AN ABSTRACT OF THE THESIS OF

Lei Luo for the degree of Master of Science in Electrical and Computer Engineering Presented on December 6, 1989.

Title: **POWER SPECTRUM ESTIMATION OF SINUSOIDS FROM WHITE AND COLORED NOISES USING THE CANONICAL CORRELATION ANALYSIS METHOD**

Abstract approved : Redacted for privacy

The purpose of this thesis is to apply the Canonical Correlation Analysis (CCA) which belongs to the parametric methods for power spectral estimation in the environments of either white noise or colored noise. It is shown that optimal state space variables belong to the range space of the canonical vectors weighted by the canonical correlation coefficients and can be obtained by using canonical analysis of the past and the future of the observation data. It is also established that this technique produces satisfactory results when applied for ARMA power spectrum estimation. Finally, in order to illustrate the effectiveness of the CCA method, some simulation results are presented.

POWER SPECTRUM ESTIMATION OF SINUSOIDS FROM WHITE AND COLORED NOISES USING THE CANONICAL CORRELATION ANALYSIS METHOD

by

Lei Luo

A THESIS

submitted to

Oregon State University

in partial fulfillment of

the requirements for the

degree of

Master of Science

Completed December 6, 1989

Commencement June 1990

APPROVED:

Redacted for privacy

Assistant Prof. of Electrical & Computer Engineering in charge of major

Redacted for privacy

Chairman of the Department of Electrical & Computer Engineering

Redacted for privacy

Dean of Graduate School

Date thesis is presented December 6, 1989

Typed by Lei Luo for Lei Luo

ACKNOWLEDGEMENT

I gratefully acknowledge to Dr. Sayfe Kiaei, my major professor and thesis advisor, for his overall guidance, many helpful suggestions, encouragement and criticisms throughout my graduate study. His generous advice, assistance and friendship are greatly appreciated.

I would like to express my deepest appreciation and respect to Professor Ronald R. Mohler for his valuable assistance and constructive guidance.

I must also express my sincere thanks and gratitude to Dr. David J. Allstot, Dr. James H. Herzog and Dr. Ernest G. Wolff for being in my graduate committee and giving me a lot of valuable advice.

I thank all the professors who have taught me during my graduate study.

TABLE OF CONTENTS

LIST OF FIGURES

Figures page

POWER SPECTRUM ESTIMATION OF SINUSOIDS FROM WHITE AND COLORED NOISES USING THE CANONICAL CORRELATION ANALYSIS METHOD

1. INTRODUCTION

One of the important applications in digital signal processing techniques is the estimation of the autocovariance and the power spectrum of a random sequence. Power spectral estimation is typically used for distinguishing and tracking signals corrupted by noise. For a majority of modern signal-processing applications, such as radar, sonar and passive arrays, the spectral analysis problem often involves estimating the locations of frequencies or peaks of sinusoidal signals added with noise. Furthermore, in many cases, the spectral analysis is based on short data records and low signal-noise ratio.

Spectral analysis is a method for estimating the spectral density function of the observed data. The spectrum of a signal is defined as the discrete Fourier transform of an infinite autocorrelation sequence with the following relationship:

$$S_y(e^{jw}) = \sum_{n=-\infty}^{\infty} r_y(n)\, e^{-jwn} \tag{1.1}$$

where

$$r_y(k) = E[\, y(i+k)\, y(i)\,] \tag{1.2}$$

is the autocovariance lags of observation $\{y(i)\}$.

is the autocovariance lags of observation $\{y(i)\}$.

The relationship between the power spectral density and the autocorrelation function is considered as a nonparametric description of the second-order statistics of a random process.

Given a stationary time series $\{y(i)\}$ corrupted by noise, the autocorrelation lag sequence of $\{y(i)\}$ is the natural tool for considering the evolution of the process through time. The autocovariance matrix of the observation sequence $\{y(i)\}$ can be estimated as follows under the assumption that $\{y(i)\}$ is ergodic:

$$r_y(k) \approx 1/N \sum_{i=1}^{N-k} y(i+k)\, y(i) \qquad k = 0,1,\ldots N \tag{1.3}$$

where N is the data length of the observation.

The simplest approach for estimating the spectrum is the Fast Fourier transform (FFT) analysis which has played a primary role in much of the earlier effort on spectral estimation. It is essentially a modification of the Fourier analysis suitable for stochastic time series. Using the FFT, the associated power spectral density function is given by

$$S_y(e^{jw}) = 1/N \left| \sum_{n=0}^{N-1} y(n)\, e^{-jwn} \right|^2 \tag{1.4}$$

in which the normalized frequency variable w takes values in $[-\pi, \pi]$ and $S_y(e^{jw})$ is a positive semi-definite symmetric function (if the time series is real). The Fourier transform pair (1.1) and (1.2) show that knowledge of the autocorrelation sequence is equivalent to that of the power spectral density function and vice versa [32]. This method, known as the periodogram method is computationally efficient and produces reasonable results for a large class of signals. Despite these advantages, there exist some limits in the above algorithm. The first problem is that the frequency resolution is roughly the reciprocal of the time interval length of the sampled data. Moreover the estimated spectrum fluctuates rather wildly about the true spectrum such that small signals might be masked by the noise peaks [21,33]. These problems could be overcome by using parametric approaches which are explained in the next section.

1.1 Parametric Methods for Spectrum Estimation

The modern approaches for overcoming the fundamental limits of the periodogram method and achieving high resolution are by extrapolating the observed data as a model. In this case the observations can be modeled as the output of a linear system driven by a white noise sequence w(i). The problem of spectral estimation is then equivalent to estimating the model parameters and obtaining the corresponding spectrum estimates from the model. The degree of improvement in the spectrum resolution is determined by the ability to fit the model with a few parameters to the measured data.

In general, three kinds of linear models which describe the input/output relationship of a system are used. They are the autoregressive (AR) process model, the moving average (MA) process model, and the autoregressive-moving average (ARMA) process model. The following section briefly introduces each model.

(a) The Autoregressive (AR) Models

A spectral model is said to be an autoregressive model of order p (i.e. AR(p)) if its spectrum satisfies the following equation

$$S_{AR}(e^{jw}) = | 1/(1 + a(1)\ e^{-jw} +..+ a(p)\ e^{-jpw} |^2 \qquad (1.1.1)$$

where {a(i)} are parameters of the AR model.

The Autoregressive (AR) model has been used in many applications and there are many efficient algorithms for estimation of the model parameters. These include the Yule-Walker equations [32], the linear prediction (LP) method [22], the Burg's method [18], the Music method [18] and recursive estimation algorithm like Levinsion algorithm [32]. The main disadvantage of the AR spectral estimator is that it is sensitive to the additional observation noise [21].

If a higher model order is selected relative to the number of signals, the AR spectral estimates tend to exhibit spurious peaks which are generated by the additional noise. Lowering the selected model order to prevent spurious peaks reduces the resolution. Thus, this method will result in poor performance at low signal-noise ratio (SNR) values (below 20dB), such that the spectral peaks are broadened and displaced from their true positions. For the heavy noisy environment, the AR model for the spectral analysis is no longer valid. Assume y(k) an AR process with order n corrupted by noise:

$$y(i) = s(i) + n(i)$$

where s(i) is the AR(n) process and n(i) is the observation noise. If n(i) is white with variance $\sigma^2{}_n$, zero mean and is uncorrelated with s(i), then its spectrum in terms of z transform is

$$Sy(z) = \frac{\sigma_w{}^2}{A(z)\ A(z^{-1})} + \sigma_n{}^2 = \frac{\sigma_w{}^2 + \sigma_n{}^2 A(z) A(z^{-1})}{A(z)\ A(z^{-1})} = \frac{\sigma^2 B(z) B(z^{-1})}{A(z)\ A(z^{-1})}$$

From the above, it can be seen that the power spectrum of y(i) is characterized by an ARMA(p,q) model and should be analyzed by the ARMA method.

(b) Autoregressive-Moving Average (ARMA) Models

A spectral model is said to be an autoregressive-Moving Average model of order p and q (i.e. ARMA(p,q)) if its spectrum satisfies the following equation

$$S_{ARMA}(e^{jw}) = \left| \frac{b(0) + b(1)\, e^{-jw} + .. + b(q)\, e^{-qjw}}{1 + a(1)\, e^{-jw} + .. + b(p)\, e^{-pjw}} \right| \qquad (1.1.2)$$

where { a(i)} {b(i)} are AR and MA parameters of the model , respectively.

The autoregressive-moving average model has more degrees of freedom than the autoregressive model [21]. Due to the nonlinear nature of the algorithm for simultaneously estimating the AR and the MA parameters of the ARMA model, iterative techniques based on maximum likelihood estimation are often used to obtain the model parameters. These techniques involve significant computations and are not guaranteed to converge to the true solution.

Because of the above problems, suboptimum techniques have been developed to significantly reduce the computational complexity. These algorithms are usually based on a least square criterion and require solution of linear equations [12-13] [18-25].

1.2 ARMA Power Spectrum Estimation Algorithms

In general, there are several methods for estimating the ARMA parameters. These include the Principal Component (PC) method [22], the Singular Value Decomposition (SVD) method [12], the Canonical Correlation Analysis (CCA) method and so on. The last two approaches will be introduced in the following sections.

(a) The Singular Value Decomposition (SVD) Method

One of the suboptimal ARMA model estimation methods is based on the singular value decomposition (SVD) [12-13]. This method estimates the AR and the MA parameters separately, rather than jointly, as required of optimal parameter estimation. The AR paramethers are estimated first, independently of the MA parameters, by using the modified Yule-Walker equation and then the MA parameters are estimated. This method will be discussed in detail in II.

(b) State Space Realization and The Canonical Correlation Analysis (CCA) Method

State-space representation for linear stochastic systems has been proved very useful in a variety of problems arising in estimation theory, stochastic and deterministic control theory and time series modeling. The state-space representation of an ARMA model, is shown in the following state variable equations:

$$
\begin{aligned}
x(i+1) &= A\,x(i) + B\,u(i) \\
y(i) &= C\,x(i) + D\,u(i)
\end{aligned}
\qquad (1.2.1)
$$

where x(i) is a state vector, u(i) is a white Gaussian process with zero mean and variance of $\sigma^2{}_n$. The poles of an ARMA model for a discrete-time process y(i) are the eigenvalues of A matrix in state-space notation and the zeros are the eigenvalues of (A-BC). The transfer function of this system is:

$$H(z) = C\,(zI-A)^{-1}\,B + D$$

From the above discussion, it is seen that the stochastic realization can be viewed as a problem of studying the relationship between the past and the future. This can be efficiently done by using the theory of canonical variables, which by now is well established in time series analysis. Later on in this section we will introduce the power spectrum estimation algorithm which uses the canonical analysis of time series.

The Canonical Correlation Analysis (CCA) of time series

The criterion of canonical correlation was first proposed in statistics by Hotelling in 1936 [9]. It was proved that variates x(1), x(2).. x(p) and x(p+1).. x(p+q) can be transformed linearly into variates $\alpha(1)$.. $\alpha(p)$ and $\alpha(p+1)$.. $\alpha(p+q)$ so that

(a) the members of each group are independent among themselves.

(b) each member of one group is independent of all but one member of the other.

(c) the non-vanishing correlations between members of different groups which are called **Canonical Correlation Coefficients** are maximized. The quantities $\alpha(1)..\alpha(p+q)$are called **canonical variates.**

The CCA has been used for analyzing the stochastic state-space modeling problem by Akaike [4-8], White [11], Desai [26-28], Kung [24-25] and Yohai [29]. It has been shown that using the canonical correlation analysis in state-space (ARMA) model estimation yields significantly better results and improved resolution for low signal-noise ratio (SNR) [17]. The main advantages of this method are that it estimates the optimal AR and MA parameters simultaneously and is relatively stable to perturbations.

For most of the modern spectral estimation approaches mentioned above, the assumption is that the additive observation noise is spatially white (stationary, uncorrelated with zero mean). However, in many practical situations, this fundermental assumption is not satisfied. In chapter IV, the CCA method is also applied to the spectral estimation in an additive colored noise environment where the noise could be an AR, MA, ARMA or a purely non-deterministic random process. In the previous studies [35-39], it was considered that the covarniance matrix of the colored noise is known priori as either an AR model or an MA model. For the AR color noise case one needs to solve non-linear equations [36] which requires many computations. It can be seen that the algorithm suggested here utilizes the CCA method for colored noise case and is relatively more efficent than the methods mentioned above.

1.3 Outline of the Thesis

Following the introduction, chapter II will review the Yule-Walk equations and the SVD estimation method. The performance of the SVD method for power spectrum estimation of multiple sinusoids in white noise will be studied.

Chapter III describes the CCA algorithm for stochastic state-space realization and the method will be used for power spectral estimation. The performance of the CCA technique is compared with that of the SVD method by simulation.

Chapter IV introduces a modefied CCA algorithm developed for power spectral estimation with unknown colored noise environment. The performance and its advantages will be discussed and compared with those of the methods in [34-35]. The simulation results and the algorithm performances are described in each section and in the conclusion.

2. ARMA POWER SPECTRAL ESTIMATION FOR WHITE NOISE

The objective for developing parametric models for spectral estimation is to achieve higher resolution spectrum density estimators. From the output observations, some knowledge about the process from which the data samples are taken is often available. This information may be used to construct a model that approximates the discrete-time proecsses encountered by the following autoregressive moving average (ARMA) model with AR order p and MA order q

$$u(i) \longrightarrow \boxed{H(z) = \frac{B(z)}{A(z)}} \longrightarrow y(i)$$

with

$$A(z) = 1 + a(1)\, z^{-1} + a(2)\, z^{-2} + \,..\, + a(p)\, z^{-p}$$

$$B(z) = b(0) + b(1)\, z^{-1} + b(2)\, z^{-2} + \,..\, + b(q)\, z^{-q}$$

and

$$y(i) = - \sum_{k=1}^{p} a(k)\, y(i-k) + \sum_{k=0}^{q} b(k)\, u(i-k) \qquad (2.1)$$

where y(i) is the output and u(i) is the white noise input. The autocorrelation sequence of {y(i)} can be calculated using (1.3) and the model parameters are determined usually based on the Yule-Walker equations.

2.1 The Yule-Walker Equations

Given the stationary process (2.1), multiplying both sides of equation (2.1) by y(i-k) and taking the expectations, we find

$$E[y(i)y(i-k)] = -\sum_{m=1}^{p} a(k)\ r_y(k-m) + \sum_{m=0}^{q} b(k)\ r_{uy}(k-m) \tag{2.1.1}$$

the relationship between the ARMA parameters and the autocorrelation of process y(i) is given by:

$$r_y(k) = \begin{cases} r_y(-k) & k < 0 \\ -\sum_{m=1}^{p} a(m)\ r_y(k-m) + \sigma_w^2 \sum_{m=k}^{q} b(m)\ r_{uy}(k-m) & 0 < k < q \\ -\sum_{m=1}^{p} a(m)\ r_y(k-m) & k > q \end{cases} \tag{2.1.2}$$

In the above expressions, E denotes expected value operation. $r_y(k)$ is the autocorrelation of y and $r_{uy}(k)$ is the cross correlation between the input u(i) and the output y(i). When k > q, the autoregressive paremeters of an ARMA model are related by a set of linear equations to the autocorrelation sequence.

$$\begin{bmatrix} r_y(q) & r_y(q+1) & \cdots & r_y(q-p+1) \\ r_y(q+1) & r_y(q+2) & \cdots & r_y(q-p+2) \\ \vdots & & & \\ r_y(q+p-1) & r_y(q+p-2) & \cdots & r_y(q) \end{bmatrix} \begin{bmatrix} a(1) \\ a(2) \\ \vdots \\ a(p) \end{bmatrix} = - \begin{bmatrix} r_y(q+1) \\ r_y(q+2) \\ \vdots \\ r_y(q+p) \end{bmatrix}$$

or

$$R\,\mathbf{a} = \mathbf{r} \tag{2.1.3}$$

If the autocorrelation sequence for lags q-p+1 to q+p is given, then the autoregressive parameters can be found separately from the moving average parameters as the solution to the simultaneous equations (2.1.3). The relationship of (2.1.3) is known as the modified Yule-Walker equations.

Assume that the signal is composed of h sinusoids

$$y(i) = \sum_{k=1}^{h} A_k \sin(w_k\, i + \theta_k) + w(i) \tag{2.1.4}$$

where w(i) is the observation noise, A_k, w_k and θ_k (k=1..h) are signal magnitudes, signal frequencies and initial phases, respectively. This signal can be thought of as a limiting form AR process with narrow-band peaks at the h sinusoidal frequency locations. The frequencies can be estimated by solving the roots of the transfer function A(z) of AR(n) (n = 2 h)

$$A(z) = 1 + a(1)\, z^{-1} + a(2)\, z^{-2} + \,..\, + a(n)\, z^{-n} = 0$$

that are on the unit circle located at the positions $\exp(\pm jw_k)$ (k=1..n). The singular value decomposition (SVD) method, which is the topic of the following section, is based on the modified Yule-Walker equations.

2.2 The Singular Value Decomposition Method

The singular value decomposition (SVD) method is based on singular value analysis of an autocorrelation matrix. It has been shown that singular value decomposition provides accurate estimates of frequencies for a process consisting of multiple sinusoids in noise [2-3,16].

From the calculation of the singular value decomposition of the autocorrelation matrix R in (2.1.3), it is seen that the effect of noise is to introduce noise eigenvectors with small eigenvalues. These nonprincipal eigenvectors exhibit large variances thus cannot be estimated reliably. In addition, These noisy eigenvectors bias the autocorrelation matrix estimate by adding components that are not present in the noiseless case.

Hence, in order to immunize the influence of the noise, it is necessary to separate the information in R or in the data matrix into two subspaces; signal subspace and noise subspace. Using this decomposition the principal eigenvectors which span the signal subspace of the autocorrelation matrix can be seperated from the non-principal noise eigenvalues. The non-noise poles will be on the unit circle at the sinusoidal frequency locations $\exp(j2\pi f_k)$ (k = 1..h, where h is the number of sinusoidal signals). The rest of the noise poles will be approximately uniformly displaced inward [16].

This technique introduced in the analysis of the autocorrelation matrix has been discussed in [14,16,22].

Although the ARMA model is seen to have a good frequency characteristic, most of the existing ARMA parameter estimation techniques usually involve complicated matrix computations and/or iterative optimization techniques. The SVD method, however, calculates the a(k) and b(k) parameters seperately. This algorithm proposed by Cadzow in 1982 has been used to analyze a larger dimension autocorrelation matrix estimated for ARMA modeling. The algorithm is as follows.

