STRONG CONSISTENCY OF MAXIMUM LIKELIHOOD PARAMETER ESTIMATION OF SUPERIMP (U) PITTSBURGH UNIV PA CENTER FOR MULTIVARIATE ANALYSIS Z D BAI ET AL JUN 87
UNCLASSIFIED TR-87-17 AFOSR-TR-87-0973 F49620-85-C-0088 F/G 12/3
Strong consistency of maximum likelihood parameter estimation of superimposed exponential signals in noise

Z.D. Bai, X.R. Chen, P.R. Krishnaiah, Y.H. Wu and L.C. Zhao

Center for Multivariate Analysis
515 Thackeray Hall
University of Pittsburgh, Pittsburgh, PA 15260

Consider the model of multiple superimposed exponential signals in additive Gaussian noise

\[ y_j(t) = \sum_{i=1}^{p} s_{ij} \lambda_i^t + e_j(t), \quad t = 0, 1, \ldots, n-1, \quad j = 1, \ldots, N \]
where $N$ is fixed and $n + \infty$, $\lambda_i = \exp(\sqrt{-1}\omega_i)$, $\omega_i \in [0, 2\pi)$, $i = 1, \ldots, p$, $\omega_i$, $s_{ij}$ are unknown parameters and $p$ is known. Further, $e_j(t) = e_{j1}(t) + \sqrt{2}e_{j2}(t)$, and $e_{j1}(t)$, $e_{j2}(t)$, $t = 0, 1, 2, \ldots$, $j = 1, \ldots, N$, are mutually independent and identically distributed real random variables with a common distribution $N(0, \sigma^2/2)$, $0 < \sigma^2 < \infty$, $\sigma^2$ is unknown. It is shown that if $\omega_i \neq \omega_j$ when $i \neq j$ and $\sum_{j=1}^{N} |s_{ij}| > 0$ for $i = 1, \ldots, p$, then the Maximum Likelihood estimate $(\hat{\lambda}_1, \ldots, \hat{\lambda}_p)$ is strongly consistent. Moreover, it is shown that $\hat{\lambda}_i$ converges to $\lambda_i$ with an exponential rate.
STRONG CONSISTENCY OF MAXIMUM LIKELIHOOD PARAMETER ESTIMATION OF SUPERIMPOSED EXPONENTIAL SIGNALS IN NOISE

Z. D. Bai, X. R. Chen, P. R. Krishnaiah  
Y. H. Wu and L. C. Zhao

Center for Multivariate Analysis  
University of Pittsburgh
STRONG CONSISTENCY OF MAXIMUM
LIKELIHOOD PARAMETER ESTIMATION OF
SUPERIMPOSED EXPONENTIAL SIGNALS IN NOISE*

Z. D. Bai, X. R. Chen, P. R. Krishnaiah
Y. H. Wu and L. C. Zhao

Center for Multivariate Analysis
University of Pittsburgh

June 1987

Technical Report No. 87-17

Center for Multivariate Analysis
Fifth Floor Thackeray Hall
University of Pittsburgh
Pittsburgh, PA 15260

* Research sponsored by the Air Force Office of Scientific Research
under Contract F49620-85-C-0008. The United States Government is
authorized to reproduce and distribute reprints for governmental
purposes notwithstanding any copyright notation hereon.
STRONG CONSISTENCY OF MAXIMUM LIKELIHOOD PARAMETER ESTIMATION OF SUPERIMPOSED EXPONENTIAL SIGNALS IN NOISE

Z. D. Bai, X. R. Chen, P. R. Krishnaiah
Y. H. Wu and L. C. Zhao

ABSTRACT

Consider the model of multiple superimposed exponential signals in additive Gaussian noise

\[ Y_j(t) = \sum_{i=1}^{p} s_{ij} \lambda_i^t + e_j(t), \quad t = 0, 1, \ldots, n-1, \quad j = 1, \ldots, N \]

where N is fixed and \( n \to \infty \), \( \lambda_i = \exp(\sqrt{-1} \omega_i) \), \( \omega_i \in [0, 2\pi) \), \( i = 1, \ldots, p \), \( \omega_i, s_{ij} \) are unknown parameters and p is known. Further, \( e_j(t) = e_{j1}(t) + \sqrt{-1} e_{j2}(t) \), and \( e_{j1}(t), e_{j2}(t), t = 0, 1, 2, \ldots, j = 1, \ldots, N \), are mutually independent and identically distributed real random variables with a common distribution \( N(0, \sigma^2/2) \), \( 0 < \sigma^2 < \infty \), \( \sigma^2 \) is unknown. It is shown that if \( \omega_i \neq \omega_j \) when \( i \neq j \) and \( \sum_{j=1}^{N} |s_{ij}| > 0 \) for \( i = 1, \ldots, p \), then the Maximum Likelihood estimate \((\hat{\lambda}_1, \ldots, \hat{\lambda}_p)\) is strongly consistent. Moreover, it is shown that \( \hat{\lambda}_i \) converges to \( \lambda_i \) with an exponential rate.


