

PARAMETER ESTIMATION OF CHIRP SIGNALS IN PRESENCE OF STATIONARY NOISE

Debasis Kundu and Swagata Nandi

Indian Institute of Technology, Kanpur and Indian Statistical Institute

Abstract: The problem of parameter estimation of the chirp signals in presence of stationary noise has been addressed. We consider the least squares estimators and observe that they are strongly consistent. The asymptotic distributions of the least squares estimators are obtained. The multiple chirp signal model is also considered and we obtain the asymptotic properties of the least squares estimators of the unknown parameters. We perform some small sample simulations to observe how the proposed estimators work.

Key words and phrases: Asymptotic distributions, chirp signals, least squares estimators, multiple chirp signals.

1. Introduction

In this paper we consider the estimation procedure of the parameters of the following signal processing model:

$$y(n) = A^0 \cos(\alpha^0 n + \beta^0 n^2) + B^0 \sin(\alpha^0 n + \beta^0 n^2) + X(n); \quad n = 1, \dots, N. \quad (1)$$

Here $y(n)$ is the real-valued signal observed at $n = 1, \dots, N$, A^0 and B^0 are real-valued amplitudes, and α^0 and β^0 are the frequency and frequency rate, respectively. The “chirp signal” model (1) does not have a constant frequency like the sinusoidal frequency model and the initial frequency changes over time at the rate β . The error $\{X(n)\}$ is a sequence of random variables with mean zero and finite fourth moment. The error random variable $X(n)$ satisfies the following.

Assumption 1. *One can write*

$$X(n) = \sum_{j=-\infty}^{\infty} a(j)e(n-j), \quad (2)$$

where $\{e(n)\}$ is a sequence of independent and identically distributed (i.i.d.) random variables with mean zero and finite fourth moment, and

$$\sum_{j=-\infty}^{\infty} |a(j)| < \infty. \quad (3)$$

The signals as described in (1) are known as chirp signals in the statistical signal processing literature (Djurić and Kay (1990)). They are quite common in various areas of science and engineering, specifically in sonar, radar, communications, etc. Several authors have considered (1) when the $X(n)$'s are i.i.d. random variables. See for example, the work of Abatzoglou (1986) Kumaresan and Verma (1987), Djurić and Kay (1990), Gini, Montanari and Verrazzani (2000), Nandi and Kundu (2004), etc. Different approaches to the estimation of chirp parameters in similar kinds of models are found in Giannakis and Zhou (1995), Zhou, Giannakis and Swami (1996), Shamsunder, Giannakis and Friedlander (1995), Swami (1996) and Zhou and Giannakis (1995). It is well known that in most practical situations, the errors are not independent. We assume stationarity through Assumption 1 to incorporate the dependence structure and make the model more realistic.

Assumption 1 is a standard assumption for a stationary linear process, and any finite dimensional stationary AR, MA or ARMA process can be represented as at (2) when the $a(j)$'s satisfy (3).

In this paper, we discuss the problem of parameter estimation of the chirp signal model in presence of stationary noise. We consider least squares estimators (LSEs) and study their properties. It is known, see Kundu (1997), that the sum of sinusoidal model does not satisfy the sufficient conditions of Jennrich (1969) or Wu (1981) for the LSEs to be consistent. While the results of Wu or Jennrich cannot be applied directly to establish strong consistency or asymptotic normality of the LSEs, the structure of the model allows us to obtain these properties. It is found that the asymptotic variances of the amplitudes, frequency and frequency rate estimators are $O(N^{-1})$, $O(N^{-3})$ and $O(N^{-5})$, respectively. Based on the asymptotic distributions, asymptotic confidence intervals can be constructed.

The paper is organized as follows. In Section 2, we provide the asymptotic properties of the LSEs. The multiple chirp model is discussed in Section 3. Some numerical results are presented in Section 4 and we conclude the paper in Section 5. The proofs of the results of Section 2 are provided in the Appendix.

2. Asymptotic Properties of LSEs

Write $\boldsymbol{\theta} = (A, B, \alpha, \beta)$, $\boldsymbol{\theta}^0 = (A^0, B^0, \alpha^0, \beta^0)$. Then the LSE of $\boldsymbol{\theta}^0$, say $\hat{\boldsymbol{\theta}} = (\hat{A}, \hat{B}, \hat{\alpha}, \hat{\beta})$, can be obtained by minimizing

$$Q(A, B, \alpha, \beta) = Q(\boldsymbol{\theta}) = \sum_{n=1}^N [y(n) - A \cos(\alpha n + \beta n^2) - B \sin(\alpha n + \beta n^2)]^2, \quad (4)$$

with respect to A , B , α and β . In the following, we state the consistency property of $\hat{\boldsymbol{\theta}}$.