The SVD Algorithm

Given stationary zero mean observation sequence y(1), y(2)..y(N),

Step 1. Calculate the sample autocorrelation matrix

$$Re = \begin{bmatrix} r_y(Qe+1) & r_y(Qe) & \cdots & r_y(Qe-Pe+t) \\ r_y(Qe+2) & r_y(Qe+1) & \cdots & r_y(Qe-Pe+t+1) \\ \vdots & \vdots & \cdots & \vdots \\ r_y(Qe+t) & r_y(Qe+t+1) & \cdots & r_y(Qe-Pe+2t) \end{bmatrix} \qquad (2.2.1)$$

where Pe >> p, Qe >> q

It has been shown in [2] that those extra autocorrelation lags used will provide the flexibility needed to reduce the sensitivity of system poles to perturbation and numerical ill-condition in power spectrum density calculation.

The autocovariance lags $r_y(i)$ are estimated using

$$r_y(k) = 1/N \sum_{i=1}^{N-k} y(i+k)\, y(i) \qquad k = 0,1,..\ N \tag{2.2.2}$$

Step 2. Compute SVD of Re: Re = Ue Σe Ve′,

Let

$$Rp = Ue\ \Sigma n\ Ve'$$

be rank n approximation of Re

$$Rp = [\ \mathbf{r}\ Ra\ ] \tag{2.2.3}$$

where $\mathbf{r}$ is the left most tx1 column vector of Rp and Ra is a t by Pe matrix composed of the Pe right most t by 1 column vectors of Rp

$$Rp\ \mathbf{a} = 0 \qquad \mathbf{a} = [\ 1\ a(1)\ a(2)\ ..\ a(n)\ ]' \tag{2.2.4}$$

$$\begin{bmatrix} a(1) \\ a(2) \\ . \\ . \\ a(n) \end{bmatrix} = -\ [Ra]^{+}\ \mathbf{r} \tag{2.2.5}$$

where " + " denotes pseudo-inverse operation

Step 3. Obtain $c(i) = \sum_{j=1}^{i} a(j)\, r_y(i-j)$ (2.2.6)

Step 4. Compute the power spectrum

$$D(e^{jw}) = \frac{c(1)\, e^{-jw} + c(2)\, e^{-j2w} + .. + c(n)\, e^{-jnw}}{1 + a(1)\, e^{-jw} + a(2)\, e^{-j2w} + .. + a(n)\, e^{-jnw}} \tag{2.2.7}$$

$$Sy(e^{jw}) = r_y(0) + 2\, D(e^{jw}) \tag{2.2.8}$$

However, a straightforward implementation of the Cadzow's method sometimes will lead to negative values of the power spectral estimates which is clearly unreasonable. The reason for this is that in practice, the autocorrelation function of y(i) must be estimated and the use of an autocorrelation estimate (even a positive-semidefinite estimate) in (2.1.6) may result in negative spectral estimates. Moreover, because of the non-symmetry of the Toeplitz matrix in (2.1.1), the autocorrelation matrix Re is not guaranteed to be either positive-definite or nonsingular. It has been proved by S.M. Kay in [20] that the lack of a nonsingularity constraint on the autocorrelation matrix estimate may lead to an increase in the variance of the AR parameter estimate.

2.3 The SVD Algorithm Simulation and Analysis

The simulation example used is the linear summation of two sinusoids in additive noise.

$$y(i) = A \sin(0.4\pi i) + A \sin(0.43\pi i) + n(i) \qquad i = 1 \ldots N$$

where n(i) is observation white noise with $\sigma^2_n = 1$. Using the SVD approach with a reduced-order approximation, the simulation results are shown in Fig. 1 to Fig. 6 for 64 data points and different signal-noise ratio (SNR) values. Five independent realizations were used in each case. The signal-noise ratios used were 0dB, 3dB, and 7dB, respectively. Model orders selected in the simulation were either n = 4 or n = 5 and AR order p = MA order q = n. The dimension of the autocorrelation matrix Re in (2.2.1) is t x Pe (with Qe = Pe, where Pe and Qe are extended orders which are assumed to be much larger than the true orders of the model p and q). Different matrix dimensions t and Pe were utilized in the simulation. In the pole plots, '+' indicates the true pole locations and '*' shows the estimated pole locations.

Fig. 1-2 are the estimated power spectrum and the poles of the model using the SVD method for p = q = n = 5, t = 45, Pe = Qe = 15, SNR = 0 dB. It can be seen that the power spectrum has some negative values and since signal-noise ratio is relatively small such that only one of the signal poles is detected.

Fig. 3-4 are power spectrum and the system poles estimated with p = q = n = 4, t = 40, Pe = Qe = 15, SNR = 3 dB. The fluctuation in the Spectrum reduces a little because of increase of SNR but still has negative peaks. From Fig.4, it can be seen that the estimated poles start to resolve the two signals.

Fig 5-6 show the spectrum and the system poles using p = q = n = 5, t = 40, Pe = Qe = 15, SNR = 7 dB. Again, a negative peak is shown at the signal frequency f = 0.2 Hz which should be a positive peak although SNR is increased to 7 dB. However, the signals can be detected correctly from the estimated poles in Fig 6.

It is shown that for the above three cases, the estimated power spectrum using the SVD method has negative values. Thus the estimation leads to invalid results. The frequencies of the two signals can be detected correctly only from the poles of the model not from the power spectrum when SNR are larger than 3dB. When the SNR is small, the poles may be out of the unit circle, which leads to an unstable system. Reasonable results can be obtained for larger autocorrelation matrix dimension (i.e. longer data length) and relatively higher signal-noise ratios.

Fig. 1 Power Spectrum (SVD), f1=0.2Hz, f2=0.215Hz, N=64, t=45
Pe=Qe=15, n=5, SNR=0dB (white noise case)

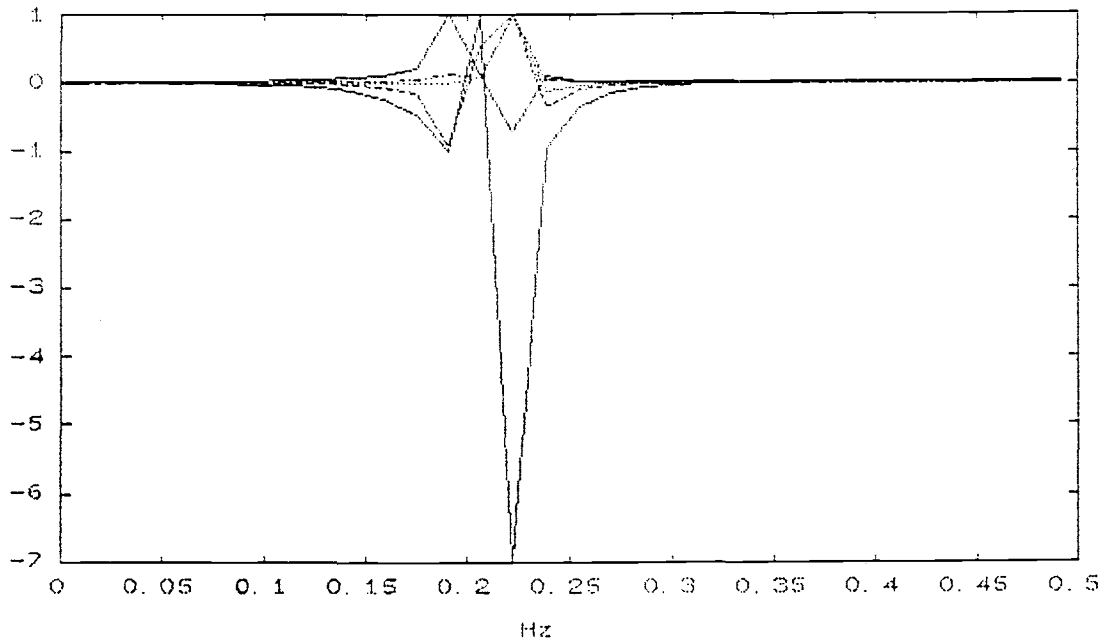


Fig. 2 Pole Plot (SVD), f1=0.2Hz, f2=0.215Hz, N=64 t=45 Pe=Qe=15
n=5, SNR=0dB (white noise case)

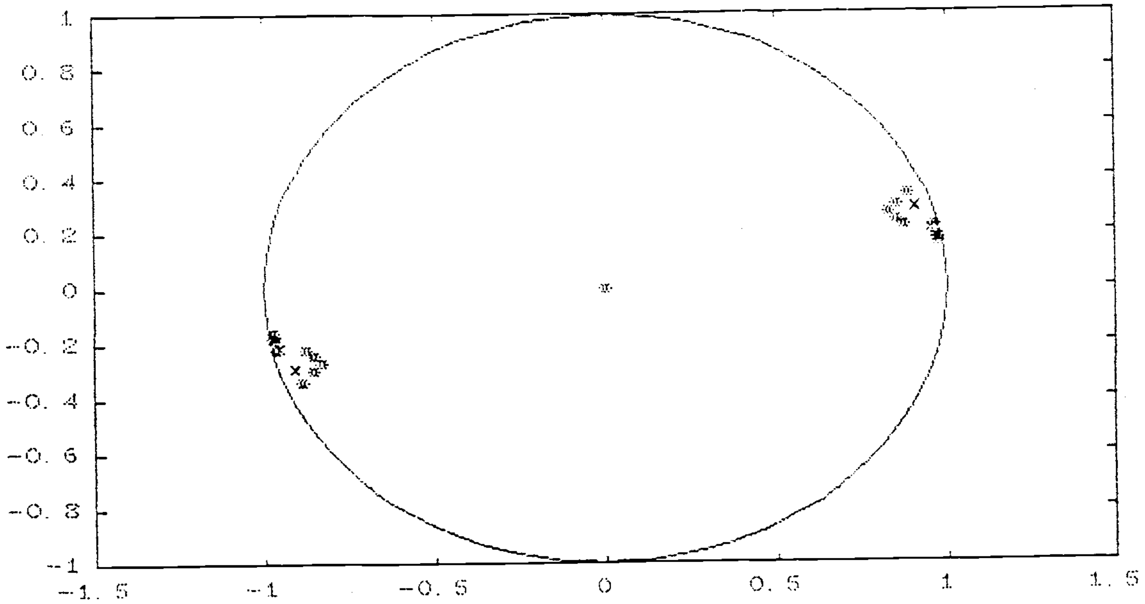


Fig. 3 Power Spectrum (SVD), f1=0.2Hz, f2=0.215Hz, N=64, t=45 Pe=Qe=15, n=4, SNR=3dB (white noise case)

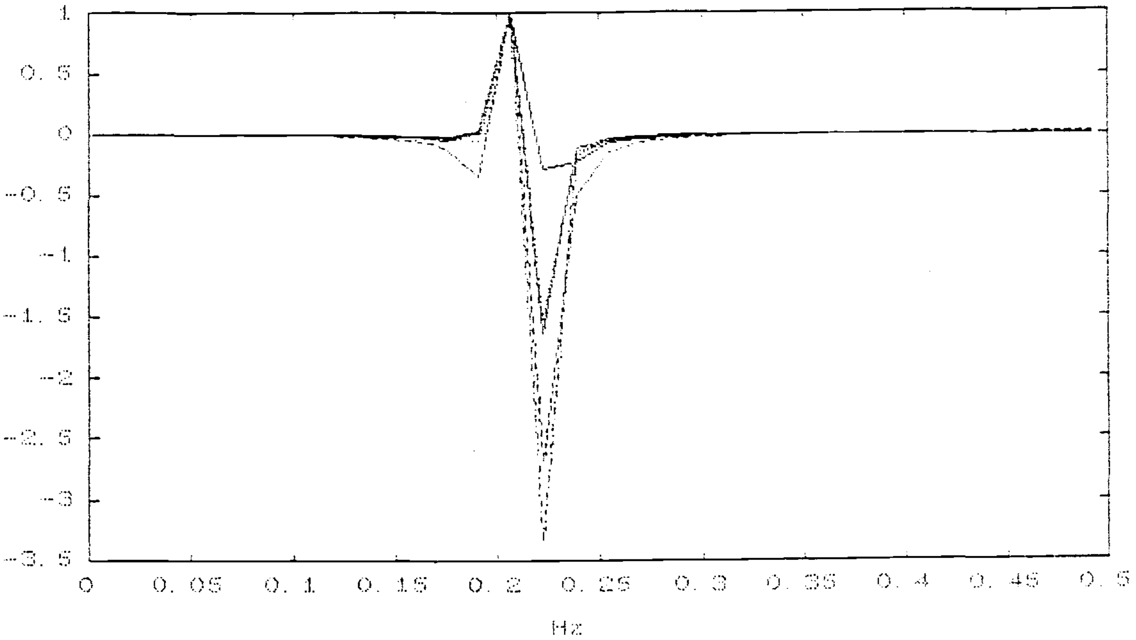


Fig. 4 Pole Plot (SVD), f1=0.2Hz, f2=0.215Hz, N=64 t=45 Pe=Qe=15 n=4, SNR=3dB (white noise case)

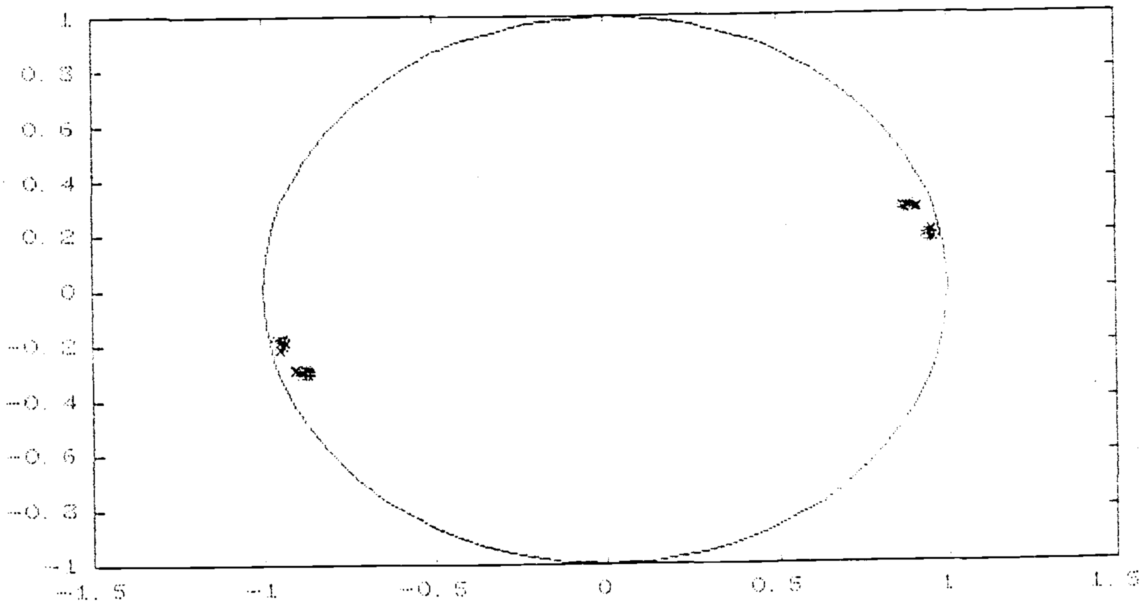


Fig. 5 Power Spectrum (SVD), f1=0.2Hz, f2=0.215Hz, N=64, t=40 Pe=Qe=15, n=5, SNR=7dB (white noise case)

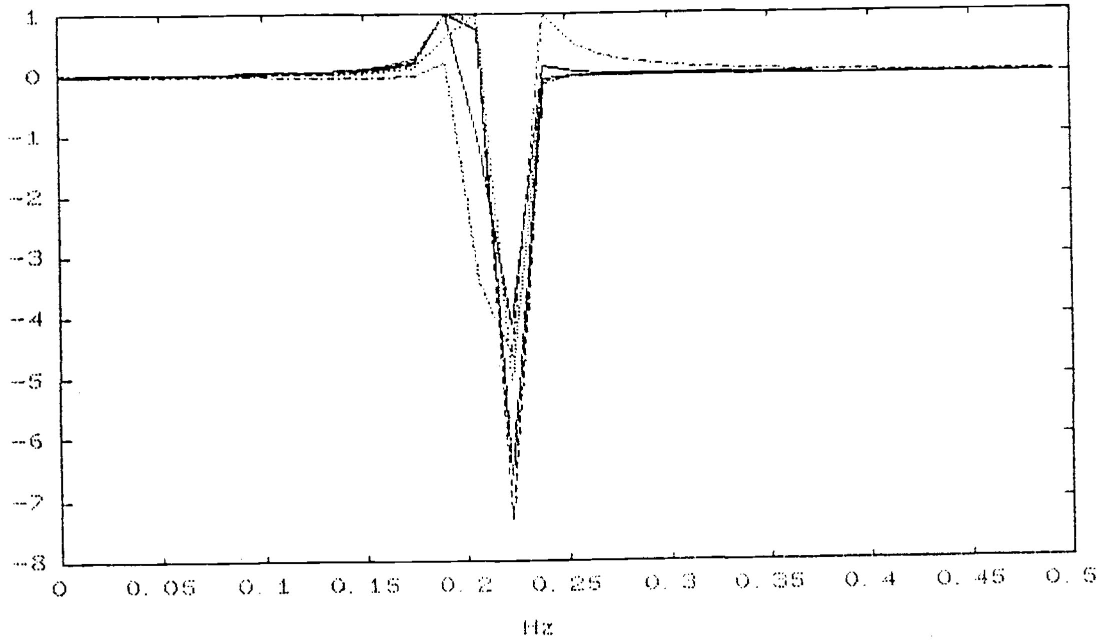
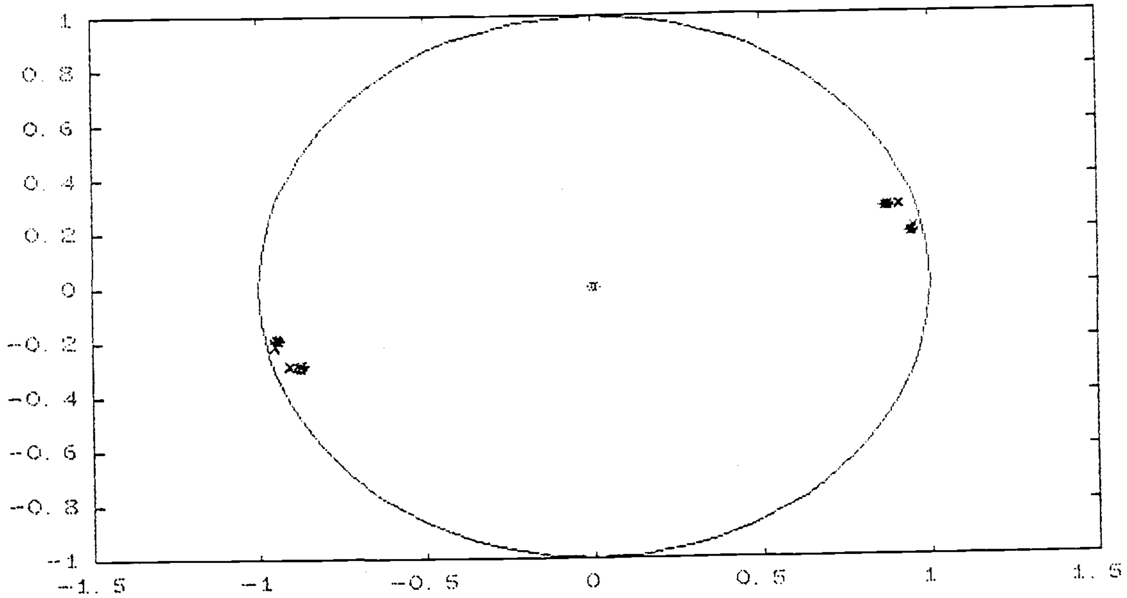


Fig. 6 Pole Plot (SVD), f1=0.2Hz, f2=0.215Hz, N=64 t=40 Pe=Qe=15 n=5, SNR=7dB (white noise case)



3. STOCHASTIC REALIZATION AND THE CCA METHOD

This chapter deals with the stochastic realization approach for spectrum estimation. Stochastic realization problem can be defined as: Given the covariance sequence for a stationary Gaussian stochastic process, determine a state space model which is driven by a white noise sequence. This is a procedure that obtains the relationship between the present and past input or observation and the future output, which is called predictor space.