Key words and phrases: consistency, exponential rate, Maximum Likelihood estimate, signal processing.

* Research sponsored by the Air Force Office of Scientific Research under Contract F49620-85-C-0008. The United States Government is authorized to reproduce and distribute reprints for governmental purposes notwithstanding any copyright notation hereon.
1. INTRODUCTION

Consider the following model of multiple superimposed exponential signals in additive Gaussian noise

\[ Y_j(t) = \sum_{i=1}^{p} s_{ij} \lambda_i^t + e_j(t), \quad t = 0, 1, \ldots, n-1, \quad j = 1, \ldots, N \quad (1) \]

where \( \lambda_i = \exp(\sqrt{-1} \omega_i), \ i = 1, \ldots, p, \ \omega_i \in [0, 2\pi), \ \omega_i, s_{ij} \) are unknowns and \( p \) is assumed to be known. Further, \( e_j(t) = e_{j1}(t) + \sqrt{\pi} e_{j2}(t) \) and \( e_{j1}(t), e_{j2}(t), \ t = 0, 1, 2, \ldots, j = 1, \ldots, N, \) are mutually independent and identically distributed (iid.) real random variables with a common distribution \( N(0, \sigma^2/2), \ 0 < \sigma^2 < \infty, \sigma^2 \) is unknown.

Quite a number of papers appeared dealing with the estimation of parameters in this model, which is important in problems related to signal processing and time series analysis. When \( \lambda_1, \ldots, \lambda_p \) are known, (1) reduces to an ordinary linear regression model in which \( s_{ij} \)'s are usually estimated by the Least Squares method. Therefore a conceivable way to handle the estimation problem in (1) is as follows: Obtain by some way an estimator \( (\hat{\lambda}_1, \ldots, \hat{\lambda}_p) \) of \( (\lambda_1, \ldots, \lambda_p) \). Substitute \( \hat{\lambda}_i \) for \( \lambda_i \) in (1), consider the \( \hat{\lambda}_i \)'s as known constants and use the LS method to yield an estimate for \( s_{ij} \). This seemingly reasonable procedure has the drawback that the estimate of \( s_{ij} \) thus obtained is usually non-consistent, as indicated in [1].

For the more important problem of estimating \( \lambda_1, \ldots, \lambda_p \), several methods have been proposed in the literature. Bresler and Macovski derived in [2] the LS criterion in the form of minimizing some function not involving \( s_{ij} \). Under the normality assumption here, it is the same as the Maximum Likelihood criterion. Their method consists in introduc-
ing a polynomial $b_0 + b_1z + \ldots + b_pz^p$ having $\lambda_1, \ldots, \lambda_p$ as roots, thus reducing the problem of estimating $\lambda_1, \ldots, \lambda_p$ to that of estimating the coefficient vector $b = (b_0, \ldots, b_p)^\top$. Specifically, define the set

$$B = \left\{ b : \sum_{i=1}^{p} |b_i|^2 = 1, \ b_p \geq 0 \right\}$$

and the $(n-p) \times n$ matrix

$$B_n(b) = \begin{pmatrix} b_0 & b_1 & \cdots & b_p & 0 \\ b_0 & b_1 & \cdots & b_p & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & b_0 & b_1 & \cdots & b_p \end{pmatrix},$$

$$D_n(b) = B_n(b)B_n^*(b),$$

where $B_n^*(b)$ denotes the conjugate transpose of $B_n(b)$. Also,

$$Y(j,n) = (Y_j(0), Y_j(1), \ldots, Y_j(n-1))^\top, \ j = 1, \ldots, N$$

$$Q_n(Y,b) = \sum_{j=1}^{N} Y_j^*(j,n)B_n^*(b)D_n^{-1}(b)B_n(b)Y(j,n)$$

Bresler and Macovski showed in [2] that the vector $\hat{b} = \hat{b}(n)$ minimizing $Q_n$ on $B$, that is to say,

$$Q_n(Y,\hat{b}) = \min_{b \in B} Q_n(Y,b)$$

is the ML estimate of $b^{(0)} = (b_0^{(0)}, \ldots, b_p^{(0)})^\top \in B$, where

$$b_0^{(0)} + b_1^{(0)}z + \ldots + b_p^{(0)}z^p = 0.$$
has roots $\lambda_1, \ldots, \lambda_p$. Bresler and Macovski suggested an iterative process to compute $\hat{b}$. No proof is given for the convergence of this process. Nor is one guaranteed that when the process does converge, the limit is indeed an overall minimizing point of $Q_n$, and not a local minimum.