Theorem 1. *Let the true parameter vector $\boldsymbol{\theta}^0 = (A^0, B^0, \alpha^0, \beta^0)$ be an interior point of the parameter space $\Theta = (-\infty, \infty) \times (-\infty, \infty) \times (0, \pi) \times (0, \pi)$ and $A^{0^2} + B^{0^2} > 0$. If the error random variables $X(n)$ satisfy Assumption 1, then $\hat{\boldsymbol{\theta}}$ is a strongly consistent estimator of $\boldsymbol{\theta}^0$.*

In this section we compute the asymptotic joint distribution of the LSEs of the unknown parameters. We use $Q'(\boldsymbol{\theta})$ and $Q''(\boldsymbol{\theta})$ to denote the 1×4 vector of first derivatives of $Q(\boldsymbol{\theta})$ and the 4×4 second derivative matrix of $Q(\boldsymbol{\theta})$, respectively. Expanding $Q'(\hat{\boldsymbol{\theta}})$ around the true parameter value $\boldsymbol{\theta}^0$ in a Taylor series, we obtain

$$Q'(\hat{\boldsymbol{\theta}}) - Q'(\boldsymbol{\theta}^0) = (\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}^0)Q''(\bar{\boldsymbol{\theta}}), \quad (5)$$

where $\bar{\boldsymbol{\theta}}$ is a point on the line joining the points $\hat{\boldsymbol{\theta}}$ and $\boldsymbol{\theta}^0$. Suppose \mathbf{D} is the 4×4 diagonal matrix $\mathbf{D} = \text{diag}\{N^{-1/2}, N^{-1/2}, N^{-3/2}, N^{-5/2}\}$. Since $Q'(\hat{\boldsymbol{\theta}}) = 0$, (5) can be written as

$$(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}^0)\mathbf{D}^{-1} = -[\mathbf{Q}'(\boldsymbol{\theta}^0)\mathbf{D}][\mathbf{D}\mathbf{Q}''(\bar{\boldsymbol{\theta}})\mathbf{D}]^{-1}, \quad (6)$$

as $[\mathbf{D}\mathbf{Q}''(\bar{\boldsymbol{\theta}})\mathbf{D}]$ is an invertible matrix *a.e.* for large N . Using Theorem 1, it follows that $\hat{\boldsymbol{\theta}}$ converges *a.e.* to $\boldsymbol{\theta}^0$ and, since each element of $Q''(\boldsymbol{\theta})$ is a continuous function of $\boldsymbol{\theta}$, $\lim_{N \rightarrow \infty} [\mathbf{D}\mathbf{Q}''(\bar{\boldsymbol{\theta}})\mathbf{D}] = \lim_{N \rightarrow \infty} [\mathbf{D}\mathbf{Q}''(\boldsymbol{\theta}^0)\mathbf{D}] = \mathbf{2}\boldsymbol{\Sigma}(\boldsymbol{\theta}^0)$, say, where $\boldsymbol{\Sigma}(\boldsymbol{\theta}) = (\sigma_{jk}(\boldsymbol{\theta}))$.

We write

$$\lim_{N \rightarrow \infty} \frac{1}{N^{p+1}} \sum_{n=1}^N n^p \cos^k(\alpha n + \beta n^2) = \delta_k(p, \alpha, \beta), \quad (7)$$

$$\lim_{N \rightarrow \infty} \frac{1}{N^{p+1}} \sum_{n=1}^N n^p \sin^k(\alpha n + \beta n^2) = \gamma_k(p, \alpha, \beta), \quad (8)$$

where k takes values 1 and 2. We compute the elements of $\boldsymbol{\Sigma}(\boldsymbol{\theta})$ in this notation.

The 4×1 random vector $[\mathbf{Q}'(\boldsymbol{\theta}^0)\mathbf{D}]$ is

$$\begin{bmatrix} -\frac{2}{\sqrt{N}} \sum_{n=1}^N X(n) \cos(\alpha^0 n + \beta^0 n^2) \\ -\frac{2}{\sqrt{N}} \sum_{n=1}^N X(n) \sin(\alpha^0 n + \beta^0 n^2) \\ \frac{2}{N^{\frac{3}{2}}} \sum_{n=1}^N n X(n) [A^0 \sin(\alpha^0 n + \beta^0 n^2) - B^0 \cos(\alpha^0 n + \beta^0 n^2)] \\ \frac{2}{N^{\frac{5}{2}}} \sum_{n=1}^N n^2 X(n) [A^0 \sin(\alpha^0 n + \beta^0 n^2) - B^0 \cos(\alpha^0 n + \beta^0 n^2)] \end{bmatrix}.$$