The state space model of the process is in the following form:

$$\begin{aligned} x(i+1) &= A\,x(i) + B\,u(i) \\ y(i) &= C\,x(i) + D\,v(i) \end{aligned} \tag{3.1}$$

where $u(i)$ is a zero mean white Gaussian process, $x(i)$ is the $n \times 1$ state vector and A, B, C and D are constant matrices of sizes $n \times n$, $n \times 1$ and $1 \times n$, respectively. The poles of the ARMA model are the eigenvalues of A, whereas the zeros are the eigenvalues of $(A - BC)$. Generally the structure of matrices A, B, C and D in the Markovian state-space representation is not unique. The problem of finding the optimal state variables and parameter matrices can be resolved by using the canonical correlation analysis [27].

3.1 Canonical Correlation Analysis

Canonical Correlation Analysis was proposed by Hotelling (1936) [9] and later was introduced into system identification by Akaik [4-8] and developed by Larimore [30], White [11] and U.B. Desai [26-28]. It is useful for studying relations between two time series **a** = [a(1).. a(n)]′ and **b** = [b(1) .. b(m)]′. Here **a** and b are stationary Gaussian time series with zero mean and covariance

$$R_a = E[\, \mathbf{a}\, \mathbf{a}'] \quad \text{and} \quad R_b = E[\, \mathbf{b}\, \mathbf{b}'] \tag{3.1.1}$$

and cross covariance

$$R_{ab} = E[\, \mathbf{a}\, \mathbf{b}'] \tag{3.1.2}$$

Define

$$\epsilon_a = R_a^{-\frac{1}{2}}\, \mathbf{a} \qquad \epsilon_b = R_b^{-\frac{1}{2}}\, \mathbf{b} \tag{3.1.3}$$

then we have

$$E[\, \epsilon_a\, \epsilon_a'] = I_n, \quad E[\, \epsilon_b\, \epsilon_b'] = I_m$$

where I_n and I_m are identity matrices with dimensions n and m.

Then calculate the singular value factorization of the normalized cross covariance

$$E[\, \epsilon_a\, \epsilon_b'] = U\, \Sigma\, V' \tag{3.1.4}$$

with $U\,U' = I$, $V\,V' = I$ and $\Sigma = \mathrm{diag}(\sigma_1, \sigma_2 \,..\, \sigma_q)$, $\sigma_1 \geq \sigma_2 \geq .. \geq 0$

then the canonical variables α and β are defined as

$$\alpha = [\alpha_1, \alpha_2 .. \alpha_n] = U'\epsilon_a \ , \ \beta = [\beta_1, \beta_2 .. \beta_m] = V'\epsilon_b \tag{3.1.5}$$

where $E[\alpha\, \alpha'] = E[\beta\, \beta'] = I$ and $E[\alpha\, \beta'] = \Sigma$.

The columns of U and V are called the canonical vectors and the diagonal entries of Σ are called the canonical correlation diagonal entries of Σ are called the canonical correlation coefficients. The canonical variables $\boldsymbol{\alpha}$ ($\boldsymbol{\beta}$) which is a orthonormal set of random variables reflects all the information about **b** (**a**) that is present in **a** (**b**). Also the angles between the subspaces spanned by **a** and **b** can be provided using the canonical correlations σ_i's defined by

$$\cos^{-1} \sigma_i = (\alpha_i , \beta_i) \tag{3.1.6}$$

3.2 Canonical Correlation Analysis for State-Space Modeling

There has been an increasing attention on applications of the this technique for system identification, power spectrum estimation, balanced state-space parameterization of the ARMA models etc. [2,14,24-25,29-31]. Unlike Gaussian maximum likelihood and one step linear prediction techniques for state-space modeling, the CCA algorithm is based on a form of least-square multi-step linear prediction.

Let {y(i)} be a zero mean stationary process generated by the state-space model

$$\begin{aligned} x(i+1) &= A\, x(i) + B\, u(i) \\ y(i) &= C\, x(i) + D\, v(i) \end{aligned} \tag{3.2.1}$$

where u(i) is the input, v(i) is the observation noise. Both u(i) and v(i) are zero mean white noises uncorrelated with state variables x(i) such that

$$E[\ x(i)\ u(i)'] = 0, \qquad E[\ x(i)\ u(i)'] = 0$$

Intuitively, the state of a minimal order system is a summary of the information in the past output history that is both necessary and sufficient to predict the future output. The primary objective of the state-space modeling is to find the linear dependence between a local past $Y_{\_}(i)$ with lenth p and a local future $Y_{+}(i)$ with length f of observation y(i) for i = 1,2...N to estimate the parameter matrices and the model order n.

In this approach, first the output data is seperated into two sequences, the past

$$Y_{\_}(i) = [\ y(i-1)\ y(i-2)\ \ldots\ y(i-p)\ ]$$

and the future

$$Y_{+}(i) = [\ y(i)\ y(i+1)\ \ldots\ y(i+f-1)\ ]$$

To solve the approximation problem, we replace the time series a and **b** by the future Y_{+} and the past $Y_{\_}$. The autocovariance of the time series

$$R_y(k) = E[\ y(i+k)\ y(i)']$$

is calculated using (1.3). Thus we have

$$R_{+} = E[\ Y_{+}\ Y_{+}'] \ , \quad R_{\_} = E[\ Y_{\_}\ Y_{\_}'] \qquad (3.2.2)$$

The matrices R_{+} and $R_{\_}$ are block Toeplitz matrices, and when y(i) is real and p = f, $R_{+} = R_{\_}$.

The cross covariance between Y_+ and Y_- is

$$R_{\pm} = E[\ Y_+\ Y_-'] = \begin{bmatrix} R_y(1) & R_y(2) & \dots & R_y(f) \\ R_y(2) & R_y(3) & \dots & R_y(f+1) \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ R_y(f) & R_y(f+1) & \dots & R_y(2f) \end{bmatrix} \tag{3.2.3}$$

Following the procedure in the above section, we define

$$\epsilon_+ = R_+^{-\frac{1}{2}}\ Y_+\ , \qquad \epsilon_- = R_-^{-\frac{1}{2}}\ Y_- \tag{3.2.4}$$

with

$$E[\ \epsilon_+\ \epsilon_+'] = I,\ E[\ \epsilon_-\ \epsilon_-'] = I$$

Compute

$$H = E[\ \epsilon_+\ \epsilon_-'] = R_+^{-\frac{1}{2}}\ R_{\pm}\ (R_-^{-\frac{1}{2}})' = U\ \Sigma\ V' \tag{3.2.5}$$

The columns of U and V are called canonical vectors and canonical correlations are the diagonal entries of

$$\Sigma = \text{diag}(\ \sigma_1,\ \sigma_2,\ \ ..\ ,\sigma_f)$$

which are ordered as

$$1 \geq \sigma_1 \geq \sigma_2 \geq\ ..\ \geq \sigma_f > 0$$

When the sequence $\{y(i)\}$ is stationary, its autocovariance $R_y(k)$ admits the factorization

$$R_y(k) = \begin{cases} C\ P\ C' + D\ D' & k = 0 \\ C\ A^{k-1}\ G & k > 0 \end{cases} \tag{3.2.6}$$

where

$$G = A\ P\ C' + B\ D' \tag{3.2.7}$$

and the P is the nxn state covariance matrix satisfying the Lyapunov equation

$$P = E[\ x(i)\ x(i)'] = A\ P\ A' + C\ C' \tag{3.2.8}$$

From (3.2.3) and (3.2.6), the block Hankel matrix

$R_{\pm} = R_{+}^{\frac{1}{2}} \, U \, \Sigma \, V' \, R_{-}^{\frac{1}{2}}$

can be expressed as

$$R_{\pm} = \begin{bmatrix} CG & CAG & \cdots \\ CAG & CA^2G & \cdots \\ CA^2G & & \cdots \\ \cdot & \cdot & \\ \cdot & \cdot & \end{bmatrix} = \begin{bmatrix} C \\ CA \\ CA^2 \\ \cdot \\ \cdot \end{bmatrix} \begin{bmatrix} G & AG & A^2G & \cdots \end{bmatrix} = \mathbf{O} \, \mathbf{C} \qquad (3.2.9)$$

where $\mathbf{O}$ is called observability matrix having rank f and $\mathbf{C}$ is called controllability matrix having rank f. The observability matrix $\mathbf{O}$ maps the state vector into the future output.

The canonical variables $\alpha(i)$ and $\beta(i)$ are

$$\alpha = V' \, R_{-}^{-\frac{1}{2}} \, Y_{-}, \quad \beta = U' \, R_{+}^{-\frac{1}{2}} \, Y_{+} \qquad (3.2.10)$$

The canonical variables have the following characteristics:

(i) $E[\,\alpha\,\alpha'\,] = E[\,\beta\,\beta'\,] = I, \qquad E[\,\alpha\,\beta'\,] = \Sigma$

The sub-space $E[Y_{+}|Y_{-}]$ that projects the future Y_{+} onto the past Y_{-}, i.e. the optimal prediction space in the least-square sense is

$$E[Y+|Y-] = R_{\pm} \, R_{-}^{-1} \, Y_{-} = \mathbf{O} \, \mathbf{C} \, R_{-} \, Y_{-} \qquad (3.2.11)$$

This sub-space is spanned by the orthonormal basis $\{\alpha(i)\}$.

(ii) The canonical correlation coefficients are the principal angles between sub-spaces of the future Y_{+} and the past Y_{-} given by

$$\theta_i = \cos^{-1}\sigma_i \qquad (\; i = 1, \; 2 \; \ldots \; f \;)$$

These angles provide us with a useful tool to characterize the difference between two sub-spaces spanned by Y_+ and Y_-. Thus it seems that a natural choice for the components of the partial state is the canonical components of Y_- that have the largest canonical correlation coefficients with respect to Y_+. The optimal state space vector is then defined as

$$x(i) = \Sigma^{\frac{1}{2}} \alpha(i) = \mathbf{C}\, R_-^{-1}\, Y_-(i) \qquad (3.2.12)$$

Hence the state vector x(i) that describes the given data is the canonical vector {α(i)} weighted by canonical correlations σ(i)'s. These states contain all the information available from the entire past $Y_-(i)$ about the entire future Y_+. The state vector in (2.3.11) is expressed as a linear combination of the local past and pridicted future is the linear combination of states

$$Y_+(i) = \mathbf{O}\, x(i)$$

(iii) The entire relationship between $Y_+(i)$ and $Y_-(i)$ is expressed in terms of only f parameters σ_1, σ_2 ... σ_f. α(i) containing all the information about Y_+ that is present in Y_-. The mutual information between Y_+ and Y_- can be expressed in terms of

$$I(Y_+,Y_-) \approx -\tfrac{1}{2} \sum_{i=1}^{f} \log(1 - \sigma_i{}^2) \approx -\tfrac{1}{2} \sum_{i=1}^{n} \log(1 - \sigma^2{}_i)$$

$$\sigma_{n+i} \ll \sigma_i \quad (i=1..\ f\text{-}n) \qquad (3.2.13)$$

Usually, the Hankel matrix formed from the noisy covariance sequence will have full rank. However, the singular values of H that ideally should be zero will be much smaller than the principal singular values. From (3.2.13), the reduced order model is given by deleting those very small σ_i's when $\sigma_{n+i} \ll \sigma_n < 1$. $(i = 1, 2, \ldots, f-n)$. Therefore, the reduced order state variables x(i) become

$$x(i) = \Sigma_n^{\frac{1}{2}} \alpha(i) \tag{3.2.14}$$

where Σ_n is the nth order approximation of Σ which is obtained by deleting those small singular values $\sigma_{n+1}, \sigma_{n+2} \ldots \sigma_f$.

3.3 Identification Procedure

The steps of stochastic realization using CCA method are listed as follows:

1. Calculate the autocorrelation matrices

$$R_- = E[Y_-Y_-'] = \begin{bmatrix} R_y(0) & R_y(1) & \ldots & R_y(f-1) \\ R_y(1) & R_y(0) & \ldots & R_y(f-2) \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ R_y(f-1) & R_y(f-2) & & R_y(0) \end{bmatrix} = R_+ = E[Y_+Y_+']$$

$$R_\pm = E[Y_+Y_+'] = \begin{bmatrix} R_y(0) & R_y(1) & \ldots & R_y(f-1) \\ R_y(1) & R_y(2) & \ldots & R_y(f) \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ R_y(f-1) & R_y(f) & \ldots & R_y(2f-2) \end{bmatrix}$$

where $r_y(i)$ is computed using (1.1).

2. Calculate SVD of

$$H = R_+^{-\frac{1}{2}} R_\pm (R_-^{-\frac{1}{2}})' = U \Sigma V' \tag{3.3.1}$$

3. Order n approximates of

$$H = U_n \Sigma_n V_n' \tag{3.3.2}$$

assoiated with largest n singular values.

where Σ_n is a diagonal matrix of $\sigma_1, \sigma_2 \ldots \sigma_n$ and U_n, V_n corespond to the first n columns of U and V.

4. Identify the system parameter matrices

$$\mathbf{O} = R_+^{\frac{1}{2}} U_n \Sigma_n^{\frac{1}{2}}, \quad C = \text{first p rows of } \mathbf{O} \tag{3.3.3}$$

$$\mathbf{C} = \Sigma_n^{\frac{1}{2}} V_n' R_-^{\frac{1}{2}}, \quad G = \text{first p columns of } \mathbf{C} \tag{3.3.4}$$

$$A = \mathbf{O}^+ R^{\sim}_{\pm} \mathbf{C}^+ = \mathbf{O}^+ \mathbf{O}^{\sim} = \mathbf{C}^{\sim} \mathbf{C}^+ \tag{3.3.5}$$

$$P = E[\, x(k)\, x(k)'] = \Sigma \tag{3.3.6}$$

From (3.2.6) and (3.2.7), we have

$$D = [\, R_y(0) - C\, \Sigma\, C' \,]^{\frac{1}{2}} \tag{3.3.7}$$

$$B = A\, P\, C' + B\, D' \tag{3.3.8}$$

$$\sigma^2_n = D^{-1} [\, R_y(0) - H\, \Sigma\, H']\, D'^{-1} \tag{3.3.9}$$

where $\mathbf{C}^+$, $\mathbf{O}^+$ are pseudo-inverse of $\mathbf{C}$ and $\mathbf{O}$

$R^{\sim}_{\pm}$ is $R_{\pm}$ left shift p rows

$\mathbf{O}^{\sim}$ is $\mathbf{O}$ left shift p rows

$\mathbf{C}^{\sim}$ is $\mathbf{C}$ left shift p rows

5. Compute the power spectrum

$$P(e^{jw}) = |\, C\, [\, e^{jw} I - A\,]^{-1} B + D\, |^2 \sigma^2_n \tag{3.3.10}$$

3.4 Simulation and Results

In this section, we present some simulation results shown in

Fig. 7 to Fig. 12 using the same group of data with length of 64 as those for the SVD method to illustrate the application of the CCA method. The dimensions of the matrices R_+, R_- and $R_\pm$ selected were f = 25, 20, and 15 respectively for independent five realizations.

Fig. 7-8 are power spectrum and the pole plot estimated using the CCA method with autocorrelation matrix R_+ and R_- and $R_\pm$ dimension p = f = 25, model order n = 5, SNR = 0 dB. The signals can be estimated correctly from both power spectrum and estimated poles even though the signal-noise ratio = 0 dB is quite low.

Fig. 9-10 show power spectrum and the pole plot for p = f = 20 and system order n = 4, SNR = 3 dB. From the power spectrum in Fig.9, it is seen that the CCA method gives satisfactory results with two sharp peaks at the locations of the signal and the poles in Fig.10 converge to the true signal poles.

Fig.11-12 depict power spectrum and the pole plot estimated when autocorrelation matrix dimension was decreased to p = f = 15, model order n = 5, SNR = 7 dB. The CCA method is able to detect the two frequencies well with relatively small matrix dimension f.

When the covariance lags are known exactly, the poles and zeros of the state model are identified exactly by the CCA technique. In the presence of white noise, the stochastic process can be modelled by a reduced order approximation model that has the same poles as those of the original ARMA model. Compare with the results of the SVD method, it is apparent that the CCA method can distinguish the two sinusoids well from both the power spectrums and the pole plots for all the three cases with smaller dimension autocorrelation matrices and has better performance than that ofthe SVD method. Unlike the results from the SVD approach, the estimated power spectrums in Fig 7, 9 and 11 are guaranteed to be positive. In addition, all the poles fall inside the unit circle such that the system is always stable. It can be seen that the canonical variate approach provides a powerful and general procedure for reduced-order modeling, filtering and system identification. The procedure is also numerically accurate and stable.