Theoretically it is interesting to give a close study of the statistical properties of the ML estimate $\hat{b}$. For, although the ML method usually gives statistical procedures with satisfactory performance, in particular when the normality assumption is in force as here, the complexity of the model (from the point of view that the unknowns of the model appear in rather complicated expressions) makes it unclear how good the ML estimate is under the present situation. As mentioned earlier, under model (1) the ML estimate of $s_{1j}$ is not even consistent. So also the good performance of the ML estimate of $(\lambda_1, \ldots, \lambda_p)$ cannot be taken for granted.

This paper is devoted to a basic problem of the asymptotic theory of the ML estimate of $(\lambda_1, \ldots, \lambda_p)$ — its consistency. On reducing the problem to the estimation of $b^{(0)}$ as described earlier, it is seen that the problem is equivalent to the consistency of the ML estimate $\hat{b} = \hat{b}(n)$ of $b^{(0)}$. Our main result is the following theorem:

**THEOREM 1.** Suppose the following conditions are satisfied:

1. $|\lambda_1| = \ldots = |\lambda_p| = 1$, $\lambda_i \neq \lambda_j$ for $i \neq j$.
2. For each $k = 1, \ldots, p$, $\sum_{j=1}^{N} |s_{kj}| > 0$.
3. $\{e_j(t)\}$ satisfies the conditions elaborated at the beginning of this section.

Then for arbitrarily given $\epsilon > 0$, there exists constant $c > 0$ inde-
pendent of \( n \) such that

\[
P(\| \hat{\beta}(n) - \beta^{(0)} \| \geq \epsilon) \leq e^{-cn}
\]

(7)

for \( n \) large, where \( \| \hat{\beta} \| \) denotes the Euclidean length of the vector \( \hat{\beta} \).

(7) entails, in view of the well-known Borel-Cantelli lemma, that \( \hat{\beta} = \hat{\beta}(n) \) is a strongly consistent estimate of \( \beta^{(0)} \).
2. LEMMAS

Some facts concerning mainly with the matrix \( D_n(b) \) will be needed in proving the theorem. For convenience we shall write

\[ m = n - p. \]  

(8)

LEMMA 1. For any \( b \in B \), we have

\[ D_n^{-1}(b) \geq \frac{1}{p+1} I_m, \]

(9)

\[ D_n^{-1}(b) \leq 2^{-p(p+1)(p+1)} p_n \]

(10)

where \( I_m \) is the identity matrix of order \( m \).

Proof. (9) follows from \( \text{tr}(B_n(b)B_n(b)) = p + 1 \).

To prove (10), we proceed to find the minimum

\[ H = \min_{b \in B} \min_{u \in A} u^* D_n(b) u \]

(11)

where \( A \) is the set \( \{ u = (u_0, ..., u_{m-1})' : \sum_{i=0}^{m-1} |u_i|^2 = 1 \} \).

Introduce the \((p+1) \times n\) matrix \( U^*(u) \):

\[
U^*(u) = 
\begin{pmatrix}
\bar{u}_0 & \bar{u}_1 & \cdots & \cdots & \bar{u}_{m-1} & 0 \\
\bar{u}_0 & \bar{u}_1 & \cdots & \cdots & \bar{u}_{m-1} \\
\vdots & \vdots & \ddots & \cdots & \vdots \\
0 & \bar{u}_0 & \bar{u}_1 & \cdots & \cdots & \bar{u}_{m-1}
\end{pmatrix}
\]

One sees easily that

\[ u^* B_n(b) = b' U^*(u). \]
Hence,
\[ u^rD_n(b)u = b^rV(u)b, \quad V(u) = U^*(u)U(u) \]
and
\[ H = \min_{u \in A} \min_{b \in B} b^r V(u)b. \] \hspace{1cm} (12)

Now we prove that for any \( u \in A \), we have
\[ \det V(u) \geq 2^{p(p+1)n-3(p+1)^2}. \] \hspace{1cm} (13)

Since (13) is true when \( p = 0 \), suppose that \( p \geq 1 \) and \( n \geq p + 1 \geq 2 \).