Using a central limit theorem (see Fuller (1976, p. 251)), it follows that

$$[\mathbf{Q}'(\boldsymbol{\theta}^0)\mathbf{D}] \xrightarrow{d} \mathcal{N}_4(\mathbf{0}, \mathbf{G}(\boldsymbol{\theta}^0)), \quad (9)$$

where the matrix $\mathbf{G}(\boldsymbol{\theta}^0)$ is the asymptotic dispersion matrix of $[\mathbf{Q}'(\boldsymbol{\theta}^0)\mathbf{D}]$. If $\mathbf{G}(\boldsymbol{\theta}) = ((g_{jk}(\boldsymbol{\theta})))$ then, for $k \geq j$,

$$g_{11}(\boldsymbol{\theta}) = \lim_{N \rightarrow \infty} \frac{4}{N} E[S_1]^2, \quad g_{12}(\boldsymbol{\theta}) = \lim_{N \rightarrow \infty} \frac{4}{N} E[S_1 S_2], \quad (10)$$

$$g_{13}(\boldsymbol{\theta}) = \lim_{N \rightarrow \infty} \frac{4}{N^2} E[S_1 S_3], \quad g_{14}(\boldsymbol{\theta}) = \lim_{N \rightarrow \infty} \frac{4}{N^3} E[S_1 S_4], \quad (11)$$

$$g_{22}(\boldsymbol{\theta}) = \lim_{N \rightarrow \infty} \frac{4}{N} E[S_2]^2, \quad g_{23}(\boldsymbol{\theta}) = \lim_{N \rightarrow \infty} \frac{4}{N^2} E[S_2 S_3], \quad (12)$$

$$g_{24}(\boldsymbol{\theta}) = \lim_{N \rightarrow \infty} \frac{4}{N^3} E[S_2 S_4], \quad g_{33}(\boldsymbol{\theta}) = \lim_{N \rightarrow \infty} \frac{4}{N^3} E[S_3]^2, \quad (13)$$

$$g_{34}(\boldsymbol{\theta}) = \lim_{N \rightarrow \infty} \frac{4}{N^4} E[S_3 S_4], \quad g_{44}(\boldsymbol{\theta}) = \lim_{N \rightarrow \infty} \frac{4}{N^5} E[S_4]^2, \quad (14)$$

where

$$S_1 = - \sum_{n=1}^N X(n) \cos(\alpha n + \beta n^2), \quad S_2 = - \sum_{n=1}^N X(n) \sin(\alpha n + \beta n^2),$$

$$S_3 = \sum_{n=1}^N n X(n) [A \sin(\alpha n + \beta n^2) - B \cos(\alpha n + \beta n^2)],$$

$$S_4 = \sum_{n=1}^N n^2 X(n) [A \sin(\alpha n + \beta n^2) - B \cos(\alpha n + \beta n^2)].$$

For $k < j$, $g_{jk}(\boldsymbol{\theta}) = g_{kj}(\boldsymbol{\theta})$. The limits given in (10) to (14) exist for fixed value of $\boldsymbol{\theta}$ because of (7) and (8). Therefore, from (6) the following theorem holds.

Theorem 2. *Under the same assumptions as in Theorem 1,*

$$(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}^0) \mathbf{D}^{-1} \xrightarrow{d} \mathcal{N}_4 \left[\mathbf{0}, \frac{1}{4} \boldsymbol{\Sigma}^{-1}(\boldsymbol{\theta}^0) \mathbf{G}(\boldsymbol{\theta}^0) \boldsymbol{\Sigma}^{-1}(\boldsymbol{\theta}^0) \right]. \quad (15)$$

Remark 1. When the $X(n)$'s are *i.i.d.* random variables, the covariance matrix takes the simplified form $\boldsymbol{\Sigma}^{-1}(\boldsymbol{\theta}^0) \mathbf{G}(\boldsymbol{\theta}^0) \boldsymbol{\Sigma}^{-1}(\boldsymbol{\theta}^0) = \sigma^2 \boldsymbol{\Sigma}^{-1}(\boldsymbol{\theta}^0)$.

Remark 2. Although we cannot prove it, it has been observed in extensive numerical computations that the limits at (7) and (8) for $k = 1, 2$, do not depend on α . If we assume that these quantities are independent of their second argument, we can write them as $\delta_k(p; \beta) = \delta_k(p, \alpha, \beta)$, $\gamma_k(p; \beta) = \gamma_k(p, \alpha, \beta)$.