Fig. 7 Power Spectrum (CCA), f1=0.2Hz, f2=0.215Hz, N=64, f=25 n=5, SNR=0dB (white noise case)

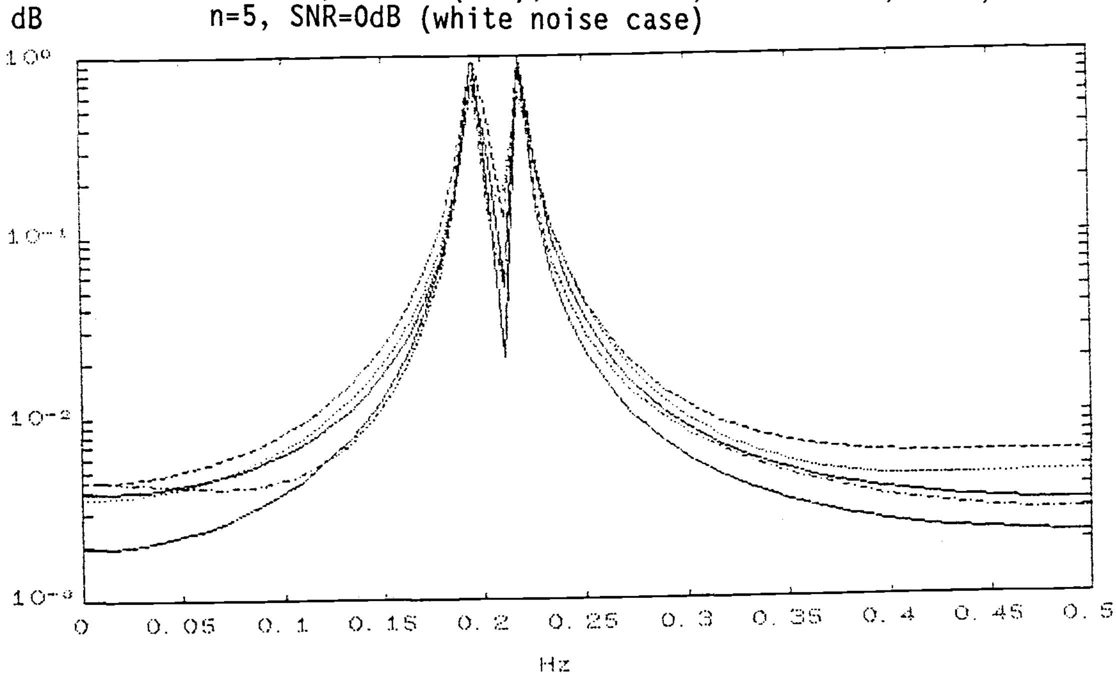


Fig. 8 Pole Plot (CCA), f1=0.2Hz, f2=0.215Hz, N=64, f=25, n=5 SNR=0dB (white noise case)

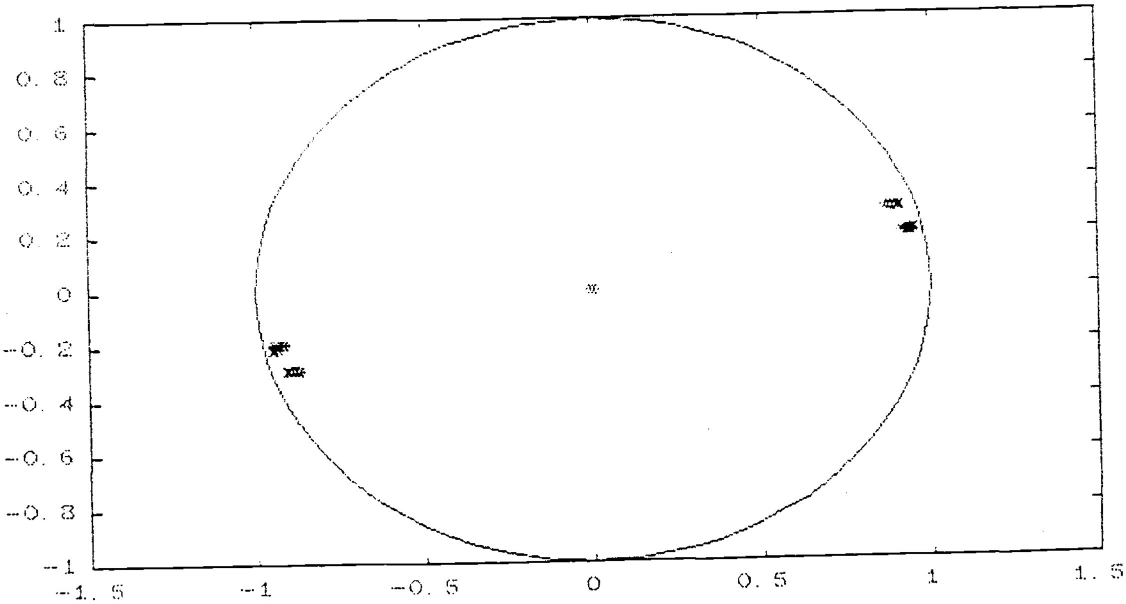


Fig. 9 Power Spectrum (CCA), f1=0.2Hz, f2=0.215Hz, N=64, f=20
n=4, SNR=3dB (white noise case)

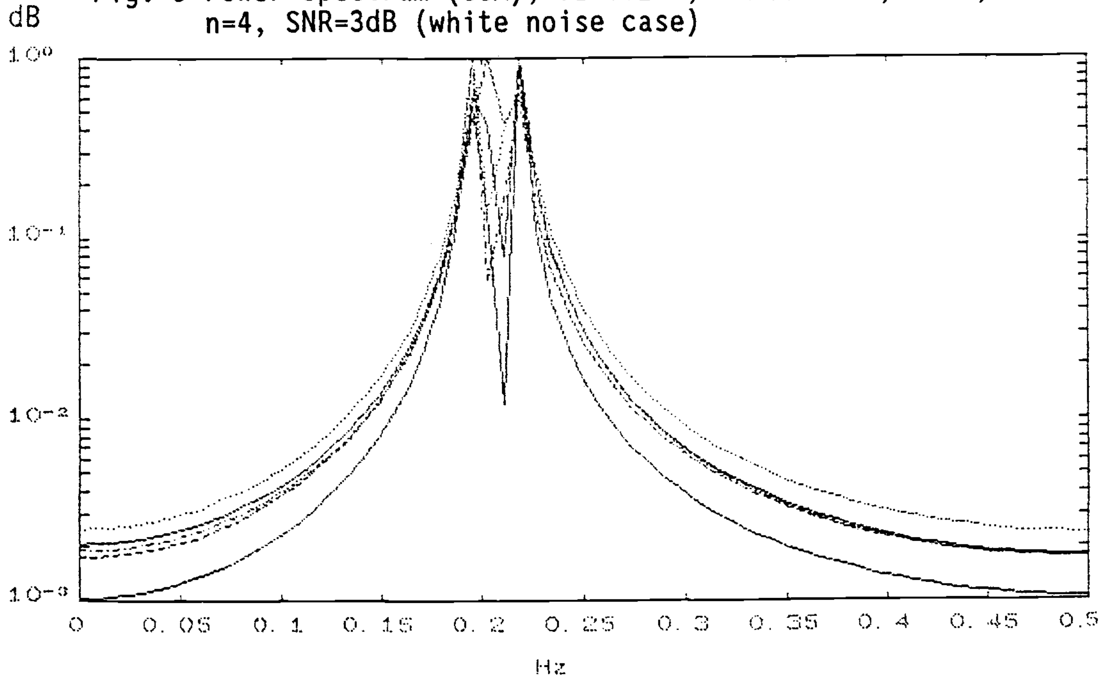


Fig. 10 Pole Plot (CCA), f1=0.2Hz, f2=0.215Hz, N=64, f=20, n=4
SNR=3dB (white noise case)

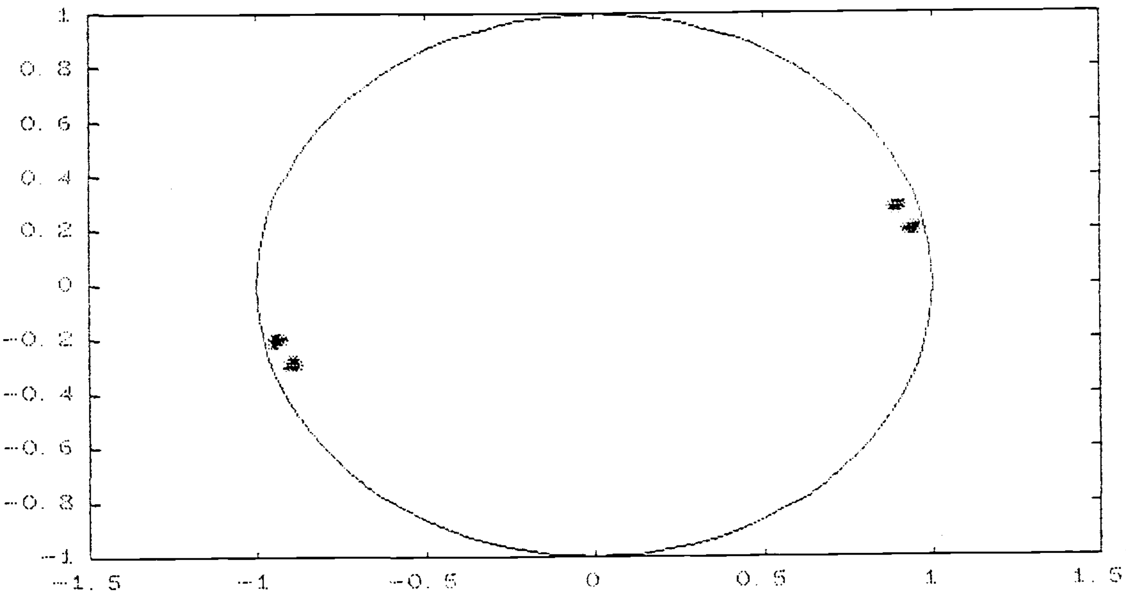


Fig. 11 Power Spectrum (CCA), f1=0.2Hz, f2=0.215Hz, N=64, f=15
n=5, SNR=7dB (white noise case)

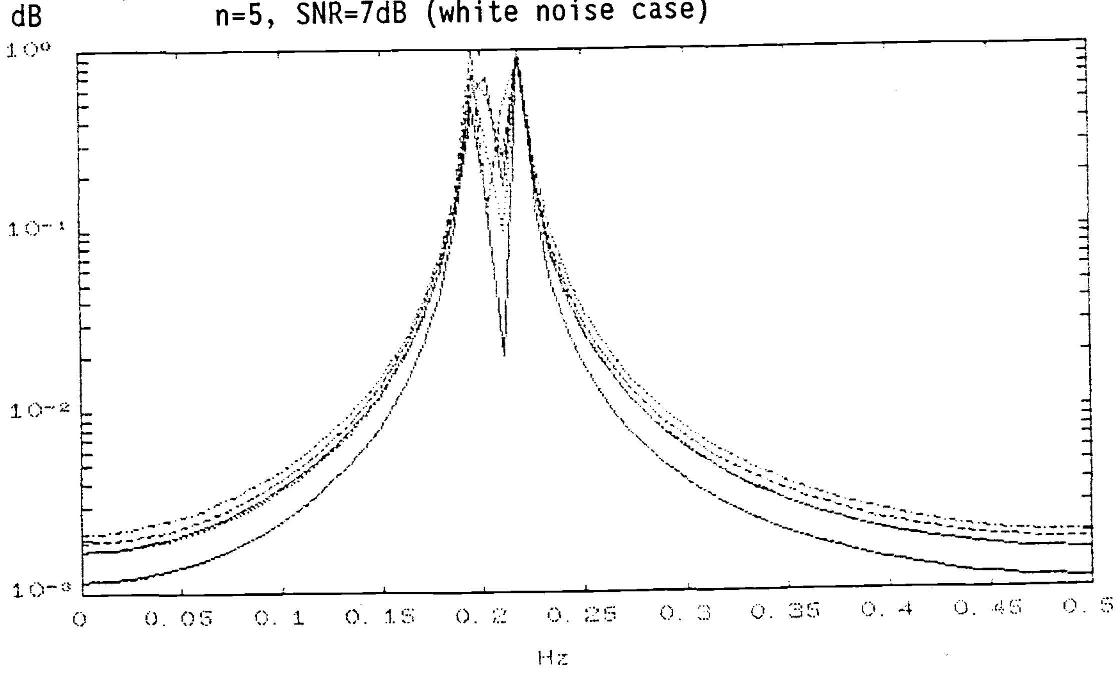
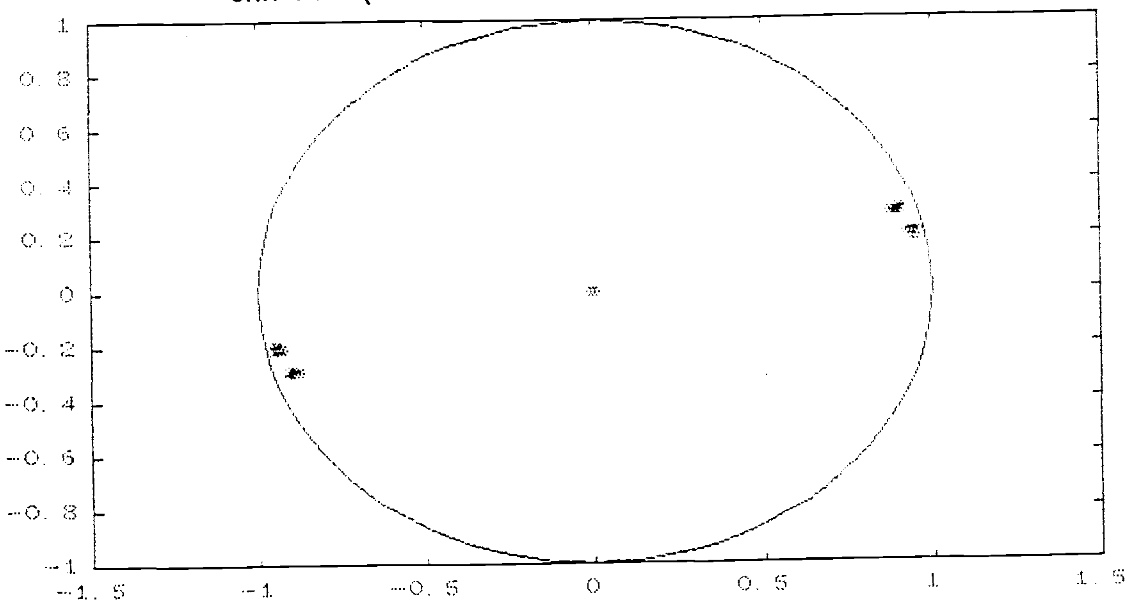


Fig. 12 Pole Plot (CCA), f1=0.2Hz, f2=0.215Hz, N=64 f=15, n=5
SNR=7dB (white noise case)



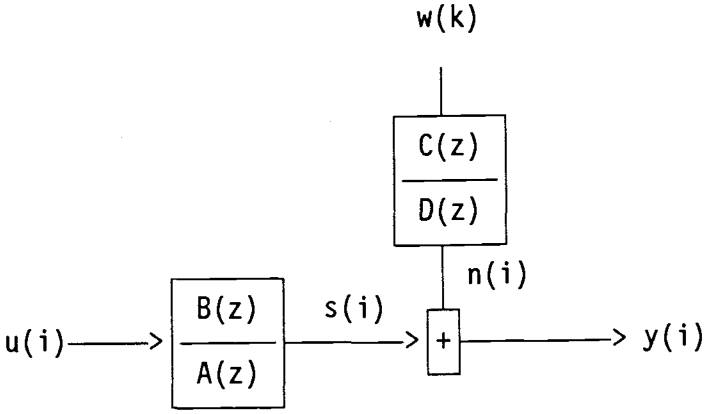
4. ARMA POWER SPECTRUAL ESTIMATION USING THE CCA METHOD FOR COLORED NOISE

Most of the conventional eigenstructure parametric spectral esitmation approaches developed assume that the additive noise is white. In many practical situations, however, the noise is colored, resulting large performance degradations in terms of bias, spurious peaks and nondetection of close-spaced or small signals [25]. Some previous studies for colored noise case have been done lately and can be found in [34-39]. The algorithm in [34] can be regarded as a generalization of the Pisarenko's harmonic retrieval method for the MA noise case. For the AR noise case, in [36], X.D. Zhang presents an ARMA spectral estimation. In [37], A.Sano and K. Hashimoto use an adaptive technique. When the noise is an ARMA(p,q) process or a purely non-deterministic process, the SVD method was used directly by simply setting initial model orders $Pe \gg p$, $Qe \gg q$ in [35].

The Canonical Correlation Analysis (CCA) method has some advantages for white noise corrupted signal estimation. In this section we will consider the application of the CCA method for spectral estimation in an additive colored noise environment. Here the noise could be an AR process, an MA process or a purely non-deterministic process. The suggested CCA algrithm realizes the ARMA model estimates from the output data by approximating the technique shows outstanding ability of suppressing the colored noise. The simulation results for MA colored noise and ARMA colored noise cases are compared with the bootstrap Pisarenko method in [34] and the SVD method in [35].

4.1 Problem Formulation

Consider the discrete single input-output system show below



Assume that the signal s(i) is a sum of h sinusoids,

$$s(i) = \sum_{j=1}^{h} A_i \cos(\, 2\pi f_j i + \theta_i \,) = \frac{B(z)}{A(z)} \qquad 0< 2\pi f_j <\pi \qquad (4.1.1)$$

Where

$$A(z) = 1 + a(1)\, z^{-1} + a(2)\, z^{-2} \ldots a(n)\, z^{-n} \qquad (4.1.2a)$$

$$B(z) = b(0) + b(1)\, z^{-1} + b(2)\, z^{-2} \ldots b(n)\, z^{-n} \qquad (4.1.2b)$$

and the noise is generated by an ARMA(p,q) process

$$\sum_{i=0}^{p} d(i)\, n(k-i) = \sum_{j=0}^{q} c(j)\, w(k-j) \qquad (4.1.3)$$

which represents the narrowband background colored noise generated by the above ARMA(p,q) rational process. Usually, we assume p = q.

w(k) denotes a white noise with zero mean and finite variance. The output observation of the system is given by

$$y(i) = s(i) + n(i) = \frac{B(z)}{A(z)} u(i) + n(i)$$

or

$$A(z)y(i) = B(z) u(i) + A(z)n(i) \qquad (4.1.4)$$

Let

$$A(z) n(i) = 1 + a(1) n(i-1) + ..+ a(n) n(i-n) = v(i)$$

$$n(i) = \frac{1}{A(z)} v(i) \qquad (4.1.5)$$

then (4.1.4) becomes:

$$A(z) y(i) = B(z) u(i) + v(i) \qquad (4.1.6)$$

The following assumptions are made for the estimation problem:

(a) The output noisy observation sequence is stationary.

(b) The colored noise n(i) is uncorrelated with the signal s(i) such that E[s(i) n(j)′] = 0.

(c) The number of sinusoids, i.e. the system order is known priori.

Assuming state space model for the signal s(i) is

$$\begin{aligned} x(i+1) &= A\, x(i) + B\, u(i) \\ s(i) &= C\, x(i) \end{aligned} \qquad (4.1.7)$$

where u(i) is the driven white noise with

$$E[u(i)] = 0, \quad E[u(i)\, u(j)'] = Q\, \delta(i-j) \qquad (4.1.8)$$

and the autocorrelation function of the noisy output is

$$Ry(k) = E[\, y(i+k)\, y(i)'] \qquad (4.1.10)$$

n(.) is the weakly stationary ARMA color noise sequence with

$$E[n(i)] = 0, \; E[n(i+k)\, n(i)'] = Rn(k) \qquad (4.1.11)$$

From the assumption (b) $E[\, s(i)\, n(j)'] = 0$, it is easy to show that

$$Ry(k) = Rs(k) + Rn(k) \qquad (4.1.12)$$

with

$$Rs(k) = E[\, s(i+k)\, s(i)']$$

Then the power spectrum of the output is equal to

$$Py(e^{jw}) = Ps(e^{jw}) + Pn(e^{jw}) \qquad (4.1.13)$$

4.2 The CCA Algorithm for the Colored Noise Case

In this section, the CCA technique is used for frequency detection from color noise. The idea is based on the combination of the bias compensation method [40] and the CCA method.

