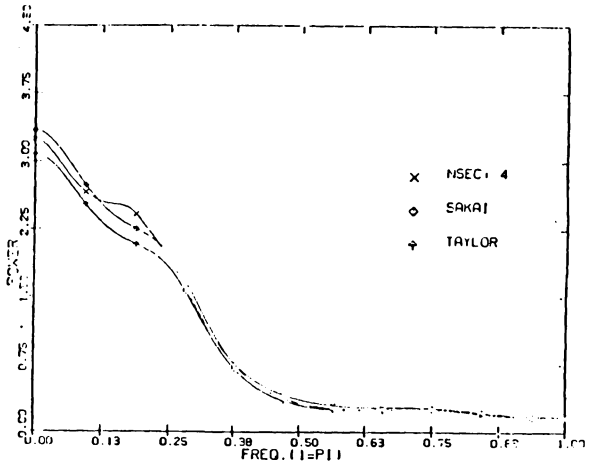
a.



b.



c.



d.

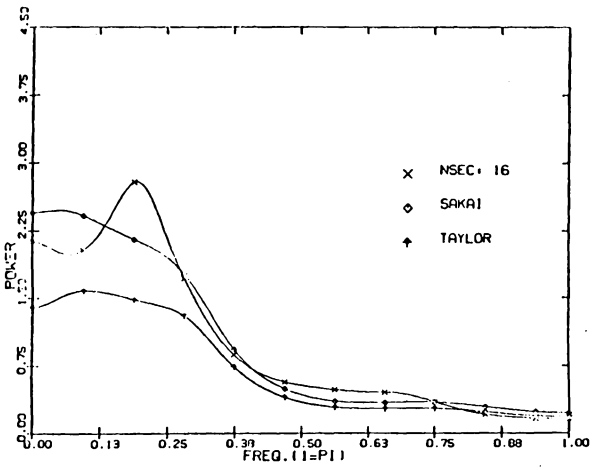
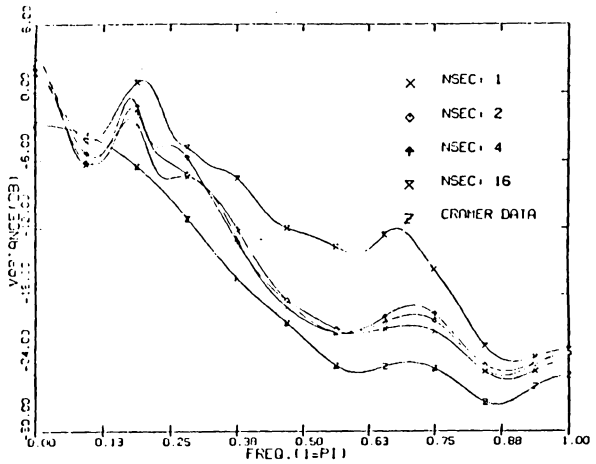
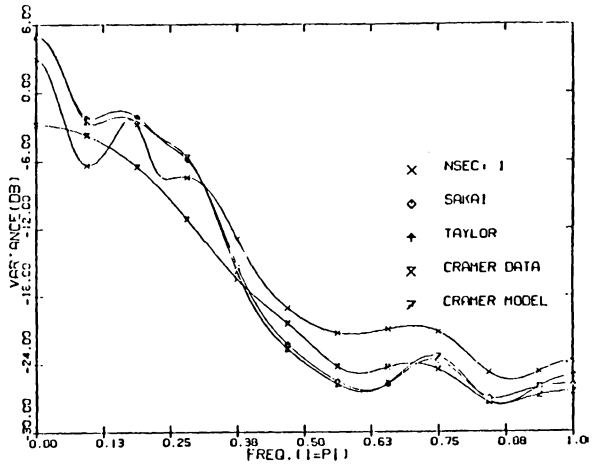


Fig. 3 Mean PSD for MA(3) data using AVA on AR(5) model

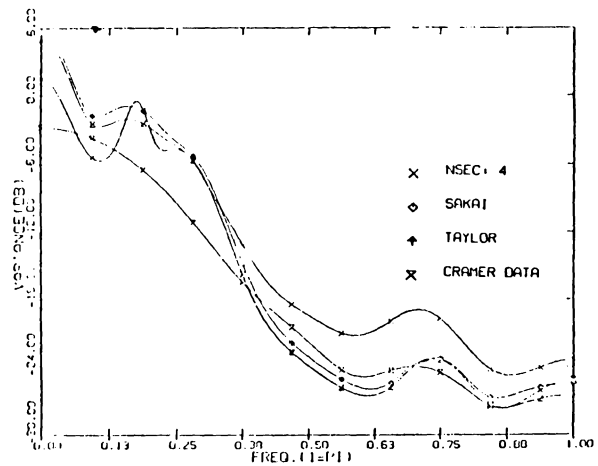
a.



b.



c.



d.

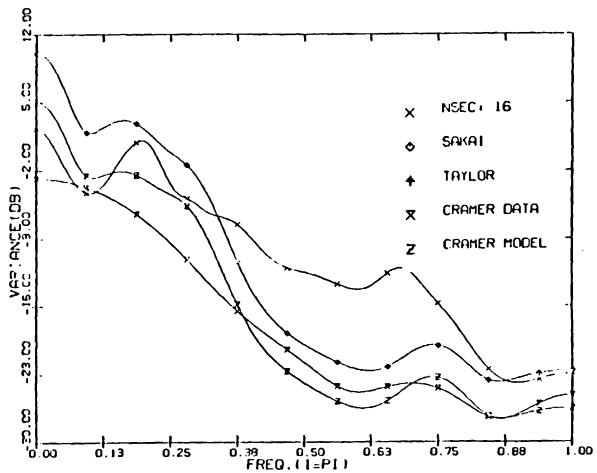
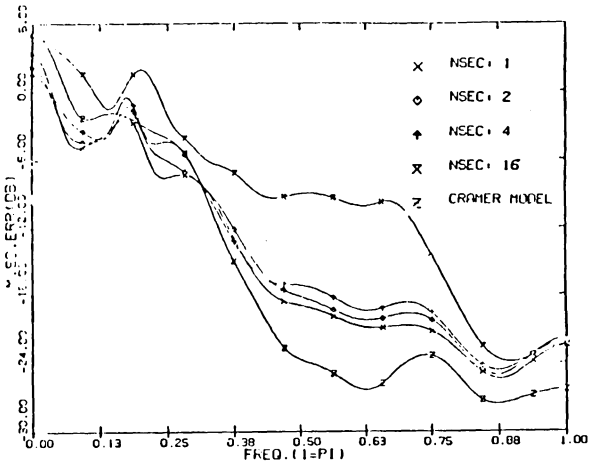
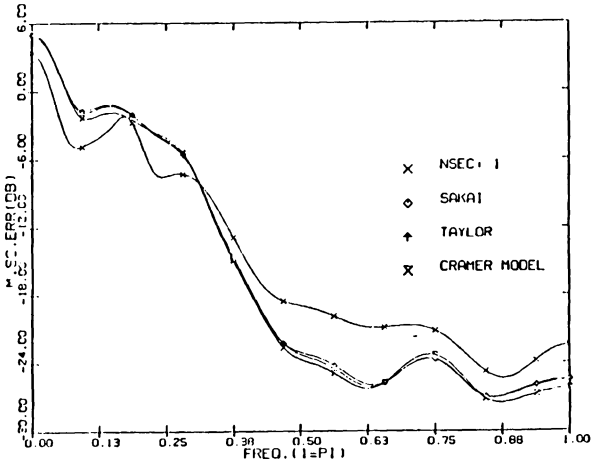


Fig. 4 Variance of PSD for MA(3) data using AVA on AR(5) model

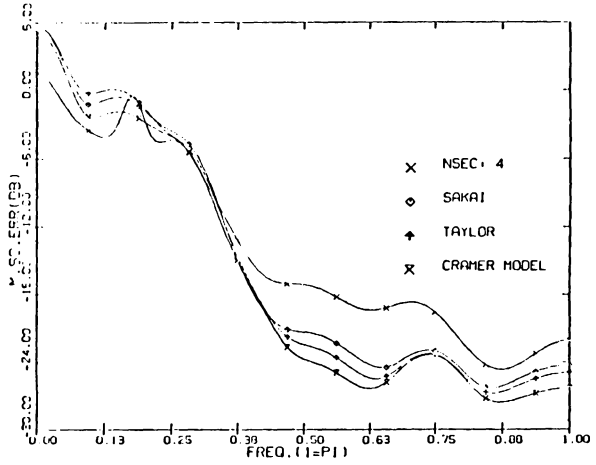
a.



b.



c.



d.

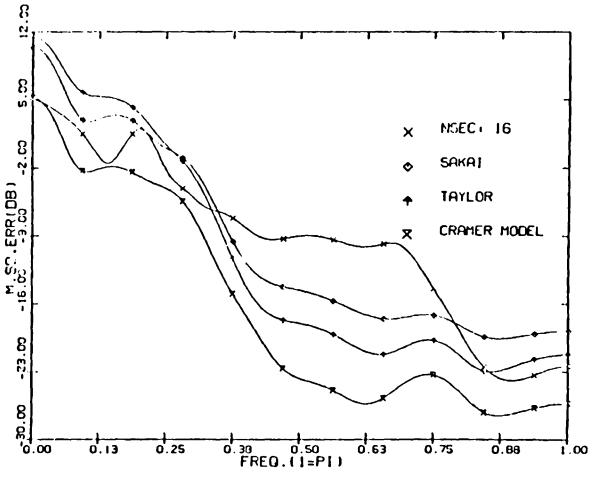
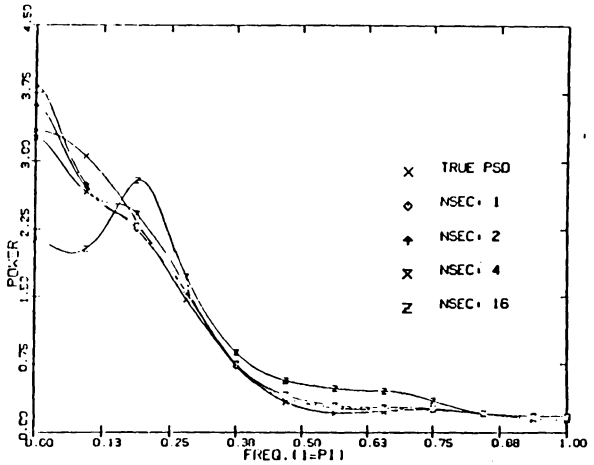
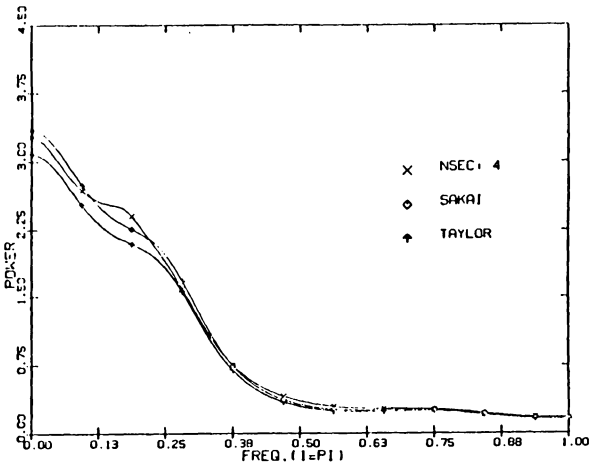


Fig. 5 Mean square error of PSD for MA(3) data using AVA on AR(5) model

a.



b.



c.

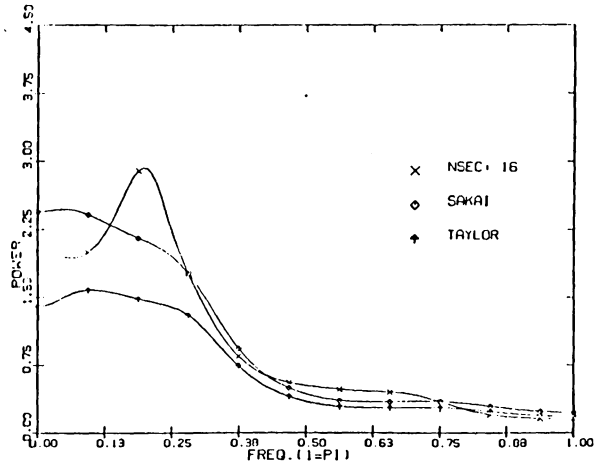
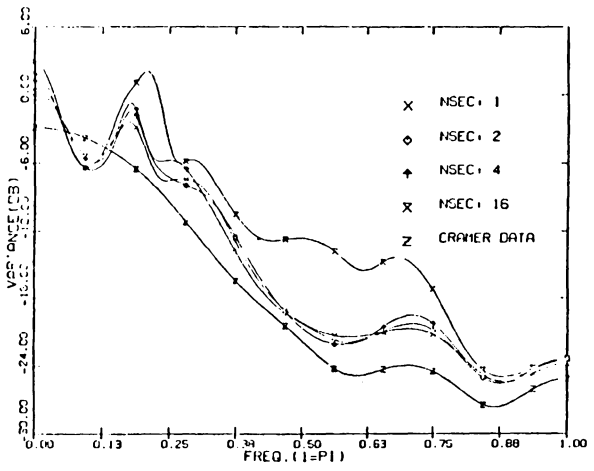
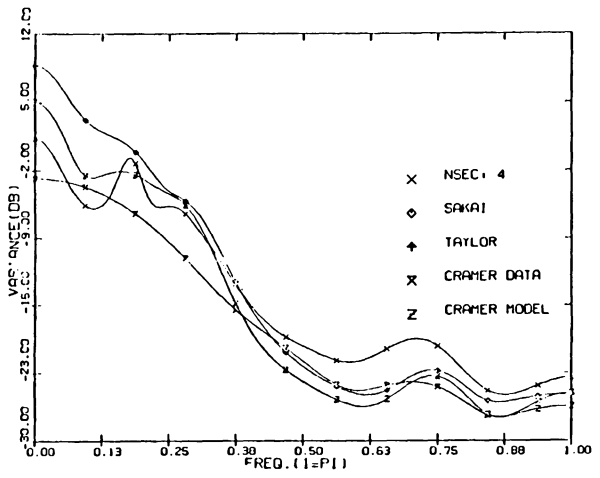


Fig. 6 Mean PSD for MA(3) data using AVK on AR(5) model

a.



b.



c.

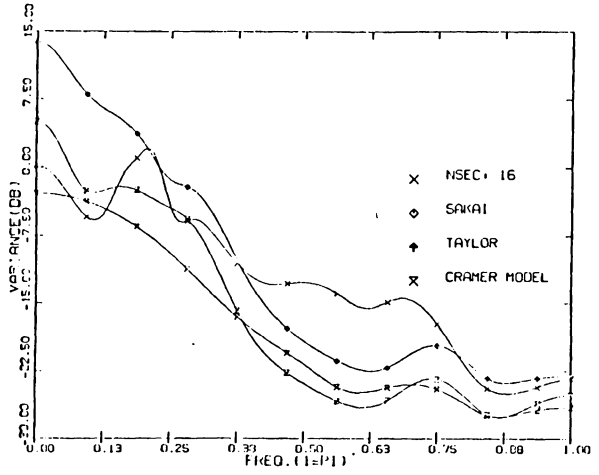


Fig. 7 Variance of PSD for MA(3) data using AVK on AR(5) model

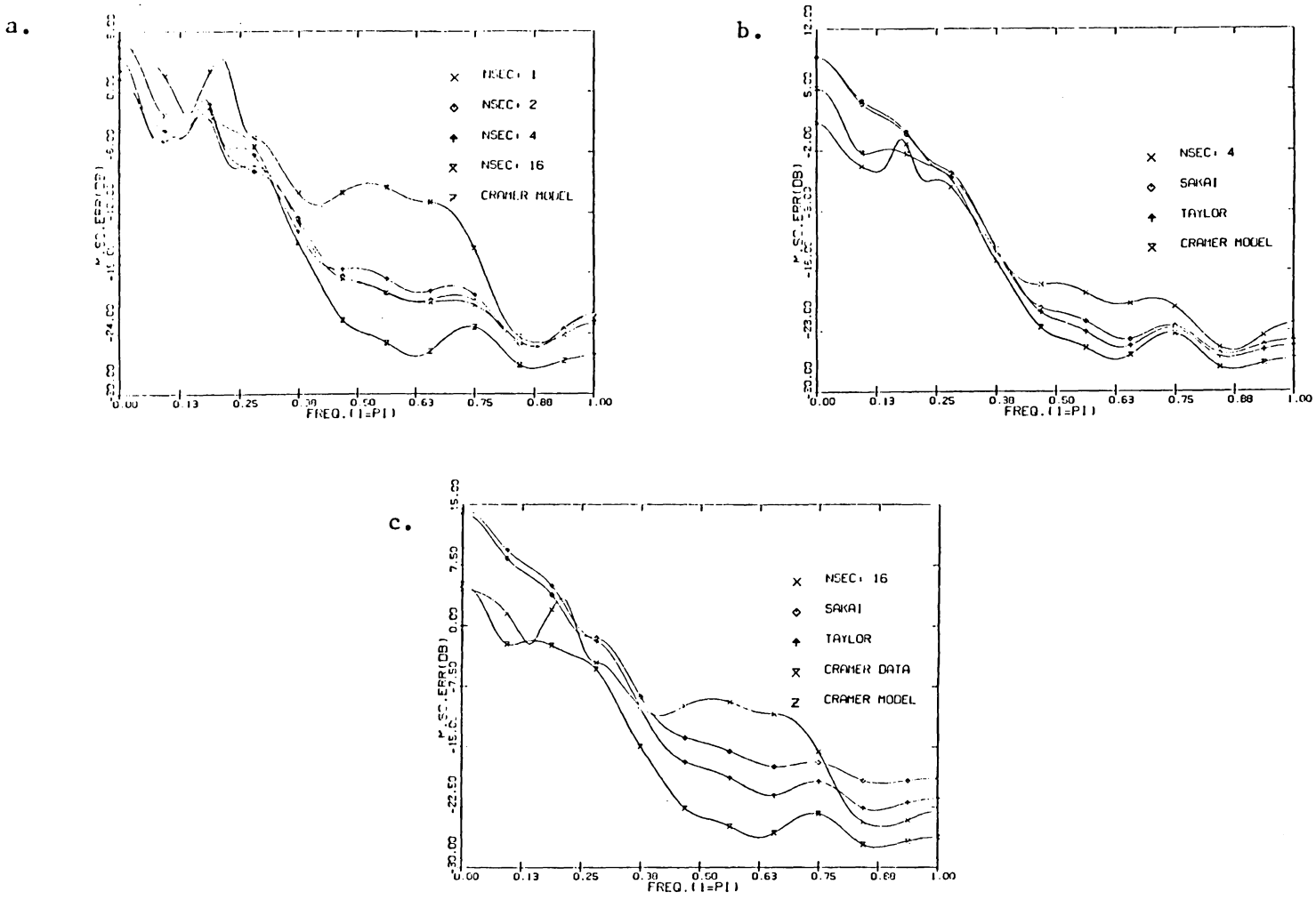
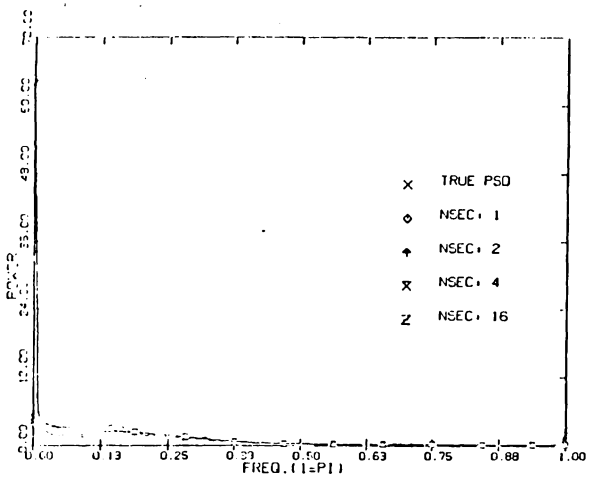
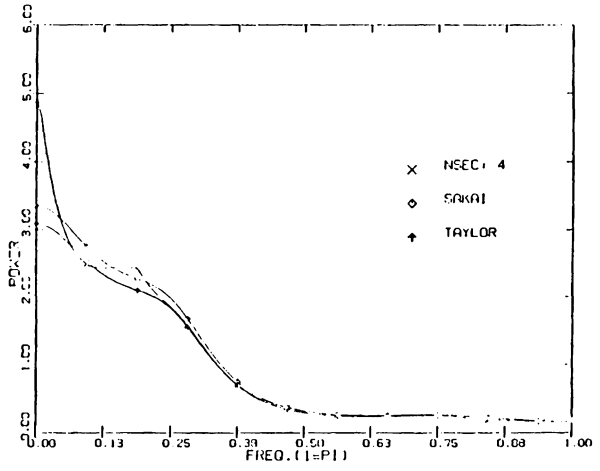


Fig. 8 Mean square error of PSD for MA(3) data using AVK on AR(5) model

a.



b.



c.

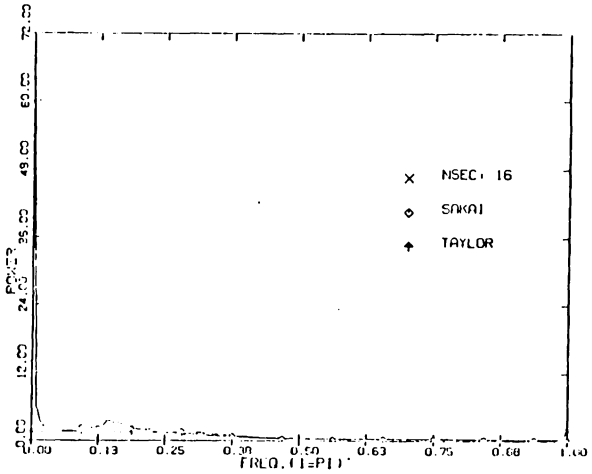
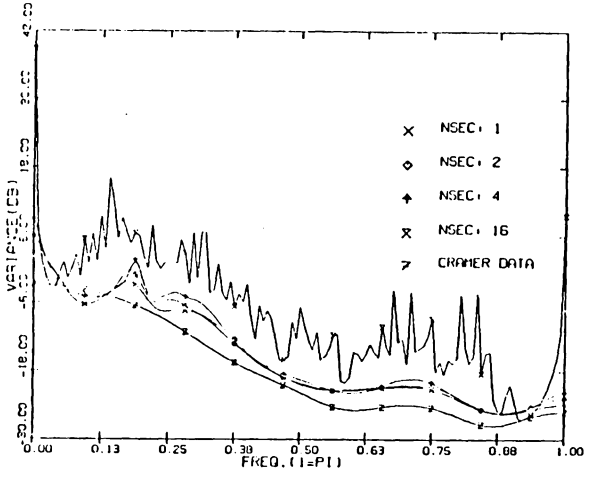
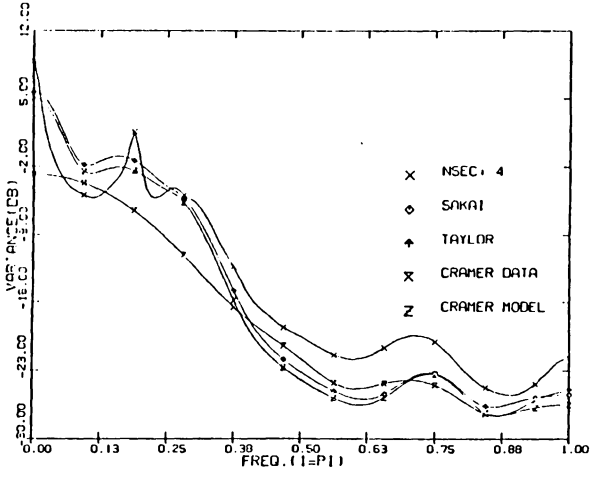


Fig. 9 Mean PSD for MA(3) data using AVP on AR(5) model

a.



b.



c.

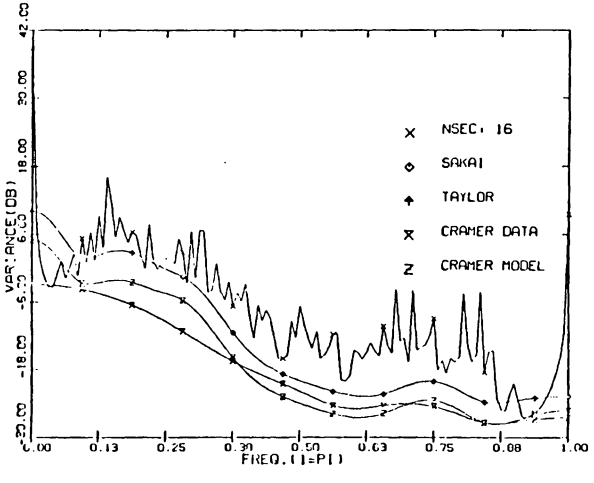
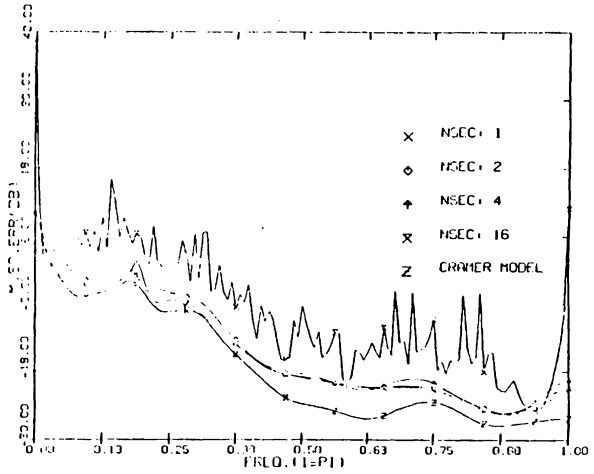
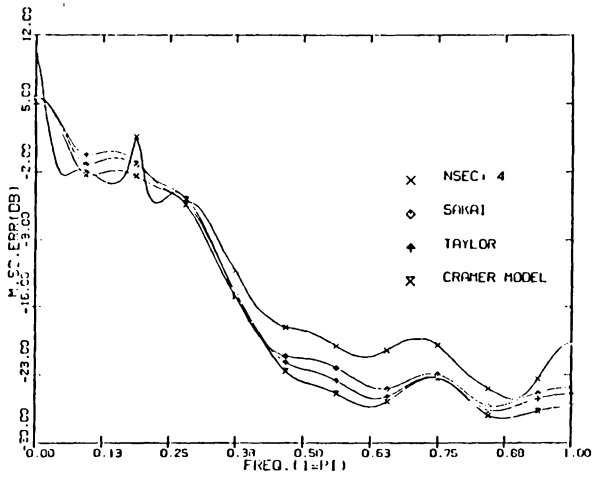


Fig. 10 Variance of PSD for MA(3) data using AVP on AR(5) model

a.



b.



c.