Define
\[ f_n(x) = \sum_{m=0}^{p} + \sum_{m=0}^{p} x^m \]
and \( \omega = \exp(\sqrt{-1} \, 2\pi/n) \), \( \Omega = \{\omega, \omega^2, \ldots, \omega^n\} \). There exist at least \( p + 1 \) elements \( \omega_1, \ldots, \omega_{p+1} \) in \( \Omega \), such that
\[ |f_n(\omega_k)| \geq n^{-(p+1)}, \quad k = 1, 2, \ldots, p+1. \] \hspace{1cm} (14)

Indeed, supposing in the contrary that
\[ \phi_k \in \Omega, \quad |f_n(\phi_k)| < n^{-(p+1)}, \quad k = 1, \ldots, n-p (=m), \] \hspace{1cm} (15)
then, on putting
\[ \Omega = \{\phi_1, \ldots, \phi_m\} = \{\phi_{m+1}, \ldots, \phi_n\}, \]
\[ \Lambda = \{1, 2, \ldots, m\}, \quad \Lambda_{\perp} = \{m+1, \ldots, n\}, \]
and using Lagrange interpolation formula, we have
\[ f_n(e^{\sqrt{-1} \theta}) = \sum_{j=1}^{m} f_n(\phi_j) \prod_{k \in \Lambda \setminus \{j\}} (e^{\sqrt{-1} \theta} - \phi_k) / \prod_{k \in \Lambda \setminus \{j\}} (\phi_j - \phi_k). \] \hspace{1cm} (16)
For fixed $j \in \Lambda$, we have
\[ \prod_{k \in \Lambda \setminus \{j\}} (\phi_j - \phi_k) \prod_{k \in \Lambda_1} (\phi_j - \phi_k) = \lim_{x \to \phi_j} \frac{x^n - 1}{x - \phi_j} = n\phi_j^n. \]

Hence,
\[ \left| \prod_{k \in \Lambda \setminus \{j\}} (\phi_j - \phi_k) \right| \leq \frac{1}{n} \left| \prod_{k \in \Lambda_1} (\phi_j - \phi_k) \right| \leq 2^p/n. \quad (17) \]

Two cases are possible: First,
\[ |\theta - \arg \phi_k| \geq \pi/n \text{ (mod 2\pi), } k \in \Lambda_1 U\{j\}. \]

In this case we have
\[ \left| \prod_{k \in \Lambda \setminus \{j\}} (e^{\sqrt{T} \theta - \phi_k}) \right| \leq \sin \frac{\pi}{n} |p+1|. \]

Since $\sin \frac{\pi}{n} > \frac{2}{n}$ when $n \geq 2$, we have
\[ \left| \prod_{k \in \Lambda \setminus \{j\}} (e^{\sqrt{T} \theta - \phi_k}) \right| = \left| e^{\sqrt{T} n \theta} - 1 \right| \left| \prod_{k \in \Lambda_1 U\{j\}} (e^{\sqrt{T} \theta - \phi_k}) \right|^{-1} \leq 2 |\sin \frac{\pi}{n}|^{-p-1} \leq 2^{-p \cdot n^{p+1}}. \quad (18) \]

Second,
\[ |\theta - \arg \phi_k| < \pi/n (\text{mod 2\pi}) \text{, for some } k \in \Lambda_1 U\{j\}. \]

(Note that there are at most one such $k$.) In this case, noticing that
\[ |\theta - \arg \phi_k| \geq \pi/n (\text{mod 2\pi}) \text{ for any } k \in \Lambda_1 U\{j\}, k \neq k, \text{ and that} \]
\[ |e^{\sqrt{T} n \theta} - 1|/|e^{\sqrt{T} \theta - \phi_k}| < n, \text{ we have} \]
\[
\left| \prod_{\kappa \in \Lambda^{-}(j)} (e^{\sqrt{-1} \theta - \phi_k}) \right| = \left| e^{\sqrt{-1} \theta} n - 1 \right| \left| \prod_{\kappa \in \Lambda_{1} \cup \{j\}} (e^{\sqrt{-1} \theta - \phi_k}) \right|^{-1} \\
\leq n \left| \prod_{\kappa \in \Lambda_{1} \cup \{j\}, k \neq \ell} (e^{\sqrt{-1} \theta - \phi_k}) \right|^{-1} \\
\leq n (\sin \frac{\pi}{n})^{-p} \leq 2^{-p} n^{p+1}.
\]

From (15)-(19), we obtain
\[
\left| f_n(e^{\sqrt{-1} \theta}) \right| < n^{-(p+1)} (2^{p} n^{-1}) 2^{-p} n^{p+1} = 1, \text{ for all } \theta \in [-\pi, \pi].
\]

Therefore,
\[
\left| \sum_{j=0}^{m-1} |u_j|^2 = \frac{1}{2\pi} \int_{-\pi}^{\pi} |f_n(e^{\sqrt{-1} \theta})|^2 d\theta < 1,
\]
contradicting the fact that \( \sum_{j=0}^{m-1} |u_j|^2 = 1 \). This proves (14).