If we write

$$c_c = \sum_{k=-\infty}^{\infty} a(k) \cos(\alpha^0 k + \beta^0 k^2), \quad c_s = \sum_{k=-\infty}^{\infty} a(k) \sin(\alpha^0 k + \beta^0 k^2),$$

c_c and c_s are functions of α^0 and β^0 . They are not made explicit here to keep the notation simple.

Now if the δ 's and γ 's are independent of α , we can explicitly compute the elements of $\mathbf{G}(\boldsymbol{\theta})$ matrix for a given $\boldsymbol{\theta}$. For different entries of the matrix $\mathbf{G}(\boldsymbol{\theta})$ in terms of δ 's and γ 's, see <http://www.isid.ac.in/~statmath/eprints/isid/ms/2005/08>.

Obtaining the explicit expression of different entries of the variance-covariance matrix $(1/4)\boldsymbol{\Sigma}^{-1}(\boldsymbol{\theta}^0)\mathbf{G}(\boldsymbol{\theta}^0)\boldsymbol{\Sigma}^{-1}(\boldsymbol{\theta}^0)$ of $(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}^0)\mathbf{D}^{-1}$ is possible by inverting the matrix $\boldsymbol{\Sigma}(\boldsymbol{\theta})$ at $\boldsymbol{\theta}^0$. These are not provided here due to the complex (notational) structure of matrices $\boldsymbol{\Sigma}(\boldsymbol{\theta})$ and $\mathbf{G}(\boldsymbol{\theta})$. If the true value of β is zero (i.e., frequency does not change over time) and if this information is used in the model, then (1) is nothing but the usual sinusoidal model. In that case, the asymptotic distribution can be obtained in compact form and the amplitude is asymptotically independent of the frequency. This has not been observed in the case of the chirp signal model.

3. Multiple Chirp Signal

In this section, we introduce the multiple chirp signal model in stationary noise. The complex-valued single chirp model was generalized to a superimposed chirp model by Saha and Kay (2002). The following is a similar generalization of model (1). We assume

$$y(n) = \sum_{k=1}^p [A_k^0 \cos(\alpha_k^0 n + \beta_k^0 n^2) + B_k^0 \sin(\alpha_k^0 n + \beta_k^0 n^2)] + X(n); \quad n = 1, \dots, N. \quad (16)$$

As with the single chirp model, the parameters $\alpha_k^0, \beta_k^0 \in (0, \pi)$ are the frequency and frequency rate, respectively; the A_k^0 's and B_k^0 's are real-valued amplitudes. Again our aim is to estimate the parameters and study their properties. We assume that the number of components, p , is known and that the $X(n)$'s satisfy Assumption 1. Estimation of p is an important problem and will be addressed elsewhere. Now take $\boldsymbol{\theta}_k = (A_k, B_k, \alpha_k, \beta_k)$ and $\boldsymbol{\nu} = (\boldsymbol{\theta}_1, \dots, \boldsymbol{\theta}_p)$. The LSE of the parameters are obtained by minimizing the objective function, say $R(\boldsymbol{\nu})$ (defined as was $Q(\boldsymbol{\theta})$; see (4)). Let $\hat{\boldsymbol{\nu}}$ and $\boldsymbol{\nu}^0$ denote the LSE and the true value of $\boldsymbol{\nu}$. The consistency of $\hat{\boldsymbol{\nu}}$ follows as did the consistency of $\hat{\boldsymbol{\theta}}$, considering the parameter vector as $\boldsymbol{\nu}$. We state the asymptotic distribution of $\hat{\boldsymbol{\nu}}$ here. The proof involves routine calculations and use of the multiple Taylor series expansion and a central limit theorem.

For the asymptotic distribution of $\hat{\boldsymbol{\nu}}$, write $\boldsymbol{\psi}_k^N = (\hat{\boldsymbol{\theta}}_k - \boldsymbol{\theta}_k^0)\mathbf{D}^{-1} = (N^{1/2}(\hat{A}_k - A_k^0), N^{1/2}(\hat{B}_k - B_k^0), N^{3/2}(\hat{\alpha}_k - \alpha_k^0), N^{5/2}(\hat{\beta}_k - \beta_k^0))$ and let c_c^k and c_s^k be obtained from c_c and c_s by replacing α^0 and β^0 by α_k^0 and β_k^0 , respectively. Let $\beta_j + \beta_k =$