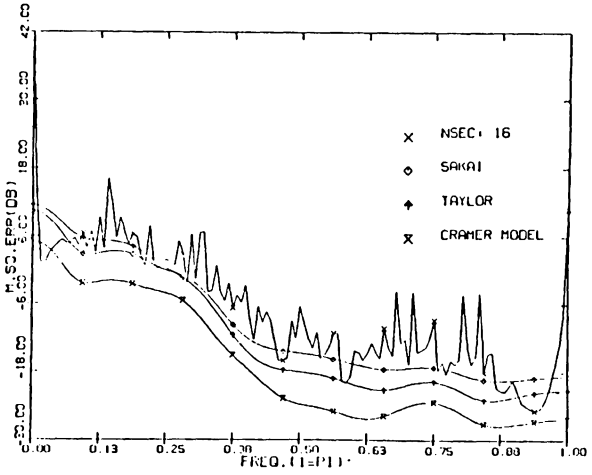


Fig. 11 Mean square error of PSD for MA(3) data using AVP on AR(5) model

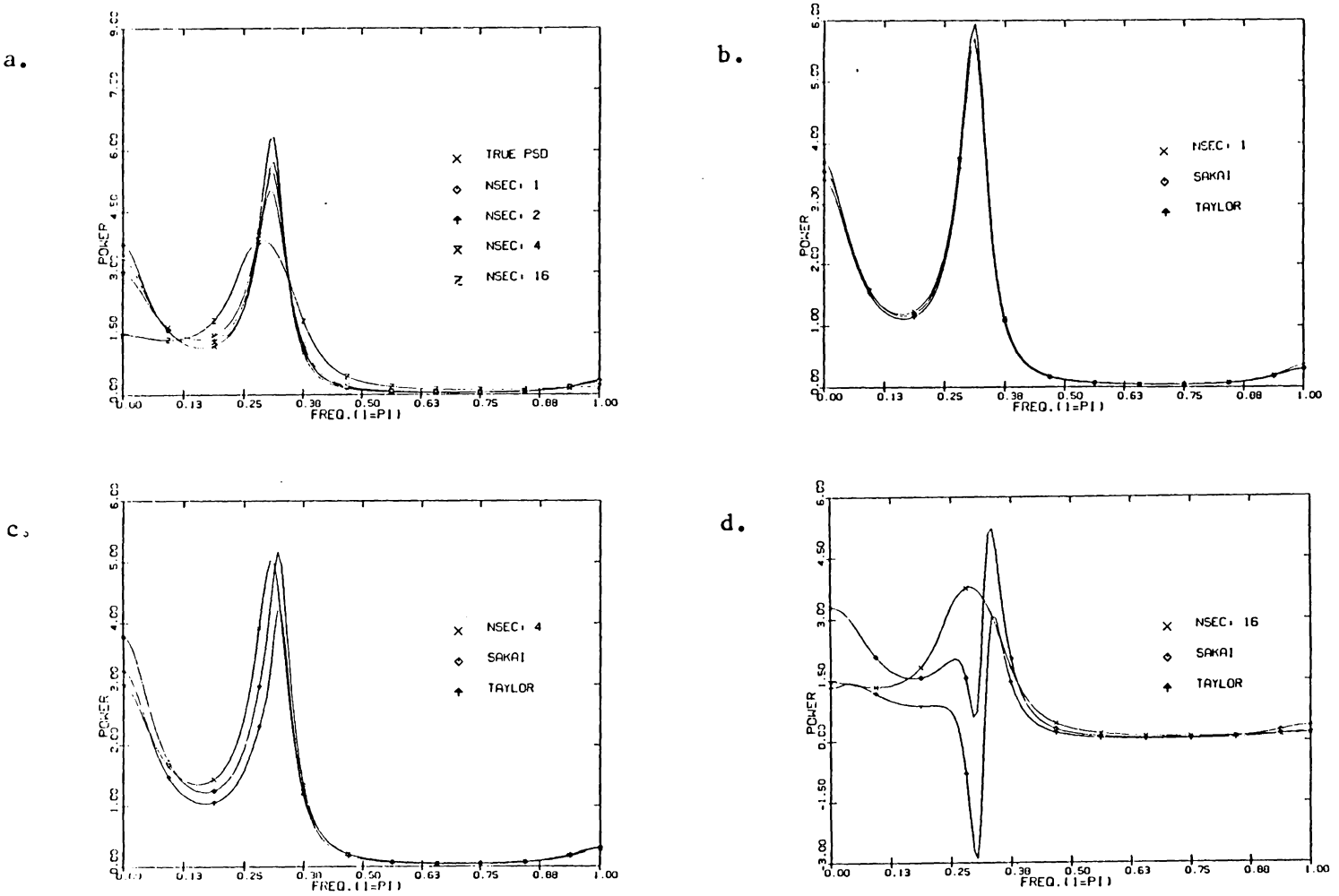
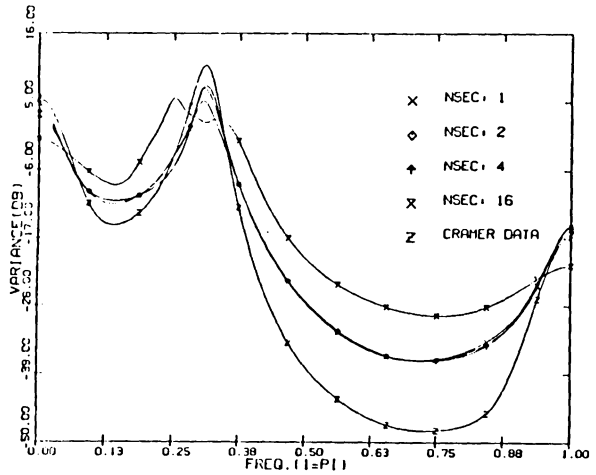
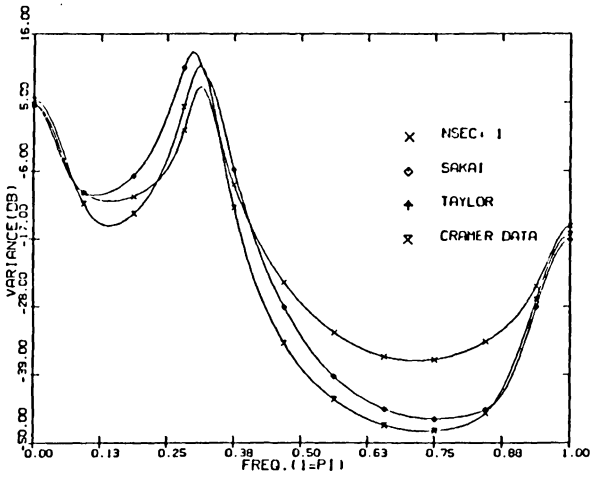


Fig. 12 Mean PSD for AR(4) data using AVA on AR(4) model

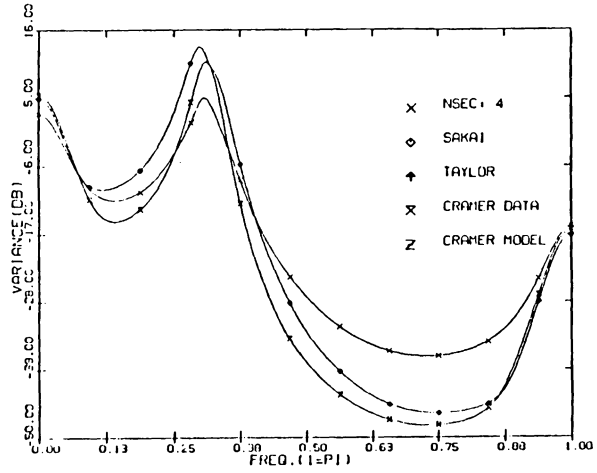
a.



b.



c.



d.

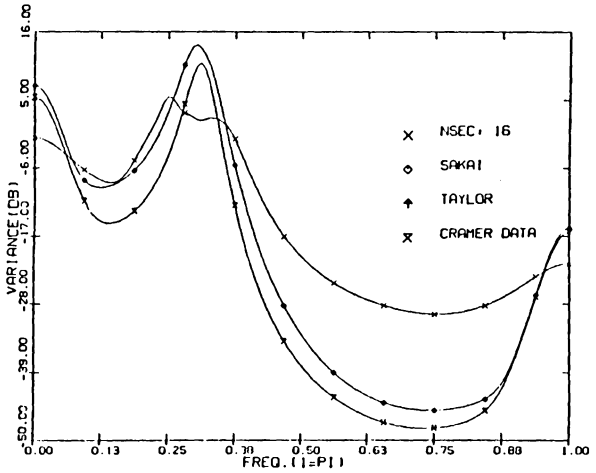
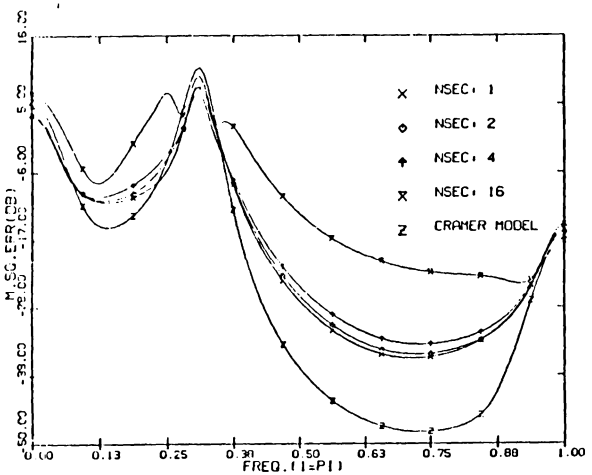
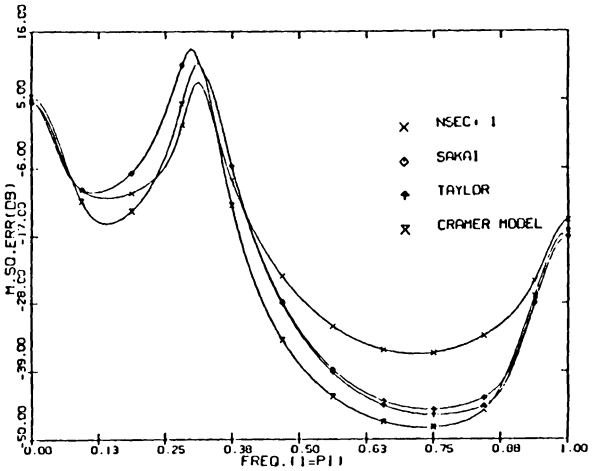


Fig. 13 Variance of PSD for AR(4) data using AVA on AR(4) model

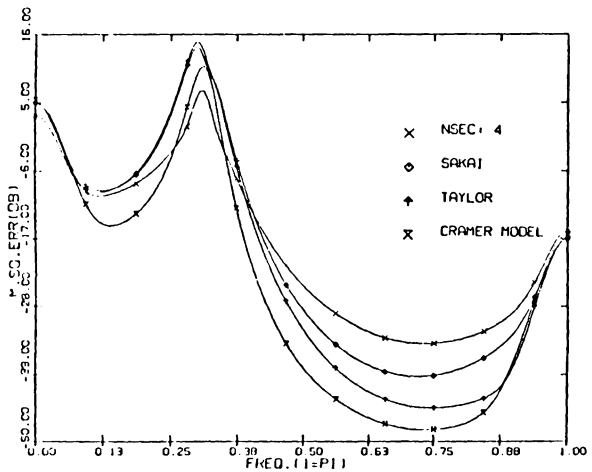
a.



b.



c.



d.

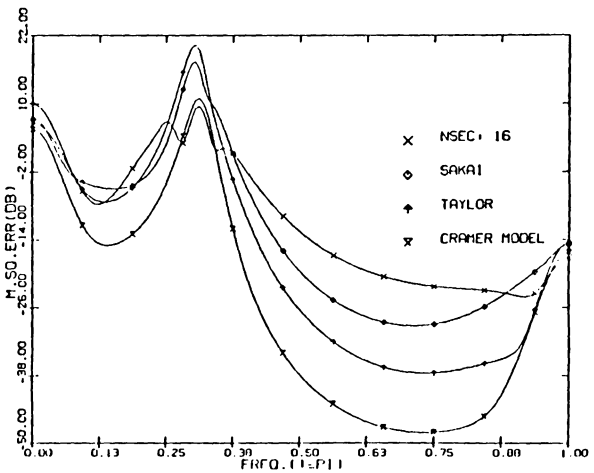
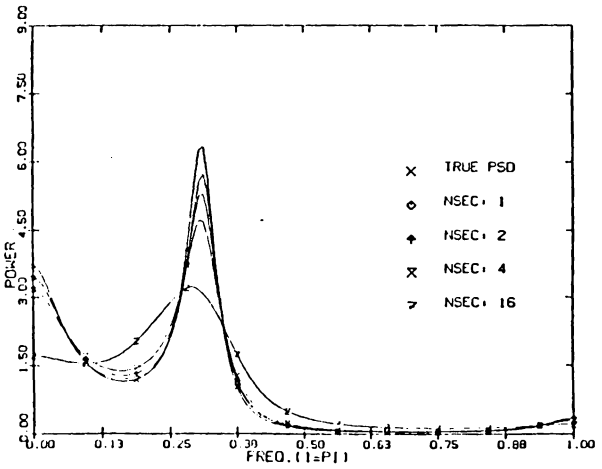
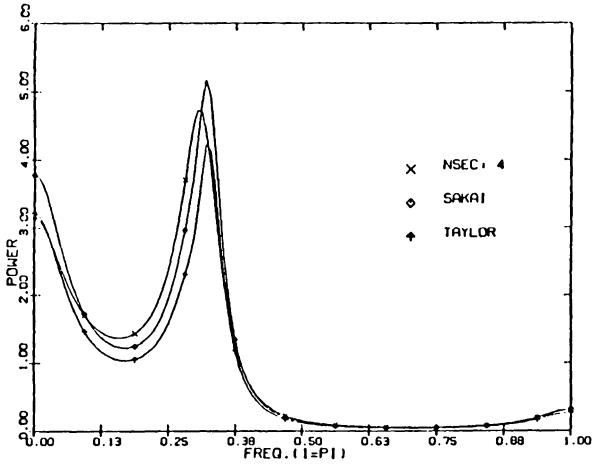


Fig. 14 Mean square error of PSD for AR(4) data using AVA on AR(4) model

a.



b.



c.

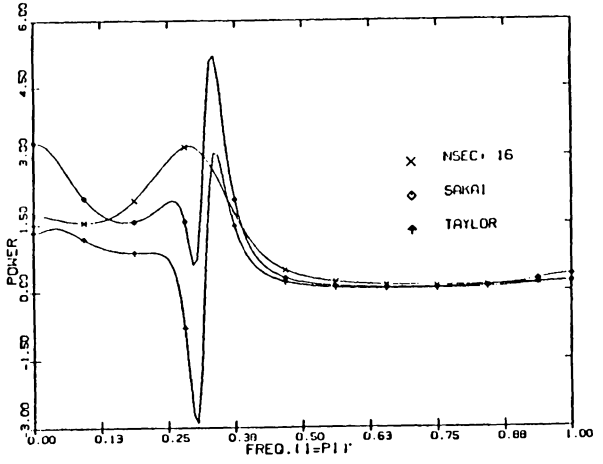
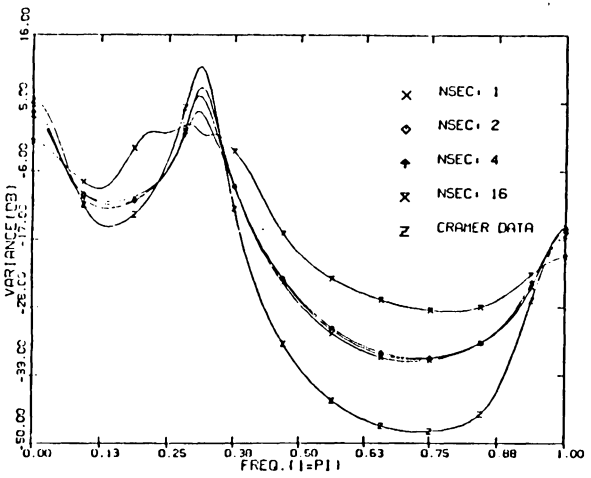
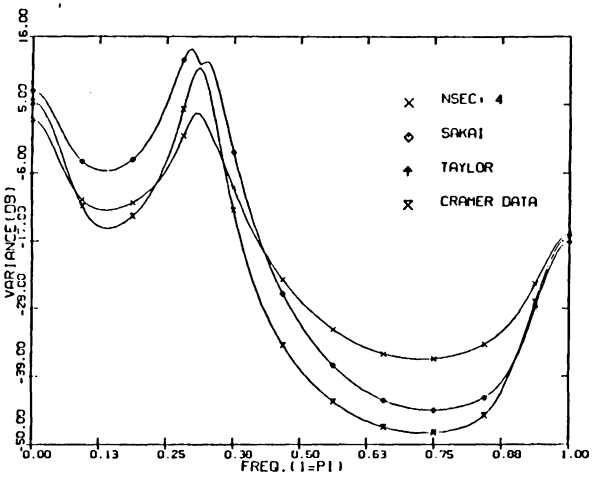


Fig. 15 Mean PSD for AR(4) data using AVK on AR(4) model

a.



b.



c.

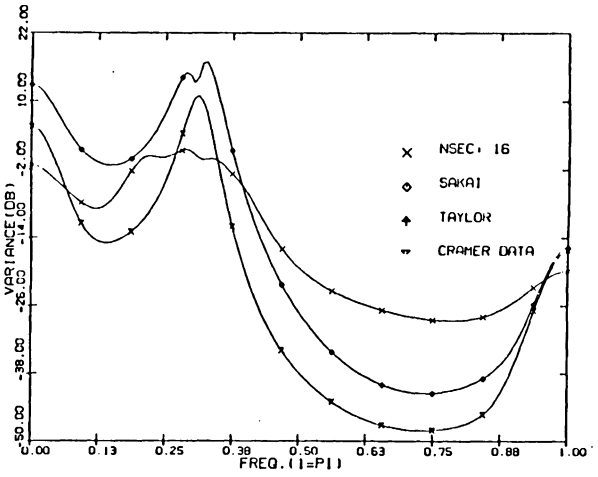
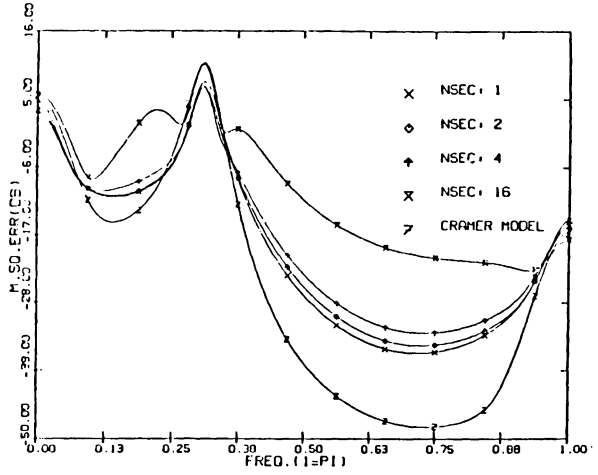
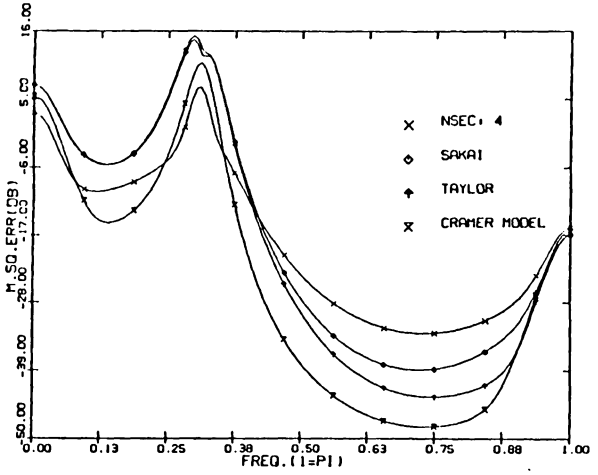


Fig. 16 Variance of PSD for AR(4) data using AVK on AR(4) model

a.



b.



c.

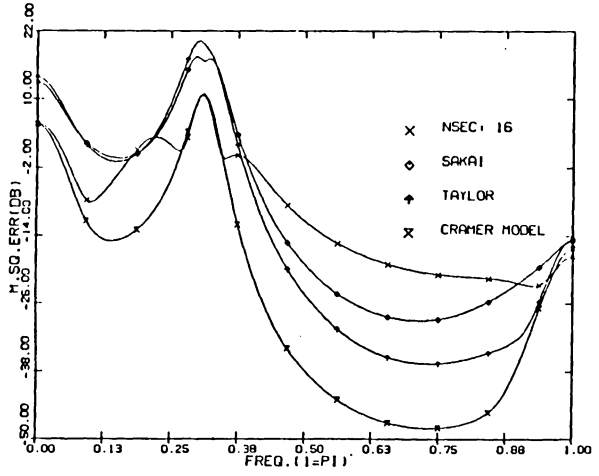
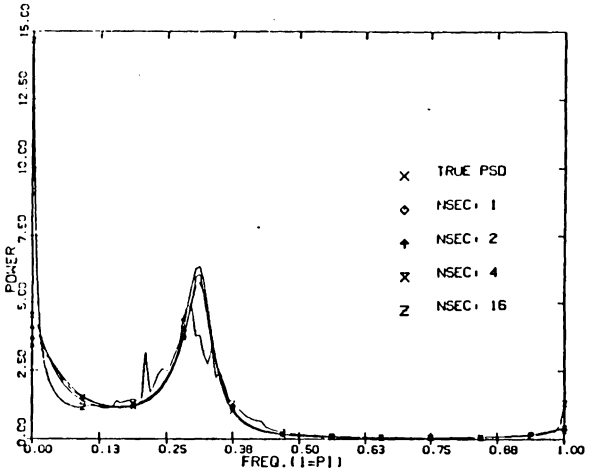
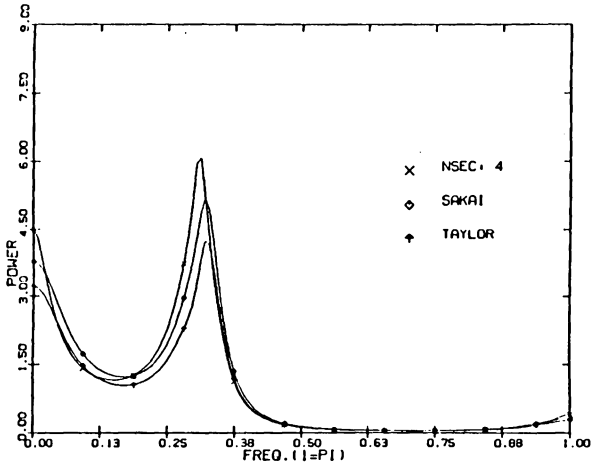


Fig. 17 Mean square error of PSD for AR(4) data using AVK on AR(4) model

a.



b.



c.

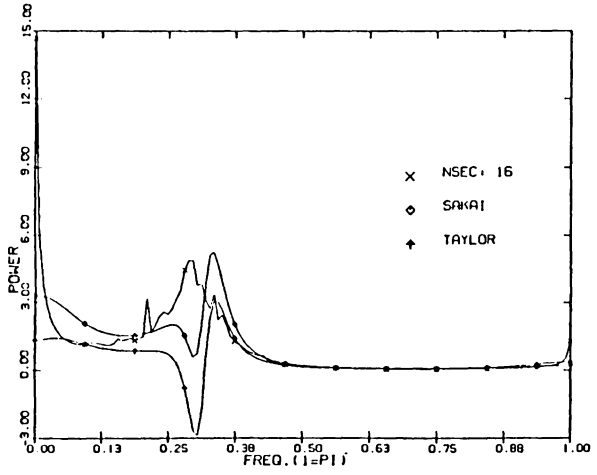
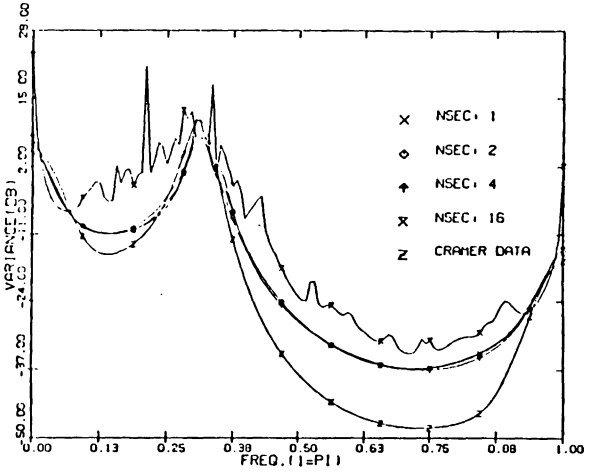
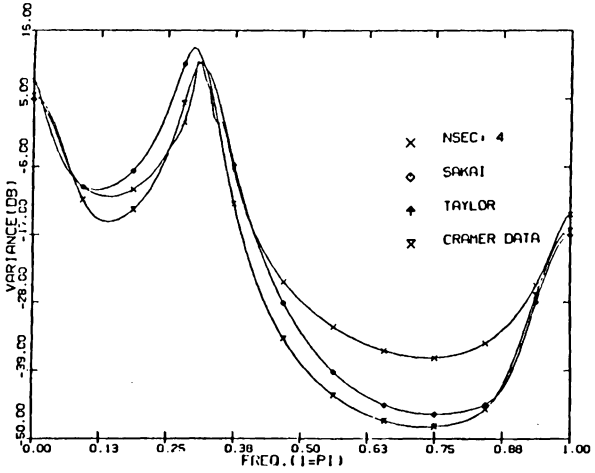


Fig. 18 Mean PSD for AR(4) data using AVP on AR(4) model

a.



b.



c.