Now put (remember that \( \omega = e^{\sqrt{-1} \theta} \))

\[
G = \begin{pmatrix}
1 & 1 & 1 & \cdots & 1 \\
1 & \omega & \omega^2 & \cdots & \omega^{n-1} \\
1 & \omega^2 & \omega^4 & \cdots & \omega^{2(n-1)} \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
1 & \omega^{n-1} & \omega^{2(n-1)} & \cdots & \omega^{(n-1)^2}
\end{pmatrix}_{(n \times n)}
\]
\[
F_{((p+1) \times n)} = \begin{pmatrix}
  f_n(1) & f_n(\omega) & \cdots & f_n(\omega^{n-1}) \\
  f_n(1) & \omega f_n(\omega) & \cdots & \omega^{n-1} f_n(\omega^{n-1}) \\
  \vdots & \vdots & \ddots & \vdots \\
  f_n(1) & \omega^P f_n(\omega) & \cdots & \omega^{p(n-1)} f_n(\omega^{n-1})
\end{pmatrix},
\]

\[
F_1_{((p+1) \times (p+1))} = \begin{pmatrix}
  f_n(\omega_1) & f_n(\omega_2) & \cdots & f_n(\omega_{p+1}) \\
  \omega_1 f_n(\omega_1) & \omega_2 f_n(\omega_2) & \cdots & \omega_{p+1} f_n(\omega_{p+1}) \\
  \vdots & \vdots & \ddots & \vdots \\
  \omega_1^P f_n(\omega_1) & \omega_2^P f_n(\omega_2) & \cdots & \omega_{p+1}^P f_n(\omega_{p+1})
\end{pmatrix}.
\]

Then
\[
U^*(u)U(u) = \frac{1}{n} U^*(u)GG^*U(u) = \frac{1}{n} FF^* \geq \frac{1}{n} F_1 F_1^*.
\]

Hence,
\[
\det V(u) = \det(U^*(u)U(u)) \geq n^{-(p+1)} |\det F_1|^2
\]
\[
= n^{-(p+1)^p+1} \prod_{k=1} \left| f_n(\omega_k) \right|^2 \prod_{1 \leq q < \omega_{p+1}} \left| \lambda_q - \omega_k \right|^2
\]
\[
\geq n^{-(p+1)n-2(p+1)} \left| \sin \frac{\pi}{n} \right|^{p(p+1)}
\]
\[
\geq n^{-(p+1)n-2(p+1)^2} \left( \frac{2}{n} \right)^{p(p+1)} = 2^{p+1} (p+1)^n - 3(p+1)^2
\]
and (13) is proved.

Denote by $L(u)$ and $\lambda(u)$ the largest and smallest eigenvalue of $V(u)$, respectively. Obviously we have $L(u) \leq p + 1$, for any $u \in A$. From this and (13), we have

$$\lambda(u) \geq C_p n^{-3(p+1)^2}, \quad C_p = 2^p(p+)(p+)^{-p}.$$  

From this and (12), we obtain

$$H \geq C_p n^{-3(p+1)^2}.$$ (20)

Now denote by $d_n(b)$ the smallest eigenvalue of $D_n(b)$. By (11), (20), we see that $d_n(b) \geq C_p n^{-3(p+1)^2}$, for any $b \in B$, which amounts to the same thing as (10). Lemma 1 is proved.

**Lemma 2.** For arbitrarily given $h > 0$, there exists $h_1 > 0$ such that for any $b \in B$, $\tilde{b} \in B$ with $|b - \tilde{b}| \leq n^{-h_1}$, we have

$$|D_n^{-1}(\tilde{b}) - D_n^{-1}(b)| < n^{-h}, \quad n \geq n_0$$ (21)

for some $n_0$ not depending on $b$, $\tilde{b}$. Where for any vector or matrix $C$, $|C|$ denotes the maximum module of the elements of $C$.

**Proof.** Since

$$D_n^{-1}(\tilde{b}) - D_n^{-1}(b) = D_n^{-1}(\tilde{b})(D_n(b) - D_n(\tilde{b}))D_n^{-1}(b),$$

(21) follows easily from Lemma 1.

In the following we use $\chi^2_n(\delta)$ to denote the noncentral Chi-square distribution with degree of freedom $n$ and noncentrality parameter $\delta$. $\chi^2_n(0)$ will be abbreviated to $\chi^2_n$. 