From the expression (4.1.12), it can be seen that each signal related correlation Rs(k) is added with an extra value of color noise correlation function Rn(k). It is known that Rs(k) is in the form

$$Rs(k) = \sum_{i=1}^{h} (A_i)^2/2 \cos(\, 2\pi f_i k \,) \qquad (4.2.1)$$

<u>Obtain the Generalized Modified Yule-Walker Equations</u>

From (4.1.6), we have

$$A(z)\, y(i) = B(z)\, u(i) + v(i)$$

Multiply both sides of this equation by y(i-k) and compute expected values, we will obtain

$$E[\ y(i-k)\ \{\ y(i) + a(1)y(i-1) + .. + a(n)y(i-n)\ \}\ ]$$

$$= E[\ y(i-k)\ \{\ b(0)u(i) + .. + b(n)u(i-n)\}\ ] + E[\ y(i-k)\ v(i)\ ]$$

which is

$$Ry(k) + a(1)Ry(k-1) + .. + a(n)Ry(k-n)$$

$$= b(0)Ryu(-k) + b(1)Ryu(-k+1) + .. + b(n)Ryu(-k+n) + Ryv(-k) \qquad (4.2.2)$$

For k > n, Ryu(k), Ryu(k-1).. Ryu(k-n) are equal to zeros, we have

$$Ry(k) + a(1)Ry(k-1) + .. + a(n)Ry(k-n) = Ryv(-k) = Rvy(k) \qquad k > n \qquad (4.2.3)$$

when n(i) is not white, the cross correlation Rvy(k) unequals zero. (4.2.3) is called **generalized modified Yule-Walker equations.** The well-known modefied Yule-Walker equations in (2.1.3) have the form of

$$Ry(k) + a(1)\ Ry(k-1) + .. + a(n)\ Ry(k-n) = 0 \qquad k > n$$

<u>Calculate the Cross Correlation Function Rvy(k)</u>

Since s(i) and n(i) are uncorrlated with each other and also from (4.1.5)

$$v(i) = 1 + a(1)n(i-1) + .. + a(n)n(i-n)$$

then

$$Rvy(k) = E[\ v(i+k)\ \{\ s(i) + n(i)\ \}\ ]$$

$$= E[\ v(i+k)\ n(i)\ ]$$

$$= E[\ \{\ n(i+k) + a(1)n(i+k-1) + .. + a(n)n(i+k-n)\ \}\ n(i)]$$

$$= Rn(k) + a(1)\ Rn(k-1) + .. + a(n)\ Rn(k-n) \qquad (4.2.4)$$

thus Rvy(k) and Rn(k) satisfy an AR difference equation

$$Rn(k) = Rvy(k) - \sum_{i=1}^{n} a(i)\ Rn(k-i) \quad k = n+1..N \tag{4.2.5}$$

Estimate Noise Autocorrelation Rn(k)

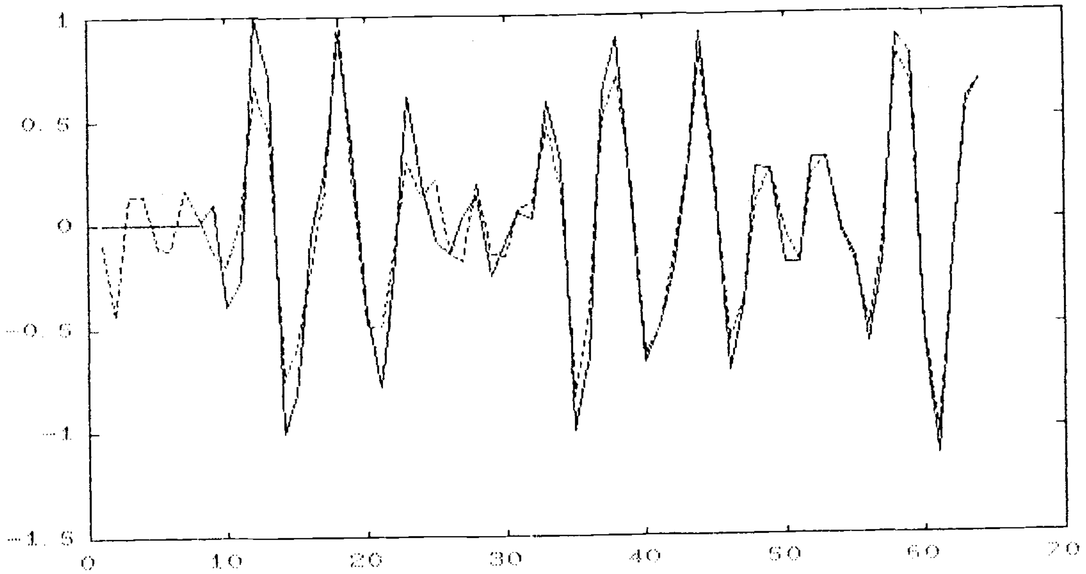
If Rvy(0).. Rvy(n) are given, Rn(0).. Rn(N) can be calculated from (4.2.5). But Rvy(0)..Rvy(n) are unknown. For estimating Rn(k), we set the initial value of Rn(n+1)= Rvy(n+1) and calculate Rn(k) using

$$Rn(k) = Rvy(k) - \sum_{i=1}^{k} a(i)\ Rn(k-i) \qquad k=n+2\ ..\ 2n \tag{4.2.6a}$$

$$Rn(k) = Rvy(k) - \sum_{i=1}^{n} a(i)\ Rn(k-i) \qquad k=2n+1..\ N \tag{4.2.6b}$$

Fig 13 shows the Rn(k) (k=n+1..N) using the true Rn(0)..Rn(n) as initial values in (3.3.5) and estimated Rn(k) using (3.3.6) for the MA(3) color noise case. It can be seen that as k increases, the estimated Rn(k) converges to the true Rn(k).

Fig. 13 Estimated and the True Noise Autocorrelation Function
(for MA(3) color noise case)
dashed line - true autocorr. function
solid line - estimated autocorr. function



Derive the Parametric Matrices and Power Spectrum Using CCA

When $Rn(k)$ is given, $Rs(k) = Ry(k) - Rn(k)$, $k = n+1 \ldots N$, $R_{\pm}$ is updated as

$$R_{\pm} = \begin{bmatrix} Rs(n+1) & Rs(n+2) & \ldots & Rs(n+f) \\ Rs(n+2) & Rs(n+3) & \ldots & Rs(n+f+1) \\ \vdots & & & \\ Rs(n+f) & Rs(n+f+1) & \ldots & Rs(n+2f+1) \end{bmatrix}$$

$$= \begin{bmatrix} CA^{n}B & CA^{n+1}B & \ldots & CA^{n+f-1}B \\ CA^{n+1}B & CA^{n+2}B & \ldots & CA^{n+f}B \\ \vdots & & & \\ CA^{n+f-1}B & CA^{n+f}B & \ldots & CA^{n+2f-2}B \end{bmatrix} \qquad (4.2.7)$$

Since $Rs(k) = C\,A^{k-1}\,B$ and $Rs(n+1) = C\,A^{n}\,B$, let n be an even integer number, $R_{\pm}$ becomes

$$R_{\pm} = \begin{bmatrix} CA^{n/2} \\ CA^{n/2+1} \\ \vdots \\ CA^{n/2+f-1} \end{bmatrix} \begin{bmatrix} A^{n/2}B & A^{n/2+1}B & \ldots & A^{n/2+f-1}B \end{bmatrix} = O\,C \qquad (4.2.8)$$

$$SVD(R_{\pm}) = U\,\Sigma\,V' = O\,C, \quad O = U\,\Sigma^{1/2}, \quad C = \Sigma^{1/2}\,V' \qquad (4.2.9)$$

$$A = O_{f}^{+}\,O_{1} = \begin{bmatrix} CA^{n/2} \\ CA^{n/2+1} \\ \vdots \\ CA^{n/2+f-2} \end{bmatrix}^{+} \begin{bmatrix} CA^{n/2+1} \\ CA^{n/2+2} \\ \vdots \\ CA^{n/2+f-1} \end{bmatrix} \qquad (4.2.10)$$

$$B = [\ A^{n/2}\ ]^{-1} \times [\text{ first column of } \mathbf{C}\ ] \tag{4.2.11}$$

$$C = [\text{ first row of } \mathbf{0}\ ] \times [\ A^{n/2}\ ]^{-1} \tag{4.2.12}$$

where $\mathbf{0}_f$ —— first f-1 rows of **0**

$\mathbf{0}_l$ —— last f-1 rows of **0**

$\mathbf{0}_f^+$ —— peudo-inverse of $\mathbf{0}_f$

The **bootstrap type algorithm** can be summarized as follows :

(1) Calculate the biased sample autocorrelation function Ry(k) from y(1) .. y(N).

$$Ry(k) = 1/N \sum_{i=1}^{N-k} y(i+k)\ y(i)$$

Let Rs(n+k)=Ry(k), n+k = k (k = 0..N) in (4.2.7) - (4.2.12) to calculate A, B, C matrices and estimate initial values of a(1), a(2), .., a(n) (select n to be an even number and larger than the true order of the system).

(2) Compute Rvy(k) for k > n from

$$Rvy(k) = Ry(k) + \sum_{i=1}^{n} a(i)\ Ry(k-i) \qquad k > n$$

Then calculate Rn(n+1) .. Rn(N) using (4.2.6).

(3) Compose Rs(k) = Ry(k)-Rn(k), for k = n+1..N and calculate new A, B, C and a(1).. a(n) using (4.2.7)-(4.2.12).

(4) Go back to step (2) to update Rvy(k) and Rn(k) as well as

system parameters and reduce the model order after one or several iterations until a(1), .., a(n) converge and the system order n is equal to 2h.

The signal poles will be the eigenvalues of final A.

From $Rs(k) = \sum_{i=1}^{h} (A_i^2/2) \cos(k\, w_i)$

Once the frequencies $\{f_i\}$ are known, the magnitude of each signal component can be obtained by solving the following equtions:

$$\begin{bmatrix} Rs(n+1) \\ Rs(n+2) \\ \cdot \\ \cdot \\ Rs(n+h) \end{bmatrix} = \{ \cos[(n+i)w_j] \}_{ij} \begin{bmatrix} A_1^2/2 \\ A_2^2/2 \\ \cdot \\ \cdot \\ A_h^2/2 \end{bmatrix} \qquad (i,j = 1..2h)$$

4.3 Hideaki Sakai's (HS) Algorithm and the SVD Method for Power Spectrum Estimation with Color Noise

In order to compare the performance of the new method with some other approaches, the methods introduced in [34] and [35] were also used for power spectrum estimation with colored noises.

(a) The Hideaki Sakai's (HS) algorithm in [34] is concerned with the estimation of frequencies of sinusoids disturbed by MA(q) colored noise based on the Pisarenko method [18]. It can be summarized as follows:

The observations are given by $y(t) = s(t) + n(t)$ with signal

$$s(t) = \sum_{i=1}^{h} A_i \sin(2 \pi f_i t + \phi_i) \tag{4.3.1}$$

and the MA(q) noise

$$n(t) = \epsilon(t) + b(1) \epsilon(t-1) +..+ b(q) \epsilon(t-q) \tag{4.3.2}$$

where $\epsilon(.)$ is white noise with zero mean.

$$y(t) + Y_{t-1} \mathbf{a} = v(t) \tag{4.3.3}$$

where $Y_{t-1} = [y(t-1),...,y(t-p)]'$ $\mathbf{a} = (a(1),...,a(p))'$

$$v(t) = n(t) + a(1) n(t-1) +..+ a(p) n(t-p) \tag{4.3.4}$$

Bootstrap type algorithm:

(i) estimate α of a and residuals e(t) from the data $\{y(t)\}$.

$$\boldsymbol{\alpha} = R^{-1} \mathbf{r} \tag{4.3.5}$$

where $R = \sum_{t=p+1}^{N} Y_{t-1} Y_{t-1}')^{-1}$, $\mathbf{r} = - \sum_{t=p+1}^{N} Y_{t-1} y(t))$

$$e(t) = y(t) + Y_{t-1}' \boldsymbol{\alpha}$$

(ii) set $\mathbf{a} = \boldsymbol{\alpha}$ so that $v(t)= e(t)$. Then obtain $r_v(k)$ and rn(k) using

$$r_v(k) = 1/(N-p) \sum_{t=p+1}^{N-k} v(t+k) v(t)$$

$$\bar{r}_n = G(a)^{-1} \bar{r}_v$$

where $\bar{r}_n = [r_n(0),..,r_n(p)]'$, $\bar{r}_v = [r_v(0),..,r_v(p)]'$ and

G(a) is a (p+1)x(p+1) matrix satisfies

$$r_v(k) = \sum_{i=0}^{p} \sum_{j=0}^{p} r_n(i-k-j) a(i) a(j) \quad \text{with } a(0) = 1$$

Obtain the new $\mathbf{a}$ of a by solving equations:

$$\mathbf{a} = \boldsymbol{\alpha} + (N-p) R^{-1} [Rn \mathbf{a} + \mathbf{r}_n]$$

i.e. $\mathbf{a} = [I_p - (N-p)\ R^{-1}\ Rn]^{-1}\ [\ \boldsymbol{\alpha} + (N-p)\ R^{-1}\ r_n\ ]$ (4.3.6)

where $Rn = \begin{bmatrix} r_n(0) & r_n(1) & \dots & r_n(p-1) \\ r_n(1) & r_n(0) & \dots & r_n(p-2) \\ . & & & . \\ r_n(p-1) & r_n(p-2) & .. & r_n(0) \end{bmatrix}$ $\quad r_n = \begin{bmatrix} r_n(1) \\ r_n(2) \\ . \\ r_n(p) \end{bmatrix}$

I_p is an identity matrix of order p.

(iii) Go back to step (ii) to recompute $\bar{r}_v$ and **a**

Repeat the successive procedures until **a** converges.

(iv) Compute the AR spectrum

$Sy(e^{jw}) = 1/\ |\ 1 + a(1)\ e^{-jw} +..+ a(p)\ e^{-pjw}\ |^2$

(b) The SVD method used in [35] for frequency analysis in ARMA(p,q) color noise is based on the same procedure introduced in chapter 2.2 by simply taking Pe >> 2h + p and Qe >> 2h + q.

4.4 Numerical Examples

In numerical simulations, we consider two sinusoids embedded in the unknown AR(MA) spectral signals n(t). Four examples are presented here to demonstrate the new algorithm that involves estimating two sinusoids with frequencies f_1 = 0.15 Hz, f_2 = 0.2 Hz from additive stationary MA(3), AR(4), ARMA(1,1) and ARMA(3,2) color noises. The signal amplitudes $A_1 = A_2$ = 1 or 2 and noise variance σ^2_n = 1 give the component signal-noise ratio SNR = 0 dB or -3 dB. The parameters vectors for the color noise are:

MA(3) : b = [-0.8411 -0.0012 0.4598]

AR(4) : a = [1.352 -1.338 0.662 -0.24]

ARMA(1,1) : a = [0.8], b = [1 0.7]

ARMA(3,2) : a = [-2 1.7 -0.5] b = [1 -1.5 0.685]

The dimension of $R_{\pm}$ was selected to be 20. The initial orders of the system were n = 8 or n = 6 and the final order was n = 4 (Assume the number of sinusoids is known priori). The true spectrums of the observations shown in Fig. 14 - 17 are characterized by two pulses at f_1 and f_2 as well as the background color noise power spectrums. The signal-noise ratios selected were equal to or smaller than 0dB.

With MA(3) colored noise, Fig. 18-19 show the results when the CCA is used directly without iterations. For n=6, the noise pole at f=0.27 Hz is even larger than the signals, making it impossible to determine the signals correctly. For n=4, the signal at f=0.2 Hz is substituted by the noise pole at 0.27 Hz. Fig. 20-23 are the initial results with n=8 and the final results with n=4 after 4 iterations. The improvement of performance is obvious. For comparison, Fig.24 is the estimated spectrum using the algorithm in [34] with the order p = q = 5. It can be seen that only one peak is shown in the graph.

For the AR(4) colored noise case with initial spectrum in Fig. 25, the new CCA algorithm produces the results shown in Fig 26 that exhibits two peaks at the correct locations without any spurious peaks.

We now compare the proposed modified CCA spectral estimator with the Cadzow's approach for Pe >> p and Qe >> q which is suggested in [35] where p and q are the AR and MA orders of the colored noise. The third example for ARMA(1,1) colored noise is depicted in Fig. 27-28. After only one time iteration, the two signals are resolved correctly and the noise component at low frequencies is completely depressed. Compare with the results in Fig. 29-30 using the SVD method, it is noted that there is little response at the signal frequency locations in both the spectrum and the pole plot.

Fig.31-32 illustrate the behavior of the CCA spectrum under the influence of ARMA(3,2) noise. Since the noise pole is very close to one of the signal poles at f=0.15 Hz (see Fig. 17), it is quite difficult to detect the signals from the noise for this case such that more iterations are used. The proposed algorithm can resolve the true frequency components and shows much better performance than that of the SVD method in Fig. 33-34.

From the above, it is apparent that after only a few iterations, the modified CCA method can estimate frequencies f_1 and f_2 correctly and gives better results for all above cases than the other two other algorithms. This method simultaneously calculates AR and MA parameters by solving a group of linear equations and increases the performance of CCA algorithm by using noise compensate technique.

