SVD-Based Information Theoretic Criteria for Detection of the Number of Damped/Undamped Sinusoids and Their Performance Analysis

V. Umapathi Reddy, Senior Member, IEEE, and L. S. Biradar

Abstract-Recently, Wax and Kailath developed information theoretic criteria for detection of the number of signals received by a sensor array. More recently, Fuchs developed a criterion, based on the perturbation analysis of the data autocorrelation matrix, for detecting the number of sinusoids. In this paper, following the information theoretic approach to model selection, we first develop criteria for detection of the number of damped/undamped sinusoids. These criteria are matched to the singular value decomposition (SVD) based methods, such as modified forward/backward and forwardbackward linear prediction, so well that the extra computations needed over and above those required for computing the SVD are marginal. Next, we develop an analytical framework for analyzing the performance of these criteria, following the assumptions made by Wang and Kaveh and the corrections given by Zhang et al. In the development of the analysis, we make some approximations which become better for large signal-tonoise ratio. Simulations are used to verify the usefulness of the analysis, and to compare the performance of our method with that of Fuchs.

I. INTRODUCTION

THE problem of estimating the parameters (i.e., frequencies and damping factors) of sinusoids (damped/ undamped) in the presence of additive white noise is an important one and several methods have been recently proposed in this regard (see [5]). At high signal-to-noise ratio (SNR), all the existing methods perform equally well asymptotically. The problem, however, becomes very difficult when the number of available data samples is small and/or the SNR is low.

The forward-backward linear prediction (FBLP) method [6] is capable of resolving closely spaced sinusoids with short data records if SNR is sufficiently high. For short data lengths and moderate SNR's, the modified FBLP (MFBLP) method for undamped sinusoids [7], modified backward linear prediction (MBLP) method for damped sinusoids [8], and a recently proposed total least

IEEE Log Number 9210127.

squares (TLS) method [9] perform much better than most of the model-based methods. A quantitative accuracy analysis of the method suggested in [8] has been developed in [17]. All these methods, i.e., MFBLP, MBLP, and TLS, require the knowledge of the number of sinusoids *a priori*. In practice, however, this *a priori* knowledge is not available. We therefore need a technique to estimate this number from the observed data.

One can use the method proposed in [10] and [11] to estimate the number of sinusoids. But, it involves a subjective decision of the designer making it difficult to implement in practice. Criteria based on the application of information theoretic principles to model selection, introduced by Akaike (AIC) [12] and by Schwartz [13] and Rissanen (MDL) [14], have been recently developed [1] for the detection of the number of signals received by a sensor array. More recently, Fuchs developed a criterion [2] based on the perturbation analysis of the data autocorrelation matrix for detecting the number of sinusoids, which is in some sense a subjective-based method.

In this paper, following the information theoretic approach, we first develop criteria for the detection of the number of damped/undamped sinusoids. Next, we develop an analytical framework for analyzing the performance of the proposed criteria. In both cases, we make some approximations which become better when SNR is large. Simulations are used to verify the usefulness of the analysis.

The paper is organized as follows. In Section II, we develop the criteria. Section III gives the performance analysis of the proposed criteria and Section IV discusses simulation results. Finally, in Section V, we conclude the paper.

II. DEVELOPMENT OF THE CRITERIA

Consider N uniformly spaced data samples y(n) of M real sinusoids corrupted by additive white Gaussian noise

$$y(n) = \sum_{i=1}^{M} a_i e^{\alpha_i n} \sin (\omega_i n + \phi_i) + v(n),$$

$$n = 1, 2, \cdots, N \qquad (2.1)$$

where a_i is the amplitude, α_i is the damping factor, ω_i is the normalized angular frequency, and ϕ_i is the initial phase (assumed to be uniformly distributed in the interval

1053-587X/93\$03.00 © 1993 IEEE

Manuscript received December 18, 1990; revised November 9, 1992. The associate editor coordinating the review of this paper and approving it for publication was Prof. Monson Hayes.

V. U. Reddy is with the Department of Electrical Communication Engineering, Indian Institute of Science, Bangalore 560 012, India.

L. S. Biradar was with the Department of Electrical Communication Engineering, Indian Institute of Science, Bangalore 560 012, India. He is now with the Department of Electronics and Communication Engineering, PDA College of Engineering, Gulbarga 585 102, India.

 $[0, 2\pi]$) of the *i*th sinusoid. $\{v(n)\}$ are zero-mean, independent and identically distributed real Gaussian random variables with variance σ^2 .

For the observed data, the set of Lth order $(2M \le L \le N - 2M)$ forward and backward linear prediction equations, in matrix form, is given by

$$y(L)$$
 $y(L-1)$
 \cdots
 $y(1)$
 g_1
 $y(L+1)$
 $y(L)$
 \cdots
 $y(2)$
 g_2
 \vdots
 \vdots
 \vdots
 \vdots
 $y(N-1)$
 $y(N-2)$
 \cdots
 $y(N-L)$
 $(y(2)$
 $y(3)$
 \cdots
 $y(L+1)$
 $y(3)$
 $y(4)$
 \cdots
 $y(L+2)$
 \vdots
 \vdots
 \vdots
 \vdots
 $y(N-L+1)$
 $y(N-L+2)$
 $y(N)$
 g_L

$$= \begin{pmatrix} y(L+1) \\ y(L+2) \\ \vdots \\ y(N) \\ \cdots \\ y(1) \\ y(2) \\ \vdots \\ y(N-L) \end{pmatrix} - \begin{pmatrix} e_f(L+1) \\ e_f(L+2) \\ \vdots \\ e_f(N) \\ \cdots \\ e_b(1) \\ e_b(2) \\ \vdots \\ e_b(N-L) \end{bmatrix}$$

which can be expressed compactly as

$$Ag = y - \xi. \tag{2.3}$$

(2.2)

For FLP/BLP, the data matrix A, the predicted data vector y, and the prediction error vector ξ will be first/second half (above/below dotted line) in (2.2) with sizes $(N - L) \times L$, (N - L), and (N - L), respectively.

In MFLP, MBLP, and MFBLP methods, the matrix A is replaced by its 2*M*-rank approximant. This requires an *a priori* knowledge of the number of sinusoids or a method which derives this information from the given data.

The Akaike's criterion for model selection selects the model which minimizes the AIC, defined by

AIC =
$$-2 \ln (f(y|\hat{\theta})) + 2d$$
 (2.4)

where $f(y|\hat{\theta})$ is the maximum of the likelihood function of the observations and d is the number of free parameters in θ . The MDL criterion selects the model which minimizes the MDL, defined by

$$MDL = -2 \ln (f(y|\hat{\theta})) + d \ln N \qquad (2.5)$$

where N denotes the number of observations in y (i.e., the length of the vector y).

A. The Detection Criteria

To apply the information theoretic approach, it is necessary to choose an appropriate family of density functions for the observed data $y(1), y(2), \dots, y(N)$. From (2.1), the data y(n) can be expressed as

$$y(n) = x(n) + v(n)$$
 (2.6)

where x(n) is deterministic for a given set of initial phases $\{\phi_i\}$.