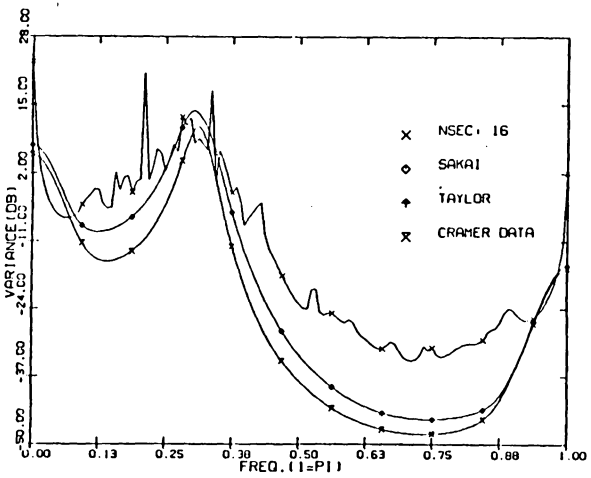
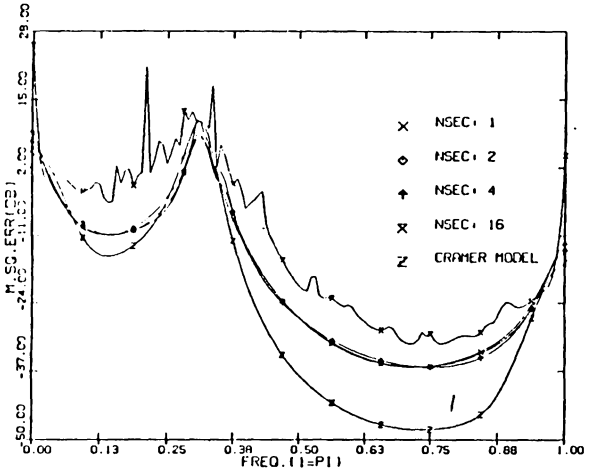
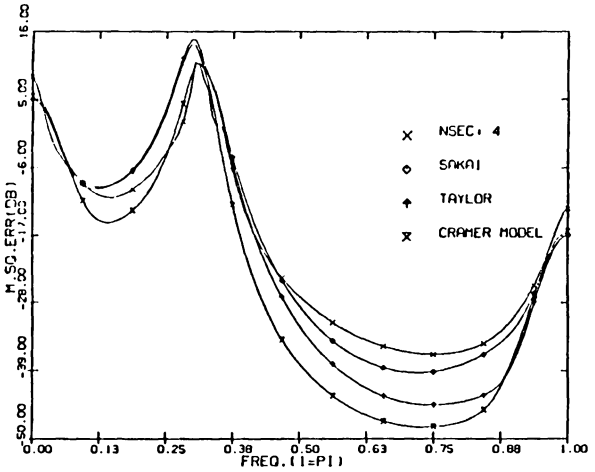


Fig. 19 Variance of PSD for AR(4) data using AVP on AR(4) model

a.



b.



c.

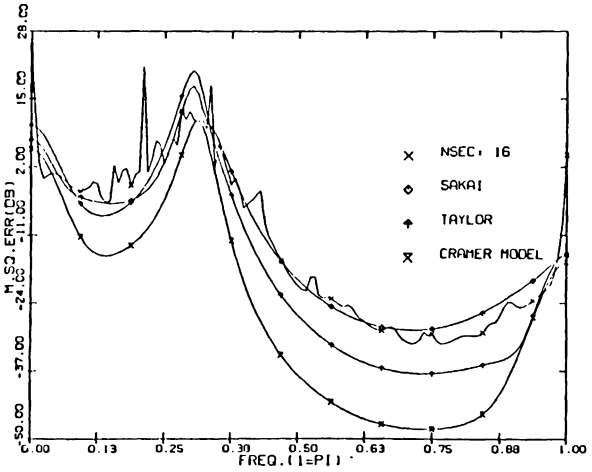
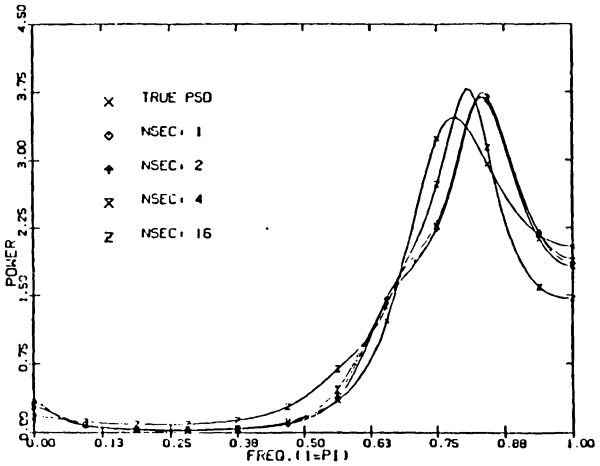
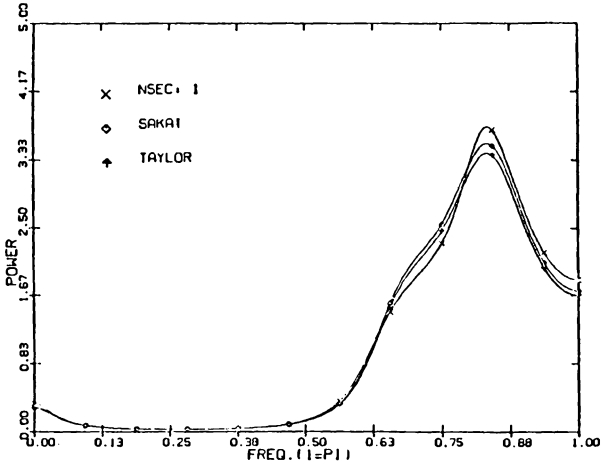


Fig. 20 Mean square error of PSD for AR(4) data using AVP on AR(4) model

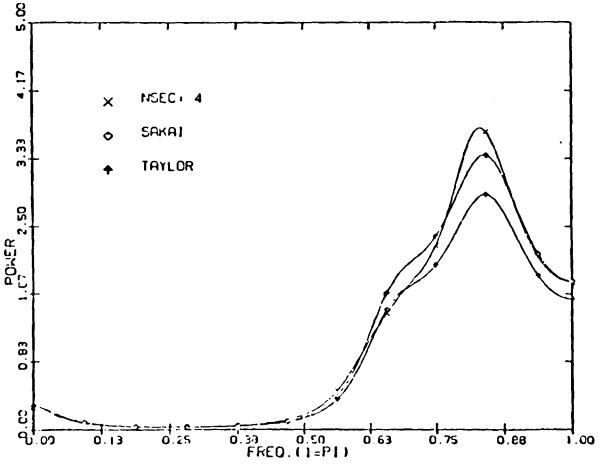
a.



b.



c.



d.

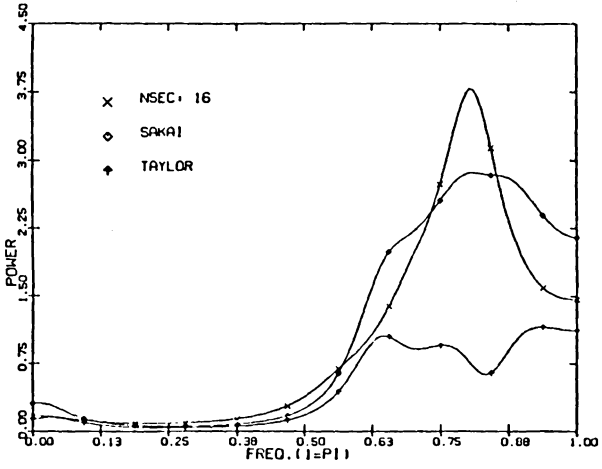
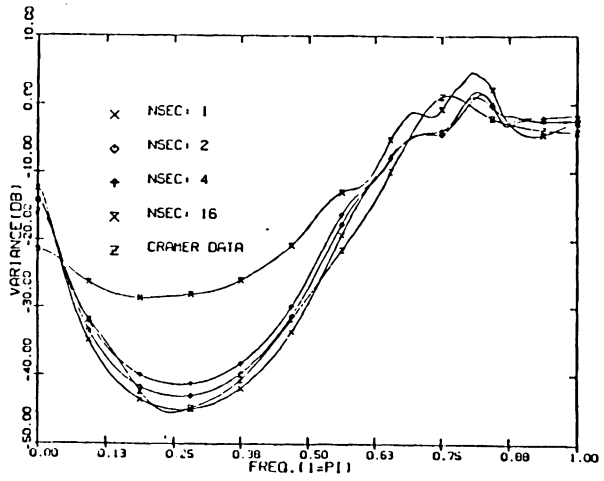
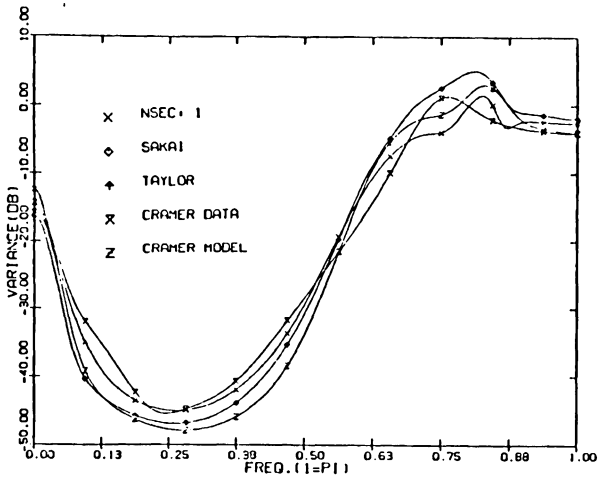


Fig. 21 Mean PSD for ARMA(3,2) data using AVA on AR(5) model

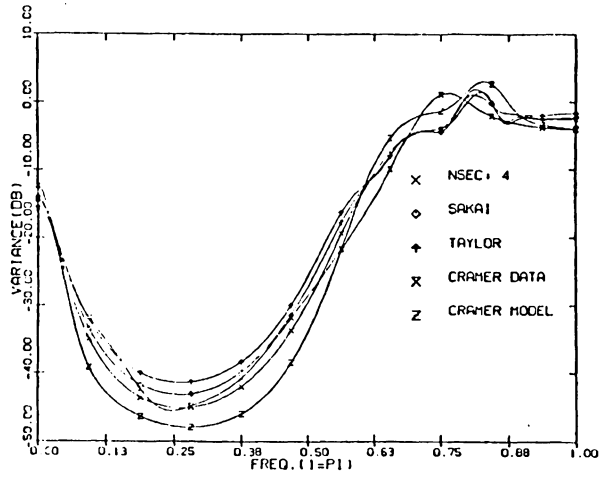
a.



b.



c.



d.

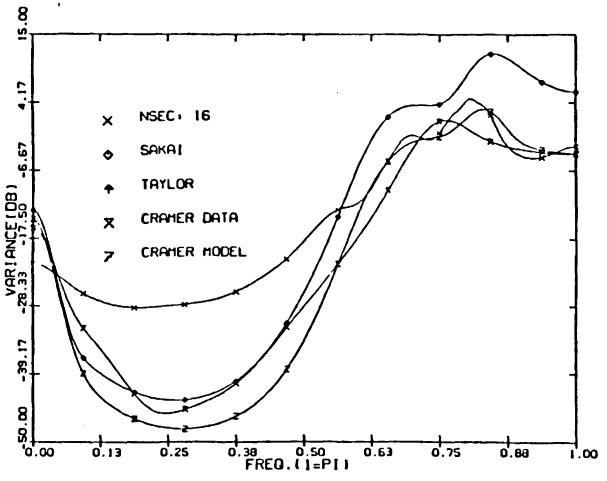
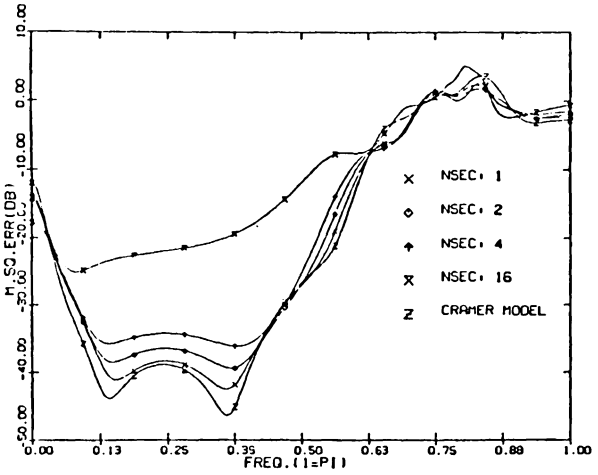
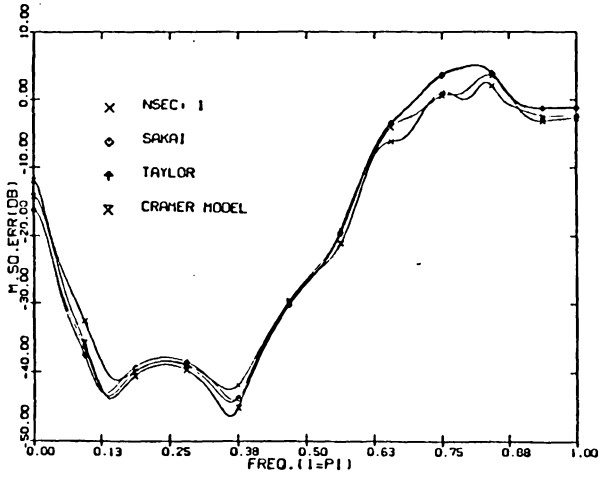


Fig. 22 Variance of PSD for ARMA(3,2) data using AVA on AR(5) model

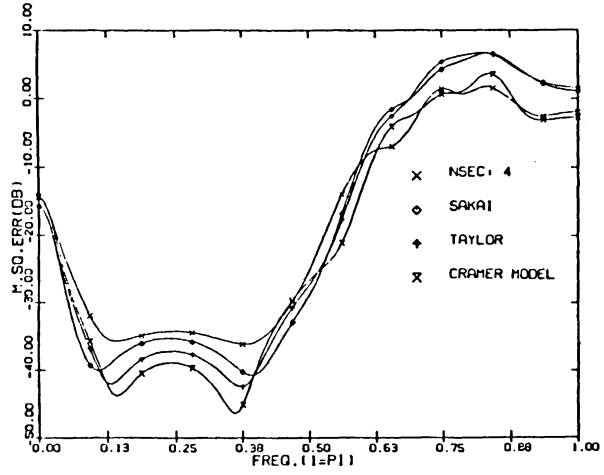
a.



b.



c.



d.

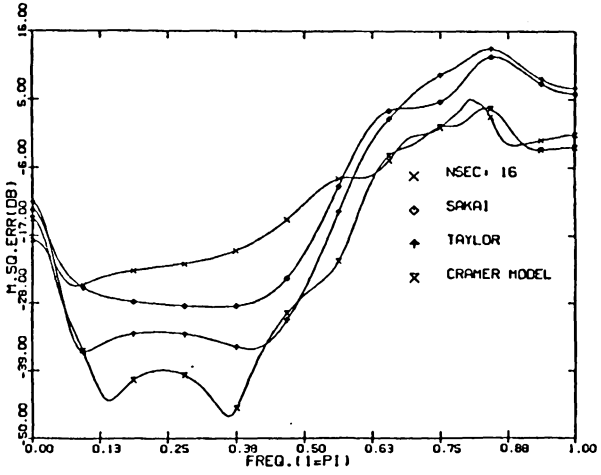
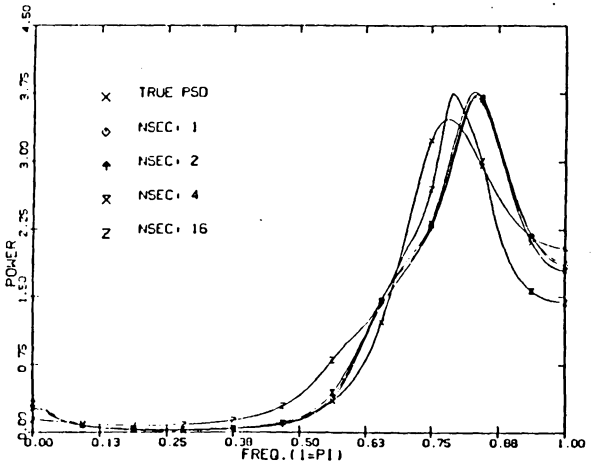
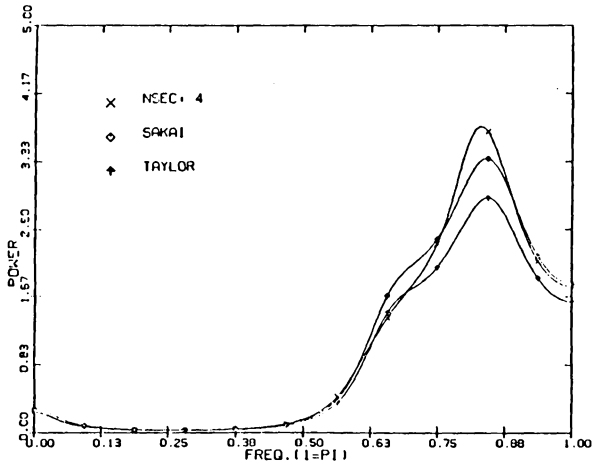


Fig. 23 Mean square error of PSD for ARMA(3,2) data using AVA on AR(5) model

a.



b.



c.

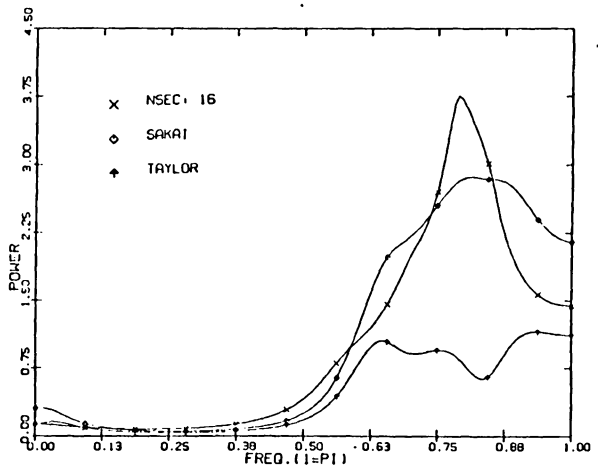
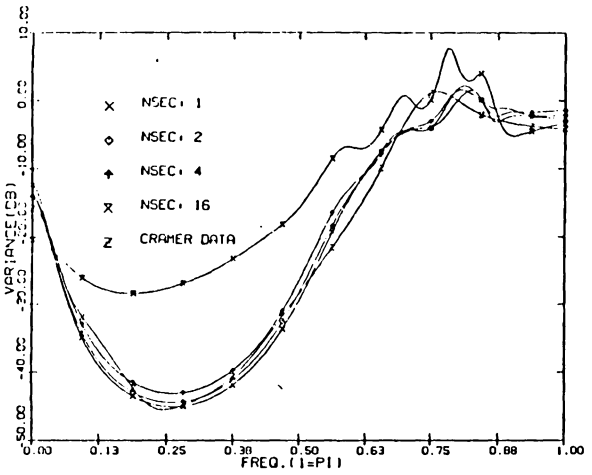
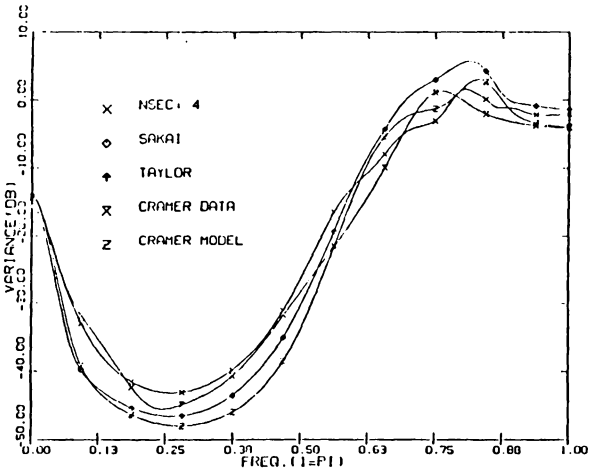


Fig. 24 Mean PSD for ARMA(3,2) data using AVK on AR(5) model

a.



b.



c.

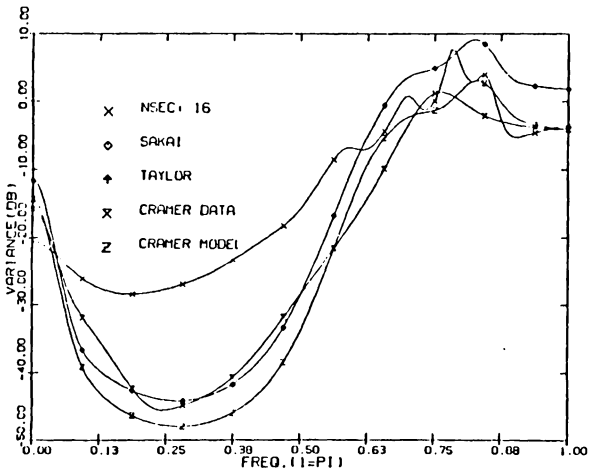
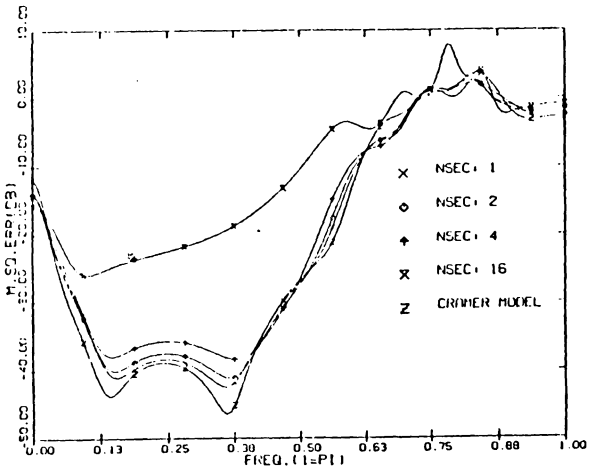
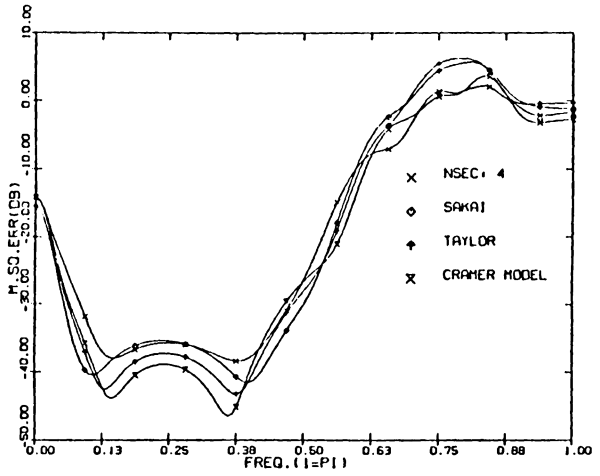


Fig. 25 Variance of PSD for ARMA(3,2) data using AVK on AR(5) model

a.



b.



c.

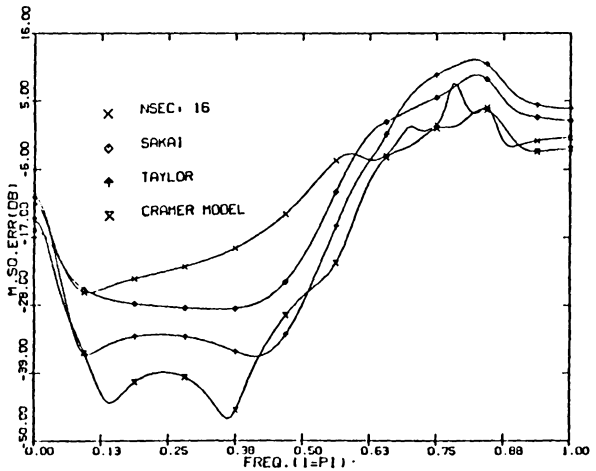
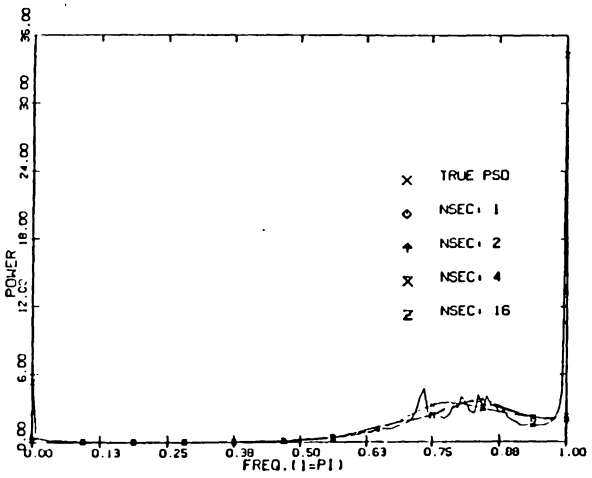
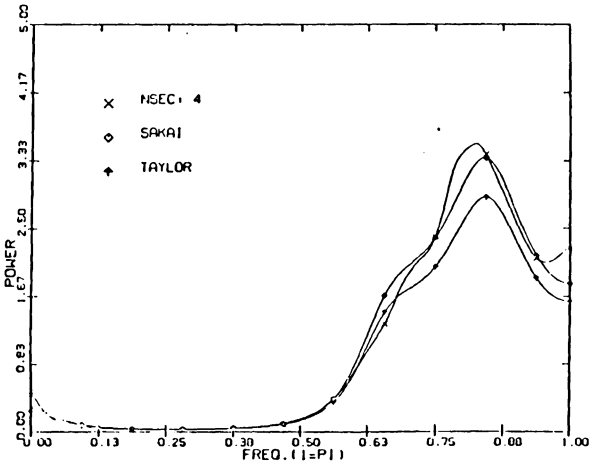


Fig. 26 Mean square error of PSD for ARMA(3,2) data using AVK on AR(5) model

a.



b.



c.