Lemma 3. Suppose that \( \{ \xi_n \} \) is a sequence of random variables, \( \xi_n \) is distributed as \( \chi^2_n(\delta_n) \), and there exists positive constants \( \eta_1 \leq \eta_2 \) such that
\[
\eta_1 n \leq \delta_n^2 \leq \eta_2 n, \quad n = 1, 2, \ldots.
\]
Then we can find positive constant \( c \) independent of \( n \), such that
\[
P(\xi_n/n > 1 + \eta_1/2) \geq 1 - e^{-cn}, \quad n = 1, 2, \ldots. \tag{22}
\]

Proof. We can find random variables \( \xi \sim \chi^2_n, Z \sim N(0,1) \), such that \( \xi_n \) is distributed as \( \xi + 2\delta_n Z + \delta_n^2 \). Choose \( \epsilon \in (0, \eta_1\eta_2^{-1/2}/8) \), we have
\[
P(|Z| > \epsilon\sqrt{n}) \leq \frac{2}{\sqrt{2\pi}\epsilon\sqrt{n}} \exp\left(-\frac{\epsilon^2 n}{2}\right) \leq e^{-c_1 n}
\]
for \( c_1 = \epsilon^2/2, n \geq 2/(\pi\epsilon^2) \). But when \( |Z| < \epsilon\sqrt{n} \), we have
\[
|2\delta_n Z| \leq 2\sqrt{n} \epsilon n \leq \eta_1 n/4.
\]
Therefore,
\[
P(\xi_n/n > 1 + \eta_1/2) \geq P(\xi/n > 1 + \eta_1/4) - P(|Z| > \epsilon\sqrt{n})
\geq 1 - P(|\xi/n - 1| > \eta_1/4) - e^{-c_1 n}, \quad n \geq 2/\epsilon^2. \tag{23}
\]
Since \( \xi \) is the sum of iid. variables \( \chi^2_1, \ldots, \chi^2_n \) with \( \chi^2_1 \sim N(0,1) \), in view of the fact that \( \chi^2_1 \) has moment generating function in some neighborhood of zero, it is well-known ([3], p.288) that there exists a constant \( c_2 > 0 \) such that
\[
P(|\xi/n - 1| > \eta_1/4) \leq e^{-c_2 n}, \quad \text{for} \quad n \geq n_1.
\]
From this and (23), we see that (22) holds for \( c = \min(c_1, c_2) \), when \( n \geq \max(2/\epsilon^2, n_1) \). Replacing \( c \) by some smaller quantity, we can make (22) true for all \( n \).
3. PROOF OF THE THEOREM

Introduce the following notations:

\[ s_j = (s_{1j}, \ldots, s_{pj})', \quad w_k(b) = b_0 + b_1 \lambda_k + \ldots + b_p \lambda_k, \]

\[
\Lambda_n = \begin{pmatrix}
1 & 1 & \ldots & 1 \\
\lambda_1 & \lambda_2 & \ldots & \lambda_p \\
& \ddots & \ddots & \ddots \\
& & \ldots & \ldots \\
& & & \lambda_{n-1} & \lambda_{n-1} & \ldots & \lambda_{n-1} & \lambda_p
\end{pmatrix},
\]

\[ \delta^2(n,b) = \sum_{j=1}^{N} s_j^{*} \Lambda_n^{*} (b) \Lambda_n^{-1}(b) B_n(b) \Lambda_n s_j. \]

For simplicity of writing and without losing generality, suppose that \( \sigma^2 = 1 \). Then from the assumptions imposed on \( (e_j(t)) \), we have

\[ Q_n(Y,b) = \chi_{2Nm}^2(\delta^2(n,b)). \]

Remember that \( m = n - p \). From (9) and the fact that \( |\lambda_1| = |\lambda_2| = \ldots = |\lambda_p| = 1 \), we have

\[
\delta^2(n,b) \geq \frac{1}{p+1} \sum_{j=1}^{N} |B_n(b) \Lambda_n s_j|^2 = \frac{1}{p+1} \sum_{j=1}^{N} \sum_{t=0}^{m-1} \sum_{k=0}^{p} |\lambda_k w_k(b)s_{kj}|^2
\]

\[
= \frac{1}{p+1} \{ \sum_{j=1}^{N} \sum_{t=0}^{m-1} \sum_{k=1}^{p} |w_k(b)|^2 s_{kj}^2 + \sum_{u \neq v} \sum_{j=1}^{N} \sum_{t=0}^{m-1} (\lambda_u \bar{\lambda}_v) t w_u(b) \bar{w}_v(b) s_u \bar{s}_v \}. \]

By assumption
\[ \alpha = \min\{ \sum_{j=1}^{N} |s_{kj}|^2 : k = 1, \ldots, p \} > 0. \]

Hence,
\[ \sum_{j=1}^{N} \sum_{t=0}^{m-1} \sum_{k=1}^{p} |w_k(b)|^2 |s_{kj}|^2 > \alpha (n-p) \sum_{k=1}^{p} |w_k(b)|^2. \quad (25) \]

Put
\[ s = \max\{|s_{kj}| : 1 \leq k \leq p, 1 \leq j \leq N\}, \]
\[ \lambda = \min\{|\lambda_i - \lambda_j| : 1 \leq i < j \leq p\}. \]