Now recall that the samples of a signal consisting of M real sinusoids (damped/undamped) satisfy a homogeneous difference equation of order $L \ge 2M$. The set of forward and backward homogeneous difference equations, in matrix form, is given by

$$A_{\mathbf{x}}\mathbf{g} = \mathbf{x} \tag{2.7}$$

where A_x and x are same as A and y, respectively, with y(n)'s replaced by x(n)'s, and g is the vector of coefficients. We note here that the g in (2.7) belongs to an affine subspace. From (2.6) and (2.7), we have

$$\mathbf{y} = A_{\mathbf{x}}\mathbf{g} + \mathbf{v} \tag{2.8}$$

where (with superscript T denoting the transpose)

$$v = [v(L + 1) \quad v(L + 2) \cdots v(N) \quad v(1) \quad v(2) \cdots \quad v(N - L)]^{T}.$$
 (2.9)

We may point out here that similar relation holds for FLP/ BLP excepting that A_x , y and v are the respective half portions of those corresponding to FBLP. Since $A_x g$ is deterministic for a given set of initial phases $\{\phi_i\}$, the family of probability density functions for the vector y can be found from the density function of the vector v.

We note from (2.9) that for $(N - L) \ge (L + 1)$, some elements in v repeat making the covariance matrix of vrank deficient. Hence, the density function of v does not exist for this case. However, no problem arises for the case (N - L) < (L + 1) in FBLP and for any L in FLP and BLP. Since the MFBLP has been shown to perform better when L is in the range (2N/3, 3N/4) [7] and these values of L correspond to the condition (N - L) < (L +1), we assume in the following development that L is such that the density function of v exists. The development of the criteria for the case $(N - L) \ge (L + 1)$ is discussed in [15].

Thus, the parametrized density function of y with parameters g and σ^2 is given by

$$f(\mathbf{y}|\mathbf{g}, \sigma^2) = (2\pi\sigma^2)^{-r/2}$$
$$\cdot \exp\left(-\frac{(\mathbf{y} - A_x\mathbf{g})^T(\mathbf{y} - A_x\mathbf{g})}{2\sigma^2}\right)$$
(2.10)

where

$$r = \begin{cases} 2(N-L) & \text{for FBLP} \\ (N-L) & \text{for FLP and BLP.} \end{cases}$$
(2.11)

In practice, we do not have the matrix A_x and what is available to us is the data-matrix A. We now make an approximation by replacing A_x with 2*M*-rank approximate of A. Note that this approximation becomes better for large SNR and indeed, it becomes exact when noise is absent. Since the number of sinusoids is not known *a priori*, we assume that there are k sinusoids and replace A_x in (2.10) by the 2k-rank approximate of A, which we denote by \tilde{A}_{2k} . Thus, the approximate family of density functions is given by

$$f(\mathbf{y}|\mathbf{g}, \sigma^2) = (2\pi\sigma^2)^{-r/2}$$

$$\cdot \exp\left(-\frac{(\mathbf{y} - \tilde{A}_{2k}\mathbf{g})^T(\mathbf{y} - \tilde{A}_{2k}\mathbf{g})}{2\sigma^2}\right).$$
(2.12)

Now, consider the negative of the logarithm of (2.12)

$$L^{(k)}(\mathbf{g}, \sigma^{2}) = -\ln \left(f(\mathbf{y} | \mathbf{g}, \sigma^{2}) \right)$$

= $\frac{r}{2} \ln 2\pi + \frac{r}{2} \ln \sigma^{2} + \frac{\|\mathbf{y} - \tilde{A}_{2k}\mathbf{g}\|^{2}}{2\sigma^{2}}$
(2.13)

where $\|\cdot\|$ denotes the 2-norm. The ML estimates of g and σ^2 are the values of g and σ^2 that minimize $L^{(k)}(g, \sigma^2)$, and are given by

$$\hat{\sigma}^2 = \frac{\|\mathbf{y} - \tilde{A}_{2k} (\tilde{A}_{2k}^T \tilde{A}_{2k})^{\#} \tilde{A}_{2k}^T \mathbf{y}\|^2}{r}$$
(2.14)

$$\hat{\boldsymbol{g}} = (\tilde{A}_{2k}^T \tilde{A}_{2k})^{\#} \tilde{A}_{2k}^T \boldsymbol{y}.$$
(2.15)

where the superscript denotes pseudoinverse. Observe that the ML estimate \hat{g} is the minimum-norm solution for the predictor coefficient vector in MFLP/MBLP/MFBLP with the assumption that the number of sinusoids is k. We may point out here that the above estimates also follow from the development in [7], [8]. From (2.14), (2.15), and (2.13), we get

$$L^{(k)}(\hat{g}, \hat{\sigma}^{2}) = \frac{r}{2} \left\{ 1 + \ln 2\pi + \ln \frac{\|y - \tilde{A}_{2k}(\tilde{A}_{2k}^{T} \tilde{A}_{2k})^{\#} \tilde{A}_{2k}^{T} y\|^{2}}{r} \right\}.$$
 (2.16)

To simplify the last term in (2.16), we use the SVD of \tilde{A}_{2k}

$$\tilde{A}_{2k} = \hat{U} \begin{pmatrix} \Sigma_{2k} & 0\\ 0 & 0 \end{pmatrix} \hat{V}^T$$
(2.17)

where \hat{U} and \hat{V} are the unitary matrices of size $r \times r$ and $L \times L$, containing the left and right singular vectors, respectively, of A and Σ_{2k} is a $2k \times 2k$ diagonal matrix whose diagonal elements are the 2k largest singular values

of A. Substituting (2.17) into the norm-squared term in (2.16) and noting that 2-norm is invariant under unitary transformation, we get

$$\|\cdot\|^{2} = \left\| \hat{U}^{T} \mathbf{y} - \hat{U}^{T} \hat{U} \begin{pmatrix} I_{2k} & 0 \\ 0 & 0 \end{pmatrix} \hat{U}^{T} \mathbf{y} \right\|^{2}$$
$$= \left\| \begin{pmatrix} \mathbf{c} \\ \mathbf{h}^{2k} \end{pmatrix} - \begin{pmatrix} I_{2k} & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} \mathbf{c} \\ \mathbf{h}^{2k} \end{pmatrix} \right\|^{2} = \|\mathbf{h}^{2k}\|^{2}$$
(2.18)

where h^{2k} is the vector containing the last (r - 2k) elements of $\hat{U}^T y$

$$\hat{U}^{T} \mathbf{y} = \begin{pmatrix} \mathbf{c} \\ \mathbf{h}^{2k} \end{pmatrix} \{2k \text{ elements} \\ \{(r - 2k) \text{ elements.} \end{cases}$$
(2.19)

Thus, (2.16) simplifies to

$$L^{(k)}(\hat{g}, \,\hat{\sigma}^2) = \frac{r}{2} \left\{ 1 \,+\, \ln 2\pi \,+\, \ln \frac{\|\boldsymbol{h}^{2k}\|^2}{r} \right\}.$$
 (2.20)