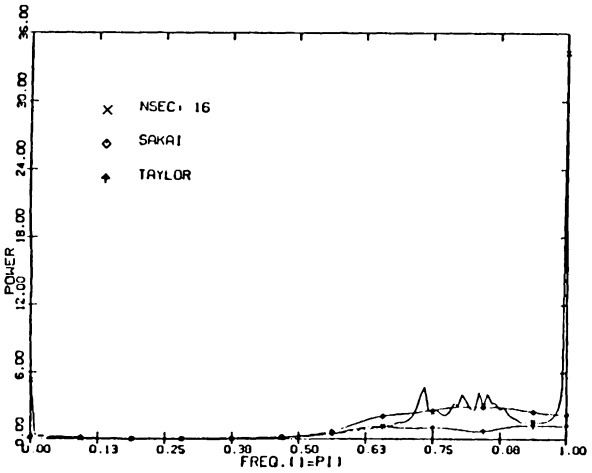
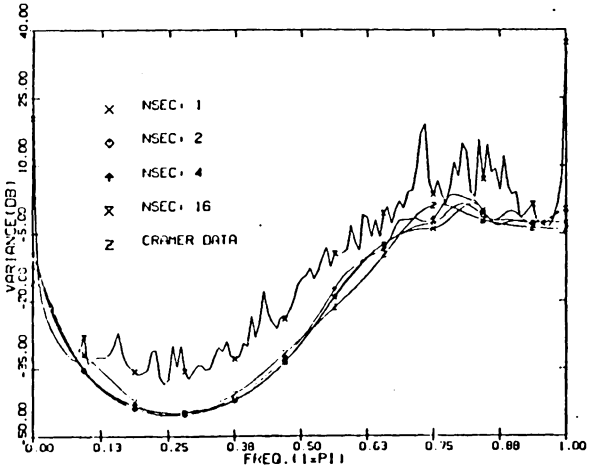
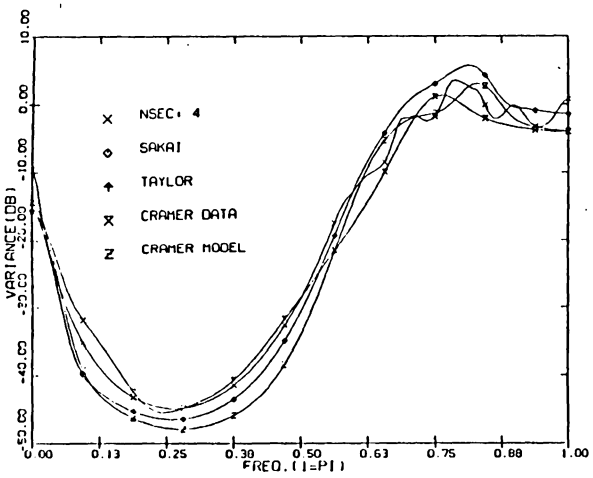


Fig. 27 Mean PSD for ARMA(3,2) data using AVP on AR(5) model

a.



b.



c.

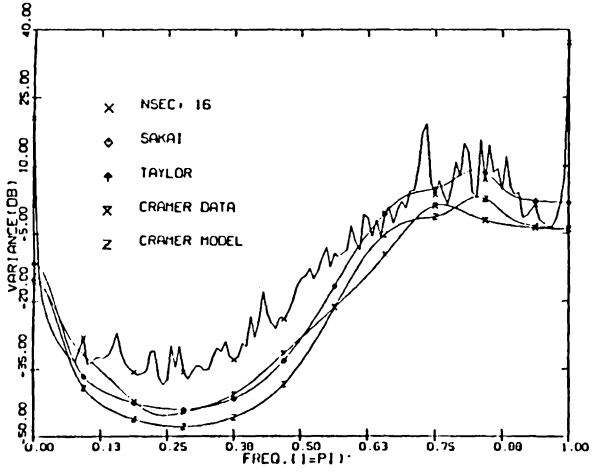


Fig. 28 Variance of PSD for ARMA(3,2) data using AVP on AR(5) model

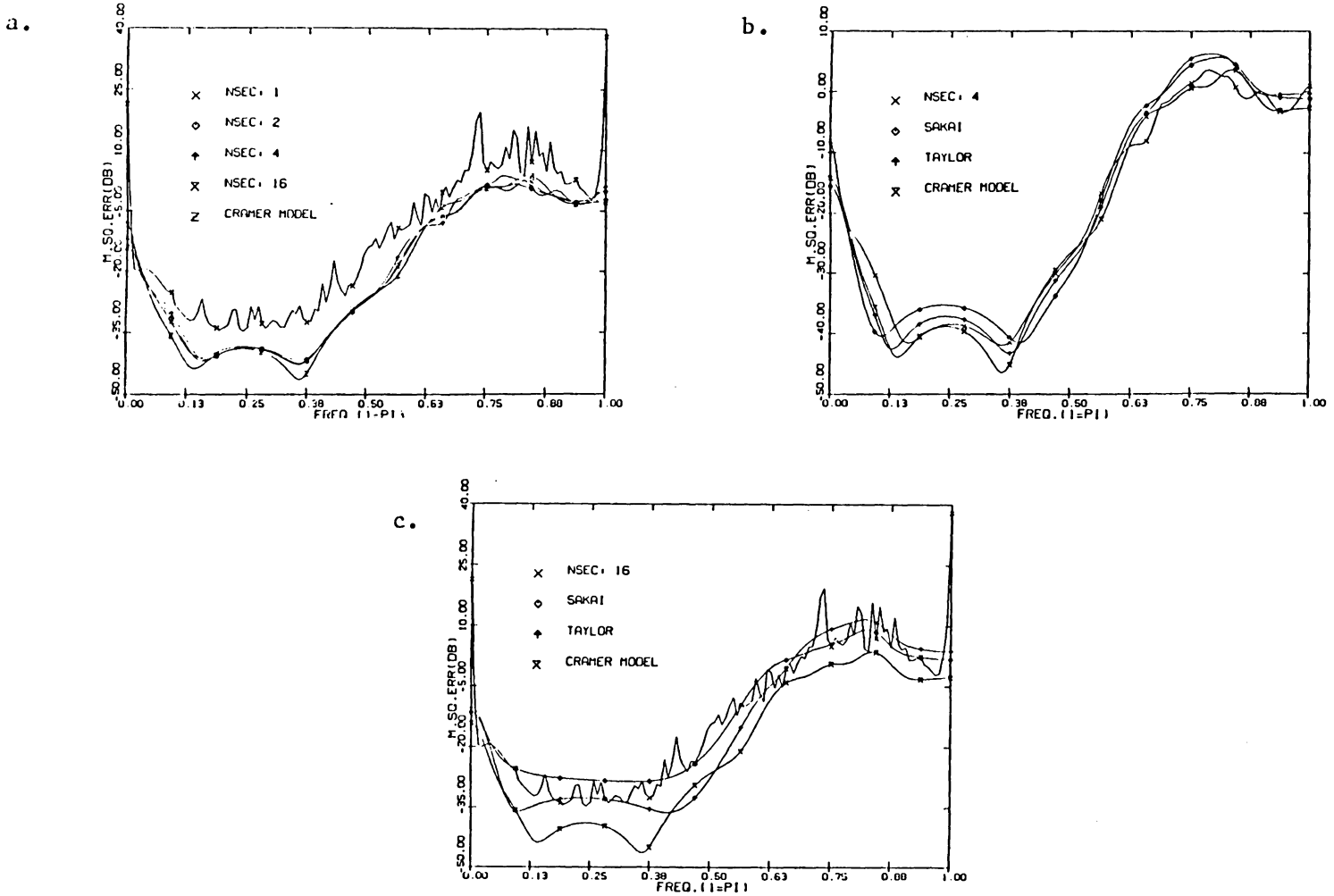
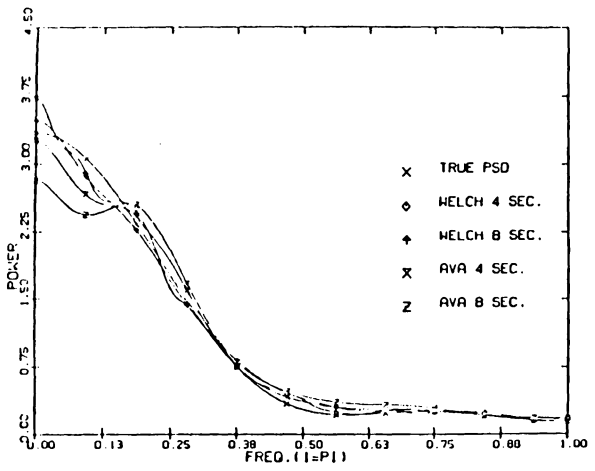
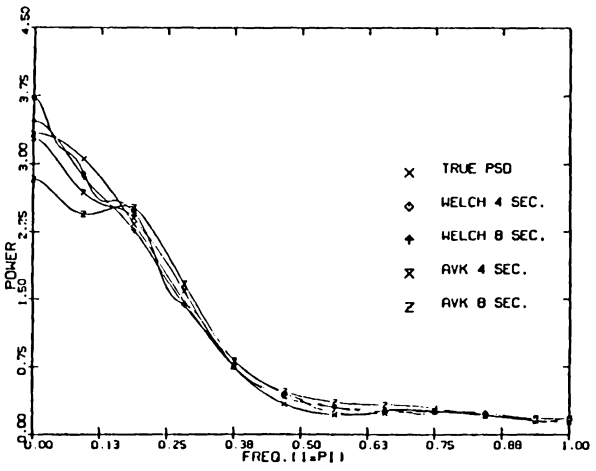


Fig. 29 Mean square error of PSD for ARMA(3,2) data using AVP on AR(5) model

a.



b.



c.

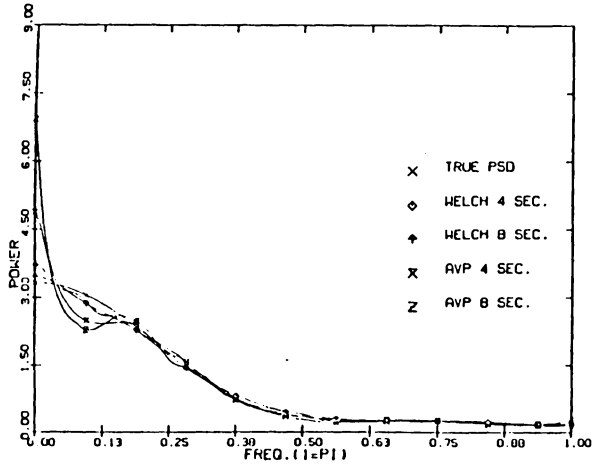


Fig. 30 Mean PSD for MA(3) data using Welch method and MBSE (AR(5) model)

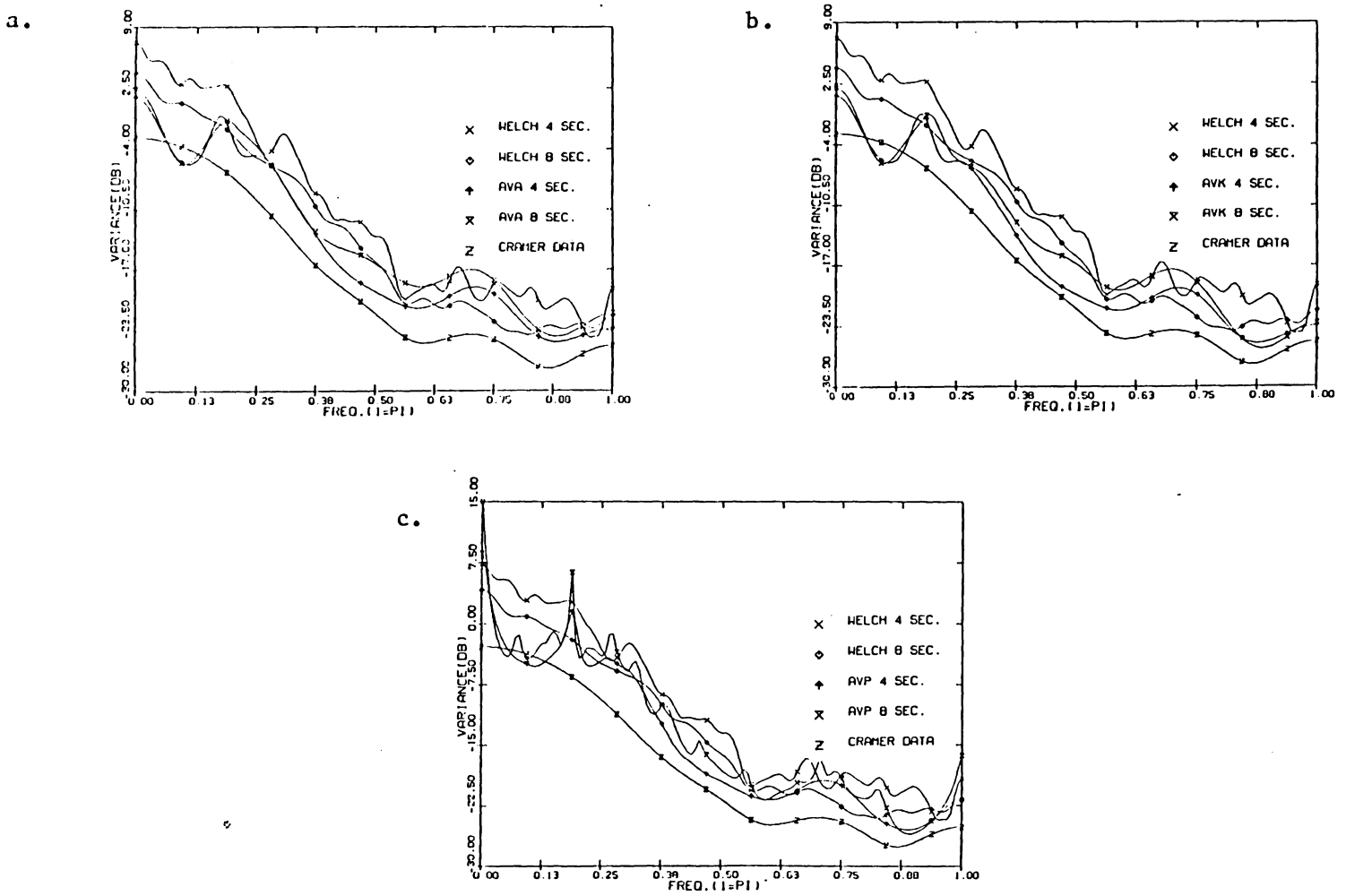
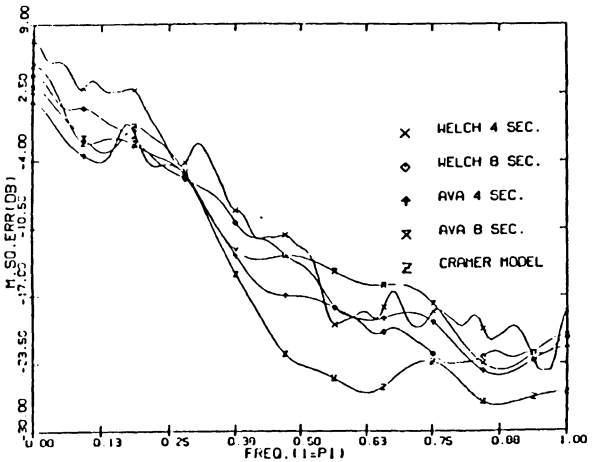
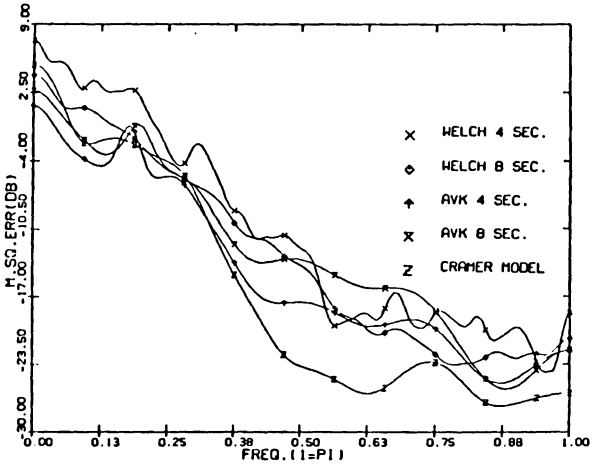


Fig. 31 Variance of PSD for MA(3) data using Welch method and MBSE (AR(5) model)

a.



b.



c.

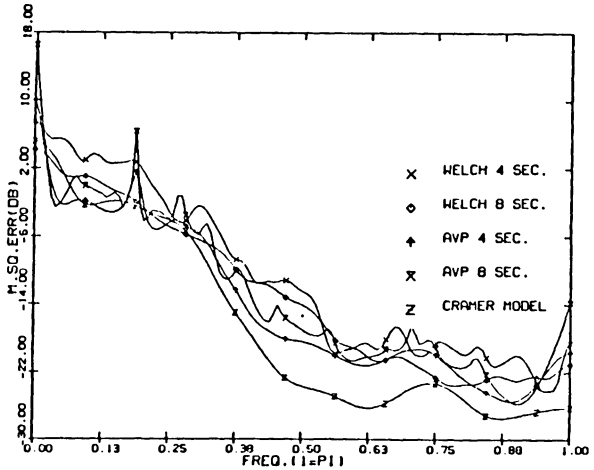
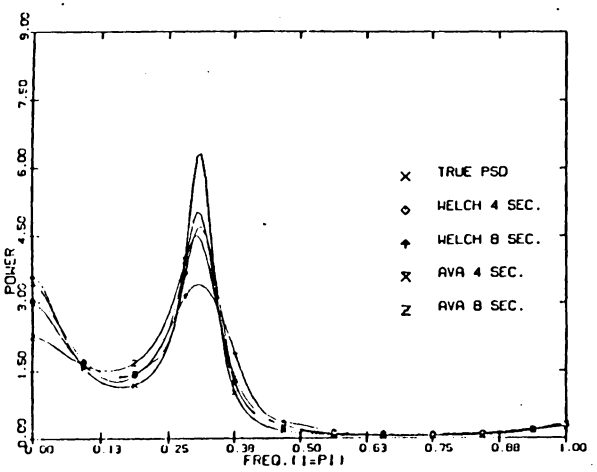
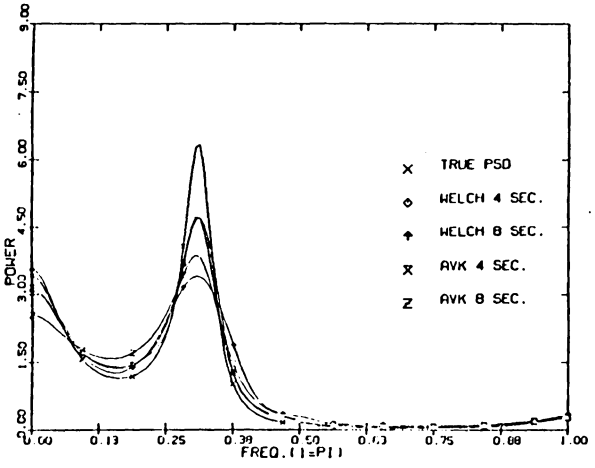


Fig. 32 Mean square error of PSD for MA(3) data using Welch method and MBSE (AR(5) model)

a.



b.



c.

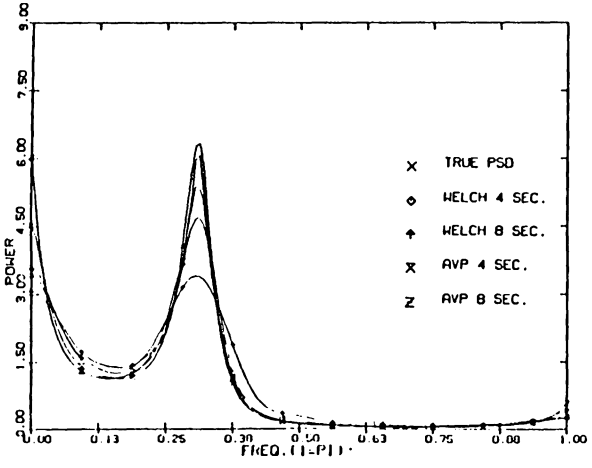
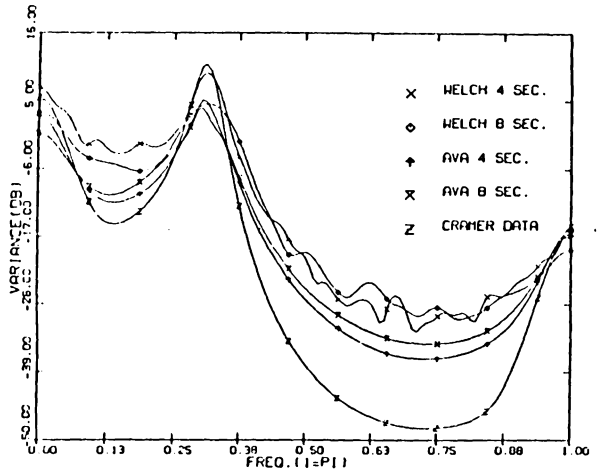
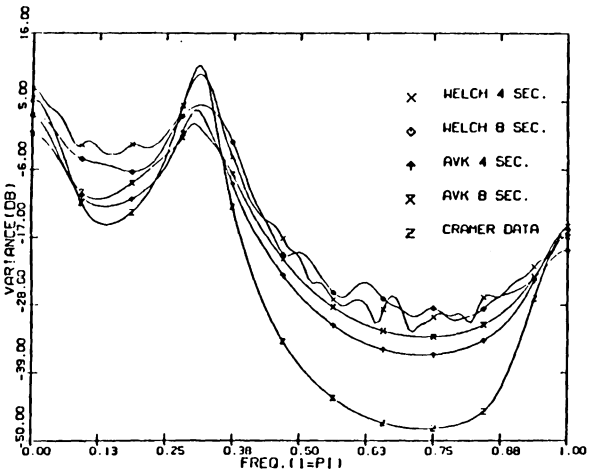


Fig. 33 Mean PSD for AR(4) data using Welch method and MBSE (AR(4) model)

a.



b.



c.

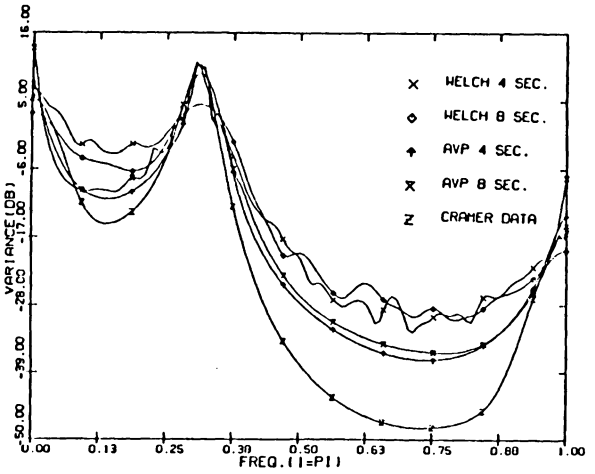
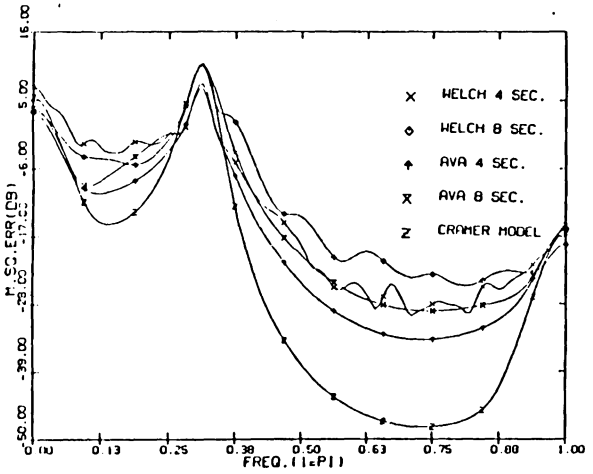
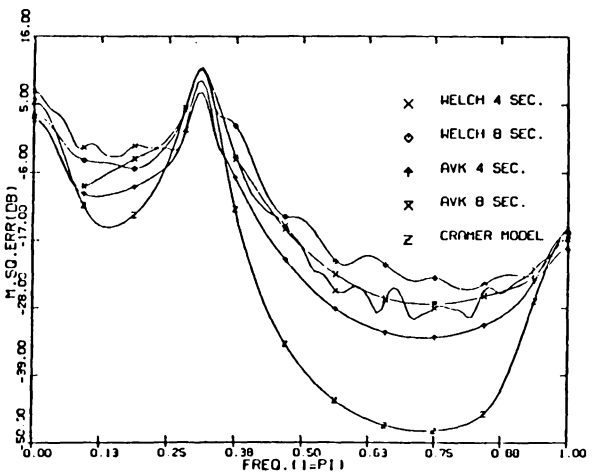


Fig. 34 Variance of PSD for AR(4) data using Welch method and MBSE (AR(4) model)

a.



b.



c.