β_{jk}^+ , $\beta_j - \beta_k = \beta_{jk}^-$, $d_1 = c_c^1 c_c^2 + c_s^1 c_s^2$, $d_2 = c_c^1 c_s^2 + c_s^1 c_c^2$, $d_3 = c_c^1 c_c^2 - c_s^1 c_s^2$ and $d_4 = c_c^1 c_s^2 - c_s^1 c_c^2$. Then

$$(\boldsymbol{\psi}_1^N, \dots, \boldsymbol{\psi}_p^N) \xrightarrow{d} \mathcal{N}_{4p}(\mathbf{0}, 2\sigma^2 \boldsymbol{\Lambda}^{-1}(\boldsymbol{\nu}^0) \mathbf{H}(\boldsymbol{\nu}^0) \boldsymbol{\Lambda}^{-1}(\boldsymbol{\nu}^0)), \quad (17)$$

$$\boldsymbol{\Lambda}(\boldsymbol{\nu}) = \begin{pmatrix} \boldsymbol{\Lambda}_{11} & \boldsymbol{\Lambda}_{12} & \cdots & \boldsymbol{\Lambda}_{1p} \\ \boldsymbol{\Lambda}_{21} & \boldsymbol{\Lambda}_{22} & \cdots & \boldsymbol{\Lambda}_{2p} \\ \vdots & \vdots & \vdots & \vdots \\ \boldsymbol{\Lambda}_{p1} & \boldsymbol{\Lambda}_{p2} & \cdots & \boldsymbol{\Lambda}_{pp} \end{pmatrix}, \quad \mathbf{H}(\boldsymbol{\nu}) = \begin{pmatrix} \mathbf{H}_{11} & \mathbf{H}_{12} & \cdots & \mathbf{H}_{1p} \\ \mathbf{H}_{21} & \mathbf{H}_{22} & \cdots & \mathbf{H}_{2p} \\ \vdots & \vdots & \vdots & \vdots \\ \mathbf{H}_{p1} & \mathbf{H}_{p2} & \cdots & \mathbf{H}_{pp} \end{pmatrix}. \quad (18)$$

The sub-matrices $\boldsymbol{\Lambda}_{jk}$ and \mathbf{H}_{jk} are square matrices of order four and $\boldsymbol{\Lambda}_{jk} \equiv \boldsymbol{\Lambda}_{jk}(\boldsymbol{\theta}_j, \boldsymbol{\theta}_k)$, $\mathbf{H}_{jk} \equiv \mathbf{H}_{jk}(\boldsymbol{\theta}_j, \boldsymbol{\theta}_k)$. $\boldsymbol{\Lambda}_{jj}$ and \mathbf{H}_{jj} can be obtained from $\boldsymbol{\Sigma}(\boldsymbol{\theta})$ and $\mathbf{G}(\boldsymbol{\theta})$ by putting $\boldsymbol{\theta} = \boldsymbol{\theta}_j$. As in the case of $\mathbf{G}(\boldsymbol{\theta})$, entries of the off-diagonal sub-matrices $\boldsymbol{\Lambda}_{jk} = ((\lambda_{rs}))$ and $\mathbf{H}_{jk} = ((h_{rs}))$ are available at <http://www.isid.ac.in/~statmath/eprints/> (isid/ms/2005/08). The elements of the matrices $\boldsymbol{\Lambda}_{jk}$ and \mathbf{H}_{jk} are non-zero, the parameters corresponding to different components, $\boldsymbol{\psi}_j^N$ and $\boldsymbol{\psi}_k^N$ for $j \neq k$, are not asymptotically independent. If the frequencies do not change over time, i.e., the β 's vanish, then (16) is equivalent to the multiple frequency model, in which case the off-diagonal matrices in \mathbf{H} and $\boldsymbol{\Lambda}$ are zero matrices and the estimators of the unknown parameters in different components are independent.

4. Numerical Experiments

In this section, we present the results of the numerical experiments. Consider a single chirp model with $A = 2.93$, $B = 1.91$, $\alpha = 2.5$ and $\beta = 0.10$. We use the sample sizes $N = 50$ and 100 . Though $\alpha, \beta \in (0, \pi)$, we take the value of β to be much less than the initial frequency α , as the frequency rate is comparatively small in general. We consider different stationary processes as the error random variables for our simulations. Errors are generated from (a) $X(t) = \rho e(t+1) + e(t)$, (b) $X(t) = \rho_1 e(t-1) + \rho_2 e(t-2) + e(t)$, and (c) $X(t) = \rho X(t-1) + e(t)$. The random variables $\{e(t)\}$ are distributed as $\mathcal{N}(0, \sigma^2)$. The processes (a), (b) and (c) are stationary $MA(1)$, $MA(2)$ and $AR(1)$ processes. For simulation purposes, $\rho = 0.5$, $\rho_1 = 0.5$, and $\rho_2 = -0.4$ have been used. We consider different values of σ^2 , and accordingly the variance of $X(t)$ differs depending on the error process model and the associated parameter values. We generate the data using (1) and the parameters as mentioned above. The LSEs of the parameters are obtained by minimizing the residual sum of squares. The starting estimates of the frequency and the frequency rate are obtained by maximizing the periodogram-like function