To determine the number of degrees of freedom in the vector \hat{g} , consider (2.15). Combining this with (2.17) and (2.19), we get

$$\hat{\boldsymbol{g}} = \hat{\boldsymbol{V}} \begin{pmatrix} \boldsymbol{\Sigma}_{2k}^{-1} \boldsymbol{c} \\ \boldsymbol{0} \end{pmatrix}$$
(2.21)

which shows that \hat{g} is the linear combination of the first 2k vectors of \hat{V} , implying that \hat{g} has 2k degrees of freedom. We note here that the columns of \hat{V} remain the same and that indeed \hat{g} is a linear combination of a varying number of them, hence the indicated number of degrees of freedom. Thus, under the assumption of k sinusoids, the number of free parameters in $\hat{\theta} = (\hat{g}, \hat{\sigma}^2)$ is (2k + 1). Using this value in place of d and r in place of N in (2.4) and (2.5), and combining with (2.20), we obtain

AIC (k) =
$$r \left\{ 1 + \ln 2\pi + \ln \frac{\|\boldsymbol{h}^{2k}\|^2}{r} \right\} + 2(2k + 1)$$

(2.22)

MDL (k) =
$$r \left\{ 1 + \ln 2\pi + \ln \frac{\|\boldsymbol{h}^{2k}\|^2}{r} \right\}$$

+ (2k + 1) ln r. (2.23)

B. Computational Load of the Criteria

The method requires one SVD computation of A and evaluation of $\|\mathbf{h}^{2k}\|^2$ for different values of k. The computation of $\|\mathbf{h}^{2k}\|^2$ can be implemented recursively as follows.

After evaluating \hat{U} through SVD, compute the squares of the elements of $\hat{U}^T y$ and store them. Let $(\hat{U}^T y)_i$ denote

2874

the *i*th term of $\hat{U}^T \mathbf{y}$. We can then use the following recursions to evaluate $\|\boldsymbol{h}^{2k}\|^2$:

$$\|\boldsymbol{h}^{0}\|^{2} = \sum_{i=1}^{r} \{ (\hat{\boldsymbol{U}}^{T} \boldsymbol{y})_{i} \}^{2}$$
(2.24)

$$\|\boldsymbol{h}^{2k}\|^2 = \|\boldsymbol{h}^{2k-2}\|^2 - \{(\hat{\boldsymbol{U}}^T\boldsymbol{y})_{2k-1}\}^2 - \{(\hat{\boldsymbol{U}}^T\boldsymbol{y})_{2k}\}^2,\$$

$$k = 1, 2, \cdots, \frac{r}{2} - 1.$$
 (2.25)

Thus, the additional computations needed over and above those required for the SVD calculation are marginal. In contrast, the method proposed by Fuchs [2] requires computations for 1) estimating the autocorrelation matrix of the observed data, 2) eigen-decomposition of the autocorrelation matrix, 3) estimating the covariance matrix of the perturbed eigenvalues of the autocorrelation matrix, and 4) computing the chi-squared (denoted as χ^2) random variable. Clearly, this method's computational load is much higher than ours. Furthermore, Fuchs method involves testing of chi-squared random variable against a threshold, thus requiring a subjective decision of the designer. However, it is the first dedicated method to this problem and the resulting test is theoretically justified asymptotically in the number of data.

Each of the criteria, (2.22) and (2.23), is evaluated for different values of k and the k for which the criterion attains minimum is taken as the estimate of the number of sinusoids as given by that criterion. Now, the issue is up to what value of k should we evaluate the criteria.

We note that as k increases, the norm square of the vector h^{2k} , $||h^{2k}||^2$, decreases and reaches zero value when k = r/2, resulting in the minimum value for both the criteria at r/2. We therefore have to restrict k to less than r/2. Normally, k should be varied up to L/2 since an Lth order predictor can handle up to L/2 real sinusoids. In practice, however, the predictor order L is chosen significantly higher than the minimum required value 2M, and it is usually selected such that the data matrix A is nearly square (when the modified methods are known to perform well). For such values of L, we suggest the evaluation of the criteria from k = 1 to L/3. In our study, we also consider a reasonably small value¹ of L which gives rise to an overdetermined set of equations in (2.3). For this value of L, we suggest the search up to L/2. We should, however, point out that when L is chosen such that (N - L) $L \ge (L + 1)$, we need to use certain linear transformations in the development of the criteria if we are working with MFBLP (see [15]). Though we have not considered such a case in the following analysis, it can be handled easily by modifying the intermediate steps.

¹This case is considered because of the remark made by one of the reviewers.

III. PERFORMANCE ANALYSIS OF THE PROPOSED CRITERIA

In this section, we develop an analytical framework for analyzing the performance of the criteria derived in Section II. To develop a common framework for both criteria, we denote the RHS of (2.22) and (2.23) as

$$\Phi(k) = r \left\{ 1 + \ln 2\pi + \ln \frac{\|\boldsymbol{h}^{2k}\|^2}{r} \right\} + p(k) \quad (3.1)$$

where p(k) is the bias correction.

Let H_M denote the hypothesis that the true number of sinusoids is M. Then the probabilities of underestimating and overestimating the number of sinusoids, given H_M , are defined, respectively, as

$$P_M = P(M < M | H_M) \tag{3.2}$$

$$P_F = P(\hat{M} > M \mid H_M) \tag{3.3}$$

where $P(\cdot)$ denotes the probability of an event and \hat{M} denotes the estimate of M. Following [3] and [4], we make the assumptions as below:

$$P(\hat{M} \le M - 2 | H_M) \iff P(\hat{M} = M - 1 | H_M)$$
 (3.4)

$$P(\hat{M} \ge M + 2 | H_M) \iff P(\hat{M} = M + 1 | H_M).$$
 (3.5)

It has been found via extensive computer simulations that the above assumptions are indeed true for the MDL criterion when the SNR is moderate and higher.

Define the probability of error P_e as

$$P_e = P_M + P_F \tag{3.6}$$

where P_M is the probability of miss and P_F is the probability of false alarm. Following the development as in [4], we can show that

$$P_e \approx P(B_1) + P(B_2) - P(B_1 \cap B_2)$$
 (3.7)

where the events B_1 and B_2 are given by

$$B_1: \Phi(M - 1) < \Phi(M)$$
 (3.8)

$$B_2: \Phi(M + 1) < \Phi(M).$$
 (3.9)

In the sequel, we do not show the conditioning on H_M as this has already been reflected in the specification of the events.

From (3.1) and (3.8), we have

$$P(B_1) = P\left(r\left\{1 + \ln 2\pi + \ln \frac{\|\boldsymbol{h}^{2(M-1)}\|^2}{r}\right\} + p(M-1)$$

< $r\left\{1 + \ln 2\pi + \ln \frac{\|\boldsymbol{h}^{2M}\|^2}{r}\right\} + p(M)\right)$ (3.10)

which can be simplified as

$$P(B_{1}) = P\left(\frac{\{(\hat{U}^{T}y)_{2M-1}\}^{2} + \{(\hat{U}^{T}y)_{2M}\}^{2}}{\sum\limits_{i=2M+1}^{r} \{(\hat{U}^{T}y)_{i}\}^{2}} < e^{(p(M) - p(M-1))/r} - 1\right).$$
(3.11)

Similarly, combining (3.1) and (3.9), we get

$$P(B_{2}) = P\left(\frac{\{(\hat{U}^{T}\mathbf{y})_{2M+1}\}^{2} + \{(\hat{U}^{T}\mathbf{y})_{2M+2}\}^{2}}{\sum_{i=2M+3}^{r} \{(\hat{U}^{T}\mathbf{y})_{i}\}^{2}} > e^{(p(M+1)-p(M))/r} - 1\right).$$
(3.12)

To evaluate $P(B_1)$ and $P(B_2)$, we need to determine the density functions of the quotients in (3.11) and (3.12).