Fig. 14 Power Spectrum of Signals and MA(3) Color Noise
b = [-0.8411 -1.0012 0.4598]

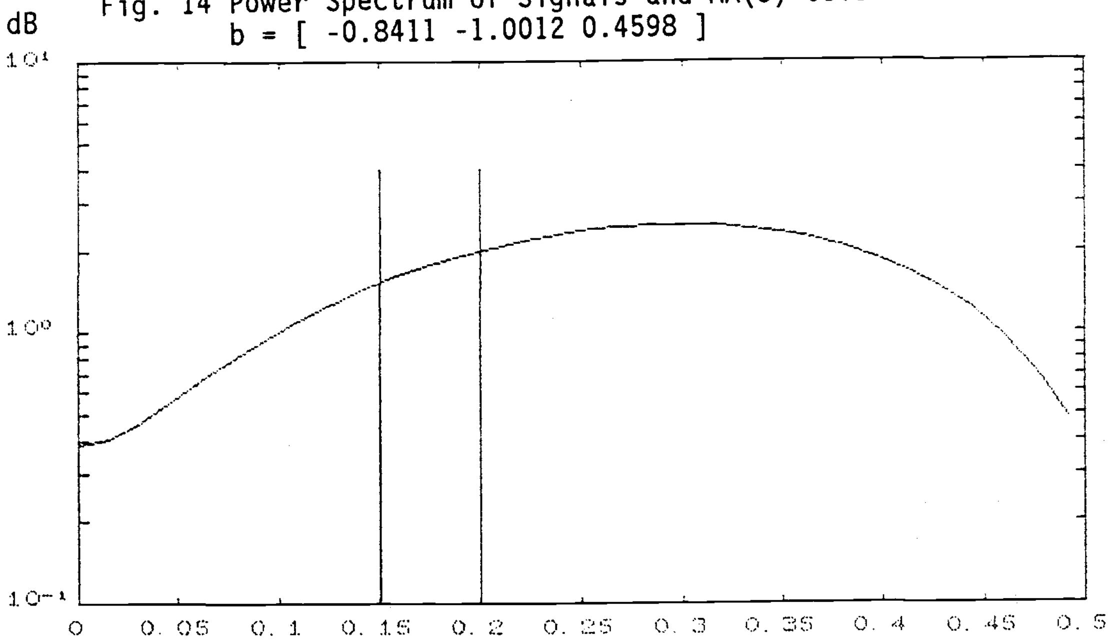


Fig. 15 Power Spectrum of Signals and AR(4) Color Noise
a = [-1.352 1.338 -0.662 0.24]

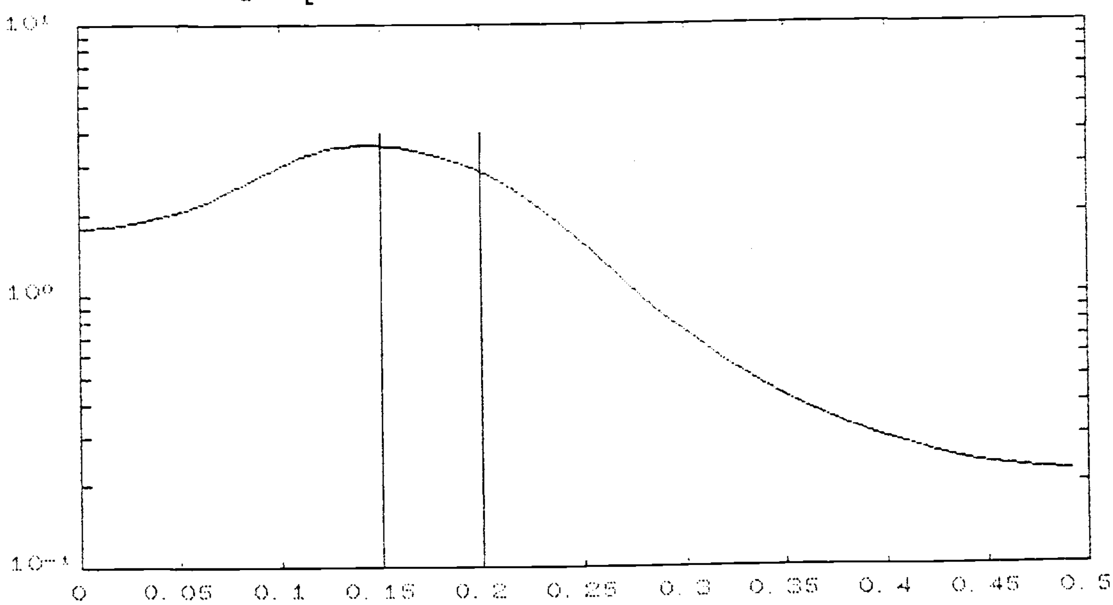


Fig. 16 Power Spectrum of Signals and ARMA(1,1) Color Noise
a = [-0.8], b = [1 0.7]

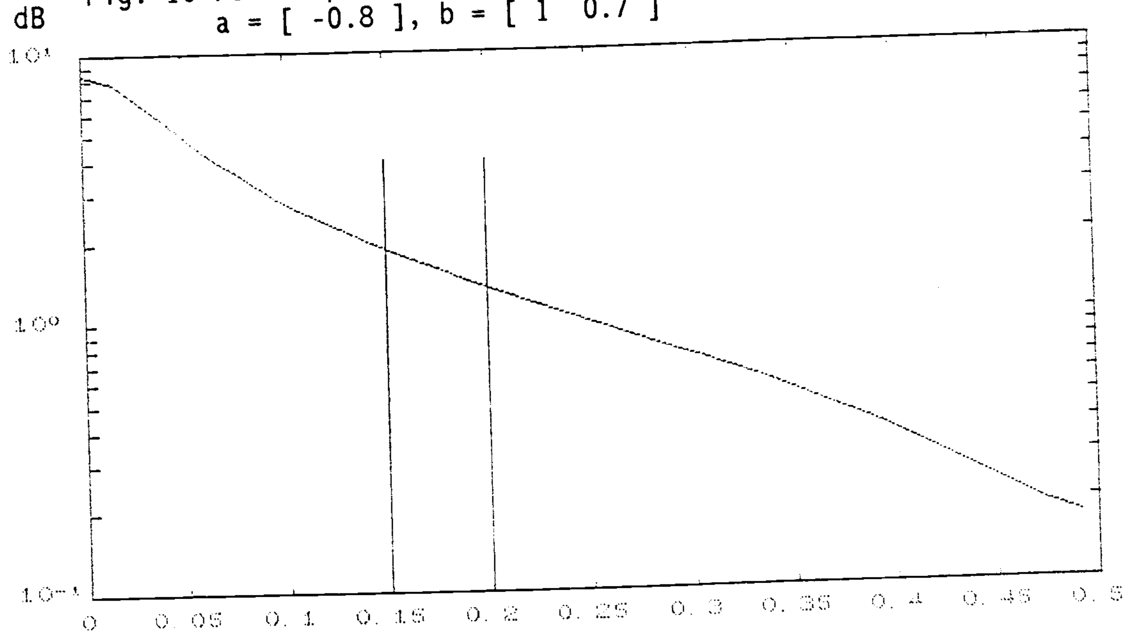


Fig. 17 Power Spectrum of Signals and ARMA(3,2) Color Noise
a = [-2 1.7 -0.5], b = [-1.5 0.685]

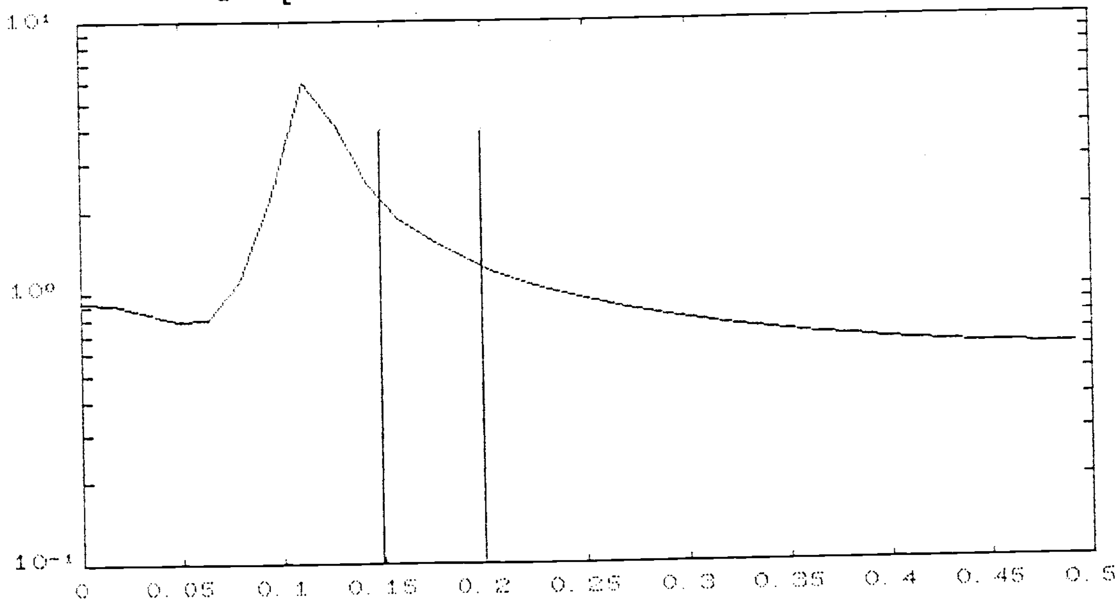


Fig. 18 Power Spectrum (dB) (CCA) fl=0.15Hz, f2=0.2Hz, N=64, f=25, n=6, SNR=-3dB, with MA(3) color noise

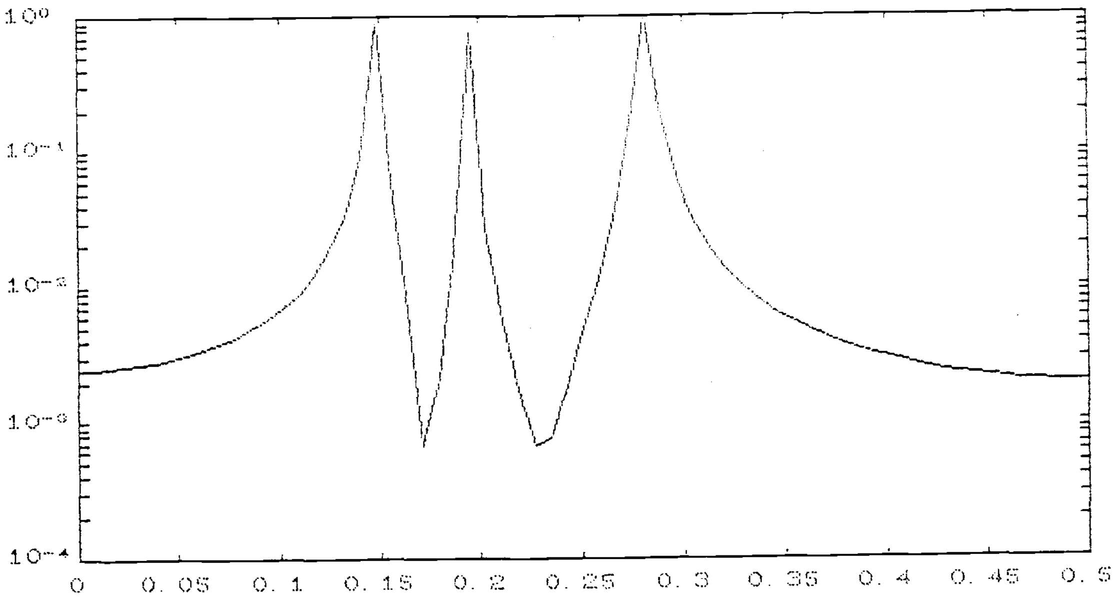


Fig. 19 Power Spectrum (dB) (CCA) fl=0.15Hz, f2=0.2Hz, N=64, f=25, n=4, SNR=-3dB, with MA(3) color noise

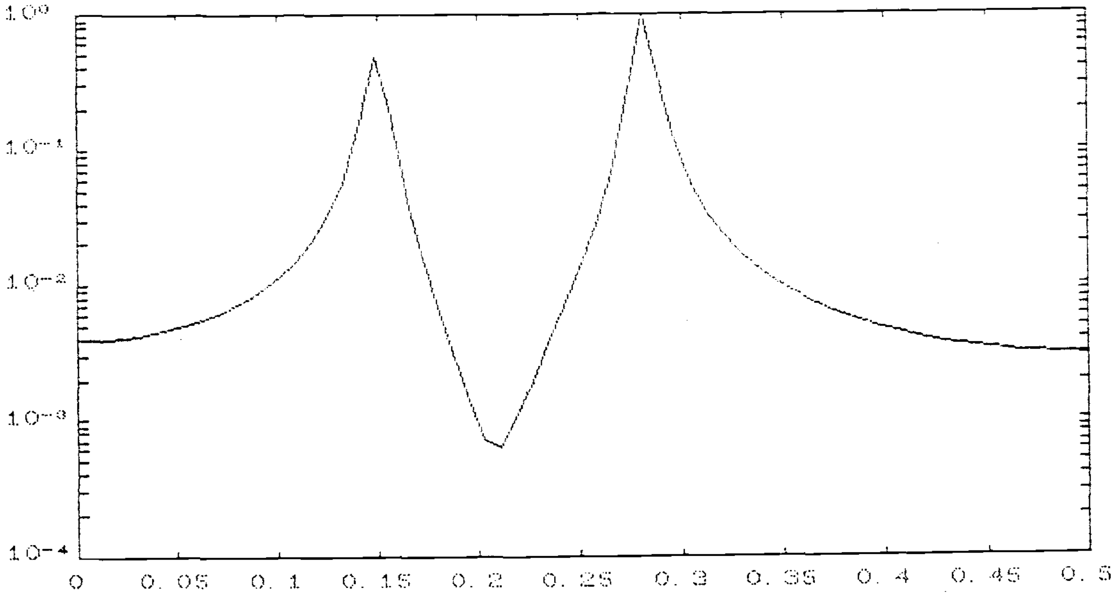


Fig. 20 Power Spectrum (CCA) f1=0.15Hz, f2=0.2Hz, N=64, f=25, n=8 iteration=1, SNR=-3dB, with MA(3) color noise

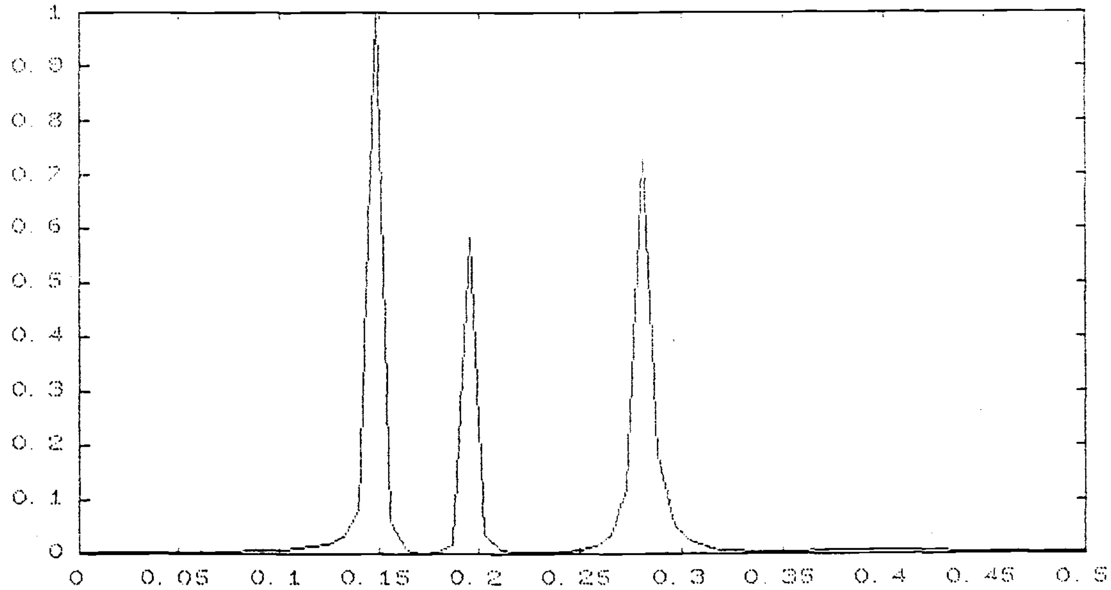


Fig. 21 Power Spectrum (dB) (CCA) f1=0.15Hz, f2=0.2Hz, N=64, f=25, n=8, iteration=1, SNR=-3dB, with MA(3) color noise

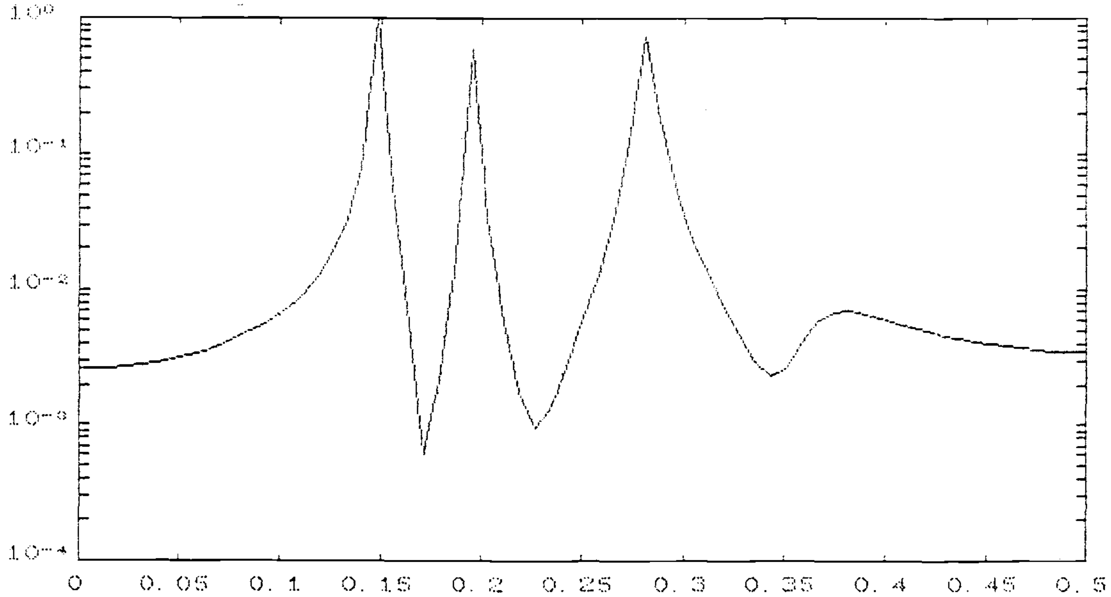


Fig. 22 Power Spectrum (CCA) fl=0.15Hz, f2=0.2Hz, N=64, f=25, n=4 iteration=4, SNR=-3dB, with MA(3) color noise

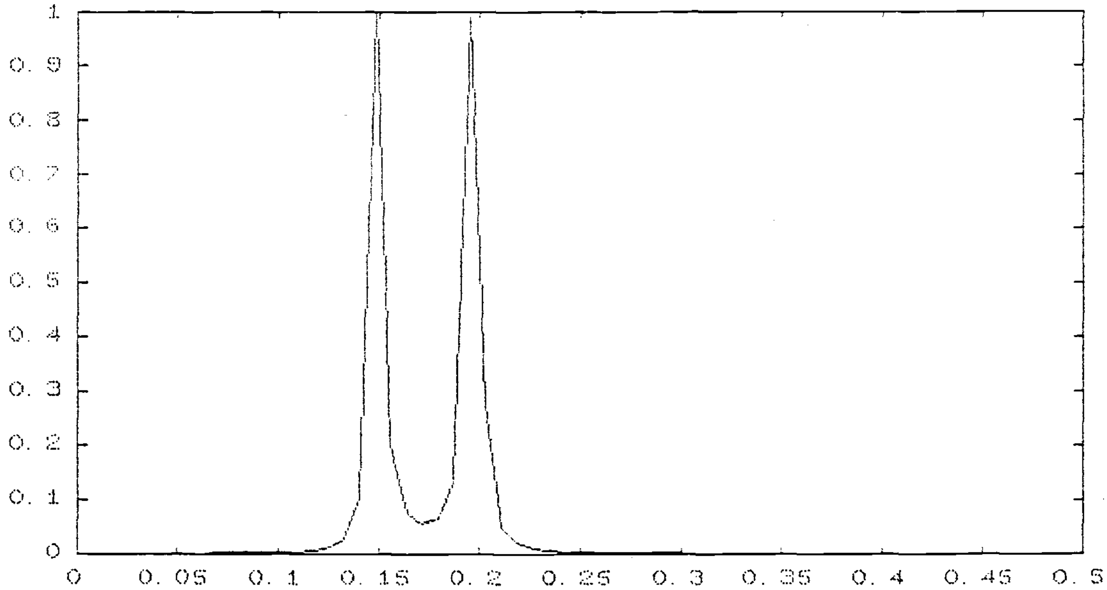


Fig. 23 Power Spectrum (dB) (CCA) fl=0.15Hz, f2=0.2Hz, N=64, f=25, n=4, iteration=4, SNR=-3dB, with MA(3) color noise