We have \( \lambda > 0 \) since by assumption \( \lambda_i \neq \lambda_j \) when \( i \neq j \). Thus
\[ \left| \sum_{t=0}^{m-1} (\lambda_i \bar{w}_v) t \right| \leq \frac{2}{1-\lambda}, \quad n > p+1, \quad u \neq v. \]

Therefore, noticing that \( |w_k(b)| \leq \sqrt{p+1} \), we have
\[ \left| \sum_{t=0}^{m-1} \sum_{j=1}^{N} (\lambda_i \bar{w}_v) t w_u(b) \bar{w}_v(b) s_{uj} \bar{s}_{vj} \right| \leq 2p^2 (p+1) N s^2 (1-\lambda)^{-1}. \quad (26) \]

Define the set
\[ B_\varepsilon = \left\{ b : b = (b_0, \ldots, b_p)' \in B, \sum_{k=1}^{p} |b_k - b_k(0)|^2 \geq \varepsilon \right\}. \]

Since \( (\lambda_1, \ldots, \lambda_p) \) is the set of all roots of \( b_0(0) + b_1(0) z + \ldots + b_p z^p = 0 \), it is easily seen that
\[ \inf\left\{ \sum_{k=1}^{p} |w_k(b)|^2 : b \in B_\varepsilon \right\} > 0. \quad (27) \]

Summing up (24)-(27), we see that there exists constants \( n_1 > 0, n_2 > 0 \) depending only on \( \varepsilon \), such that
\[ n_1 n \leq \delta^2(n,b) \leq n_2 n, \quad b \in B_\varepsilon. \quad (28) \]
To simplify the wording, in the sequel the symbol c will be used to denote any positive constant not depending on b, n, which may assume different values on each of its appearances, and the phrase "for n large" means that "for n larger than some \( n_0 \) independent of \( b \in B \)." Since \( Q_n(Y, b) \sim \chi^2_{2N_m}(\delta^2(n, b)) \), from (28) and Lemma 3, we obtain

\[
P\left( Q_n(Y, b) / nN \geq 1 + \eta_1/2 \right) \geq 1 - e^{-cn}. \tag{29}
\]

Choose \( h_1 > 0 \) according to some \( h > 0 \) as in Lemma 2. The value of \( h \) will be specified later. Choose a subset \( B_{en} \) of \( B_c \) with no more than \( 2ph_1 \) points, such that for each \( b \in B_c \), there exists \( \tilde{b} \in B_{en} \) such that \( |\tilde{b} - b| \leq n^{-h_1} \). From (44), for \( n \) large, we have

\[
P\left( \min_{b \in B_{en}} Q_n(Y, b) / nN \geq 1 + \eta_1/2 \right) \geq 1 - n^{-e^{-cn}}. \tag{30}
\]

Now choose arbitrarily \( b \in B_c \). Find \( \tilde{b} \in B_{en} \) such that \( |\tilde{b} - b| \leq n^{-h_1} \). Consider

\[
J = |Q_n(Y, b) - Q_n(Y, \tilde{b})|. \tag{31}
\]

Abbreviating \( B_n(b), B_n(\tilde{b}), \) etc. to \( B_n, \tilde{B}_n, \) etc., we have

\[
J \leq \sum_{j=1}^{N} \left| y^*(j, n)B_n^{-1}B_nD_n^{-1}B_nY(j, n) - y^*(j, n)B_n^{-1}B_nY(j, n) \right|
\]

\[
\leq \sum_{j=1}^{N} \left| y^*(j, n)B_n^{-1}(D_n^{-1} - D_n^{-1})B_nY(j, n) \right|
\]

\[
+ \sum_{j=1}^{N} \left| y^*(j, n)B_n^{-1}B_nD_n^{-1}B_nY(j, n) \right|
\]

\[
\leq J_1 + J_2. \tag{32}
\]
By (21), we have
\[ J_1 \leq n^{-h} \sum_{j=1}^{N} \|B_{n^{-}} Y(j,n)\|^2 \leq n^{-h} \sum_{j=1}^{N} \sum_{t=0}^{n-1} (p+1)|Y_j(t)|^2. \]

Put \( \bar{s} = \max(\sum_{j=1}^{p} |s_{1j}|): j = 1, \ldots, N \). Since \( |\lambda_1| = \ldots = |\lambda_p| = 1 \), we have
\[ \left| \sum_{j=1}^{p} s_{1j} \lambda_j^t \right| \leq \bar{s}, \quad j = 1, \ldots, N. \]

Hence,
\[ J_1 \leq n^{-h} \left\{ 2nN(p+1)s^2 + 2 \sum_{j=1}^{N} \sum_{t=0}^{n-1} (p+1)|e_j(t)|^2 \right\} \]
\[ \leq Cn^{-h+1} + Cn^{-h} \sum_{j=1}^{N} \sum_{t=0}^{n-1} |e_j(t)|^2. \]  (33)