Note that \hat{U} is the matrix containing the left singular vectors of A or the matrix containing the eigenvectors of AA^T/L . Now, consider

$$\frac{1}{L}AA^{T} = \frac{1}{L}A_{x}A_{x}^{T} + \frac{1}{L}A_{x}A_{v}^{T} + \frac{1}{L}A_{v}A_{x}^{T} + \frac{1}{L}A_{v}A_{v}^{T}$$
(3.13)

where A_x and A_v are the signal only and noise only parts of A, respectively. For moderate values of L, second and third terms become relatively small as the SNR increases. For a finite but moderate SNR, the fourth term is closer to $\sigma^2 I$ in the FLP/BLP case than in the FBLP case because of the following. In FBLP case, $(1/L)A_vA_v^T$ consists of four $(N - L) \times (N - L)$ block matrices where the two diagonal blocks will be closer to $\sigma^2 I$ and the other two off-diagonal blocks contain an element such as $(1/L)v^2(\cdot)$ in addition to cross terms in alternate positions. On the other hand, in FLP/BLP case, $(1/L)A_vA_v^T$ is simply the first/second diagonal block of the FBLP. In the special case of no noise, the last three terms in (3.13) vanish.

Keeping the above arguments in view, we make the following approximation:

$$\frac{1}{L}AA^{T} \approx \frac{1}{L}A_{x}A_{x}^{T} + \sigma^{2}I \qquad (3.14)$$

which is more justifiable in the FLP/BLP case than in the FBLP case. For the given initial phases of the sinusoids and noise variance, the RHS of (3.14) is known and its eigenvector matrix, say U, can be determined. Thus, U is independent of the noise realization. Thus, for the given initial phases of the sinusoids, U can be treated as constant. Now, approximating \hat{U} to U in (3.11) and (3.12), and noting that y(n) is Gaussian with mean x(n) and variance σ^2 , we have

$$E[(U^{T}\mathbf{y})_{i}] = E[\mathbf{u}_{i}^{T}\mathbf{y}] = \mathbf{u}_{i}^{T}E[\mathbf{y}]$$
$$= \mathbf{u}_{i}^{T}\mathbf{x}, \quad 1 \le i \le r \qquad (3.15)$$

$$\operatorname{var} \left((U^T \mathbf{y})_i \right) = E[\mathbf{u}_i^T \boldsymbol{v} \boldsymbol{v}^T \boldsymbol{u}_i] = \mathbf{u}_i^T E[\boldsymbol{v} \boldsymbol{v}^T] \mathbf{u}_i$$
$$= \sigma^2, \quad 1 \le i \le r. \quad (3.16)$$

Note that x, like A_x , depends on the initial phases of the sinusoids.

Under the hypothesis H_M , the eigenvectors u_1 , u_2 , \cdots , u_{2M} span the true signal subspace and the remaining u_{2M+1} , u_{2M+2} , \cdots , u_r span the true noise subspace. It therefore follows that

$$E[(U^T \mathbf{y})_i] = \mathbf{u}_i^T \mathbf{x}, \qquad 1 \le i \le 2M \qquad (3.17)$$

and

$$E[(U^T \mathbf{y})_i] = \mathbf{u}_i^T \mathbf{x} = 0, \quad 2M + 1 \le i \le r.$$
 (3.18)

Further, from the orthogonality of the singular vectors, we get

$$\operatorname{cov} ((\boldsymbol{U}^{T}\boldsymbol{y})_{i}, (\boldsymbol{U}^{T}\boldsymbol{y})_{j})$$

= $E[\boldsymbol{u}_{i}^{T}\boldsymbol{v}\boldsymbol{v}^{T}\boldsymbol{u}_{j}] = \boldsymbol{u}_{i}^{T}E[\boldsymbol{v}\boldsymbol{v}^{T}]\boldsymbol{u}_{j} = 0,$
 $1 \leq i, j \leq r, \quad i \neq j.$ (3.19)

Since y is Gaussian, (3.19) implies that the components of $(U^T y)$ are independent.

Now, consider

$$Z_{1} = \frac{\{(U^{T} \mathbf{y})_{2M-1}\}^{2} + \{(U^{T} \mathbf{y})_{2M}\}^{2}}{\sum_{i=2M+1}^{r} \{(U^{T} \mathbf{y})_{i}\}^{2}} = \frac{S_{1}}{T_{1}} \quad (\text{say}).$$
(3.20)

It is easy to see that S_1 is a noncentral χ^2 random variable with 2 degrees of freedom. Replacing q, λ , and n in (A.4) (see Appendix A) with s_1 , λ_1 , and 2, respectively, the density function of S_1 is obtained as

$$f_{S_{1}}(s_{1}) = \frac{e^{-s_{1}/2\sigma^{2}}e^{-\lambda_{1}/2\sigma^{2}}}{2\sigma^{2}} \left\{ 1 + \frac{1}{2} \left(\frac{s_{1}\lambda_{1}}{2\sigma^{4}} \right) + \frac{1}{2 \cdot 4} \frac{1}{2!} \left(\frac{s_{1}\lambda_{1}}{2\sigma^{4}} \right)^{2} + \cdots \right\}$$

$$0 \le s_{1} \le \infty$$
(3.21)

where

$$\lambda_1 = \{ (U^T \mathbf{x})_{2M-1} \}^2 + \{ (U^T \mathbf{x})_{2M} \}^2.$$
 (3.22)

Now, consider the random variable T_1 . From (3.16), (3.18), and (3.19), it follows that T_1 is a χ^2 random variable with (r - 2M) degrees of freedom and, hence, its density function is given by

$$f_{T_{1}}(t_{1}) = \frac{t_{1}^{(r-2M)/2-1}e^{-t_{1}/2\sigma^{2}}}{(2\sigma^{2})^{(r-2M)/2}}\Gamma\left(\frac{r-2M}{2}\right)$$
$$0 \le t_{1} < \infty.$$
(3.23)

Since S_1 and T_1 are independent, the density function of $Z_1 = S_1/T_1$ is given by

$$f_{Z_1}(z_1) = \int_0^\infty t_1 f_{S_1}(t_1 z_1) f_{T_1}(t_1) dt_1 \qquad (3.24)$$

which, using (3.21) and (3.23), can be simplified to

$$f_{Z_{1}}(z_{1}) = \frac{r - 2M}{2} e^{-\lambda_{1}/2\sigma^{2}} (1 + z_{1})^{-((r - 2M)/2 + 1)}$$

$$\cdot \left\{ 1 + \frac{r - 2M}{2} + \frac{1}{2} \left(\frac{z_{1}\lambda_{1}}{\sigma^{2}} \right) (1 + z_{1})^{-1} + \frac{\left(\frac{r - 2M}{2} + 1 \right) \left(\frac{r - 2M}{2} + 2 \right)}{2 \cdot 4} + \frac{1}{2!} \left(\frac{z_{1}\lambda_{1}}{\sigma^{2}} \right)^{2} (1 + z_{1})^{-2} + \cdots \right\}$$

$$0 \le z_{1} < \infty. \qquad (3.25)$$

Observe that $f_{Z_1}(z_1)$ resembles Rician distribution whose peak location from the origin is proportional to the non-centrality parameter λ_1/σ^2 .