$$I(\omega_1, \omega_2) = \frac{1}{N} \left| \sum_{t=1}^N y(t) e^{-i(\omega_1 t + \omega_2 t^2)} \right|^2$$

over a fine two-dimensional grid of $(0, \pi) \times (0, \pi)$. The linear parameters, A and B are expressible in terms of α and β . The minimization of $Q(\boldsymbol{\theta})$ with respect to $\boldsymbol{\theta}$ involves a 2-d search; once the non-linear parameters α and β are estimated, A and B are estimated using linear regression. We replicate this procedure of data generation and estimation 1,000 times, then we calculate the average estimate (AVEEST), the bias (BIAS) and the mean squared error (MSE) of each parameter. We summarize results when the errors are of type (c) in Table 1 ($N = 50$) and Table 2 ($N = 100$). We do not report results with errors of types (a) and (b); they are available at <http://www.isid.ac.in/~statmath/eprints/> (isid/ms/2005/08). In Section 2, we obtained approximate confidence intervals for the unknown parameters based on fixed finite length data, using Theorem 2. Due to the complexity involved in the distributions, these are hard to implement in practice. In the numerical experiments, the convergence of sequences of δ 's and γ 's depends on the parameters and, in many cases, we need a very large value to stabilize the convergence. For this reason, we have used the percentile bootstrap method for interval estimation of the different parameters, as suggested by Nandi, Iyer and Kundu (2002). In each replication of our experiment, we generate 1,000 bootstrap resamples using the estimated parameters and then the bootstrap confidence intervals using the bootstrap quantiles at the 95% nominal level. Then we have 1,000 intervals for each parameter from the replicated experiment. We estimate the 95% bootstrap coverage probability by calculating the proportion covering the true parameter value. We report them as B-COVP in Tables 1 and 2. We also report the average length of the bootstrap confidence interval as B-AVEL. So, in each table, we report the average estimate, its bias and mean squared error, and the 95% bootstrap coverage probability and average length. We have seen in simulations, that the maximizer of the periodogram-like function defined above over a fine grid provides reasonably good initial estimates of α and β in most of the cases.

In these experiments, we collected the LSEs of all parameters estimated in all replications with $N = 50$, type (c) error, and $\sigma^2 = 0.1$. The type (c) error, being an AR(1) process, has variance $\sigma^2/(1 - \rho^2) = 0.13$. We plot the histograms of the LSEs of A and B in Figure 1 and the histograms of the LSEs of α and β in Figure 2. To see how the fitted signal looks, we generated a realization using the same type of error and $\sigma^2 = 0.1$. The fitted signal is plotted in Figure 3, along with the original one.

We observe that average estimates are quite good, which is reflected in the fact that the biases are quite small in absolute value. The MSEs are reasonably small and we observe that they are in decreasing order of (A, B) , α and β . Findings are similar in case of the average bootstrap confidence lengths: the average lengths decrease with (A, B) , α , β . The asymptotic distribution suggests rates of convergence of $N^{-1/2}$, $N^{-3/2}$, $N^{-5/2}$, respectively. These are reflected

Table 1. Average estimates, biases, MSEs, coverage probabilities and average lengths using bootstrap technique when errors are of type (c) and sample size $N = 50$.

		Parameters			
σ^2		A	B	α	β
0.1	AVEEST	2.92884326	1.91418457	2.50015926	0.0999963731
	BIAS	-1.15680695e-3	4.18460369e-3	1.59263611e-4	-3.62843275e-6
	MSE	8.06691125e-3	1.39705129e-2	2.45984756e-5	1.25075923e-8
	B-COVP	0.976	0.973	0.897	0.848
	B-AVEL	0.435647398	0.585830092	1.85300075e-2	3.61990707e-4
0.5	AVEEST	2.92720795	1.91918254	2.50034285	0.0999920592
	BIAS	-2.79211998e-3	9.18257236e-3	3.42845917e-4	-7.94231892e-6
	MSE	4.05863188e-2	6.96346015e-2	1.23846767e-4	6.29799075e-8
	B-COVP	0.977	0.971	0.895	0.849
	B-AVEL	0.965029001	1.29620469	4.16616127e-2	8.12350831e-4
1.0	AVEEST	2.92532682	1.92366946	2.50049949	0.0999882892
	BIAS	-4.67324257e-3	1.36694908e-2	4.99486923e-4	-1.17123127e-5
	MSE	8.10585618e-2	0.138886198	2.47357559e-4	1.25821842e-7
	B-COVP	0.977	0.972	0.894	0.852
	B-AVEL	1.36005151	1.81864798	5.9380278e-2	1.15502137e-3

Table 2. Average estimates, biases, MSEs, coverage probabilities and average lengths using bootstrap technique when errors are of type (c) and sample size $N = 100$.