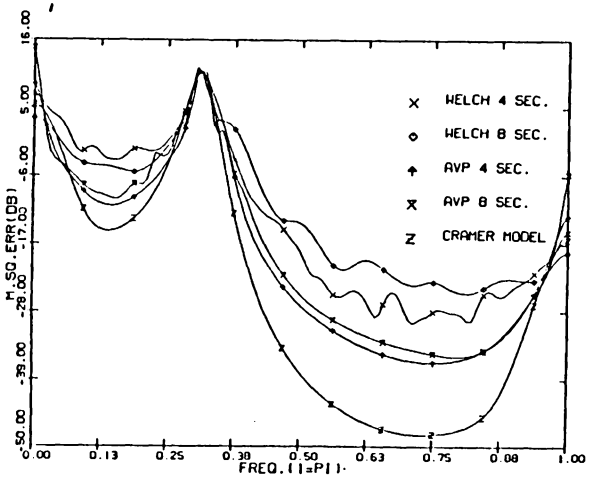
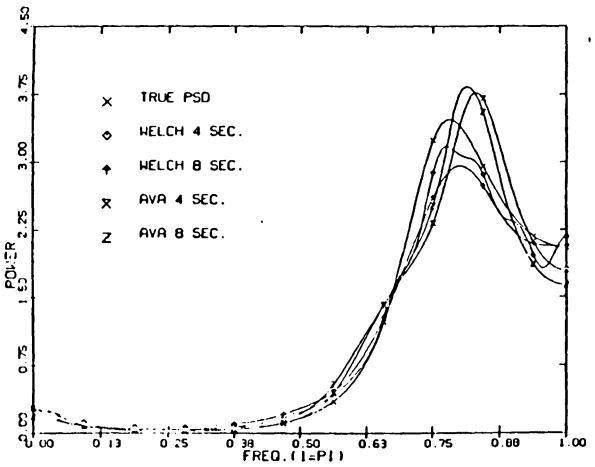
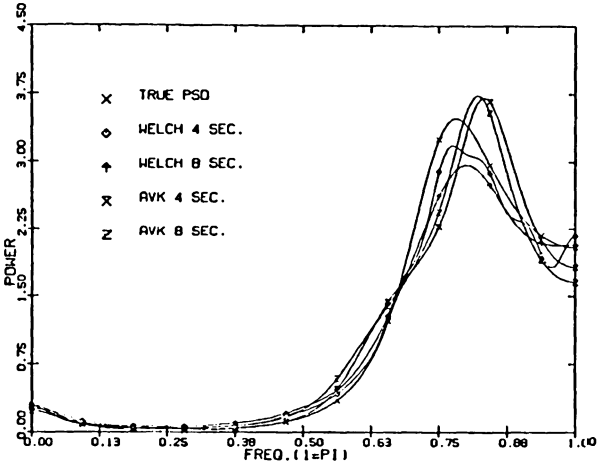


Fig. 35 Mean square error of PSD for AR(4) data using Welch method and MBSE (AR(4) model)

a.



b.



c.

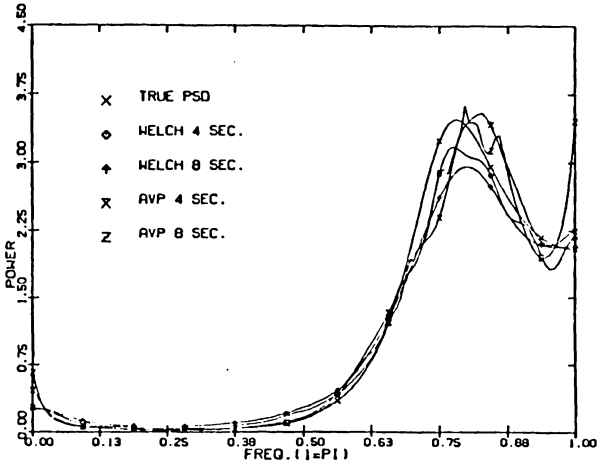
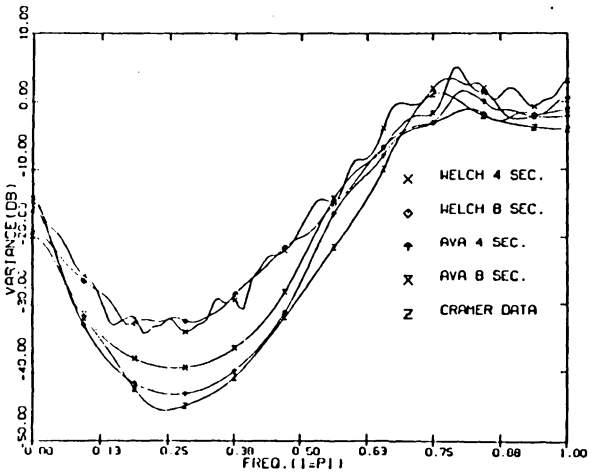
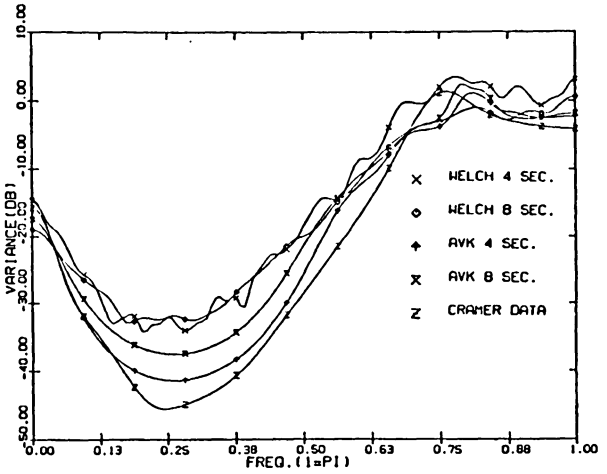


Fig. 36 Mean PSD for ARMA(3,2) data using Welch method and MBSE (AR(5) model)

a.



b.



c.

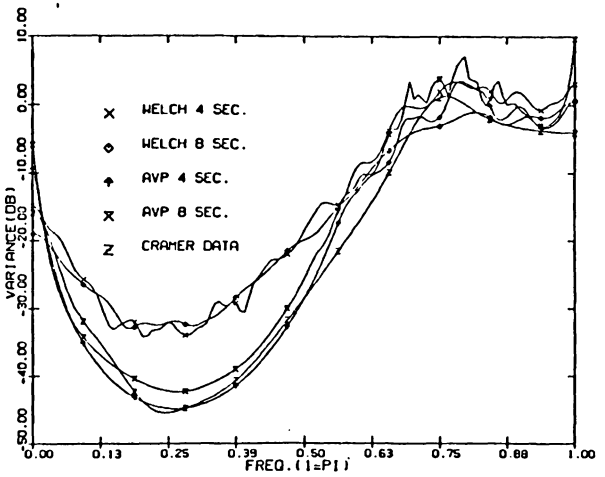
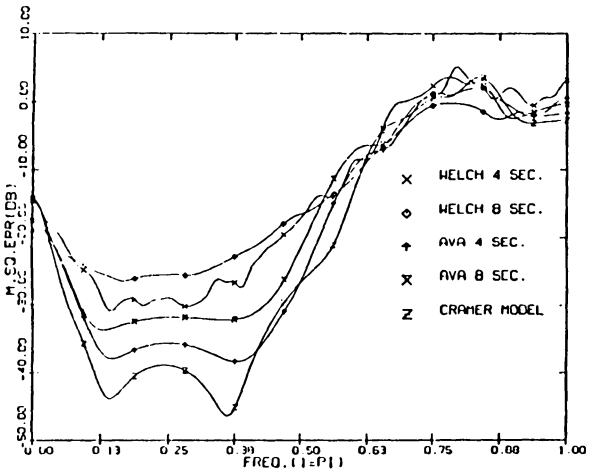
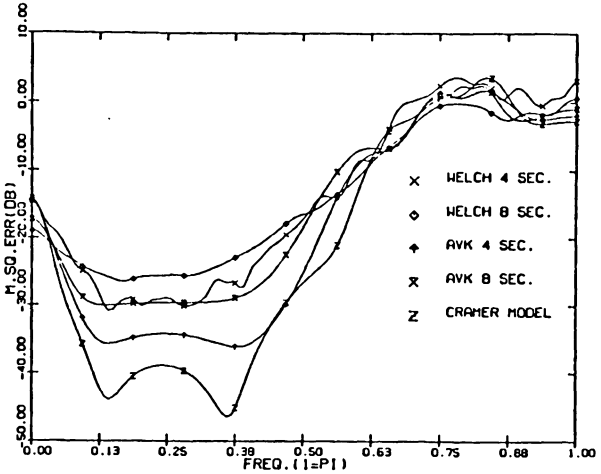


Fig. 37 Variance of PSD for ARMA(3,2) data using Welch method and MBSE (AR(5) model)

a.



b.



c.

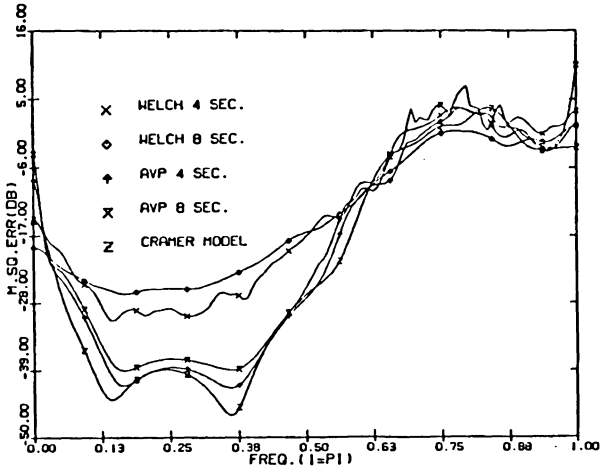


Fig. 38 Mean square error of PSD for ARMA(3,2) data using Welch method and MBSE (AR(5) model)

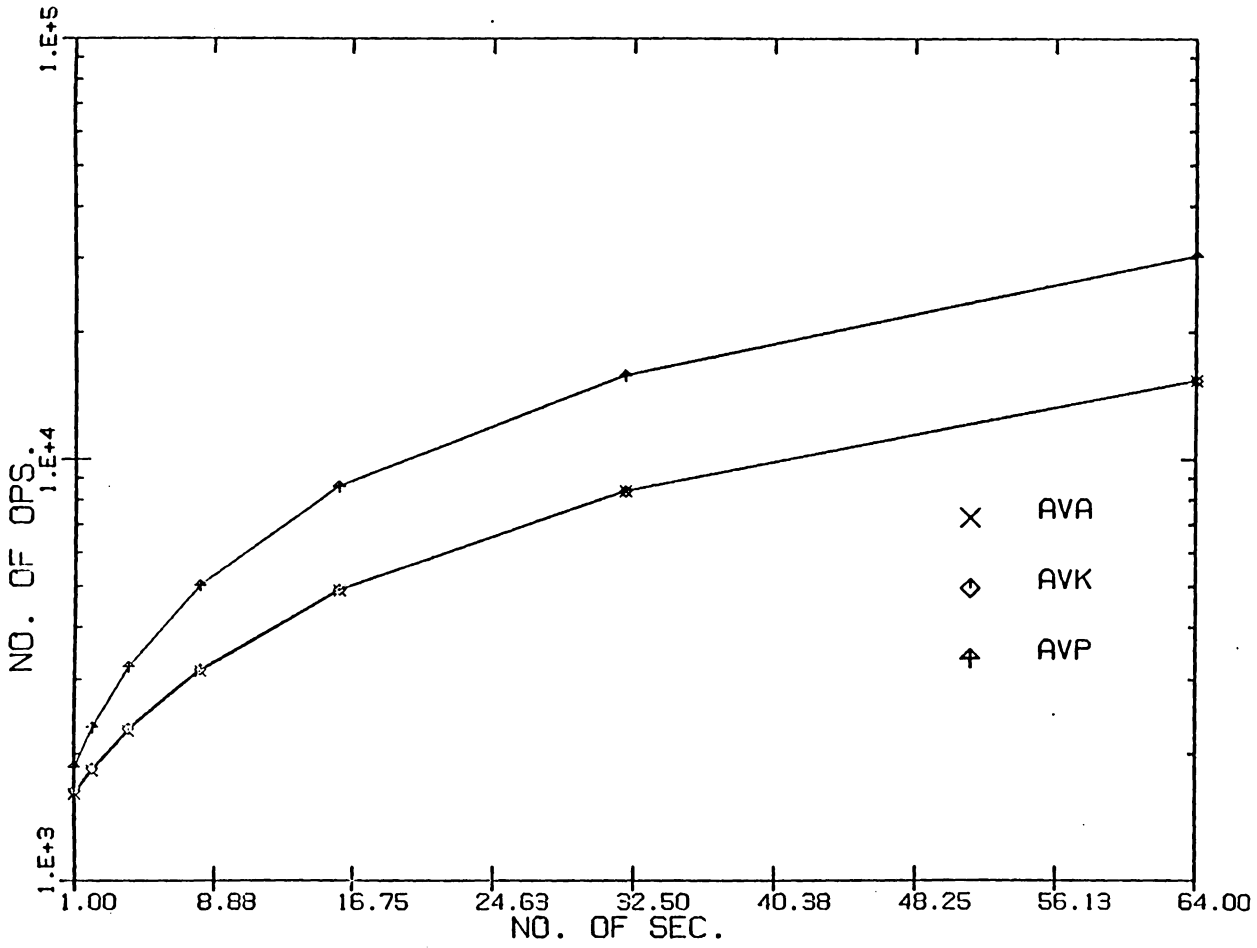
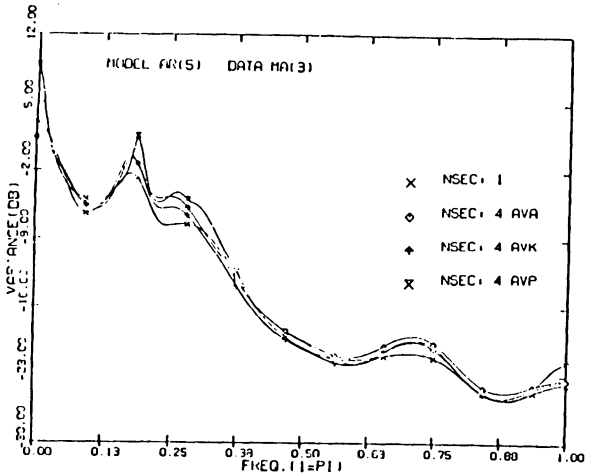
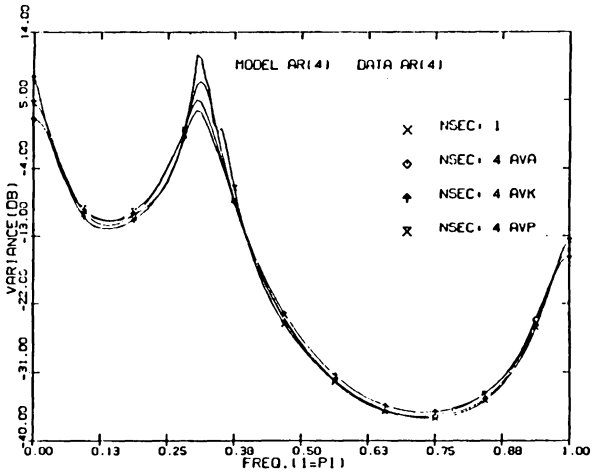


Fig. 39 Number of operations versus number of segments

a.



b.



c.

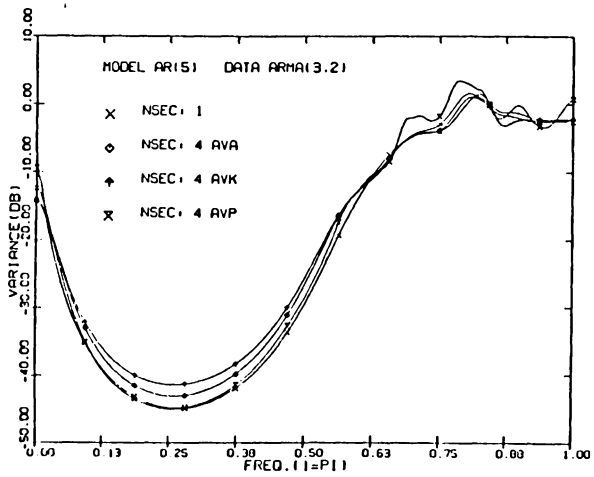
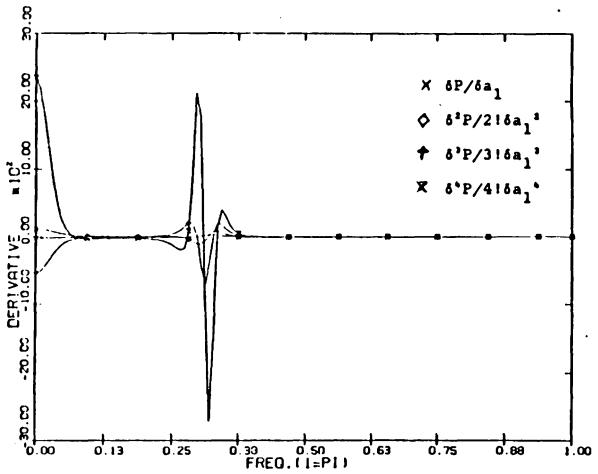
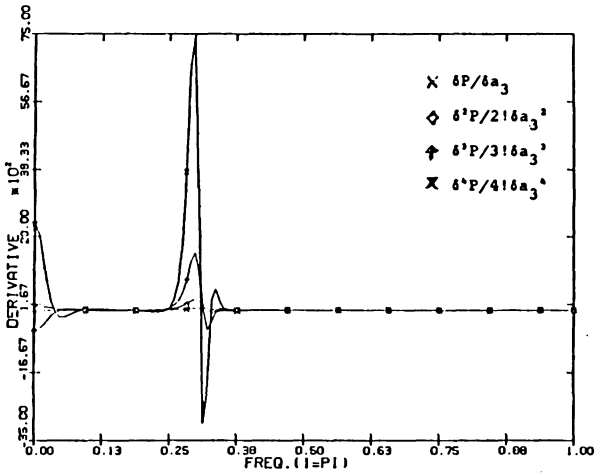


Fig. 40 Variance of PSD estimates of modified Burg spectral estimators

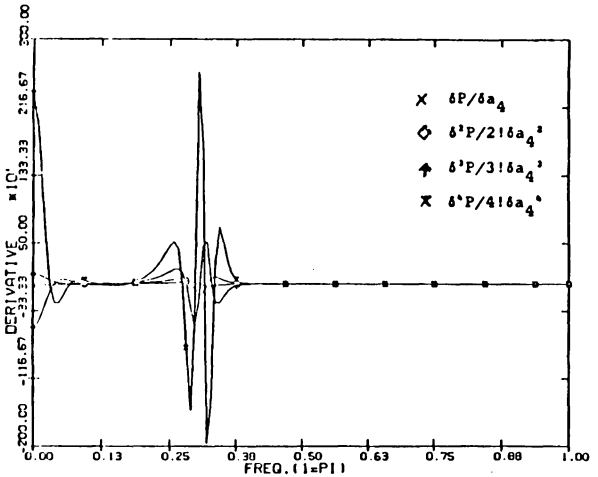
a.



b.



c.



d.

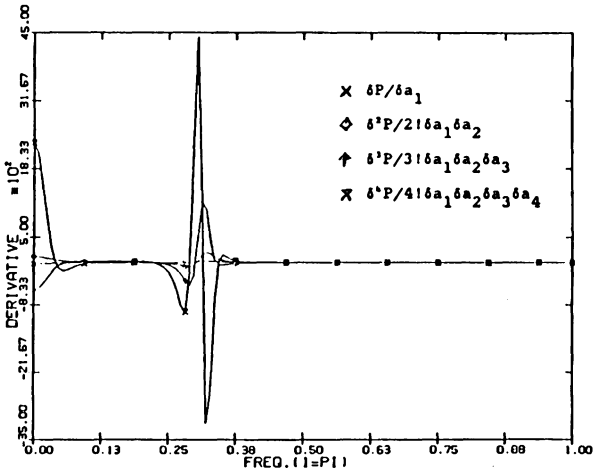


Fig41 Partial derivatives of PSD with respect to AR parameters (AR(4) model)

Chapter V

CONCLUSION

The effect of segment averaging on the quality of the Burg spectral estimator has been analyzed. When a segment averaging method is applied to the Burg spectral estimator, we define the resulting estimator as a modified Burg spectral estimator (MBSE). Three different types of segment averaging are considered. The first one (AVA) averages the autoregressive coefficients computed from each section. The resulting averaged autoregressive parameters then yield the power spectral density estimate. The second one (AVK) averages the reflection coefficients evaluated from each segment. These averaged reflection coefficients are then used to compute the corresponding spectral density estimate. The final approach (AVP) evaluates the power spectral density estimate associated with each segment, and then averages these directly.

Our numerical simulation studies reveal that the AVA and AVK methods are better than the AVP method, particularly with respect to the variances of the power spectral density estimates. The AVP method gives the largest variance and the other two methods give almost the same but considerably lower variance than the AVP method, especially around the

peak frequencies of the actual power spectral density. Moreover, the power spectral density estimates of the modified Burg spectral estimator using the AVA and AVK methods are comparatively smoother than the corresponding estimates of the same estimator employing the AVP method. It is observed that the variance of the MBSE estimates follows the theoretical Cramer-Rao lower bound even for a moderate number of data points. When the number of samples per segment is a few times the order of the estimated autoregressive model this amount of samples is considered moderate. As the number of sections exceeds a certain level so that the number of observations per section is small relative to the dynamics of the data, the efficiency of the modified Burg spectral estimator declines rapidly. Comparative study shows that the AVA and AVK methods give better estimates in terms of variance than the Welch procedure. We note that the AVP method and the Welch procedure use the same averaging technique but different spectral estimators. The latter must explains therefore that the AVP method gives relatively poor estimates at radian frequencies 0, ω_p, and π. ω_p is the peak frequency of the actual power spectral density.

To analyze the statistical properties of the modified Burg spectral estimator, two approximation methods are

formulated, namely the Sakai approximation and the Taylor approximation. The basic difference between these two is that in the Taylor approximation we neglect the bias and variance of the gain factor of the power spectral density estimate. The expected mean and mean square error predicted by the Sakai method follow the corresponding sample mean and sample mean square error more closely than the prediction from the Taylor method. Both approximation methods, however, give the same value for the variances of the power spectral density estimates. These results imply that the variance of the gain factor has little effect on the variance of the power spectral density estimates whereas the bias of the gain factor has a considerable effect on the expected mean of the estimates. Both methods employ a recursive technique to find the approximate mean and variance of the estimates of the modified Burg spectral estimator. The following assumptions are made in the derivation of the approximation methods:

a) the given process is zero mean, gaussian, wide sense stationary, and ergodic,

b) the gain factor and any one of the autoregressive parameters of the power spectral density estimate are uncorrelated,

c) reflection coefficient at mth stage is weakly correlated with any one element of the corresponding autoregressive coefficient vector of (m-1)th stage,

d) samples in successive segments are uncorrelated,

e) the probability density function of any reflection coefficient is concentrated near its mean value.

These assumptions were found to be practically tenable for a moderate number of data points and hence both the Sakai approximation and the Taylor approximation methods give good predictions. For very short records however, both these methods fail to give the expected mean and variance of the estimates since the last three assumptions are not valid. The algorithms, used for evaluating the approximate mean and variance of the modified Burg spectral estimators, may be modified to compute the same for any autoregressive spectral estimator. Only the derivation of the mean and variance of the reflection coefficients will change from one type of autoregressive estimator to another. This is a result of the different reflection coefficient definition for each type of autoregressive estimator.

A lower bound for the variance of a reflection coefficient using the Burg method has been derived. This bound found to hold for the first order autoregressive model. For reflection coefficients of higher order stages

however, it is difficult to confirm experimentally the validity of this theoretical lower bound, because the exact statistics of the forward and backward error sequences are not known. Further investigation is needed to determine whether this theoretical bound provides a tight bound.

Considering all the aspects of the three averaging methods we like to recommend the AVK method for the modified Burg spectral estimator. The AVA method has almost the same characteristics as the AVK method. We do not suggest to use the AVP method. It is found that when the number of data per section is moderate the performances of the modified Burg spectral estimator do not differ too much from that of the Burg spectral estimator. However, the total number of operations is higher for the modified Burg spectral estimator. So, for moderate amount of data it is better to use the Burg spectral estimator. Finally, from this investigation it is found that segment averaging does not appreciably reduce the variance of the modified Burg spectral estimator. This is logical because the modified Burg spectral estimator performs close to the theoretical Cramer-Rao bound even for a moderate number of samples. The only advantage of modified Burg spectral estimator therefore lies in operating on a smaller number of samples per segment, but doing it more often, instead of operating once on the full length of data.

REFERENCES

1. Kay, S. M. and Marple, S. L. (Jr.), *"Spectral Analysis- A Modern Perspective"*, Proceedings of the IEEE, Vol.69, No.11, Nov. 1981, pp 1380-1419.

2. Brillinger, D. R., Time Series Data Analysis and Theory, Holt, Rinehart and Winston, Inc., 1975.

3. Bartlett, M. S., An Introduction to Stochastic Processes with Special Reference to Methods and Applications, Cambridge University Press, New York, 1953.