From (3.11) and (3.20), we have

$$P(B_1) = P(Z_1 < e^{(p(M) - p(M-1))/r} - 1)$$

=
$$\int_0^{e^{(p(M) - p(M-1))/r} - 1} f_{Z_1}(z_1) d_{z_1}.$$
 (3.26)

Substituting (3.25) in (3.26) and evaluating the integral numerically, we obtain $P(B_1)$ for a given set of initial phases of the sinusoids.

Now, consider the density function of the quotient in (3.12). Let

$$Z_{2} = \frac{\{(U^{T}\mathbf{y})_{2M+1}\}^{2} + \{(U^{T}\mathbf{y})_{2M+2}\}^{2}}{\sum_{i=2M+3}^{r} \{(U^{T}\mathbf{y})_{i}\}^{2}} = \frac{S_{2}}{T_{2}} \quad (\text{say}).$$
(3.27)

Here, S_2 and T_2 are χ^2 random variables with 2 and (r - 2(M + 1)) degrees of freedom, respectively. Hence, the density functions of S_2 and T_2 are given by

$$f_{s_2}(s_2) = \frac{e^{-s_2/2\sigma^2}}{2\sigma^2} \quad 0 \le s_2 < \infty$$
 (3.28)

and

$$f_{T_2}(t_2) = \frac{t_2^{(r-2(M+1))/2-1} e^{-t_2/2\sigma^2}}{(2\sigma^2)^{(r-2(M+1))/2} \Gamma\left(\frac{r-2(M+1)}{2}\right)}$$
$$0 \le t_2 < \infty.$$
(3.29)

Using the fact that S_2 and T_2 in (3.27) are independent, the density function of Z_2 can be shown to be

$$f_{Z_2}(z_2) = \left(\frac{r-2(M+1)}{2}\right)(1+z_2)^{-(r-2M)/2}$$
$$0 \le z_2 < \infty.$$
(3.30)

From (3.12), (3.27), and (3.30), the probability of the event B_2 is given by

$$P(B_2) = P(Z_2 > e^{(p(M+1)-p(M))/r} - 1)$$

= $e^{-(r-2(M+1))(p(M+1)-p(M))/2r}$. (3.31)

Note that $P(B_2)$ does not depend on the signal and noise powers; it depends only on the penalty and r (cf. (2.11)).

Consider now the probability of the event $(B_1 \cap B_2)$. From Appendix B,

$$P(B_1 \cap B_2) \le P(Z_3 < e^{(p(M) - p(M-1))/r})P(B_2)$$
 (3.32)
where

$$Z_{3} = \frac{S_{1}}{S_{2}} = \frac{\{(U^{T}y)_{2M-1}\}^{2} + \{(U^{T}y)_{2M}\}^{2}}{\{(U^{T}y)_{2M+1}\}^{2} + \{(U^{T}y)_{2M+2}\}^{2}}.$$
 (3.33)

Since S_1 and S_2 are independent, the density function of Z_3 can be shown to be

$$f_{Z_{3}}(z_{3}) = e^{-\lambda_{1}/2\sigma^{2}}(1+z_{3})^{-2} \left\{ 1 + \frac{2}{2} \left(\frac{z_{3}\lambda_{1}}{\sigma^{2}} \right) (1+z_{3})^{-1} + \frac{3}{2 \cdot 4} \left(\frac{z_{3}\lambda_{1}}{\sigma^{2}} \right)^{2} (1+z_{3})^{-2} + \frac{4}{2 \cdot 4 \cdot 6} \left(\frac{z_{3}\lambda_{1}}{\sigma^{2}} \right)^{3} (1+z_{3})^{-3} + \cdots \right\}$$

$$0 \le z_{3} < \infty.$$
(3.34)

From (3.31) and (3.34), (3.32) can be evaluated numerically for a given set of initial phase angles of the sinusoids.

Combining (3.7), (3.26), (3.31), (3.32), and (3.34), the bounds for the probability of detection $(P_D = 1 - P_e)$, given by

$$1 - P(B_1) - P(B_2)$$

$$\leq P_D \leq 1 - P(B_1) - P(B_2)$$

$$+ P(Z_3 < e^{(p(M) - p(M-1))/r})P(B_2), \quad (3.35)$$

can be evaluated.

Note that the bounds for P_D are functions of the initial phase angles of the sinusoids, which are assumed to be uniformly distributed in the interval $[0, 2\pi]$. We evaluated these bounds choosing various sets of initial phase angles in the interval $[0, 2\pi]$ and then determined the unconditional probability of detection through numerical integration. To clarify this further, consider the case of two sinusoids. Without loss of generality, let the initial phase of one of the sinusoids be zero and the phase of the other sinusoid be ϕ . Let $P_{D|\phi}$ denote the probability of detection with these initial phases. Then, the unconditional probability of detection, denoted as $P_{D(uc)}$, is given by

$$P_{D(\mathrm{uc})} = \frac{1}{2\pi} \int_0^{2\pi} P_{D|\phi} \, d\phi \qquad (3.36)$$

since ϕ is assumed to be uniformly distributed in $[0, 2\pi]$. To obtain $P_{D(uc)}$ numerically, we vary the phase ϕ in steps of $\Delta \phi$ (say) covering the range 0 to 2π , and determine the probability of detection at each of these steps, and then use the numerical integration technique to evaluate the definite integral part of (3.36).

IV. SIMULATION RESULTS

In this section, we present some simulation results to demonstrate both the effectiveness of our method and the usefulness of the analysis.

To demonstrate the effectiveness of our method compared to that of Fuchs [2], we considered the same example as the one in [2]. The data samples were generated from

$$y(n) = \sqrt{20} \sin 2\pi f_1 n + \sqrt{2} \sin (2\pi f_2 n + \phi) + v(n),$$

$$n = 1, 2, \cdots, N$$

choosing the normalized frequencies f_1 and f_2 as 0.2 and $0.2 + \delta$, respectively, with $\delta = 1/N$ or 1/2N, and the noise variance, σ^2 , as unity. The number of samples N was chosen as 64. We used the FBLP formulation with L= 32 (Fuchs used the same formulation for estimating the autocorrelation matrix). The phase of the second sinusoid was varied randomly from trial to trial in the interval [0, 2π and the seed for the noise sequence was selected randomly in each trial. We performed 50 Monte Carlo trials and determined the number of times the criteria attained the minimum value at different values of k (see Tables I and II). Note from the tables that the MDL criterion detected correctly 42 times out of 50 when $\delta = 1/64$ and 20 times out of 50 when $\delta = 1/128$, while the Fuchs method gives 18 and 9 times out of 20, respectively, for $\delta = 1/64$ and $\delta = 1/128$ (these results are taken from [2]). This shows that the two methods perform similarly. The simulation results also show that the detection performance of the MDL is much superior to that of AIC, and that the penalty (or the bias correction) used in AIC is too low to hold the criterion from leaning towards higher model order.