		Parameters			
σ^2		A	B	α	β
0.1	AVEEST	2.92976546	1.91153288	2.50007677	0.0999990776
	BIAS	-2.34603882e-4	1.53291225e-3	7.67707825e-5	-9.23871994e-7
	MSE	5.03146602e-3	1.00822281e-2	2.46111904e-6	2.25312574e-10
	B-COVP	0.946	0.955	0.962	0.960
	B-AVEL	0.309637368	0.458883166	7.0124059e-3	6.7486304e-5
0.5	AVEEST	2.92800331	1.91039991	2.50014496	0.0999983251
	BIAS	-1.9967556e-3	3.99947166e-4	1.44958496e-4	-1.67638063e-6
	MSE	2.51741707e-2	5.0748501e-2	1.23751706e-5	1.13128373e-9
	B-COVP	0.946	0.957	0.962	0.959
	B-AVEL	0.686493933	1.01766443	1.56924874e-2	1.51613291e-4
1.0	AVEEST	2.92434359	1.90953135	2.50021315	0.0999974459
	BIAS	-5.65648079e-3	-4.68611717e-4	2.1314621e-4	-2.55554914e-6
	MSE	5.05070463e-2	0.101672225	2.4809211e-5	2.26738139e-9
	B-COVP	0.943	0.957	0.965	0.961
	B-AVEL	0.967028975	1.42990315	2.22291183e-2	2.15437249e-4

in the bootstrap intervals to some extent. Moreover, the order of the MSEs approximately match the order given in the asymptotic distribution of the LSEs as expected for finite samples of moderate size. For each type of error, the average lengths of intervals, biases and MSEs increase as the error variance increases for

all the parameters. By comparing the different types of errors and σ^2 , we see that with $N = 50$ the coverage probabilities do not attain the nominal level, mainly for the frequency rate β , except for type (b) error. However, when $N = 100$, the bootstrap coverage probabilities are quite close to the nominal level. In some cases, mainly for the linear parameters, the bootstrap method overestimates the coverage probabilities. We understand that using the given sample size in calculating the limiting δ 's and r 's may cause the overestimation.

We have plotted the histograms of the LSEs in Figures 1 and 2. It is clear from the plots that the LSEs are distributed symmetrically around the true value for all parameters, and the histograms gives a very good idea of the variability of the estimates. In Figure 3, the fitted signal have been plotted with the observed one for a particular case. We see that the fitted one matches reasonably well with the observed one.

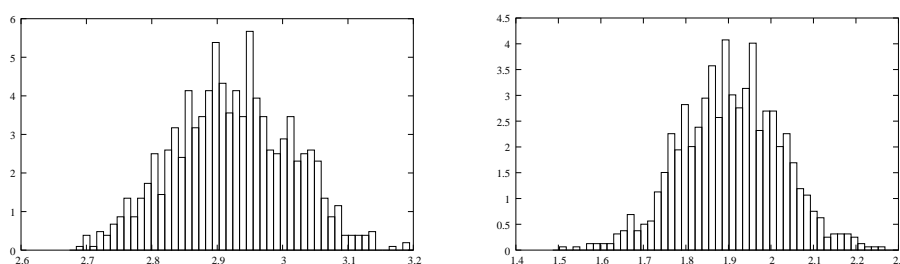


Figure 1. Plot of the histograms of LSEs of A (left plot) and B (right plot).

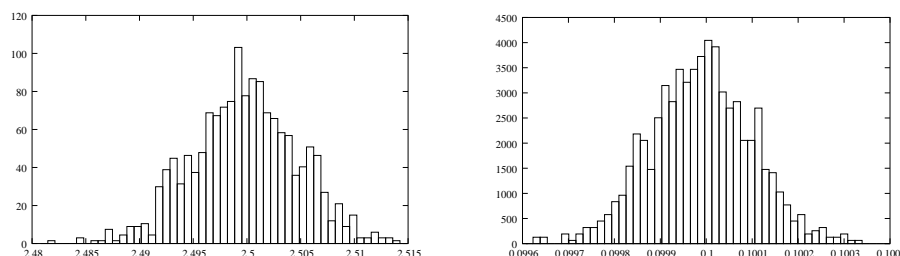


Figure 2. Plot of the histograms of LSEs of α (left plot) and β (right plot).