4. Grenander, U. and Rosenblatt, M., Statistical Analysis of Stationary Time series, John Wiley & Sons, Inc., New York, 1957.

5. Blackman, R. B. and Tukey, J. W., The Measurement of Power Spectra from the Point of View of Communication Engineering, Dover Press, New York, 1959.

6. Jenkins, G. M. and Watts, D. G., Spectral Analysis and Its Applicatons, Holden-Day, Inc., San Francisco, 1968.

7. Koopmans, L. H., Spectral Analysis of Time Series, Academic Press, New York, 1974.

8. Burg, J. P., *"Maximum Entropy Spectral Analysis"*, Ph.D. Dissertation, Department of Geophysics, Stanford University, Stanford, California, May 1975.

9. Parzen, E., Multiple Time Series Modelling, Multivariate Analysis-II, P. R. Krishnaiah (editor), Academic Press, New York, 1969.

10. Tretter, S. A. and Steiglitz, K., *"Power Spectrum Identification in terms of Rational Models"*, IEEE Transactions on Automatic Control, Vol.AC-12, Apr. 1967, pp. 185-188.

11. Gutowski, P. R., Robinson, E. A. and Treitel, S., *"Spectral Estimation: Fact or Fiction"*, IEEE Transactions on Geoscience Electronics, Vol.GE-16, No.2, Apr. 1978, pp. 80-84.

12. Modern Spectral Analysis, D. G. Childers (editor), IEEE Press, Inc., New York, 1978.

13. Special Issue on Spectral Estimation, Proceedings of IEEE, Vol.70, No.9, Sep. 1982.

14. Oppenheim, A. V. and Schafer, R. W., Digital Signal Processing, Prentice-Hall, Inc., Englewood Cliffs, New Jersey, 1975.

15. Akaike, H., *"Power Spectrum Estimation through Autoregressive Model Fitting"*, Annals of the Institute of Statistical Mathematics, Vol.21, 1969, pp. 407-419.

16. Kromer, R. E., *"Asymptotic Properties of the Autoregressive Spectral Estimator"*, Ph.D. Dissertation, Stanford University, 1970.

17. Berk, K. N., *"Consistent Autoregressive Spectral Estimates"*, Annals of Statistics, Vol.2, May 1974, pp. 489-502.

18. Sakai, H., *" Statistical Properties of AR Spectral Analysis"*, IEEE Transactions on Acoustics, Speech, and Signal Processing, Vol.ASSP-27, No.4, Aug. 1979, pp. 402-409.

19. Mann, H. B. and Wald, A., *"On the Statistical Treatment of Linear Stochastic Difference Equations"*, Econometrica, Vol.11, 1943, pp. 173-220.

20. Anderson, T. W. and Walker, J. W., *"On the Asymptotic Distribution of the Autocorrelations of a Sample from a Linear Stochastic Process"*, Annals of Mathematical Statistics, Vol.35, 1964, pp. 1296-1303.

21. Baggeroer, A. B., *"Confidence Intervals for Regression (MEM) Spectral Estimates"*, IEEE Transactions on Information Theory, Vol.IT-22, Sep. 1976, pp. 534-545.

22. Huzzi, M., *"On a Spectral Estimate obtained by an Autoregressive Model Fitting"*, Annals of the Institute of Statistical Mathematics, Vol.29, 1977, pp. 415-431.

23. Grenander, U. and Szego, G., *"Toeplitz forms and their Applications"* University of California Press, Berkely, 1958, pp. 64.

24. Gersch, W. and Sharpe, D. R., *"Estimation of Power Specta with Finite Order Autoregressive Models"*, IEEE Transactions on Automatic control, Vol.AC-18, 1973, pp. 367-369.

25. van den Bos, A., *"Alternative Interpretation of Maximum Entropy Spectral Analysis"*, IEEE Transactions on Information Theory, Vol.IT-17, 1971, pp. 493-494.

26. Burg, J. P., *"The Relationship between Maximum Entropy Spectra and Maximum Likelihood Spectra"*, Geophysics, Vol.37, Apr. 1972, pp. 375-376.

27. Pusey, L. C., *"High Resolution Spectral Estimates"*, Technical Note, Lincoln Laboratory, M.I.T., Jan. 1975.

28. Burg, J. P., *"A New Analysis Technique for Time Series Data"*, Presented at the NATO Advanced Study Institute on Signal Processing, Enschede, Netherlands, Aug. 1968.

29. Makhoul, J., *"Stable and Efficient Lattice Methods for Linear Prediction"*, IEEE Transactions on Acoustics, Speech, and Signal Processing, Vol.ASSP-25, No.5, Oct. 1977, pp. 423-428.

30. Dugre, J. P., Beex, A. A. and Scharf, L. L., *"Generating Covariance Sequences and the Calculation of Quantization and Rounding Error Variances in Digital Filters"*. IEEE Transactions on Acoustics, Speech, and Signal Processing, Vol.ASSP-28, No.1, Feb. 1980, pp. 102-104.

31. Sakai, H., Soeda, T. and Tokumaru, H., *"On the Relation between Fitting Autoregression and Periodograms with Applications"*, Annals of Statistics, Vol.7, Jan. 1979, pp. 96-107.

32. Papoulis, A., Probability, Random Variables, and Stochastic Processes, McGraw-Hill Inc., 1965, pp. 210.

33. Kay, S. and Makhoul, J., *"On the Statistics of the Estimated Reflection Coefficients of an Autoregressive Process"*, IEEE Transactions on Acoustics, Speech, and Signal Processing, Vol.ASSP-31, No.6, Dec. 1983, pp. 1447-1455.

34. Cramer, H., Mathematical Methods of Statistics, Princeton University Press, New Jersey, 1951.

35. Rao, C. R., Linear Statistical Inference and Its Applications, John Wiley & Sons, New York, 1965.

36. Eykhoff, P., System Identification, John Wiley & Sons, New York, 1974.

37. Loffler, H. E., *"About the Quality of Parametric Power Spectral Estimates"*, M.S. Thesis, Department of Electrical Engineering, Virginia Polytechnic Institute and State University, Dec. 1983.

38. Whittle, P., *"The Analysis of Multiple Stationary Time Series"*, Journal of Royal Statistical Society, Series B, Vol.15, No.1, 1953, pp. 125-129.

39. Friedlander, B., *"The Asymptotic Performance of the Modified Yule-Walker Estimator"*, Proceedings International Federation on Automatic Control Symposium, 1983, pp. 22-26.

40. Friedlander, B., *"On the Computation of the Cramer Rao Bound for ARMA Parameter Estimation"* Proceedings International Conference on Acoustics, Speech, and Signal Processing, Mar. 1984, pp. 14.1.1-14.1.4.

41. Welch, P. D., *"The Use of Fast Fourier Transform for the Estimation of Power Spectra: A Method based on Time Averaging over short, modified Periodogram"* IEEE Transactions on Audio Electroacoustic, Vol. AU-15, Jun. 1967, pp. 70-73.

Appendix A

DERIVATION OF EQUATION (3-6)

$$\Delta\hat{P}(\omega) = \hat{P}(\omega) - \tilde{P}(\omega)$$

$$= \frac{\hat{\sigma}^2}{|A(e^{j\omega})|^2} - \frac{\tilde{\sigma}^2}{|A(e^{j\omega})|^2}$$

$$= \frac{\Delta\hat{\sigma}^2|\tilde{A}(e^{j\omega})|^2 - \tilde{\sigma}^2\Delta|\hat{A}(e^{j\omega})|^2}{|\hat{A}(e^{j\omega})|^2.|\tilde{A}(e^{j\omega})|^2} \qquad \text{(A-1)}$$

where

$$\Delta\hat{\sigma}^2 = \hat{\sigma}^2 - \tilde{\sigma}^2$$

$$\Delta|\hat{A}(e^{j\omega})|^2 = |\hat{A}(e^{j\omega})|^2 - |\tilde{A}(e^{j\omega})|^2$$

Now

$$|\hat{A}(e^{j\omega})|^2 = |1+ \sum_{i=1}^{p} \hat{a}_{i,p}e^{j\omega i}|^2$$

$$= |1+ \sum_{i=1}^{p} \tilde{a}_{i,p}e^{j\omega i} + \sum_{i=1}^{p} \Delta\hat{a}_{i,p}e^{j\omega i}|^2$$

$$= |\tilde{A}(e^{j\omega}) + \underline{\xi}^T(e^{j\omega})\Delta\underline{\hat{a}}|^2$$

$$= |\tilde{A}(e^{j\omega})|^2 + \tilde{A}(e^{-j\omega})\underline{\xi}^T(e^{j\omega})\Delta\underline{\hat{a}}$$

$$+ \tilde{A}(e^{j\omega})\underline{\xi}^T(e^{-j\omega})\Delta\underline{\hat{a}} + \underline{\xi}^T(e^{j\omega})\Delta\underline{\hat{a}}\Delta\underline{\hat{a}}^T\underline{\xi}(e^{-j\omega}) \qquad \text{(A-2)}$$

where

$$\Delta\hat{a}_{i,p} = \hat{a}_{i,p} - \tilde{a}_{i,p}$$

$$\underline{\xi}(e^{j\omega}) = [e^{j\omega}, e^{j2\omega}, \ldots, e^{j\omega p}]^T$$

So,

$$\Delta|\hat{A}(e^{j\omega})|^2 = 2Re\{\tilde{A}(e^{j\omega})\underline{\xi}^T(e^{-j\omega})\Delta\underline{\hat{a}}\} + \underline{\xi}^T(e^{j\omega})\Delta\underline{\hat{a}}\Delta\underline{\hat{a}}^T\underline{\xi}(e^{-j\omega}) \tag{A-3}$$

Assuming

$$2Re\{\tilde{A}(e^{j\omega})\underline{\xi}^T(e^{-j\omega})\Delta\underline{\hat{a}}\} >> \underline{\xi}^T(e^{j\omega})\Delta\underline{\hat{a}}\Delta\underline{\hat{a}}^T\underline{\xi}(e^{-j\omega}) \tag{A-4}$$

we get

$$\Delta|\hat{A}(e^{j\omega})|^2 \simeq 2Re\{\tilde{A}(e^{j\omega})\underline{\xi}^T(e^{-j\omega})\Delta\underline{\hat{a}}\} \tag{A-5}$$

Again we assume that

$$1/|\hat{A}(e^{j\omega})|^2 \simeq 1/|\tilde{A}(e^{j\omega})|^2 \tag{A-6}$$

Substituting (A-5) and (A-6) in (A-7), we have

$$\Delta\hat{P}(\omega) \simeq \frac{1}{|\tilde{A}(e^{j\omega})|^2}\left[\Delta\hat{\sigma}^2 - \frac{2\tilde{\sigma}^2 Re\{\tilde{A}(e^{j\omega})\underline{\xi}^T(e^{-j\omega})\Delta\underline{\hat{a}}\}}{|\tilde{A}(e^{j\omega})|^2}\right]$$

$$= \tilde{P}(\omega)\left[\frac{\Delta\hat{\sigma}^2}{\tilde{\sigma}^2} - \left\{\frac{\underline{\xi}^T(e^{j\omega})}{\tilde{A}(e^{j\omega})} + \frac{\underline{\xi}^T(e^{-j\omega})}{\tilde{A}(e^{-j\omega})}\right\}\Delta\underline{\hat{a}}\right] \tag{A-7}$$

Therefore,

$$E\{\Delta\hat{P}(\omega)\} \simeq \tilde{P}(\omega)[\tilde{\sigma}^{-2}E(\Delta\hat{\sigma}^2) - \underline{H}^T(e^{j\omega})E(\Delta\underline{\hat{a}})] \tag{A-8}$$

where

$$\underline{H}^T(e^{j\omega}) = \frac{\underline{\xi}^T(e^{j\omega})}{\tilde{A}(e^{j\omega})} + \frac{\underline{\xi}^T(e^{-j\omega})}{\tilde{A}(e^{-j\omega})}$$

$$= 2Re\left[\frac{\underline{\xi}^T(e^{j\omega})}{A(e^{j\omega})}\right]$$

Appendix B

TAYLOR SERIES EXPANSION OF PSD FUNCTION

Expanding P(a) in a Taylor series

$$P(\hat{\underline{a}}) = P(\underline{a}) + \sum_{n=1}^{\infty} \frac{1}{n!} (\Delta a_1 \frac{\delta}{\delta\phi_1} + \ldots + \Delta a_p \frac{\delta}{\delta\phi_p})^n P(\underline{\phi})|_{\underline{\phi}=\underline{a}} \tag{B-1}$$

where

$$\underline{a} = [a_1, a_2, \ldots, a_p]^T$$

$$\Delta a_i = \hat{a}_i - a_i$$

Hence,

$$\Delta P(\underline{a}) = P(\hat{\underline{a}}) - P(\underline{a})$$

$$= \sum_{i1=1}^{p} \Delta a_{i1} \frac{\delta P}{\delta\phi_{i1}} + \sum_{i1=1}^{p} \sum_{i2=1}^{p} \Delta a_{i1} \Delta a_{i2} \frac{\delta^2 P}{2! \delta\phi_{i1} \delta\phi_{i2}}$$

$$+ \ldots + \sum_{i1=1}^{p} \sum_{ip=1}^{p} \Delta a_{i1} \ldots \Delta a_{ip} \frac{\delta^n P}{n! \delta\phi_{i1} \ldots \delta\phi_{ip}}|_{\underline{\phi}=\underline{a}} \tag{B-2}$$

Here

$$P(\underline{a}) = \frac{\sigma^2}{A(z)A(z^{-1})} \tag{B-3}$$

So,

$$\frac{\delta P}{\delta a_i} = -2P[\mathrm{Re}\{\frac{z^i}{A(z)}\}] \tag{B-4}$$

$$\frac{\delta^2 P}{2!\,\delta a_i \delta a_j} = (4P)\mathrm{Re}\{\frac{z^i}{A(z)}\}\mathrm{Re}\{\frac{z^j}{A(z)}\} - \frac{P^2}{\sigma^2}\mathrm{Re}\{z^{i-j}\} \tag{B-5}$$

$$\begin{aligned}\frac{\delta^3 P}{3!\,\delta a_i \delta a_j \delta a_k} = &-(8P)\mathrm{Re}\{\frac{z^i}{A(z)}\}\mathrm{Re}\{\frac{z^j}{A(z)}\}\mathrm{Re}\{\frac{z^k}{A(z)}\} \\ &+ \frac{4P^2}{3\sigma^2}\mathrm{Re}\{z^{i-j}\}\mathrm{Re}\{\frac{z^k}{A(z)}\} \\ &+ \frac{4P^2}{3\sigma^2}\mathrm{Re}\{z^{k-i}\}\mathrm{Re}\{\frac{z^j}{A(z)}\} \\ &+ \frac{4P^2}{3\sigma^2}\mathrm{Re}\{z^{j-k}\}\mathrm{Re}\{\frac{z^i}{A(z)}\}\end{aligned} \tag{B-6}$$

and so on. If Δa is within the region of convergence and all the partial derivatives of $P(\underline{a})$ with respect to the AR coefficients exist, the Taylor series expansion of $P(\hat{\underline{a}})$ converges. Since the poles of a stable AR model are not on the unit circle in the z-domain, the above conditions of convergence are satisfied. However, when the poles approach the unit circle the rate of convergence goes down. The partial derivatives of $P(\underline{a})$ with respect to the AR coefficients are shown in Fig. 41. It is observed that the derivative increases with the order near the location of poles. So, if $(\hat{\underline{a}}-\underline{a})$ is not sufficiently small the series will not converge monotonically; the series will have damped oscillatory behavior. Here we observed that the rate of decrement of $(\hat{\underline{a}}-\underline{a})^n$ is higher than the rate of increment of

the nth order derivative. As a result the series converges. But when we take the expected value on both sides of (B-2), the series does not converge monotonically. Because even if $(a=a)^n$ is small it does not guarantee that $E(a-a)^n$ will be small. For this reason we find that the third term on the right hand side of (3-10) overestimates the error $\{\Delta P(w)\}$ near the location of poles which are close to the unit circle in the z-domain.

Appendix C

DERIVATION OF EQUATIONS (3-19) AND (3-20)

Let $g(y,z)$ be a function of two random variables y and z. Expanding $g(y,z)$ into a Taylor series about the mean values of y and z and then taking the expected value [33] yields

$$E\{g(y,z)\} \simeq g(\eta_y,\eta_z)+\mu_{yO}\frac{\delta^2 g}{2\delta y^2}+\mu_{yz}\frac{\delta^2 g}{\delta y\delta z}+\mu_{Oz}\frac{\delta^2 g}{2\delta z^2} \tag{C-1}$$

and

$$\begin{aligned} E\{g^2(y,z)\} \simeq\ & g^2(\eta_y,\eta_z) + \mu_{yO}[(\frac{\delta g}{\delta y})^2 + g\frac{\delta^2 g}{\delta y^2}] \\ & + 2\mu_{yz}[\frac{\delta g}{\delta y}\frac{\delta g}{\delta z} + g\frac{\delta^2 g}{\delta y\delta z}] \\ & + \mu_{Oz}[(\frac{\delta g}{\delta y})^2 + g\frac{\delta^2 g}{\delta z^2}] \end{aligned} \tag{C-2}$$

where

$$\eta_y = E(y), \qquad \eta_z = E(z),$$

$$\mu_{yO} = Var(y), \qquad \mu_{Oz} = Var(z), \qquad \mu_{yz} = Cov(yz) \tag{C-3}$$

Let

$$\hat{K}_m = g(y,z) = y/z \tag{C-4}$$

So,

$$\frac{\delta\hat{K}_m}{\delta y} = -1/z, \qquad \frac{\delta^2\hat{K}_m}{\delta y^2} = 0$$

$$\frac{\delta\hat{K}_m}{\delta z} = -y/z^2, \qquad \frac{\delta^2\hat{K}_m}{\delta z^2} = 2y/z^3, \qquad \frac{\delta^2\hat{K}_m}{\delta y\delta z} = -1/z^2 \tag{C-5}$$

Substituting (C-4) in (C-1) and (C-2) we get

$$E(\hat{K}_m) \simeq \frac{\eta_y}{\eta_z} - \frac{\mu_{y,z}}{\eta_z^2} + \frac{\eta_y \mu_{0z}}{\eta_z^3} \tag{C-6}$$

and

$$E(\hat{K}^2_m) = \frac{\eta_y^2 + \mu_{y0}}{\eta_z^2} + 3\frac{\eta_y^2 \mu_{0z}}{\eta_z^4} - 4\frac{\eta_y \mu_{yz}}{\eta_z^3} \tag{C-7}$$

Let us now derive equation (3-20).

$$E(y) = -\frac{2}{N} E\Big[\sum_{t=m}^{N-1} f_{t,m-1} b_{t-1,m-1}\Big]$$

$$= -\frac{2}{N} \sum_{t=m}^{N-1} E\Big\{\sum_{i=0}^{m-1}\sum_{j=0}^{m-1} \hat{a}_{i,m-1}\hat{a}_{j,m-1} x_{t-i} x_{t-m+j}\Big\}$$

$$= -\frac{2(N-m)}{N}\Big\{\sum_{i=0}^{m-1}\sum_{j=0}^{m-1} E(\hat{a}_{i,m-1})E(\hat{a}_{j,m-1}) r_{m-i-j}$$

$$= -2(N-m)\gamma_1/N = \eta_{y0} \tag{C-8}$$

where

$$\gamma_1 = \sum_{i=0}^{m-1}\sum_{j=0}^{m-1} E(\hat{a}_{i,m-1})E(\hat{a}_{j,m-1}) r_{m-i-j}$$

Similarly

$$\eta_z = 2(N-m)\gamma_2/N \tag{C-9}$$

where

$$\gamma_2 = \sum_{i=0}^{m-1} \sum_{j=0}^{m-1} E(\hat{a}_{i,m-1}) E(\hat{a}_{j,m-1}) r_{i-j}$$

In the above derivation we have treated the predictor coefficient $\hat{a}_{i,m}$ to behave as the constant $E(\hat{a}_{i,m})$.

$$E(y^2) = \frac{4}{N^2} E \sum_{t1=m}^{N-1} \sum_{t2=m}^{N-1} f_{t1,m-1} b_{t1-1,m-1} f_{t2,m-1} b_{t2-1,m-1}$$

$$= \frac{4}{N^2} E \sum_{t1=m}^{N-1} \sum_{t2=m}^{N-1} \sum_{i=0}^{m-1} \sum_{j=0}^{m-1} \sum_{k=0}^{m-1} \sum_{p=0}^{m-1} [a_{i,m-1} a_{j,m-1} a_{k,m-1} a_{p,m-1} \cdot$$

$$\cdot (x_{t1-i} x_{t1-m+j} x_{t2-k} x_{t2-m+p})]$$

(C-10)

Using the fourth order moment rule for the gaussian process, we get

$$E(p^2) = \frac{4}{N^2} \sum_{t1=m}^{N-1} \sum_{t2=m}^{N-1} \sum_{i=0}^{m-1} \sum_{j=0}^{m-1} \ldots \sum [E(\hat{a}_{i,m-1}) \ldots E(\hat{a}_{p,m-1}) \cdot$$

$$\cdot \{r_{m-i-j} r_{m-k-p} + r_{t1-t2-i+k} r_{t1-t2+j-p}$$

$$+ r_{t1-t2+m-i-p} r_{t1-t2-m+j+k}\}]$$

(C-11)

Therefore

$$\mu_{y0} = E(y^2) - E^2(y)$$

$$= \frac{4}{N^2} \sum_{t=-(N-1-m)}^{N-1-m} \sum_{i=0}^{m-1} \sum_{j=0}^{m-1} \ldots \sum [(N-m-|t|)E(\hat{a}_{i,m-1}) \ldots E(\hat{a}_{p,m-1}) \cdot$$

$$\cdot \{r_{t-i+k} r_{t+j-p} + r_{t+m-i-p} r_{t-m+j+k}\}] \qquad \text{(C-12)}$$

Simplifying we get

$$\mu_{y0} = 4\{\kappa_y + \sum_{t=1}^{N-m-1} 2(N-m-t)(\gamma_3^2 + \gamma_4\gamma_5)\}/N^2 \qquad \text{(C-13)}$$

where

$$\gamma_3 = \sum_{i=0}^{m-1} \sum_{j=0}^{m-1} E(\hat{a}_{i,m-1})E(\hat{a}_{j,m-1}) r_{t+i-j}$$

$$\gamma_4 = \sum_{i=0}^{m-1} \sum_{j=0}^{m-1} E(\hat{a}_{i,m-1})E(\hat{a}_{j,m-1}) r_{t+m-i-j}$$

$$\gamma_5 = \sum_{i=0}^{m-1} \sum_{j=0}^{m-1} E(\hat{a}_{i,m-1})E(\hat{a}_{j,m-1}) r_{t-m+i+j}$$

$$\kappa_y = (N-m)(\gamma_1^2 + \gamma_2^2)$$

In a similar way equations (3-20c) and (3-20d) can be derived.