We now present some simulation results to demonstrate the usefulness of the analysis developed in Section III. The undamped sinusoids case is considered first.

A. Case i): Undamped Sinusoids

The data model used is given by

$$y(n) = \sqrt{20} \sin 2\pi f_1 n + \sqrt{20} \sin (2\pi f_2 n + \phi) + v(n),$$

$$n = 1, 2, \cdots, N$$

where the frequencies f_1 and f_2 were selected as in the previous example with $\delta = 1/64$, and the variance of the

 TABLE 1

 DETECTION PERFORMANCE OF THE PROPOSED CRITERIA WITH

 FBLP FORMULATION (N = 64, L = 32, $f_1 = 0.2$, $f_2 = 0.2 + 1/64$,

 MONTE CASH OF UNS = 50)

		MONTE CAREO RONS = 50)							
k	0	1	2	3	4	5	6	7	8
AIC MDL	0 0	0 1	14 42	7 7	3 0	7 0	7 0	6 0	6 0

TABLE II
DETECTION PERFORMANCE OF THE PROPOSED CRITERIA WITH
FBLP FORMULATION ($N = 64, L = 32, f_1 = 0.2, f_2 = 0.2 + 1/128$,
MONTE CARLO RUNG $= 50$

	MONTE CARLO KUNS = 50)								
k	0	1	2	3	4	5	6	7	8
AIC MDL	0 0	6 26	12 20	8 1	5 0	4 2	6 1	4 0	5 0

TABLE III DETECTION PERFORMANCE OF THE CRITERIA FOR N = 64 and L = 22 $(f_1 = 0.2, f_2 = 0.2 + 1/64$, Monte Carlo Runs = 500)

		$P_{D(uc)}$ (in p	ercentage)	
	SNR	= 5 dB	SNR	= 10 dB
	Analysis FLP	Simulation FLP	Analysis FLP	Simulation FLP
AIC	81.99	64.0	81.99	68.8
MDL	95.94	89.4	95.94	93.8

noise, σ^2 , was chosen to give the desired signal-to-noise ratio (SNR), defined as SNR = $10 \log_{10} 10/\sigma^2$. Both FLP and FBLP formulations were considered in this example. In each case, we considered two values of *L* (predictor order).

Tables III through V give the probability of detection (in percentage), obtained from the simulations and the analysis. The simulation results were obtained from 500 Monte Carlo trials. In each trial, we chose randomly the seed for the noise realization and for the initial phase in the interval [0, 2π]. For the analytical result, ϕ was varied uniformly in steps of $\pi/18$ and the conditional $P_{D|\phi}$ was computed for each setting of ϕ as discussed in the previous section. The unconditional probability of detection was then evaluated from these 36 conditional results through the numerical integration as described in the last section (see the discussion in the last paragraph of Section III). We may point out here that the value of $P(Z_3 <$ $e^{(p(M) - p(M-1))/r} P(B_2)$ was negligibly small and hence the probability of detection was essentially determined by $P(B_1)$ and $P(B_2)$, which are now (for all practical purposes) the probabilities of miss and false alarm, respectively.

We note the following from the results of the tables. The predicted value of probability of detection is closer

TABLE IVDETECTION PERFORMANCE OF THE CRITERIA FOR N = 64 and L = 32 $(f_1 = 0.2, f_2 = 0.2 + 1/64, MONTE CARLO RUNS = 500)$

				$P_{D(uc)}$ (in p	ercentage)				
		SNR = 5 dB				SNR = 10 dB			
	Ana	lysis	Sim	ulation	Ana	lysis	Sim	ulation	
	FLP	FBLP	FLP	FBLP	FLP	FBLP	FLP	FBLP	
AIC MDL	80.31 94.00	83.68 97.69	54.0 88.4	28.2 82.4	80.31 94.02	83.68 97.69	56.6 91.0	29.2 83.8	

TABLE VDETECTION PERFORMANCE OF THE CRITERIA FOR N = 64 and L = 42 $(f_1 = 0.2, f_2 = 0.2 + 1/64, MONTE CARLO RUNS = 500)$

	$P_{D(uc)}$ (in percentage)				
	= 10 dB	= 5 dB SNR			
	Simulation FBLP	Simulation Analysis FBLP FBLP			
AIC MDL	32.4 85.4	31.2 82.22 81.0 96.18			
AIC MDL		31.2 82.22 81.0 96.18			

to the simulated value when the SNR is larger. This is expected since the approximations we made in the analysis become better for larger SNR. In the case of AIC criterion, the predicted values differ widely from those of the simulated. This is because the assumption made in (3.5) is not satisfied in this case, as is evident from the results of Tables I and II.

Consider the results for FLP and FBLP, given in Tables III and IV for two different values of SNR. The analytical value of P_D remains same when the SNR is increased to 10 dB, keeping the value of L fixed. The reasons for this are as follows. Recall that $P(B_2)$ depends only on the penalty and r (which is (N - L) for FLP and 2(N - L) for FBLP), and hence, does not change with SNR. Now, refer to $P(B_1)$ given by (3.26). Recall that $f_{Z_1}(z_1)$ resembles Rician distribution whose peak location from the origin is proportional to the noncentrality parameter λ_1/σ^2 (see (3.25)). For the cases under consideration, this parameter is very large for SNR = 5 dB, implying that the location of the peak is very far from the origin. On the other hand, the upper limit of the integral in (3.26)is closer to the origin. Consequently, the value of the integral, and hence $P(B_1)$, is nearly equal to zero (as observed during the analytical evaluation). Any further increase in the SNR will not cause any noticeable change in $P(B_1)$ and as a result, P_D remains nearly the same when we increase the SNR.

In the case of FBLP with L = 42 (see Table V), the value of P_D increases when the SNR is increased to 10 dB. This increase is obviously caused due to the reduction in $P(B_1)$ since $P(B_2)$ is independent of the SNR. In this case, the noncentrality parameter is not very large for SNR = 5 dB, and hence, the increase in this parameter value

due to the enhanced SNR causes a reduction in the value of $P(B_1)$.

Now consider the effect of increase in the predictor order (L) on the theoretical value of P_D . The results of the Tables III to V show that P_D decreases with increasing L. The reasons for this are as follows. A closer examination of the expression for $P(B_2)$ (cf. (3.31)) reveals that its value increases when r decreases. A similar effect takes place with $P(B_1)$ even though it is difficult to see this from the expressions (3.25) and (3.26). Since an increase in L causes a decrease in r, it immediately follows from the above that P_D decreases with increasing L.

We may point out here that the theoretical and simulated values of P_D (for MDL) are closer in the case of FLP than in FBLP for the reasons given in the paragraph preceding the approximation (3.14), i.e., for a given L and SNR, the approximation is more justifiable in FLP/BLP case than in FBLP case. Further, the SNR of 5 dB is very close to threshold SNR of the MFLP method for the scenarios considered in Tables III and IV. To estimate the threshold SNR, we have applied the MFLP method to the same 500 data realizations as those used in the detection criteria and determined the mean square error performance in the estimates of f_1 and f_2 for different values of SNR. Though our criteria can be applied at lower values of SNR than the threshold, the analysis will not be useful in that region. Away from the threshold SNR, the predicted values are close to the simulated, particularly in the FLP/BLP case. To support this further, we considered a scenario with closer spacing, $f_1 = 0.2$ and $f_2 = 0.2 +$ 1/128. Table VI gives the results for this case which show the closeness between the theoretical and simulated values of P_D for MDL. Once again, the SNR of 15 dB is approximately equal to the threshold SNR of the MFLP method for the scenario of Table VI.