Introduce the event
\[ E_n = \left\{ \sum_{j=1}^{N} \sum_{t=0}^{n-1} |e_j(t)|^2 \leq 2nN \right\}. \]

By (33), we have
\[ E_n \subset \left\{ J_1 \leq Cn^{-h+1} \right\}. \]  (34)

For \( J_2 \), we have
\[ J_2 \leq \sum_{j=1}^{N} \left| Y^*(j,n)(\bar{B}_{n} - B_{n}^*) \overline{D}_{n}^{-1} \bar{B}_{n} Y(j,n) \right| \]
\[ + \sum_{j=1}^{N} \left| Y^*(j,n)B_{n}^* \overline{D}_{n}^{-1} (\bar{B}_{n} - B_{n}) Y(j,n) \right| \equiv J_3 + J_4. \]  (35)

By the extended Schwarz inequality,
\[ J_3^2 \leq N \sum_{j=1}^{N} \left( j, n \right) \bar{b}_n^* \bar{D}_n^{-1} b_n \bar{Y}(j, n) Y^*(j, n) (\bar{b}_n^* - b_n) \bar{D}_n^{-1} (\bar{b}_n - B_n) Y(j, n). \]  

(36)

Write \( w = 3(p + 1)^2 + 1 \). By (10), (36), we have for \( n \) large,

\[ J_3^2 \leq N \sum_{j=1}^{N} n^{2w} ||B_n Y(j, n)||^2 ||(\bar{b}_n - B_n) Y(j, n)||^2 \]

\[ \leq N n^{2w} \sum_{j=1}^{N} ||B_n Y(j, n)||^2 \sum_{j=1}^{N} ||(\bar{b}_n - B_n) Y(j, n)||^2. \]  

(37)

In the course of proving (33), we have shown that

\[ \sum_{j=1}^{N} ||B_n Y(j, n)||^2 \leq C_n + C \sum_{j=1}^{N} \sum_{t=0}^{n-1} |e_j(t)|^2. \]  

(38)

Further, in view of \( |\bar{b}_n - B_n| \leq n^{-1} \), we have

\[ \sum_{j=1}^{N} ||(\bar{b}_n - B_n) Y(j, n)||^2 \leq n^{-2h_1} (p + 1)^2 \sum_{j=1}^{N} \sum_{t=0}^{n-1} |Y_j(t)|^2 \]

\[ \leq C_n^{-h_1} + C_n^{-2h_1} \sum_{j=1}^{N} \sum_{t=0}^{n-1} |e_j(t)|^2. \]  

(39)

From (37)-(39), we see that

\[ E_n = \left\{ J_3 \leq C_n^{-(h_1 - w - 1)} \right\}. \]  

(40)

Likewise, we obtain

\[ E_n = \left\{ J_4 \leq C_n^{-(h_1 - w - 1)} \right\}. \]  

(41)

Summing up (31), (33), (35), (40) and (41), we obtain

\[ E_n = \left\{ |Q_n(Y, b) - Q_n(Y, \bar{b})| \leq C_n^{-h+1} + C_n^{-h_1+w+1} \right\}. \]  

(42)
Now we choose \( h = 1 \). Choose \( h_1 \) corresponding to this \( h \) according to Lemma 2 such that \( h_1 \geq w + 1 \). For this choice of \( h \) and \( h_1 \), from (30) and (42), we get for \( n \) large

\[
P(\min_{b \in B} Q_n(Y,b)/(nN) \geq 1 + n_1/2) \geq 1 - e^{-cn} - (1 - P(E_n)). \tag{43}
\]

On the other hand, since \( Q_n(Y,b(0)) = \chi_{2nN}^2 \), we have for \( n \) large,

\[
P(Q_n(Y,b(0)) \leq 1 + n_1/4) \geq 1 - e^{-cn}. \tag{44}
\]

Likewise, since \( \sum_{j=1}^{N} \sum_{t=0}^{n-1} |e_j(t)|^2 - \chi_{2nN}^2 \), we have for \( n \) large,

\[
P(E_n) \geq 1 - e^{-cn}. \tag{45}
\]

Summing up (43)-(45), we obtain for \( n \) large,

\[
P(\min_{b \in B} Q_n(Y,b) > Q_n(Y,b(0))) \leq e^{-cn}.
\]

which entails (7), and the theorem is proved.
REFERENCES

Asymptotic property of the EVLP estimation for superimposed ex-
potential signals in noise. Technical report. Center for Multi-
variate Analysis, University of Pittsburgh, in preparation.

parameter estimation of superimposed exponential signals in noise.
IEEE Trans. Acoustics, Speech and Signal Processing, Vol. ASSP-34,
No. 5.

Springer-Verlag, New York.
END
DATE
FILMED
JAN
1988