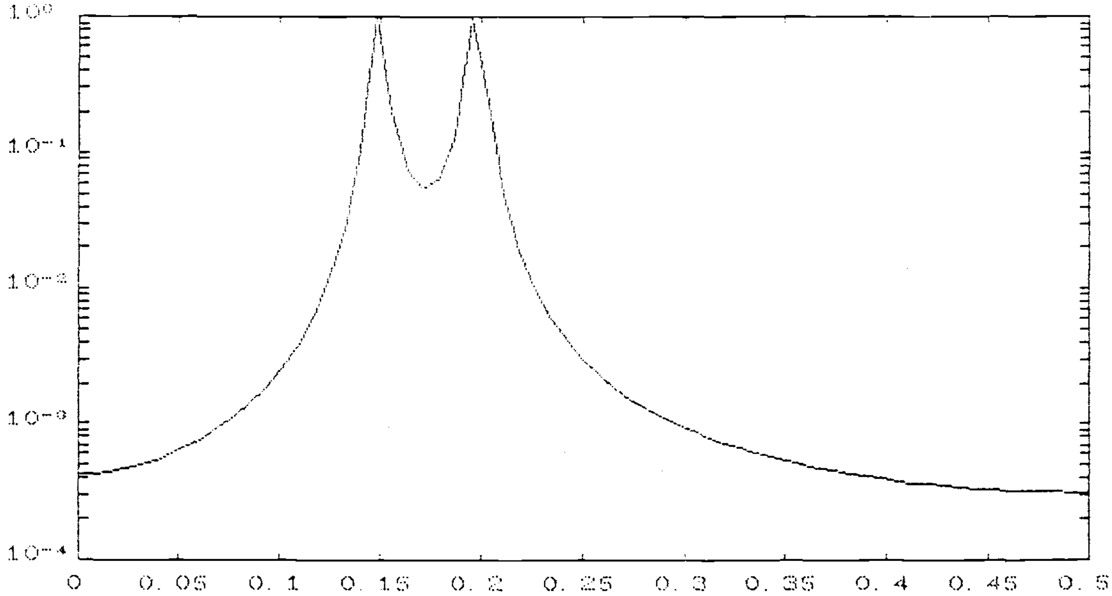


Fig. 24 Power Spectrum (dB) (HK) f1=0.15Hz, f2=0.2Hz, N=200, p=q=5 iteration=4, SNR=-3dB, with MA(3) color noise

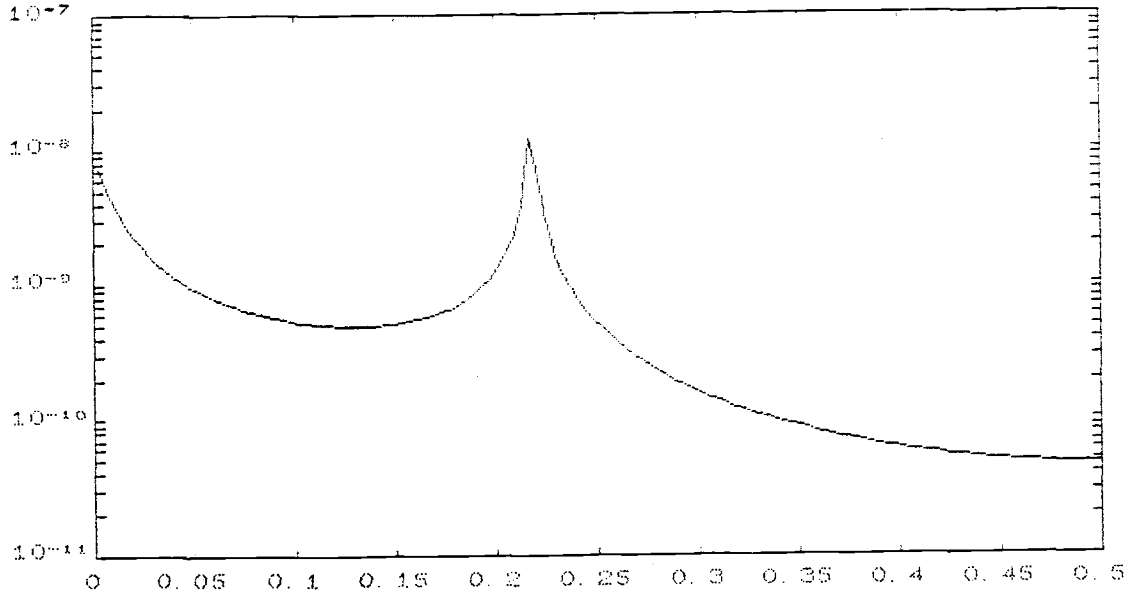


Fig. 25 Power Spectrum (dB) (CCA) f1=0.15Hz, f2=0.2Hz, N=64, f=25, n=6, iteration=1, SNR=-3dB, with AR(4) color noise

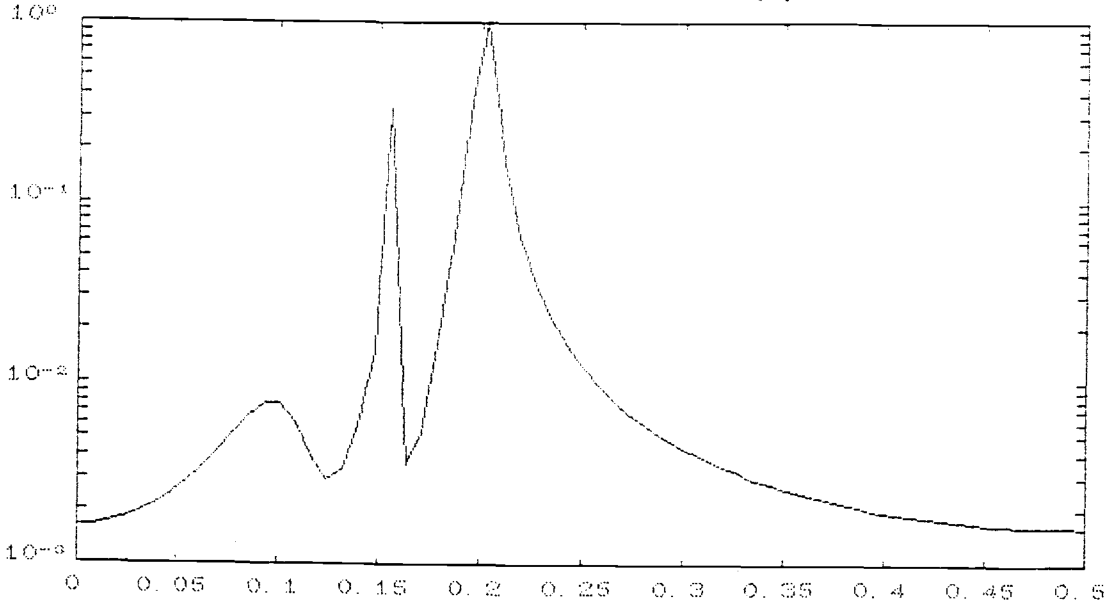


Fig. 26 Power Spectrum (dB) (CCA) f1=0.15Hz, f2=0.2Hz, N=64, f=25, n=4, iteration=4, SNR=-3dB, with AR(4) color noise

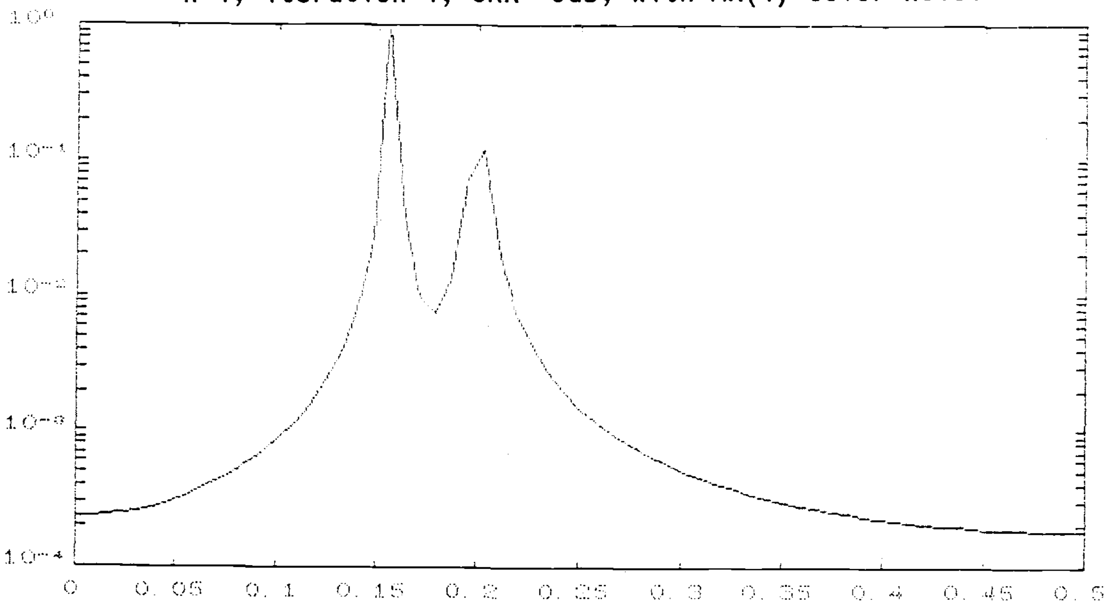


Fig. 27 Power Spectrum (dB) (CCA) f1=0.15Hz, f2=0.2Hz, N=64, f=20, n=6, iteration=1, SNR=0dB, with ARMA(1,1) color noise

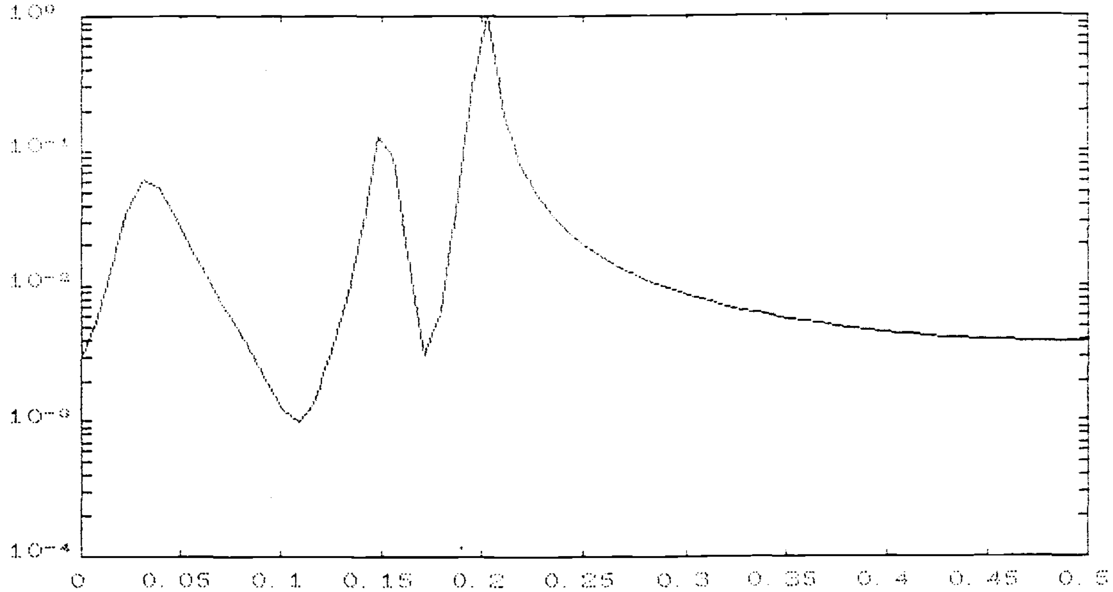


Fig. 28 Power Spectrum (dB) (CCA) f1=0.15Hz, f2=0.2Hz, N=64, f=20, n=4, iteration=4, SNR=0dB, with ARMA(1,1) color noise

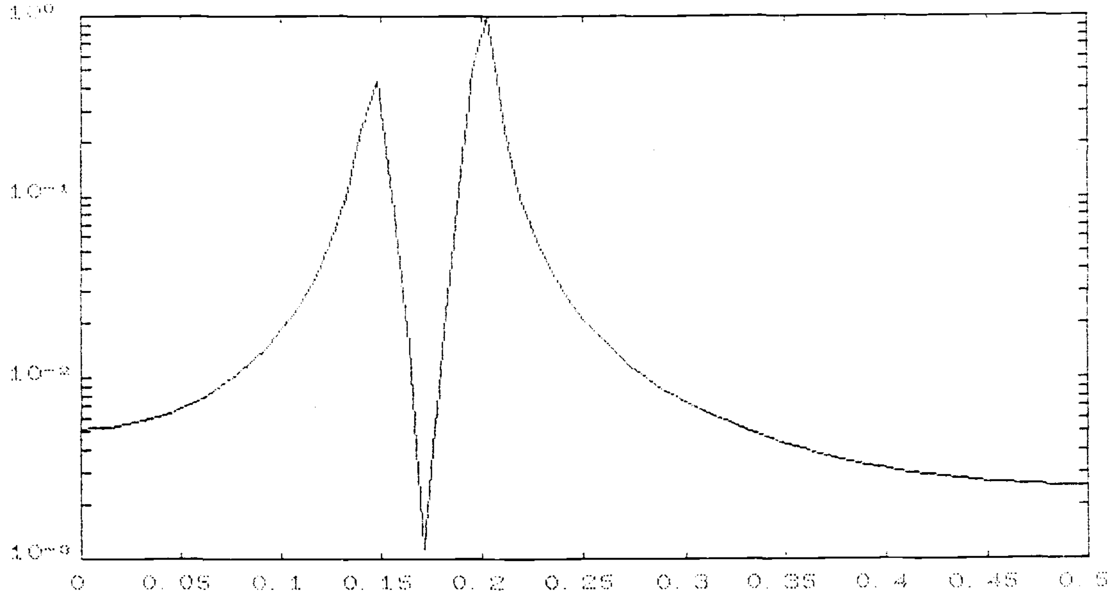


Fig. 29 Power Spectrum (SVD) f1=0.15Hz, f2=0.2Hz, N=64, Pe=Qe=15
t=45, n=5, SNR=0dB, with ARMA(1,1) color noise

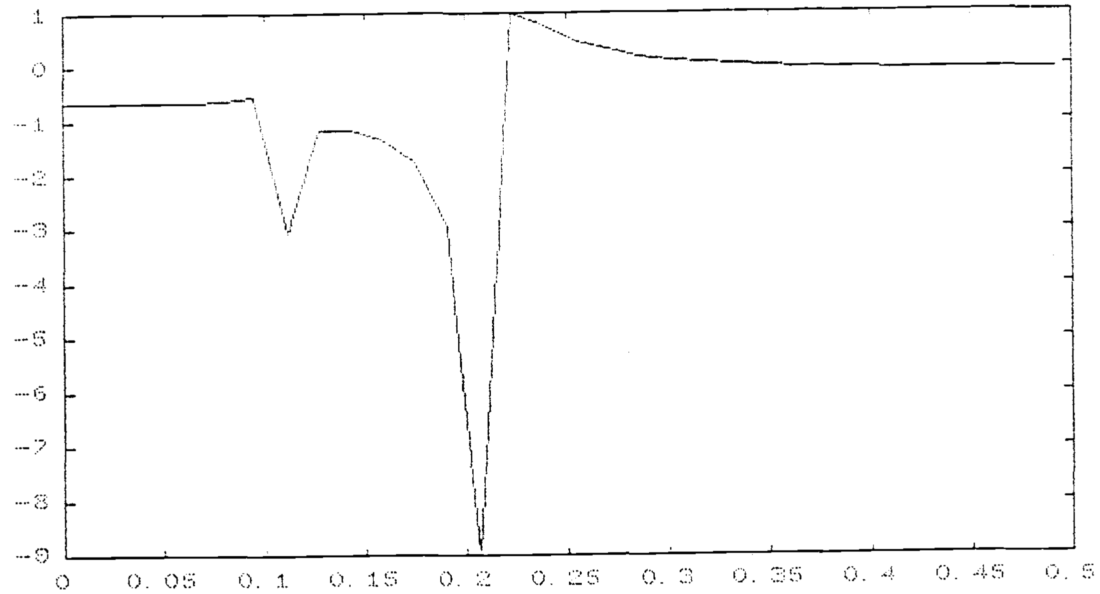


Fig. 30 Pole Plot (SVD) f1=0.15Hz, f2=0.2Hz, N=64, Pe=Qe=15, n=5
SNR=0dB, with ARMA(1,1) color noise

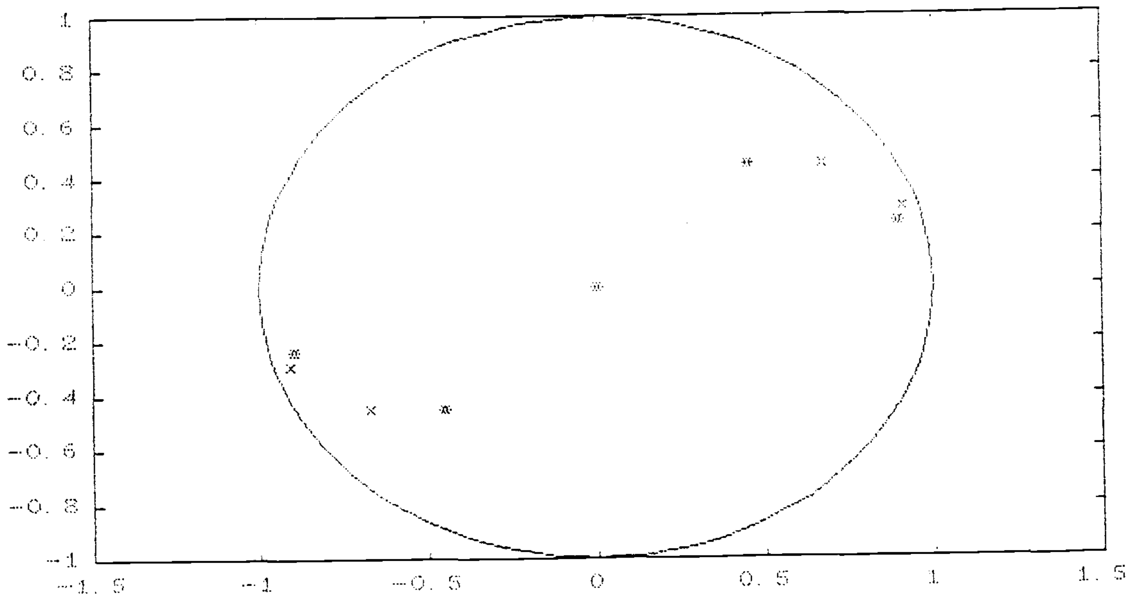


Fig. 31 Power Spectrum (dB) (CCA) f1=0.15Hz, f2=0.2Hz, N=64, f=20, n=6, iteration=1, SNR=0dB, with ARMA(3,2) color noise