B. Case ii): Damped Sinusoids

The data model used is given by

$$y(n) = \sqrt{20} e^{-0.1n} \sin 2\pi 0.2n + \sqrt{20} e^{-0.05n} \sin (2\pi 0.24n + \phi) + v(n), n = 1, 2, \cdots, N$$

where the variance of the noise, σ^2 , was chosen to give the required SNR, defined as $10 \log_{10} 10/\sigma^2$. Unlike case

TABLE VIDETECTION PERFORMANCE OF THE CRITERIA FOR N = 64 and L = 32 $(f_1 = 0.2, f_2 = 0.2 + 1/128$, MONTE CARLO RUNS = 500)

	$P_{D(uc)}$ (in SNR	percentage) = 15 dB
	Analysis FLP	Simulation FLP
AIC MDL	80.31 94.01	58.6 87.8

Detecti BLP Form	TABLE VII ON PERFORMANCE OF THE OULATION ($N = 64$, $L = 32$ RUNS = 500)	Criteria with , Monte Carlo		
	$P_{D(uc)}$ (in SNR ₀	$P_{D(uc)}$ (in percentage) SNR ₀ = 20 dB		
	Analysis	Simulation		
AIC	80.31	53.2		
MDL	94.02	89.0		

SNR₀ refers to the SNR of zeroth sample.

i), the SNR here refers to that of the zeroth sample and not to the average over the entire data record. We therefore refer to this as zeroth sample SNR and denote it by SNR₀. We considered the BLP formulation with N = 64 and L = 32.

Table VII gives the values of P_D obtained from the simulation and analysis. The results show that the predicted value is closer to that of the simulated for the MDL case and it differs widely from the simulated in the case of AIC for the same reason given in case i).

We may point out here that the SNR value of 20 dB may appear rather large. But, this is the SNR of zeroth sample computed as 10 \log_{10} (square of the sample amplitude/2)/ σ^2 . For the sinusoid with damping factor -0.1 considered here, the SNR of 32nd sample is -7.8 dB and that of the 64th sample is -35.5 dB. Also, as in the case of FLP, we have estimated the threshold SNR of the MBLP method for this scenario, and SNR₀ = 20 dB is slightly above the threshold value.

V. CONCLUSIONS

In this paper, following the information theoretic approach to model selection, we first developed criteria for detecting the number of damped/undamped sinusoids. These are so well matched to the SVD-based methods of spectral estimation that the additional computations required over and above those needed for the SVD calculation are marginal. The detection performance of these criteria is compared to that of Fuchs [2] and the results show that MDL criterion, developed in the paper, performs nearly same as the Fuchs method. However, Fuchs methods requires much more computations. Next, an analytical framework has been developed for predicting the detection performance of the criteria. In the development of the analysis, we made some approximations which be-

come better for large SNR. Simulation results show that the theoretically predicted values of probability of detection and those of simulation are close with FLP/BLP formulation even at moderate values of SNR.

Appendix A

Let $w_i \sim N(\mu_i, \sigma^2)$, $i = 1, 2, \dots, n$ and assume that they are statistically independent. Then, the density function of $w = \sum_{i=1}^{n} w_i^2 / \sigma^2$ is noncentral χ^2 , having *n* degrees of freedom with the noncentrality parameter λ / σ^2 $= \sum_{i=1}^{n} \mu_i^2 / \sigma^2$ [16], and is given by

$$f_{W}(w) = e^{-\lambda/2\sigma^{2}} \sum_{j=0}^{\infty} \frac{1}{j!} \left(\frac{\lambda}{2\sigma^{2}}\right)^{j} G\left(w|\frac{1}{2}, j+\frac{n}{2}\right)$$
$$0 \le w < \infty \qquad (A.1)$$

where

$$G\left(w|\frac{1}{2}, j+\frac{n}{2}\right) = \frac{\binom{1}{2}^{j+n/2}}{\Gamma\left(j+\frac{n}{2}\right)} e^{-w/2} w^{j+n/2-1}$$
$$0 \le w < \infty$$
(A.2)

with Γ denoting the gamma function. Substituting (A.2) into (A.1), we get

$$f_{W}(w) = \frac{e^{-w/2}e^{-\lambda/2\sigma^{2}}w^{n/2-1}}{2^{n/2}\Gamma\left(\frac{n}{2}\right)} \left\{ 1 + \frac{1}{n}\left(\frac{w\lambda}{2\sigma^{2}}\right) + \frac{1}{n(n+2)}\frac{1}{2!}\left(\frac{w\lambda}{2\sigma^{2}}\right)^{2} + \frac{1}{n(n+2)(n+4)}\frac{1}{3!}\left(\frac{w\lambda}{2\sigma^{2}}\right)^{3} + \cdots \right\}$$

$$0 \le w < \infty.$$
(A.3)

Now, the density function of $\sigma^2 w = q$ (say) immediately follows from (A.3) as

$$f_{Q}(q) = \frac{e^{-q/2\sigma^{2}}e^{-\lambda/2\sigma^{2}}q^{n/2-1}}{(2\sigma^{2})^{n/2}\Gamma\left(\frac{n}{2}\right)} \left\{ 1 + \frac{1}{n}\left(\frac{q\lambda}{2\sigma^{4}}\right) + \frac{1}{n(n+2)}\frac{1}{2!}\left(\frac{q\lambda}{2\sigma^{4}}\right)^{2} + \frac{1}{n(n+2)(n+4)}\frac{1}{3!}\left(\frac{q\lambda}{2\sigma^{4}}\right)^{3} + \cdots \right\}$$

$$0 \le q < \infty.$$
(A.4)

APPENDIX B

From (3.11), (3.12), (3.20), and (3.27), the probability of the event $(B_1 \cap B_2)$ is given by

$$P(B_1 \cap B_2) = P\left(\left\{\frac{S_1}{S_2 + T_2} < e^{(p(M) - p(M-1))/r} - 1\right\}\right)$$
$$\cap\left\{\frac{S_2}{T_2} > e^{(p(M+1) - p(M))/r} - 1\right\}\right) (B.1)$$

where we used the fact that $T_1 = S_2 + T_2$. It is easy to verify that

$$e^{(p(M)-p(M-1))/r} - 1 = e^{(p(M+1)-p(M))/r} - 1 = \mu$$
 (say).
(B.2)