Appendix D

COMPUTER PROGRAM LISTING

D.1 FORTRAN PROGRAM LISTING FOR COMPUTING POWER SPECTRAL DENSITY ESTIMATES OF THE MODIFIED BURG SPECTRAL ESTIMATOR

```
C       THIS PROGRAM COMPUTES THE POWER SPECTRAL DENSITY
C    ESTIMATES OF THE MODIFIED BURG SPECTRAL ESTIMATOR. THE
C    SAMPLE MEAN, SAMPLE VARIANCE AND SAMPLE MEAN SQUARE
C    ERROR OF THE ESTIMATES ARE ALSO EVALUATED. HERE THREE
C    TYPES OF AVERAGING TECHNIQUE ARE EMPLOYED, NAMELY AVA,
C    AVK AND AVP. ONLY ONE AVERAGING METHOD CAN BE USED AT A
C    TIME. IMSL SUBROUTINE IS USED FOR MATRIX INVERSION.
C
C       THIS IS AN INTERACTIVE PROGRAM. THE USER HAS TO ENTER
C    ALL THE VARIABLES (ORDER OF THE AR MODEL, NUMBER OF DATA
C    POINTS, NUMBER OF SECTIONS, AND NUMBER OF REALIZATIONS)
C    WHILE BEEN PROMPTED BY THE PROGRAM. THE COEFFICIENTS OF
C    THE FILTER SHOULD BE ENTERED IN THE FILE 'ARMACOEF DATA'
C    (DEVICE NUMBER 11) WITH THE FOLLOWING FORMATTING:
C
C    Q [ORDER OF MA PART (INT)]
C    B(0) B(1) ... B(Q) [MA COEFFICIENTS (FLOAT)]
C    P [ORDER OF AR PART (INT)]
C    A(0) A(1) ... A(P) [AR COEFFICIENTS (FLOAT)]
C
C
C       THE OUTPUT IS STORED IN THE FILES DEFINED AS
C    FOLLOWS:
C
C    DEVICE NO.     FILE                  DATA  STORED
C
C         1        ASP  DATA                ACTUAL PSD
C         2        MEAN DATA          ESTIMATED MEAN OF PSD
C         3        VAR  DATA      ESTIMATED VARIANCE OF PSD (DB)
C         4        MSE  DATA       ESTIMATED MEAN SQUARE ERROR
C                                  OF PSD (DB)
C
C*************************************************************
C
C    MAIN: COMPUTES THE SAMPLE MEAN, SAMPLE VARIANCE AND
C    SAMPLE MEAN SQUARE ERROR OF PSD ESTIMATES.
```

```
C
C*************************************************************
C
      DIMENSION A(0:20),B(0:20),E(512),X(512),ASP(129),
     #PSD(50,129)
      INTEGER P,Q,PQ
      REAL MEAN,MSQER
      DOUBLE PRECISION DSEED
      DATA NMAX/20/
C
C   READ INPUT DATA
C
      WRITE(6,701)
701   FORMAT(1X,'ENTER ORDER OF AR MODEL :')
      READ(5,*) NSTAGE
      WRITE(6,702)
702   FORMAT(1X,'ENTER NUMBER OF DATA POINTS :')
      READ(5,*) NX
      WRITE(6,703)
703   FORMAT(1X,'ENTER NUMBER OF SECTIONS :')
      READ(5,*) NSEC
      WRITE(6,704)
704   FORMAT(1X,'ENTER NUMBER OF REALIZATIONS :')
      READ(5,*) IRL
      WRITE(6,705)
      WRITE(6,706)
      WRITE(6,707)
      WRITE(6,708)
705   FORMAT(1X,'WHICH AVERAGING METHOD DO YOU WANT.')
706   FORMAT(1X,'1 : AVA METHOD')
707   FORMAT(1X,'2 : AVK METHOD')
708   FORMAT(1X,'3 : AVP METHOD')
      READ(5,*) IFLAG
C
      READ(11,*) Q
      READ(11,*) (B(I),I=0,Q)
      READ(11,*) P
      READ(11,*) (A(I),I=0,P)
C
C   FIND MAX(P,Q)
C
      PQ=MAX(P,Q)
      IF(P.EQ.Q) GO TO 991
      IF(PQ.EQ.P) GO TO 11
      GO TO 12
11    QP1=Q+1
      DO 13 I=QP1,PQ
      B(I)=0.
13    CONTINUE
      GO TO 991
```

```
12     PP1=P+1
       DO 14 I=PP1,PQ
       A(I)=0.
14     CONTINUE
C
C    COMPUTE ACTUAL POWER SPECTRAL DENSITY
C
       CALL PWR(A,B,P,Q,PQ,NMAX,ASP)
C
C    GENERATE NORMALIZED DATA
C
       NX1=NX+128
       DSEED=3214567.D0
C
       IR=0
999    IR=IR+1
       WRITE(6,709) IR
709    FORMAT(1X,'REALIZATION NUMBER ', I5)
       CALL GGNML(DSEED,NX1,E)
C
       RHAT=0.
       DO 21 I=1,NX1
       X(I)=B(0)*E(I)/A(0)
       DO 22 J=1,PQ
       IF(I.EQ.J) GO TO 23
       X(I)=-A(J)*X(I-J)/A(0)+B(J)*E(I-J)/A(0)+X(I)
22     CONTINUE
23     IF(I.LE.NX) GO TO 21
       RHAT=X(I)**2+RHAT
21     CONTINUE
       RHAT=RHAT/FLOAT(NX)
C
       DO 25 I=1,NX
       X(I)=X(NX+I)/SQRT(RHAT)
25     CONTINUE
C
C    SELECT THE TYPE OF AVERAGING METHOD
C
       IF(IFLAG.EQ.1) CALL AVA(X,PSD,RHAT,NSTAGE,NX,NSEC,IR,
      #NMAX)
       IF(IFLAG.EQ.2) CALL AVK(X,PSD,RHAT,NSTAGE,NX,NSEC,IR,
      #NMAX)
       IF(IFLAG.EQ.3) CALL AVP(X,PSD,RHAT,NSTAGE,NX,NSEC,IR,
      #NMAX)
C
       IF(IR.LT.IRL) GO TO 999
C
C    COMPUTE THE SAMPLE MEAN, SAMPLE VARIANCE, AND SAMPLE
C    MEAN SQUARE ERROR OF PSD ESTIMATES
C
```

```
      PI=4.*ATAN(1.)
      N=129
      AINC=PI/(N-1)
      DO 44 I=1,N
      RI=I
      W=AINC*(RI-1.)
      MEAN=0.
      DO 45 J=1,IRL
      MEAN=MEAN+PSD(J,I)
45    CONTINUE
      MEAN=MEAN/FLOAT(IRL)
      VAR=0.
      MSQER=0.
      DO 46 J=1,IRL
      VAR=(PSD(J,I)-MEAN)**2+VAR
      MSQER=(PSD(J,I)-ASP(I))**2+MSQER
46    CONTINUE
      VAR=VAR/FLOAT(IRL)
      MSQER=MSQER/FLOAT(IRL)
      VAR=10.*ALOG10(VAR)
      MSQER=10.*ALOG10(MSQER)
      W=W/PI
      WRITE(1,777) W,ASP(I)
      WRITE(2,777) W,MEAN
      WRITE(3,777) W,VAR
      WRITE(4,777) W,MSQER
777   FORMAT(2E16.6)
44    CONTINUE
C
110   STOP
      END
C
C*************************************************************
C
C   AVA: COMPUTES POWER SPECTRAL DENSITY ESTIMATES
C        USING AVERAGED AR COEFFICIENTS.
C
C*************************************************************
C
      SUBROUTINE AVA(X,PSD,RHAT,NSTAGE,NX,NSEC,IR,NMAX)
C
      DIMENSION A(20),B(20),SIG(512),X(512),PHI(20,20),
     #SCR(20),AVA(20),PSD(50,129)
      REAL K(20)
C
C   INITIALIZE VARIABLES
C
      L=0
      DO 30 I=1,NSTAGE
      AVA(I)=0.
```

```
30     CONTINUE
C
C    DIVIDE DATA SEQUENCE IN TO M SEGMENTS
C
111    NSIG=NX/NSEC
       DO 40 I=1,NSIG
       SIG(I)=X(L*NSIG+I)
40     CONTINUE
C
C    COMPUTE AUTOREGRESSIVE AND REFLECTION COEFFICIENTS
C
       DO 50 M=1,NSTAGE
       CALL COVAR(SIG,NSIG,M,PHI,NMAX)
       CALL CLHARM(PHI,NMAX,M,A,K(M),SCR)
50     CONTINUE
C
       DO 60 I=1,NSTAGE
       AVA(I)=AVA(I)+A(I)
60     CONTINUE
C
       IF(L.EQ.(NSEC-1)) GO TO 555
       L=L+1
       GO TO 111
C
C    AVERAGE AR COEFFICIENTS
C
555    DO 70 I=1,NSTAGE
       A(I+1)=AVA(I)/FLOAT(NSEC)
70     CONTINUE
       NB=0
       B(1)=1.
       A(1)=1.
C
C    FIND GAIN FACTOR SO THAT AVERAGE POWER IS UNITY
C
       CALL RSEQ(A,B,NSTAGE,NB,RO,NMAX)
C
C    COMPUTE PSD ESTIMATE
C
       SGMASQ=1./RO
       CALL PWRSPC(A,NSTAGE,NMAX,SGMASQ,IR,PSD)
C
       RETURN
       END
C
C*************************************************************
C
C    CLHARM: COVARIANCE LATTICE ROUTINE FOR HARMONIC MEAN
C            METHOD. IT COMPUTES THE REFLECTION COEFFICIENT
C            OF STAGE M, GIVEN THE COVARIANCE MATRIX OF THE
```

```
C            SIGNAL AND THE PREDICTOR COEFFICIENTS UP TO
C            STAGE M-1.
C    (R. VISWANATHAN AND J. MAKHOUL, "EFFICIENT LATTICE
C    METHODS FOR LINEAR PREDICTION", PROGRAMS FOR DIGITAL
C    SIGNAL PROCESSING, IEEE PRESS)
C
C*************************************************************
C
      SUBROUTINE CLHARM(PHI,NMAX,M,A,K,ERROR,SCR)
      DIMENSION PHI(NMAX,NMAX),A(NMAX),SCR(NMAX)
      REAL K
C
      IF(M.GT.1) GO TO 20
C
C    EXPLICIT COMPUTATION OF THE FIRST STAGE REFLECTION
C    COEFFICIENT
C
      FPLUSB=PHI(1,1)+PHI(2,2)
      C=PHI(1,2)
      K=0.
      IF(C.EQ.0.0) GO TO 10
      K=-2.*C/FPLUSB
10    A(1)=K
      GO TO 90
C
C    RECURSIVE COMPUTATION OF THE M-TH STAGE (M.GE.2)
C    REFLECTION COEFFICIENT
C
20    MP1=M+1
      MM1=M-1
      SUM1=0.
      SUM3=0.
      SUM4=0.
      SUM6=0.
      DO 30 I=1,MM1
      IP1=I+1
      SCR(I)=A(I)
      MP1MI=MP1-I
      SUM1=SUM1+A(I)*(PHI(1,IP1)+PHI(MP1,MP1MI))
      SUM3=SUM3+A(I)*(PHI(1,MP1MI)+PHI(IP1,MP1))
      Y=A(I)**2
      SUM4=SUM4+Y*(PHI(IP1,IP1)+PHI(MP1MI,MP1MI))
      SUM6=SUM6+Y*PHI(IP1,MP1MI)
30    CONTINUE
      SUM7=0.
      SUM9=0.
      IF(M.EQ.2) GO TO 60
      MM2=M-2
      DO 50 I=1,MM2
      IP1=I+1
```

```
      MP1MI=MP1-I
      DO 40 J=IP1,MM1
      Y=A(I)*A(J)
      MP1MJ=MP1-J
      SUM7=SUM7+Y*(PHI(IP1,J+1)+PHI(MP1MI,MP1MJ))
      SUM9=SUM9+Y*(PHI(IP1,MP1MJ)+PHI(J+1,MP1MI))
40    CONTINUE
50    CONTINUE
60    FPLUSB=PHI(1,1)+PHI(MP1,MP1)+2.*(SUM1+SUM7)+SUM4
      C=PHI(1,MP1)+SUM3+SUM6+SUM9
      K=0.
      IF(C.EQ.0.0) GO TO 70
      K=-2.*C/FPLUSB
70    CONTINUE
C
C   RECURSION TO CONVERT REFLECTION COEFFICIENT TO
C   PREDICTOR COEFFICIENT
C
      DO 80 I=1,MM1
      MMI=M-I
      A(I)=SCR(I)+K*SCR(MMI)
80    CONTINUE
      A(M)=K
C
      RETURN
      END
C
C*************************************************************
C
C   COVAR: COMPUTES (M+1)X(M+1) COVARIANCE MATRIX
C          CORRESPONDING TO THE LATTICE STAGE M.
C   (R. VISWANATHAN AND J. MAKHOUL, "EFFICIENT LATTICE
C   METHODS FOR LINEAR PREDICTION", PROGRAMS FOR DIGITAL
C   SIGNAL PROCESSING, IEEE PRESS)
C
C*************************************************************
C
      SUBROUTINE COVAR(SIG,NSIG,M,PHI,NMAX)
      DIMENSION SIG(NSIG),PHI(NMAX,NMAX)
C
      IF(M.GT.1) GO TO 20
C
C   FIRST STAGE: M=1
C
      TEMP1=0.
      TEMP2=0.
      DO 10 K=2,NSIG
      TEMP1=TEMP1+SIG(K)**2
      TEMP2=TEMP2+SIG(K)*SIG(K-1)
10    CONTINUE
```

```
      PHI(1,1)=TEMP1
      PHI(1,2)=TEMP2
      PHI(2,1)=TEMP2
      PHI(2,2)=TEMP1+SIG(1)**2-SIG(NSIG)**2
      RETURN
C
C   M-TH STAGE, M.GE.2
C
20    MP1=M+1
      NSP1=NSIG+1
      NSM=NSP1-M
      DO 30 J=2,MP1
      NSJ=NSP1+1-J
      PHI(MP1,J)=PHI(M,J-1)-SIG(NSM)*SIG(NSJ)
      PHI(J,MP1)=PHI(MP1,J)
30    CONTINUE
      TEMP1=0.
      DO 40 K=MP1,NSIG
      KMM=K-M
      TEMP1=TEMP1+SIG(K)*SIG(KMM)
40    CONTINUE
      PHI(MP1,1)=TEMP1
      PHI(1,MP1)=TEMP1
      DO 60 I=1,M
      MP1MI=MP1-I
      DO 50 J=I,M
      MP1MJ=MP1-J
      PHI(I,J)=PHI(I,J)-SIG(MP1MI)*SIG(MP1MJ)
      PHI(J,I)=PHI(I,J)
50    CONTINUE
60    CONTINUE
      RETURN
      END
C
C*************************************************************
C
C   PWRSPC: COMPUTE POWER SPECTRAL DENSITY FOR AR MODEL
C
C*************************************************************
C
      SUBROUTINE PWRSPC(A,NSTAGE,NMAX,BETASQ,IR,P1)
      DIMENSION A(NMAX),P1(50,129)
      COMPLEX Z,H
C
      PI=4.*ATAN(1.)
      N=129
      AINC=PI/(N-1)
C
      DO 30 I=1,N
      RI=I
```

```
      W=AINC*(RI-1.)
      H=(0.,0.)
      DO 40 J=1,NSTAGE
      RJ=J
      Z=CMPLX(0.,-W*RJ)
      Z=CEXP(Z)
      H=A(J+1)*Z+H
40    CONTINUE
      H=H+1
      P1(IR,I)=BETASQ/(CABS(H)**2)
30    CONTINUE
C
      RETURN
      END
C
C*************************************************************
C
C   RSEQ: COMPUTE THE CORRELATION SEQUENCE GENERATED BY ARMA
C         FILTER
C   (DUGRE, J. P., BEEX, A. A. AND SCHARF, L. L.,
C   "GENERATING COVARIANCE SEQUENCES AND THE CALCULATION .."
C   IEEE TRANS., VOL. ASSP-28, FEB. '80.)
C
C*************************************************************
C
      SUBROUTINE RSEQ(A,B,N,M,RO,NMAX)
      DIMENSION A(NMAX),B(NMAX),H(20),D(20),R(20),AR(20,20),
     #WKAREA(20)
C
      MP1=M+1
      NP1=N+1
C
C   COMPUTE IMPULSE RESPONSE
C
      DO 10 I=1,MP1
      H(I)=B(I)
      IF(NP1.EQ.1) GO TO 10
      DO 20 J=2,NP1
      IF((I-J+1).LT.1) GO TO 10
      H(I)=H(I)-A(J)*H(I-J+1)
20    CONTINUE
10    CONTINUE
C
C   COMPUTE 'D' VECTOR
C
      ND=NP1
      IF(NP1.LT.MP1) ND=MP1
      DO 30 I=1,ND
      D(I)=0.
      IF(I.GE.(MP1+1)) GO TO 31
```

```
      MJ=MP1-I+1
      DO 40 J=1,MJ
      D(I)=D(I)+H(J)*B(J+I-1)
40    CONTINUE
31    R(I)=D(I)
30    CONTINUE
C
C   COMPUTE MATRIX 'AR'
C
      IF(NP1.EQ.1) GO TO 89
      DO 55 I=1,NP1
      AR(I,1)=A(I)
      AR(1,I)=A(I)
55    CONTINUE
C
      DO 60 I=2,NP1
      DO 70 J=2,NP1
      AR(I,J)=0.
      IF((I+J-1).LT.1 .OR. (I+J-1).GT.NP1) GO TO 80
      AR(I,J)=A(I+J-1)
80    IF((I-J+1).LT.1 .OR. (I-J+1).GT.NP1) GO TO 70
      AR(I,J)=AR(I,J)+A(I-J+1)
70    CONTINUE
60    CONTINUE
C
C   SOLVE AR*R=D
C
      IDGT=0
      IM=1
      CALL LEQT1F(AR,IM,NP1,NMAX,R,IDGT,WKAREA,IER)
      IF(IER.NE.0) WRITE(6,200) IER
200   FORMAT(1X,'MATRIX INVERSION NOT POSSIBLE',I5)
C
89    R0=R(1)
C
      RETURN
      END
C
C*************************************************************
C
C   PWR: COMPUTE POWER SPECTRAL DENSITY FOR A GIVEN ARMA
C        DATA
C
C*************************************************************
C
      SUBROUTINE PWR(AA,BB,NN,MM,NM,NMAX,ASP)
      DIMENSION AA(NMAX),BB(NMAX),ASP(129)
      COMPLEX Z,ANUM,DNUM
C
      PI=4.*ATAN(1.)
```

```
      N=129
      AINC=PI/(N-1)
C
      NMP=NM+1
      DO 10 I=1,N
      RI=I
      W=AINC*(RI-1.)
      ANUM=(0.,0.)
      DNUM=(0.,0.)
      DO 20 J=1,NMP
      RJ=J
      Z=CMPLX(0.,-W*(RJ-1))
      Z=CEXP(Z)
      ANUM=BB(J)*Z+ANUM
      DNUM=AA(J)*Z+ANUM
20    CONTINUE
      ASP(I)=CABS(ANUM/DNUM)
      ASP(I)=ASP(I)**2
10    CONTINUE
C
      CALL RSEQ(AA,BB,NN,MM,RO,NMAX)
      DO 12 I=1,N
      ASP(I)=ASP(I)/RO
12    CONTINUE
C
      RETURN
      END
C
C*************************************************************
C
C   AVK: COMPUTES POWER SPECTRAL DENSITY ESTIMATES
C        USING AVERAGED REFLECTION COEFFICIENTS
C
C*************************************************************
C
      SUBROUTINE AVK(X,PSD,RHAT,NSTAGE,NX,NSEC,IR,NMAX)
      DIMENSION A(20),B(20),SIG(512),X(512),PHI(20,20)
     #SCR(20),AVK(20),PSD(50,129),AP(20)
      REAL K(20)
C
C   INITIALIZE VARIABLES
C
      L=0
      DO 30 I=1,NSTAGE
      AVK(I)=0.
30    CONTINUE
C
C   DIVIDE DATA SEQUENCE IN TO M SEGMENTS
C
111   NSIG=NX/NSEC
```

```
      DO 40 I=1,NSIG
      SIG(I)=X(L*NSIG+I)
40    CONTINUE
C
C   COMPUTE AUTOREGRESSIVE AND REFLECTION COEFFICIENTS
C
      DO 50 M=1,NSTAGE
      CALL COVAR(SIG,NSIG,M,PHI,NMAX)
      CALL CLHARM(PHI,NMAX,M,A,K(M),ERROR,SCR)
50    CONTINUE
C
      DO 60 I=1,NSTAGE
      AVK(I)=AVK(I)+K(I)
60    CONTINUE
C
      IF(L.EQ.(NSEC-1)) GO TO 555
      L=L+1
      GO TO 111
C
C   AVERAGE REFLECTION COEFFICIENTS
C
555   DO 70 I=1,NSTAGE
      K(I)=AVK(I)/FLOAT(NSEC)
70    CONTINUE
C
      A(1)=K(1)
      AP(1)=K(1)
      DO 80 I=2,NSTAGE
      A(I)=K(I)
      MM1=I-1
      DO 90 J=1,MM1
      A(J)=AP(J)+K(I)*AP(I-J)
90    CONTINUE
      DO 93 J=1,I
      AP(J)=A(J)
93    CONTINUE
80    CONTINUE
C
      DO 85 I=1,NSTAGE
      A(I+1)=AP(I)
85    CONTINUE
      NB=0
      B(1)=1.
      A(1)=1.
C
C   FIND GAIN FACTOR SO THAT AVERAGE POWER IS UNITY
C
      CALL RSEQ(A,B,NSTAGE,NB,RO,NMAX)
C
C   COMPUTE PSD ESTIMATE
```

```
C
      SGMASQ=1./R0
      CALL PWRSPC(A,NSTAGE,SGMASQ,IR,PSD)
C
      RETURN
      END
C
C*************************************************************
C
C   AVP: COMPUTES POWER SPECTRAL DENSITY ESTIMATES
C        USING AVERAGED PSD ESTIMATES
C
C*************************************************************
C
      SUBROUTINE AVP(X,PSD,RHAT,NSTAGE,NX,NSEC,IR,NMAX)
      DIMENSION PHI(20,20),A(20),SCR(20),SIG(512),X(512),
     #PSD(50,129),AA(20),BB(20),PP(129),AVP(129)
      REAL K(20)
C
C   INITIALIZE VARIABLES
C
      L=0
      DO 30 I=1,129
      AVP(I)=0.
30    CONTINUE
C
C   DIVIDE DATA SEQUENCE IN TO M SEGMENTS
C
      NSIG=NX/NSEC
111   RHAT=0.
      DO 40 I=1,NSIG
      SIG(I)=X(L*NSIG+I)
      RHAT=RHAT+SIG(I)**2
40    CONTINUE
      RHAT=RHAT/FLOAT(NSIG)
C
C   COMPUTE AUTOREGRESSIVE AND REFLECTION COEFFICIENTS
C
      DO 50 M=1,NSTAGE
      CALL COVAR2(SIG,NSIG,M,PHI,NMAX)
      CALL CLHARM(PHI,NMAX,M,A,K(M),ERROR,SCR)
50    CONTINUE
C
C   FIND GAIN FACTOR SO THAT AVERAGE POWER IS UNITY
C
      NB=0
      DO 65 I=I,NSTAGE
      AA(I+1)=A(I)
65    CONTINUE
      AA(1)=1.
```

```
      B(1)=1.
C
      CALL RSEQ(AA,B,NSTAGE,NB,R0,NMAX)
      BTASQ=RHAT/R0
C
C   EVALUATE PSD ESTIMATE
C
      CALL PWRSPC(AA,NSTAGE,NMAX,BTASQ,PP)
C
      DO 60 I=1,129
      AVP(I)=AVP(I)+PP(I)
60    CONTINUE
C
      IF(L.EQ.(NSEC-1)) GO TO 555
      L=L+1
      GO TO 111
C
C   AVERAGE PSD ESTIMATES
C
555   DO 70 I=1,129
      PSD(IR,I)=AVP(I)/FLOAT(NSEC)
70    CONTINUE
C
110   STOP
      END
```