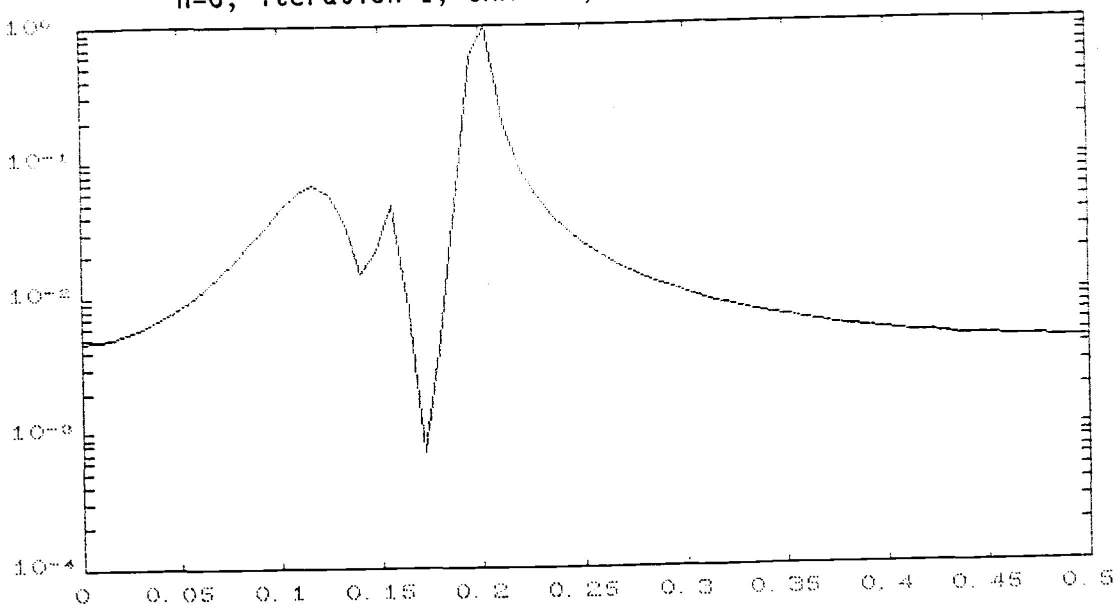


Fig. 32 Power Spectrum (dB) (CCA) f1=0.15Hz, f2=0.2Hz, N=64, f=20, n=4, iteration=7, SNR=0dB, with ARMA(3,2) color noise

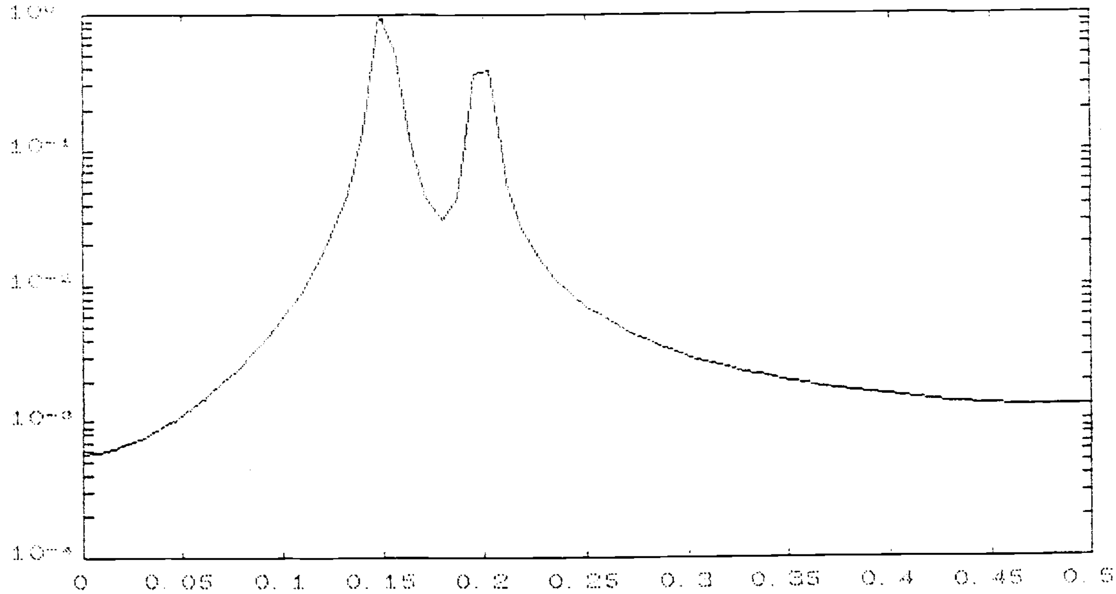


Fig. 33 Power Spectrum (SVD) fl=0.15Hz, f2=0.2Hz, N=64, Pe=Qe=15
t=45, n=5, SNR=0dB, with ARMA(3,2) color noise

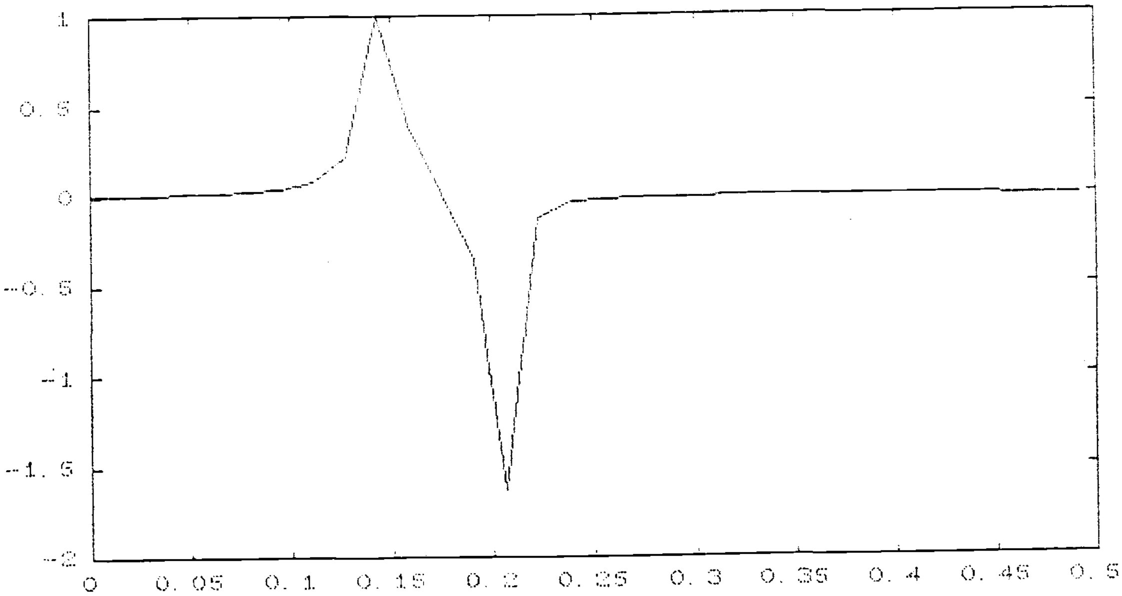
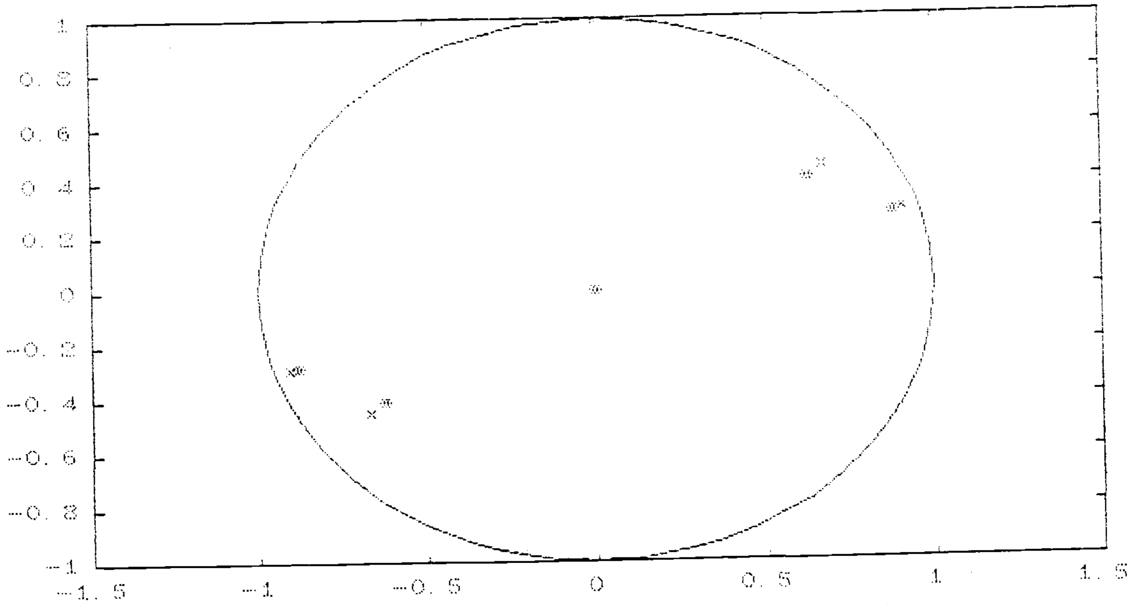


Fig. 34 Pole Plot (SVD) fl=0.15Hz, f2=0.2Hz, N=64, Pe=Qe=15, n=5
SNR=0dB, with ARMA(3,2) color noise



5. CONCLUSIONS

The Canonical Correlation Analysis can be considered as an efficient system identification technique to time series analysis and ARMA spectral estimation given observation with short data length and low SNR in the environment of either white noise or colored noise. The CCA method is based on a local past and a local future of the data to find the optimal state vector that contains all the information available from the entire past to the entire future. In other words, it is a sufficient statistic of the past for optimally predicting the future. Since the state variance is normalized using the SVD technique, which is a robust and numerically stable approximation algorithm, this makes the method insensitive to covariance perturbations, including the effect of colored noise and covariance estimation errors. Furthermore, the modified CCA method proposed here for ARMA spectral estimation in colored noise has the following properties:

(a) The method uses the CCA technique in the bootstrap type algorithm to estimate the AR and MA parameters simultaneously which leads to superior ARMA modeling performance.

(b) The optimal state space model is generated by solving a finite number of linear equations. This greatly reduces computation expenses and still gives satisfactory estimation results. The method can be used for signal detection from ARMA colored noise which is the most general case of noise.

(c) Since the parameter matrices of the CCA method are sign symmetric, the number of calculations needed is reduced by half. The approach not only has good performance ,but also is computationally simpler than other iterative methods.

6. BIBLIOGRAPHY

1. Albert Benveniste, "Single Sample Model Identification of a Non-stationary Stochastic Process", IEEE Trans. on Automatic Control vol. AC-30, No.1, pp66-74, Jan. 1985.

2. B. D.Rao," Perturbation Analysis of an SVD-based Linear Prediction Method for Estimating the Frequencies of Multiple Sinusoids " , IEEE Trans. on ASSP, vol. 36, No.7, pp1026-1035, July 1988.

3. Donald W. Tufts, Ramdas Kumaresan,"Estimation of Frequencies of Multiple Sinusoids: Making Linear Prediction Perform like Maximum Likelihood", Proc. of IEEE, vol. 70, No. 9, Sept. 1982.

4. H.Akaike, " Markovian Representation of Stochastic Processed and its Application to the Analysis of Autoregressive Moving Average Processes ", SIAM, J. Control 21, 1974, pp363-387.

5. H. Akaike,"Autoregressive Model Fitting for Control", Ann. Inst. Statist. Math. 23, 1971, pp163-180.

6. H. Akaike, " Markovian Representation of Stochastic Processes by Canonical Variables " , SIAM, J. Control 13, 1975, pp162-173.

7. H. Akaike, " Canonical Correlation Analysis of Time Series and the Use of an Information Criterion " , in System Identification: Advances and Case Studies , R.K. Mehra and D.G. Lainiotis, Eds. New York: Academic, 1976, pp27-96.

8. H. Akaike,"Stochastic Theory of Minimal Realization",IEEE Trans. on Automatic Control, vol. AC-19, No. 6, Dec. 1974 pp667-674.

9. H. Hotelling, "Relation Between Two Sets of Variates" Biometrika 1936 V28 pp321-377.

10. James L. Phillips, " The Triangular Decomposition of Hankel Matrices ", Mathematics of Computation, vol.25, No. 115, July 1971 pp599-602.

11. James V. White, " Stochastic State-space Models From Empirical Data", IEEE, ICASSP 1983, pp243-246.

12. J. A. Cadzow, "High Performance Spectral-Estimation- A New ARMA Method," IEEE Trans. Acoust. , Speech , Signal Processing, vol. ASSP-28, pp524-529, Oct. 1980.

13. J. A. Cadzow, "Spectral Estimation: An Overdetermined Rational Equation Approach ," Proc. IEEE, vol 70, pp907-939, Sept. 1982.

14. K.S. Arun, " Principal Components Algorithms for ARMA Spectrum

Estimation ", IEEE Trans. on ASSP, vol.37, No.4, pp566-571, April 1989.

15. Pham-Dinh-Tuan, "On the Fitting of Multivariate Processes of the Autoregressive-Moving Average Type",Miometrika 1978, 65,1, pp99-107.

16. R. Kumaresan and D.W. Tufts, " Estimating the Parameters of exponentially Damped Sinusoids and Pole-Zero Modeling in Noise " , IEEE trans. on ASSP-30, No. 6, Dec. 1982, pp 833-840.

17. S. Kiaei, L. Luo, "Canocical Correlation Analysis (CCA) for ARMA Spectral Estimation ", Proc. of IEEE, vol.2, 1989.

18. S.L. Marple , " Digital Spectral Analysis : with Application " Prentice Hall 1987.

19. S.M. Kay, "A New ARMA Spectral Estimator, " IEEE trans Acoustic Speech, Signal Processing, vol. ASSP-28, pp 585-588, Oct. 1980.

20. S.M. Kay, "Noise Compensation for Autoregressive Spectral Estimation", IEEE trans. on ASSP-28 No.3 June 1980, pp 292-302.

21. S.M. Kay, S.L. Marple,"Spectrum Analysis-A modern Perspective ", Proc. of IEEE vol. 69, No. 11, Nov. 1981, pp 1380-1419.

22. S.M. Kay, A.K. Shaw, " Frequency Estimation by Principal Component AR Spectral Estimation", IEEE Trans. on ASSP, vol. 36, No.1, Jan 1988, pp.95-101.

23. Surendra Prasad, K.V.S. Hari, "Improved ARMA Spectral Estimation Using the Canonical Variate Method", IEEE Trans. on Acoustics Speech and Signal Processing, vol. ASSP 35, No. 6, June 1987.

24. S.Y. Kung, K.S. Arun, D.V. Bhaskar Rao,"State-Space and Singular Value Decomposition-Based Approximation Methods for the Harmonic Retrieval Problem", J. Opt. Soc. Am. vol. 73, No.12, pp1799-1811, Dec. 1983.

25. S.Y. Kung, K.S. Arun, "A Novel Hankel Approximation Method for ARMA Pole-Zero Estimation From Noisy Covariance Data", in Tech. Dig. Topical Meet. Signal Recovery and Synthesis with Incomplete Inform. and Parcial Constraints. Incline Village , NV: Opt. Soc. Amer., 1983, pp.WA19-1--WA19-5.

26. U.B. Desai, C.S. Usu, "Minumum Phase Model Reduction of Deterministic Systems Via Stochastic Balancing ", Control-Theory and Advanced Technology, vol.4, No.3, pp349-358, 1988.

27. U.B. Desai, D. Pal,"A Realization Approach to Stochastic model Reduction", Oct. 1982.

28. U.B. Desai, " Modeling Application of Stochastic Processes ", Klawer Academic Publisher, 1987.

29. V.J. Yohai, M.S. Garcia Ben, " Canonical variables as optimal predictors ", AMS Subject Classification 1979.

30. W.E. Larimore, "System identification, reduced-order filtering and modeling via canonical variate analysis" , Proc. 1983 Amer. Cont. Conf., San Francisco, CA , June 1983, pp445-451.

31. Yoram Baram, "Realization and Reduction of Markovian Models from Nonstationary Data", IEEE Trans. on Automatic Control, vol. AC-26, No.6, pp 1225-1231, Dec. 1981.

32. C. Chatfield, "The Analysis of Time Series:an Introduction " , Chapman and Hall 1987.

33. A.V. Oppenheim, R.W. Schafer, "Digital Signal Processing", 1975.

34. Hideaki Sakai,"Estimation of Frequencies of Sinusoids in Colored Noise", IEEE ICASSP 86, pp177-180.

35. Xian-Da Zhang, "Estimation of Frequencies of Sinusoids in ARMA Colored Noise Via Singular Value Decomposition ", IEEE ISCAS 89, pp1315-1318, 1989.

36. Xian-Da Zhang, Hiroshi Takeda,"An Approach to Time Series Analysis and ARMA Spectral Estimation", IEEE trans. on ASSP vol. 35, No. 9, Sept. 1987, pp1303-1313.

37. Akira Sano, K. Hashimoto,"Adaptive Recursive Scheme for Spectral Analysis of Sinusoids in Signals with Unknown Colored Spectrum", IEEE ICASSP 1985, pp109-112.

38. A. H. Tewfik , " Harmonic Retrieval in the Presence of Colored Noise " , IEEE ICASSP 1989, pp2069-2072.

39. P.J. Sherman and A.E. Frazho, " High Resolution Spectral Estimation of Sinusoids in Colored Noise Using a Modified Pisarenko Decomposition", IEEE ICASSP 86, pp181-184.

40. S. Sagara, K. Wada, "On-line Modified Least Dynamic Systems", Int. J. Control, vol. 25, pp329-343, 1977.