Substituting (B.2) into (B.1) and simplifying, we get

$$P(B_{1} \cap B_{2}) = P\left(\frac{S_{1}}{S_{2} + T_{2}} < \mu \text{ and } \frac{S_{2}}{T_{2}} > \mu\right)$$

$$= P\left(\frac{S_{1}}{\mu S_{2}} - 1 < \frac{T_{2}}{S_{2}} \text{ and } \frac{T_{2}}{S_{2}} < \frac{1}{\mu}\right),$$

$$S_{2}, T_{2} \text{ and } \mu \text{ are } + ve$$

$$= P\left(\frac{S_{1}}{\mu S_{2}} - 1 < \frac{T_{2}}{S_{2}} \middle| \frac{T_{2}}{S_{2}} < \frac{1}{\mu}\right)$$

$$\cdot P\left(\frac{T_{2}}{S_{2}} < \frac{1}{\mu}\right)$$

$$\leq P\left(\frac{S_{1}}{\mu S_{2}} - 1 < \frac{1}{\mu}\right) P\left(\frac{S_{2}}{T_{2}} > \mu\right)$$

$$= P\left(\frac{S_{1}}{S_{2}} < \mu + 1\right) P(B_{2}). \quad (B.3)$$

Denoting $Z_3 = S_1/S_2$ and substituting for μ from (B.2) into (B.3), we get

$$P(B_1 \cap B_2) \leq P(Z_3 < e^{(p(M) - p(M-1))/r})P(B_2).$$
 (B.4)

ACKNOWLEDGMENT

The authors wish to thank the reviewers for their critical and helpful remarks which improved the quality of the paper.

REFERENCES

- M. Wax and T. Kailath, "Detection of signals by information theoretic criteria," *IEEE Trans. Acoust.*, Speech, Signal Processing, vol. 33, no. 2, pp. 387-392, Apr. 1985.
 J. J. Fuchs, "Estimating the number of sinusoids in additive white
- [2] J. J. Fuchs, "Estimating the number of sinusoids in additive white noise," *IEEE Trans. Acoust.*, *Speech, Signal Processing*, vol. 36, no. 12, pp. 1846–1853, Dec. 1988.
- no. 12, pp. 1846-1853, Dec. 1988.
 [3] H. Wang and M. Kaveh, "On the performance of signal-subspace processing part I: Narrow-band systems," *IEEE Trans. Acoust.*, *Sneech. Signal Processing*, vol. 34, no. 5, pp. 1201-1209, Oct. 1986.
- Speech, Signal Processing, vol. 34, no. 5, pp. 1201-1209, Oct. 1986.
 [4] Q. Zhang, K. M. Wong, P. C. Yip, and J. P. Reilly, "Statistical analysis of the performance of information theoretic criteria in the detection of the number of signals in array processing," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 37, no. 10, pp. 1557-1567, Oct. 1989.
- [5] S. M. Kay and S. L. Marple, "Spectrum analysis—a modern perspective," *Proc. IEEE*, vol. 69, no. 11, pp. 1380-1419, Nov. 1981.
 [6] T. J. Ulrych and R. W. Clayton, "Time series modelling and maximum an
- [6] T. J. Ulrych and R. W. Clayton, "Time series modelling and maximum entropy," *Phys. Earth Planet Interiors*, vol. 12, pp. 188-200, Aug. 1976.
- [7] D. W. Tufts and R. Kumaresan, "Estimation of frequencies of multiple sinusoids: Making linear prediction perform like maximum likelihood," Proc. IEEE, vol. 70, no. 9, pp. 975-989, Sept. 1982.
 [8] R. Kumaresan and D. W. Tufts, "Estimating the parameters of ex-
- [8] R. Kumaresan and D. W. Tufts, "Estimating the parameters of exponentionaly damped sinusoids and pole-zero modelling in noise," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 30, pp. 833-840, Dec. 1982.

- [9] M. A. Rahman and K.-B. Yu, "Total least squares approach for frequency estimation using linear prediction," *IEEE Trans. Acoust.*, *Speech, Signal Processing*, vol. 35, no. 10, pp. 1440-1454, Oct. 1987.
- [10] M. S. Bartlett, "A note on the multiplying factors for various x² approximations," J. Roy. Stat. Soc., Ser. B, vol. 16, pp. 296-298, 1954.
- [11] D. N. Lawley, "Tests of significance of the latent roots of the covariance and correlation matrices," *Biometrica*, vol. 43, pp. 128-136, 1956.
- [12] H. Akaike, "A new look at the statistical model identification," *IEEE Trans. Automat. Contr.*, vol. 19, pp. 716–723, Dec. 1974.
- [13] G. Schwartz, "Estimating the dimension of a model," Ann. Stat., vol. 6, no. 2, pp. 461-464, 1978.
 [14] J. Rissanen, "Modelling by shortest data description," Automatica,
- [14] J. Rissanen, "Modelling by shortest data description," Automatica, vol. 14, pp. 465-471, 1978.
- [15] L. S. Biradar and V. U. Reddy, "SVD based information theoretic criteria for detection of the number of sinusoids," in *Proc. IEEE ICASSP*, Albuquerque, NM, Apr. 1990, pp. 2519-2522.
- [16] C. R. Rao, Linear Statistical Inference and Its Applications. Wiley Eastern Ltd., 1973.
- [17] B. Porat and B. Friedlander, "On the accuracy of the Kumaresan-Tufts method for estimating complex damped exponentials," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 35, no. 2, pp. 231-235, Feb. 1987.



V. Umapathi Reddy (S'68-M'70-SM'82) received the B.E. and M.Tech degrees in electronics and communication engineering from Osmania University and the Indian Institute of Technology (IIT), Kharagpur, in 1962 and 1963, respectively, and the Ph.D. degree in electrical engineering from the University of Missouri in 1971.

He was an Assistant Professor at IIT, Madras, during 1972-1976 and Professor at IIT, Kharagpur, during 1976-1979. During 1979-1982 and 1986-1987, he was a Visiting Professor at the De-

partment of Electrical Engineering, Stanford University. In April 1982, he joined Osmania University as a Professor, and was the Founder-Director of the Research and Training Unit for Navigational Electronics, funded by the Department of Electronics, Government of India. Since April 1988, he has been with the Indian Institute of Science, Bangalore, as a Professor of Electrical Communication Engineering and is presently its Chairman. He has served as a consultant in signal processing to Avionics Design Bureau of Hindustan Aeronautics Limited, Hyderabad, and to Central Research Laboratory, Bharat Electronics Limited, Bangalore. His recent research interests are in sensitivity analysis of high-resolution algorithms, adaptive algorithms, adaptive arrays, and wavelet transforms.

Dr. Reddy is a Fellow of the Indian Academy of Sciences, Indian National Academy of Engineering, and Indian National Science Academy, and Fellow of the Institute of Electronics and Telecommunication Engineers (IETE), India. He received the S. K. Mitra Memorial Award (1989) from IETE for the best research paper.



L. S. Biradar was born in Gulbarga, India, on January 27, 1960. He received the B.E. degree in electronics and communication engineering from Gulbarga University in 1983, the M. Tech. degree in information and controls from the Indian Institute of Technology, Kanpur, India, in 1986, and the Ph.D. degree from the Indian Institute of Science, Bangalore, India, in 1992.

From 1983 to 1984 and 1986 to 1988 he was a Lecturer in P.D.A. Engineering College, Gulbarga. From 1988 to 1991 he was at the Indian

Institute of Science as a Research Scholar. Presently he is working as an Assistant Professor in P.D.A. Engineering College, Gulbarga. His research interests are in statistical signal processing, parameter estimation, and sensor array processing.