D.2 FORTRAN PROGRAM LISTING FOR COMPUTING APPROXIMATE MEAN AND VARIANCE OF THE MODIFIED BURG SPECTRAL ESTIMATOR

```
C      THIS PROGRAM COMPUTES THE APPROXIMATE MEAN AND
C   VARIANCE OF THE MBSE ESTIMATES USING THE SAKAI AND
C   TAYLOR APPROXIMATION METHODS. THIS PROGRAM ALSO
C   EVALUATES THE THEORETICAL CRAMER-RAO LOWER BOUND
C   ('CRAMER MODEL' BOUND) FOR THE VARIANCE OF ESTIMATES.
C   IMSL SUBROUTINE IS USED FOR MATRIX INVERSION.
C
C      THIS IS AN INTERACTIVE PROGRAM. THE USER HAS TO ENTER
C   ALL THE VARIABLES (ORDER OF THE AR MODEL, NUMBER OF DATA
C   POINTS, NUMBER OF SECTIONS, AND NUMBER OF REALIZATIONS)
C   WHILE BEEN PROMPTED BY THE PROGRAM. THE COEFFICIENTS OF
C   THE FILTER SHOULD BE ENTERED IN THE FILE 'ARMACOEF DATA'
C   (DEVICE NUMBER 11) WITH THE FOLLOWING FORMATTING:
C
C   M [ORDER OF MA PART (INT)]
C   B(0) B(1) ... B(M) [MA COEFFICIENTS (FLOAT)]
```

```
C   N [ORDER OF AR PART (INT)]
C   A(0) A(1) ... A(N) [AR COEFFICIENTS (FLOAT)]
C
C
C      THE OUTPUT IS STORED IN THE FILES DEFINED AS
C   FOLLOWS:
C
C   DEVICE NO.     FILE                DATA  STORED

C        1        ASP  DATA              ACTUAL PSD
C        2        PWR  DATA     PSD ESTIMATED FROM AR MODEL
C        3        MINS DATA     APPROX. MEAN (SAKAI) OF PSD
C        4        VARS DATA     APPROX. VARIANCE (SAKAI) OF
C                               PSD IN DB
C        5        MSES  DATA    APPROX. MEAN SQ. ERROR (SAKAI)
C                               OF PSD IN DB
C        6        MINT DATA     APPROX. MEAN (TAYLOR) OF PSD
C        7        VART DATA     APPROX. VARIANCE (TAYLOR) OF
C                               PSD IN DB
C        8        MSET DATA     APPROX. MEAN SQ. ERROR (TAYLOR)
C                               OF PSD IN DB
C        9        VARC DATA     CRLB MODEL BOUND FOR VARIANCE
C       10        MSEC DATA     CRLB MODEL BOUND FOR MEAN SQ.
C                               ERROR
C
C*************************************************************
C
C   MAIN: COMPUTES THE APPROXIMATE MEAN AND VARIANCE USING
C         THE SAKAI APPROXIMATION AND THE TAYLOR
C         APPROXIMATION METHODS. ALSO COMPUTES THE 'CRAMER
C         MODEL' BOUND FOR THE VARIANCE AND MEAN SQUARE
C         ERROR OF PSD ESTIMATES.
C
C*************************************************************
C
      DIMENSION A(0:20),B(0:20),R(512),COVA(20,20),D(20,20),
     #EDELA(20),CVDELA(20,20),P(129),H(20),VART(129),
     #AS(20),BS(20),SMEAN(20),SA(20)
      REAL MEANA(20),MEANK,MINS,MINT,ASP(129)
      DATA NMAX/20/
C
C   READ INPUT DATA
C
      NMAX=20
      READ(11,*) M
      READ(11,*) (B(I),I=0,M)
      READ(11,*) N
      READ(11,*) (A(I),I=0,N)
      WRITE(6,11)
11    FORMAT(1X,'GIVE THE VALUE OF NSTAGE:')
```

```
      READ(5,*) NSTAGE
      WRITE(6,12)
12    FORMAT(1X,'GIVE THE VALUE OF NSAMPLE:')
      READ(5,*) NSMPL
      WRITE(6,13)
13    FORMAT(1X,'GIVE THE VALUE OF NSECTION:')
      READ(5,*) NSEC
      WRITE(6,14)
14    FORMAT(1X,'SELECT AVERAGING METHOD:')
      WRITE(6,15)
15    FORMAT(1X,'AVA=1, AVK=2, AVP=3')
      READ(5,*) IFG
C
C   FIND MAX(N,M)
C
      NM=MAX(N,M)
      IF(N.EQ.M) GO TO 991
      IF(NM.EQ.N) GO TO 61
      GO TO 62
61    MP1=M+1
      DO 63 I=MP1,NM
      B(I)=0.
63    CONTINUE
      GO TO 991
62    NP1=N+1
      DO 64 I=NP1,NM
      A(I)=0.
64    CONTINUE
C
C   COMPUTE ACTUAL POWER SPECTRAL DENSITY
C
991   CALL PWR(A,B,N,M,NM,NSTAGE,NSMPL,NMAX,ASP)
C
C   EVALUATE ACTUAL AUTOCORRELATION SEQUENCE
C
      RNSCT=NSEC
      NSMPL=NSMPL/NSEC
      IF(IFG.EQ.3) NSEC=1
      IFLAG=0
      CALL RSEQ(A,B,N,M,R,NSTAGE,NSMPL,NMAX,IFLAG)
C
C   COMPUTE APPROX. BIAS AND VARIANCE OF AR COEFFICIENTS
C   ESTIMATES
C
      DO 10 NST=1,NSTAGE
      CALL KSTAT(NST,R,MEANA,MEANK,VARK,NSMPL)
      CALL AREST(AS,BS,N,M,R,NST,NMAX)
      CALL VAR(MEANA,NST,MEANK,VARK,COVA,NSEC,NMAX,IFG)
      CALL BIAS(AS,NST,EDELA,MEANA,MEANK,NMAX)
10    CONTINUE
```

```
C
      DO 32 I=1,NSTAGE
      SA(I)=AS(I+1)
32    CONTINUE
C
C   FIND ERROR COVARIANCE MATRIX OF AR COEFFICIENTS
C
      CALL VARDEL(SA,NSTAGE,MEANA,COVA,CVDELA,NSEC,NMAX,IFG)
C
C   DETERMINE PSD ESTIMATE FROM THE AR MODEL
C
      IFLAG=1
      CALL RSEQ(AS,BS,N,M,R,NSTAGE,NSMPL,NMAX,IFLAG)
      SGMSQ=1./R(1)
      CALL PWRSPC(SA,NSTAGE,SGMSQ,P,NMAX)
C
C   COMPUTE BIAS OF GAIN FACTOR
C
      DO 49 I=1,NSTAGE
      SMEAN(I+1)=MEANA(I)
49    CONTINUE
      SMEAN(1)=1.
C
      CALL RSEQ(SMEAN,BS,N,M,R,NSTAGE,NSMPL,NMAX,IFLAG)
      SGMSQT=1./R(1)
C
      DELSGM=SGMSQT-SGMSQ
C
C   COMPUTE APPROX. MEAN AND MEAN SQUARE OF  P(w)
C   USING THE SAKAI METHOD AND THE TAYLOR METHOD
C
      PI=4.*ATAN(1.)
      NPT=129
      AINC=PI/(NPT-1)
C
      DO 20 L=1,NPT
      RL=L
      W=AINC*(RL-1.)
      VARS=0.
      VART(L)=0.
      MINS=0.
      CALL HMAT(SA,NSTAGE,H,W,NMAX)
      DO 30 I=1,NSTAGE
      DO 31 J=1,NSTAGE
      VARS=VARS+H(I)*CVDELA(I,J)*H(J)
      VART(L)=VART(L)+H(I)*CVDELA(I,J)*H(J)
31    CONTINUE
      VARS=VARS-2.*H(I)*EDELA(I)*DELSGM/SGMSQ
      MINS=MINS-H(I)*EDELA(I)
30    CONTINUE
```

```
      VARS=(VARS+(DELSGM**2)/(SGMSQ**2))*(P(L)**2)
      VART(L)=VART(L)*(P(L)**2)
      MINS=(MINS+DELSGM/SGMSQ)*P(L)
      IF(IFG.EQ.3) VARS=VARS/RNSCT+(RNSCT-1.)*(MINS**2)/RNSCT
      W=W/PI
      SQERRS=(ASP(L)-P(L))**2+VARS-2.*(ASP(L)-P(L))*MINS
      VARS=SQERRS-(ASP(L)-P(L)-MINS)**2
      MINS=P(L)+MINS
      VARS=10.*ALOG10(VARS)
      SQERRS=10.*ALOG10(SQERRS)
C
      WRITE(1,100) W,ASP(L)
      WRITE(2,100) W,P(L)
      WRITE(3,100) W,MINS
      WRITE(4,100) W,VARS
      WRITE(5,100) W,SQERRS
100   FORMAT(2E16.6)
C
20    CONTINUE
C
      DO 70 L=1,NPT
      RL=L
      W=AINC*(RL-1.)
      MINT=0.
      PP=P(L)
      CALL HMAT(SA,NSTAGE,H,W,NMAX)
      DO 80 I=1,NSTAGE
      MINT=MINT-H(I)*EDELA(I)*P(L)
80    CONTINUE
      IF(IFG.EQ.3) VART(L)=VART(L)/RNSCT+(RNSCT-1.)
     #*(MINT**2)/RNSCT
      W=W/PI
      SQERRT=(ASP(L)-P(L))**2+VART(L)-2.*(ASP(L)-P(L))*MINT
      VART(L)=SQERRT-(ASP(L)-P(L)-MINT)**2
      MINT=P(L)+MINT
      VART(L)=10.*ALOG10(VART(L))
      SQERRT=10.*ALOG10(SQERRT)
      WRITE(6,100) W,MINT
      WRITE(7,100) W,VART(L)
      WRITE(8,100) W,SQERRT
70    CONTINUE
C
C   COMPUTE THE 'CRAMER MODEL' BOUND
C
      CALL CRAMER(SA,NSTAGE,CVDELA,NSMPL,RNSCT,NMAX)
C
      DO 23 L=1,NPT
      RL=L
      W=AINC*(RL-1.)
      VARC=0.
```

```
      CALL HMAT(SA,NSTAGE,H,W,NMAX)
      DO 24 I=1,NSTAGE
      DO 24 J=1,NSTAGE
      VARC=VARC+H(I)*CVDELA(I,J)*H(J)
24    CONTINUE
      VARC=(VARC+2)*(P(L)**2)
      W=W/PI
      SQERRC=(ASP(L)-P(L))**2+VARC
      VARC=10.*ALOG10(VARC)
      SQERRC=10.*ALOG10(SQERRC)
      WRITE(9,100) W,VARC
      WRITE(10,100) W,SQERRC
23    CONTINUE
C
999   STOP
      END
C
C*************************************************************
C
C   TRANS: TRANSPOSE THE MATRIX
C
C*************************************************************
C
      SUBROUTINE TRANS(A,B,N,M,NMAX)
      DIMENSION A(NMAX,NMAX),B(NMAX,NMAX)
C
      DO 10 I=1,M
      DO 10 J=1,N
      B(I,J)=A(J,I)
10    CONTINUE
C
      RETURN
      END
C
C*************************************************************
C
C   MULMAT: MULTIPLY TWO MATRICES
C
C*************************************************************
C
      SUBROUTINE MULMAT(A,B,C,N,M,L,NMAX)
      DIMENSION A(NMAX,NMAX),B(NMAX,NMAX),C(NMAX,NMAX)
C
      DO 10 I=1,N
      DO 10 J=1,L
      C(I,J)=0.
      DO 10 K=1,M
      C(I,J)=A(I,K)*B(K,J)+C(I,J)
10    CONTINUE
C
```

```
      RETURN
      END
C
C*************************************************************
C
C   RSEQ: COMPUTE AUTOCORRELATION SEQUENCE
C
C   (DUGRE, J. P., BEEX, A. A., AND SCHARF, L. L.,
C   "GENERATING COVARIANCE SEQUENCES AND THE CALCULATION .."
C   IEEE TRANS., VOL. ASSP-28, FEB. '80.)
C
C*************************************************************
C
      SUBROUTINE RSEQ(A,B,N,M,R,NSTAGE,NSMPL,NMAX,IFLAG)
      DIMENSION A(NMAX),B(NMAX),H(20),D(20),E(20),
     #C(20,20),WKAREA(20),R(512)
C
      MP1=M+1
      NP1=N+1
C
C   EVALUATE IMPULSE RESPONSE
C
      DO 10 I=1,MP1
      H(I)=B(I)
      IF(NP1.EQ.1) GO TO 10
      DO 20 J=2,NP1
      IF((I-J+1).LT.1) GO TO 10
      H(I)=H(I)-A(J)*H(I-J+1)
20    CONTINUE
10    CONTINUE
C
C   COMPUTE 'D' VECTOR
C
      ND=NP1
      IF(NP1.LT.MP1) ND=MP1
      DO 30 I=1,ND
      D(I)=0.
      IF(I.GE.(MP1+1)) GO TO 31
      MJ=MP1-I+1
      DO 40 J=1,MJ
      D(I)=D(I)+H(J)*B(J+I-1)
40    CONTINUE
31    E(I)=D(I)
30    CONTINUE
C
C   FORM MATRIX 'C'
C
      IF(NP1.EQ.1) GO TO 89
      DO 55 I=1,NP1
      C(I,1)=A(I)
```

```
      C(1,I)=A(I)
55    CONTINUE
C
C   SOLVE C*E=D
C
      DO 60 I=2,NP1
      DO 70 J=2,NP1
      C(I,J)=0.
      IF((I+J-1).LT.1 .OR. (I+J-1).GT.NP1) GO TO 80
      C(I,J)=A(I+J-1)
80    IF((I-J+1).LT.1 .OR. (I-J+1).GT.NP1) GO TO 70
      C(I,J)=C(I,J)+A(I-J+1)
70    CONTINUE
60    CONTINUE
C
      IDGT=0
      IM=1
C
      CALL LEQT1F(C,IM,NP1,NMAX,E,IDGT,WKAREA,IER)
      IF(IER.NE.0) WRITE(6,200) IER
200   FORMAT(1X, 'MATRIX INVERSION NOT POSSIBLE, IER=',I5)
C
C   COMPUTE AUTOCORRELATION SEQUENCE
C
89    N1=NSMPL+1
      DO 90 I=1,N1
      IF(I.GE.(NP1+1)) GO TO 92
      R(I)=E(I)
      GO TO 90
92    R(I)=0.
      IF(I.LE.MP1) R(I)=D(I)
      DO 101 J=2,NP1
      R(I)=R(I)-A(J)*R(I-J+1)
101   CONTINUE
90    CONTINUE
C
      IF(IFLAG.EQ.1) RETURN
C
      Q=R(1)
      DO 91 I=1,N1
      RI=I-1
      R(I)=R(I)/Q
91    CONTINUE
C
      RETURN
      END
C
C*************************************************************
C
C   AREST: DETERMINE AR COEFFS. OF PTH ORDER AR MODEL
```

```
C
C*************************************************************
C
      SUBROUTINE AREST(A,B,N,M,R,NSTAGE,NMAX)
      DIMENSION A(NMAX),B(NMAX),R(512),RMAT(20,20),RVEC(20),
     #WKAREA(20)
C
C    FORM COVARIANCE MATRIX RMAT
C
      DO 10 I=1,NSTAGE
      DO 20 J=I,NSTAGE
      RMAT(I,J)=R(J-I+1)
      RMAT(J,I)=RMAT(I,J)
20    CONTINUE
      RVEC(I)=-R(I+1)
10    CONTINUE
C
C    SOLVE FOR AR COEFFICIENTS
C
      IM=1
      IDGT=0
      CALL LEQT1F(RMAT,IM,NSTAGE,NMAX,RVEC,IDGT,WKAREA,IER)
      IF(IER.NE.0) WRITE(6,*) IER
C
      N=NSTAGE
      DO 30 I=1,NSTAGE
      A(I+1)=RVEC(I)
30    CONTINUE
      M=0
      A(1)=1.
      B(1)=1.
C
      RETURN
      END
C
C*************************************************************
C
C    PWR: EVALUATE POWER SPECTRAL DENSITY FOR A GIVEN ARMA
C         DATA
C
C*************************************************************
C
      SUBROUTINE PWR(AA,BB,NN,MM,NM,NSTAGE,NSMPL,NMAX,ASP)
      DIMENSION AA(NMAX),BB(NMAX),ASP(129),R(512)
      COMPLEX Z,ANUM,DNUM
C
      PI=4.*ATAN(1.)
      N=129
      AINC=PI/(N-1)
C
```

```
      NMP=NM+1
      DO 10 I=1,N
      RI=I
      W=AINC*(RI-1.)
      ANUM=(0.,0.)
      DNUM=(0.,0.)
      DO 20 J=1,NMP
      RJ=J
      Z=CMPLX(0.,-W*(RJ-1))
      Z=CEXP(Z)
      ANUM=BB(J)*Z+ANUM
      DNUM=AA(J)*Z+ANUM
20    CONTINUE
      ASP(I)=CABS(ANUM/DNUM)
      ASP(I)=ASP(I)**2
10    CONTINUE
C
      IFLAG=1
      CALL RSEQ(AA,BB,NN,MM,R,NSTAGE,NSMPL,NMAX,IFLAG)
      DO 12 I=1,N
      ASP(I)=ASP(I)/R(1)
12    CONTINUE
C
      RETURN
      END
C
C*************************************************************
C
C   PWRSPC: COMPUTE POWER SPECTRAL DENSITY FOR AR MODEL
C
C*************************************************************
C
      SUBROUTINE PWRSPC(A,NSTAGE,BETASQ,P1,NMAX)
      DIMENSION A(NMAX),P1(129)
      COMPLEX Z,H
C
      PI=4.*ATAN(1.)
      N=129
      AINC=PI/(N-1)
C
      DO 30 I=1,N
      RI=I
      W=AINC*(RI-1.)
      H=(0.,0.)
      DO 40 J=1,NSTAGE
      RJ=J
      Z=CMPLX(0.,-W*RJ)
      Z=CEXP(Z)
      H=A(J)*Z+H
40    CONTINUE
```

```
      H=H+1.
      P1(I)=BETASQ/(CABS(H)**2)
30    CONTINUE
C
      RETURN
      END
C
C*************************************************************
C
C   HMAT: DETERMINE 'H' MATRIX
C
C*************************************************************
C
      SUBROUTINE HMAT(A,NSTAGE,H,W,NMAX)
      DIMENSION A(NMAX),H(NMAX)
      COMPLEX Z,DNUM
C
      DNUM=(1.,0.)
      DO 10 I=1,NSTAGE
      RI=I
      Z=CMPLX(0.,W*RI)
      Z=CEXP(Z)
      DNUM=DNUM+A(I)*Z
10    CONTINUE
C
      DO 20 I=1,NSTAGE
      RI=I
      Z=CMPLX(0.,W*RI)
      Z=CEXP(Z)
      H(I)=2.*REAL(Z/DNUM)
20    CONTINUE
C
      RETURN
      END
C
C*************************************************************
C
C   BIAS: EVALUATE MEAN AND BIAS OF AR COEFFS.
C
C*************************************************************
C
      SUBROUTINE BIAS(A,N,EDELA,MEANA,MEANK,NMAX)
      DIMENSION A(NMAX),EDELA(NMAX),EAHAT(20)
      REAL MEANA(NMAX),MEANK
C
      IF(N.NE.1) GO TO 79
      MEANA(1)=MEANK
      EDELA(1)=MEANA(1)-A(2)
      RETURN
C
```

```
79     NM1=N-1
       DO 10 I=1,NM1
       EAHAT(I)=MEANA(I)+MEANK*MEANA(N-I)
10     CONTINUE
       EAHAT(N)=MEANK
C
       DO 20 I=1,N
       MEANA(I)=EAHAT(I)
20     CONTINUE
C
       DO 30 I=1,N
       EDELA(I)=MEANA(I)-A(I+1)
30     CONTINUE
C
       RETURN
       END
C
C*************************************************************
C
C    VARDEL: DETERMINE ERROR COVARIANCE MATRIX OF AR COEFFS.
C
C*************************************************************
C
       SUBROUTINE VARDEL(A,NSTAGE,MEANA,COVA,CVDELA,NSEC,
      #NMAX,IFG)
       DIMENSION A(NMAX),CVDELA(NMAX,NMAX),COVA(NMAX,NMAX),
      #TEMP1(20,20),TEMP2(20,20),COVAHT(20,20)
       REAL MEANA(NMAX),AA(20,20),MAA(20,20)
C
       RNSEC=NSEC
       DO 32 I=1,NSTAGE
       AA(I,1)=A(I)
       MAA(I,1)=MEANA(I)
32     CONTINUE
C
       CALL TRANS(AA,TEMP1,NSTAGE,1,NMAX)
       CALL MULMAT(AA,TEMP1,TEMP2,NSTAGE,1,NSTAGE,NMAX)
C
       DO 10 I=1,NSTAGE
       DO 10 J=1,NSTAGE
       IF(IFG.EQ.1.OR.IFG.EQ.3) CVDELA(I,J)=TEMP2(I,J)
      #+COVA(I,J)/RNSEC
       IF(IFG.EQ.2) CVDELA(I,J)=TEMP2(I,J)+COVA(I,J)
10     CONTINUE
C
       CALL TRANS(MAA,TEMP1,NSTAGE,1,NMAX)
       IF(IFG.EQ.2) GO TO 43
       CALL MULMAT(MAA,TEMP1,TEMP2,NSTAGE,1,NSTAGE,NMAX)
C
       DO 20 I=1,NSTAGE
```

```
      DO 20 J=1,NSTAGE
      CVDELA(I,J)=(RNSEC-1.)*TEMP2(I,J)/RNSEC+CVDELA(I,J)
20    CONTINUE
C
43    CALL MULMAT(AA,TEMP1,TEMP2,NSTAGE,1,NSTAGE,NMAX)
C
      DO 30 I=1,NSTAGE
      DO 30 J=1,NSTAGE
      CVDELA(I,J)=CVDELA(I,J)-TEMP2(I,J)-TEMP2(J,I)
30    CONTINUE
C
      RETURN
      END
C
C*************************************************************
C
C   KSTAT: COMPUTE MEAN AND MEAN SQUARE OF REFLECTION COEFF.
C
C*************************************************************
C
      SUBROUTINE KSTAT(NST,RR,MEANA,MEANK,VARK,NSMPL)
      DIMENSION RR(0:511),R(-511:511)
      REAL MEANK,MEANA(1:20),MEANAA(0:19)
      INTEGER T
C
C   INITIALIZE VARIABLES
C
      NSMPLM=NSMPL-1
      DO 9 I=1,NSMPLM
      R(I)=RR(I)
      R(-I)=RR(I)
9     CONTINUE
      R(0)=RR(0)
C
      NSTM=NST-1
      NM=NSMPL-NST-1
      RNSMPL=NSMPL
      RNST=NST
C
      IF(NSTM.EQ.0) GO TO 74
96    DO 73 I=1,NSTM
      MEANAA(I)=MEANA(I)
73    CONTINUE
74    MEANAA(0)=1.
C
      EP=0.
      EQ=0.
      EPQ=0.
      EPSQ=0.
      EQSQ=0.
```

```
C
      DO 11 I1=0,NSTM
      DO 11 I2=0,NSTM
      EP=EP+MEANAA(I1)*MEANAA(I2)*R(NST-I1-I2)
      EQ=EQ+MEANAA(I1)*MEANAA(I2)*R(I1-I2)
11    CONTINUE
      R4=EP
      R5=EQ
      EP=-2.*(RNSMPL-RNST)*EP/RNSMPL
      EQ=2.*(RNSMPL-RNST)*EQ/RNSMPL
C
      DO 51 T=1,NM
      R1=0.
      R2=0.
      R3=0.
      DO 61 I1=0,NSTM
      DO 61 I2=0,NSTM
      R1=R1+MEANAA(I1)*MEANAA(I2)*R(T+I1-I2)
      R2=R2+MEANAA(I1)*MEANAA(I2)*R(T+NST-I1-I2)
      R3=R3+MEANAA(I1)*MEANAA(I2)*R(T-NST+I1+I2)
61    CONTINUE
C
      RT=T
      EPSQ=EPSQ+2.*(RNSMPL-RNST-RT)*(R1**2+R2*R3)
      EQSQ=EQSQ+4.*(RNSMPL-RNST-RT)*(2.*R1**2+R2**2+R3**2)
      EPQ=EPQ+4.*(RNSMPL-RNST-RT)*R1*(R2+R3)
51    CONTINUE
C
      EPSQ=EPSQ+(RNSMPL-RNST)*(R4**2+R5**2)
      EQSQ=EQSQ+4.*(RNSMPL-RNST)*(R4**2+R5**2)
      EPQ=EPQ+4.*(RNSMPL-RNST)*(R4*R5)
C
      EPQ=-2.*EPQ/(RNSMPL**2)
      EPSQ=4.*EPSQ/(RNSMPL**2)
      EQSQ=EQSQ/(RNSMPL**2)
C
      MEANK=EP/EQ-EPQ/(EQ**2)+EQSQ*EP/(EQ**3)
      VARK=(EP/EQ)**2+EPSQ/(EQ**2)+EQSQ*((EP**2)/(EQ**4)
     #     +2.*(EP**2)/(EQ**4))-4.*EPQ*EP/(EQ**3)
C
97    RETURN
      END
C
C*************************************************************
C
C   VAR: EVALUATES APPROX. VARIANCE OF AR COEFFICIENTS
C
C*************************************************************
C
      SUBROUTINE VAR(MEANA,N,MEANK,VARK,COVA,NSEC,NMAX,IFG)
```

```
      DIMENSION COVA(NMAX,NMAX),UNITJ(20,20),T1(20,20),
     #T2(20,20)
      REAL MEANA(NMAX),MEANK,T3(20,20)
C
      IF(N.NE.1) GO TO 75
      COVA(1,1)=VARK
      RETURN
C
C   FORM REVERSE OPERATOR MATRIX J
C
75    NM1=N-1
      DO 10 I=1,NM1
      DO 10 J=1,NM1
      UNITJ(I,J)=0.
      IF(J.EQ.(NM1-I+1)) UNITJ(I,J)=1.
10    CONTINUE
C
C   EVALUATES COVARIANCE MATRIX OF AR COEFFICIENTS
C
      CALL MULMAT(UNITJ,COVA,T1,NM1,NM1,NM1,NMAX)
      CALL MULMAT(COVA,UNITJ,T2,NM1,NM1,NM1,NMAX)
      CALL MULMAT(T1,UNITJ,T3,NM1,NM1,NM1,NMAX)
C
      IF(IFG.EQ.1.OR.IFG.EQ.3) GO TO 73
      RNSEC=NSEC
      VARK=(VARK/RNSEC)+(RNSEC-1.)*(MEANK**2)/RNSEC
C
73    DO 20 I=1,NM1
      DO 20 J=1,NM1
      COVA(I,J)=COVA(I,J)+MEANK*T1(I,J)+MEANK*T2(I,J)
     #+VARK*T3(I,J)
20    CONTINUE
C
      DO 30 I=1,NM1
      COVA(I,N)=MEANK*MEANA(I)+VARK*MEANA(N-I)
      COVA(N,I)=COVA(I,N)
30    CONTINUE
C
      COVA(N,N)=VARK
C
      RETURN
      END
C
C*************************************************************
C
C   CRAMER: DETERMINE THEORETICAL CRAMER LOWER BOUND
C           FOR VARIANCE
C
C*************************************************************
C
```

```
      SUBROUTINE CRAMER(A,NSTAGE,COVA,NSMPL,RNSCT,NMAX)
      DIMENSION A(NMAX),COVA(NMAX,NMAX),A1(20,20),AT1(20,20),
     #B1(20,20),A2(20,20),AT2(20,20),B2(20,20)
C
      DO 10 I=1,NSTAGE
      DO 10 J=1,NSTAGE
      A1(I,J)=0.
      IF((I-J+1).LE.0) GO TO 10
      A1(I,J)=A(I-J+1)
10    CONTINUE
C
      DO 20 I=1,NSTAGE
      DO 20 J=1,NSTAGE
      A2(I,J)=0.
      IF((NSTAGE-I+J+1).GT.(NSTAGE+1)) GO TO 20
      A2(I,J)=A(NSTAGE-I+J+1)
20    CONTINUE
C
      CALL TRANS(A1,AT1,NSTAGE,NSTAGE,NMAX)
      CALL MULMAT(A1,AT1,B1,NSTAGE,NSTAGE,NSTAGE,NMAX)
      CALL TRANS(A2,AT2,NSTAGE,NSTAGE,NMAX)
      CALL MULMAT(A2,AT2,B2,NSTAGE,NSTAGE,NSTAGE,NMAX)
C
      RNSMPL=NSMPL
      DO 30 I=1,NSTAGE
      DO 30 J=1,NSTAGE
      COVA(I,J)=(B1(I,J)-B2(I,J))/(RNSMPL*RNSCT)
30    CONTINUE
C
      RETURN
      END
```

The vita has been removed from the scanned document

THE EFFECT OF SEGMENT AVERAGING ON THE QUALITY OF THE BURG SPECTRAL ESTIMATOR

by

Md. Anisur Rahman

(ABSTRACT)

The Burg spectral estimator (BSE) exhibits better peak resolution than conventional linear spectral estimators, particularly for short data records. Based on this property, the quality of the BSE is investigated with the available data record segmented and the relevant parameters or functions associated with each segment averaged. Averaging of autoregressive coefficients, reflection coefficients, or spectral density functions is used with the BSE and the corresponding performances are studied. Approximate expressions for the mean and variance of these modified Burg spectral estimators are derived. Lower bounds for the mean and variance of reflection coefficients are also deduced. Finally, the variance of the estimation errors associated with the modified power spectral density estimators is compared against the theoretical Cramer-Rao lower bound.