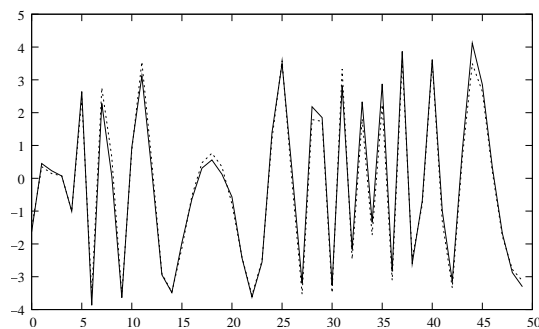


Figure 3. Plot of original signal (solid line) and estimated signals (dotted line).

5. Conclusions

In this paper, we study the problem of estimation of the parameters of the real single chirp signal model, as well as the multiple chirp signal model, in stationary noise. This is a generalization of the multiple frequency model, similar to the way the complex-valued chirp signal is a generalization of the exponential model. We propose LSEs of the unknown parameters and study their asymptotic properties. As the joint asymptotic distribution of the LSEs is quite complicated for practical purposes, we have used a parametric bootstrap method for interval estimation. The results here are quite satisfactory. In simulations, initial estimates of the frequency and the frequency rate are obtained by maximizing a periodogram-like function. It will be interesting to explore the properties of the estimators obtained by maximizing the periodogram like function defined in Section 4. Also, generalization of some of the existing iterative and non-iterative methods for the frequency model to the chirp signal model needs to be addressed, as well as the estimation of the number of chirp components for the multiple chirp model.

Appendix

We first state Lemmas 1 and 2, and then state and prove the lemmas A.1 to A.6 prior to proving Lemma 2.

Lemma 1. *Let $S_{C,M} = \{\boldsymbol{\theta}; \boldsymbol{\theta} = (A_R, A_I, \alpha, \beta), |\boldsymbol{\theta} - \boldsymbol{\theta}^0| \geq 4C, |A_R| \leq M, |A_I| \leq M\}$. If for any $C > 0$ and for some $M < \infty$, $\liminf_{N \rightarrow \infty} \inf_{\boldsymbol{\theta} \in S_{C,M}} [Q(\boldsymbol{\theta}) - Q(\boldsymbol{\theta}^0)]/N > 0$ a.s., then $\hat{\boldsymbol{\theta}}$ is a strongly consistent estimator of $\boldsymbol{\theta}^0$.*

Proof of Lemma 1. The proof can be obtained by contradiction along the lines of Lemma 1 of Wu (1981).

Lemma 2. *As $N \rightarrow \infty$, $\sup_{\alpha, \beta} |\sum_{n=1}^N X(n)e^{i(\alpha n + \beta n^2)}/N| \rightarrow 0$ a.s..*

Lemma A.1. *Let $\{e(n)\}$ be a sequence of i.i.d. random variables with mean zero and finite fourth moment. Then for $k = 2, 3, \dots, N - 2$,*

$$E \left| \sum_{n=1}^{N-2} e(n)e(n+1)^2e(n+2) \right| = O(N^{\frac{1}{2}}), \quad (19)$$

$$E \left| \sum_{n=1}^{N-k-1} e(n)e(n+1)e(n+k)e(n+k+1) \right| = O(N^{\frac{1}{2}}). \quad (20)$$

Proof of Lemma A.1. We prove (19) and then (20) follows similarly. Note that

$$E \left| \sum_{n=1}^{N-2} e(n)e(n+1)^2e(n+2) \right| \leq \left[E \left(\sum_{n=1}^{N-2} e(n)e(n+1)^2e(n+2) \right)^2 \right]^{\frac{1}{2}} = O(N^{\frac{1}{2}}).$$

Lemma A.2. For an arbitrary integer m , $E \sup_{\theta} \left| \sum_{n=1}^N e(n)e(n+k)e^{im\theta n} \right| = O(N^{3/4})$.

Proof of Lemma A.2.

$$\begin{aligned}
E \sup_{\theta} \left| \sum_{n=1}^N e(n)e(n+k)e^{im\theta n} \right| &\leq \left[E \sup_{\theta} \left| \sum_{n=1}^N e(n)e(n+k)e^{im\theta n} \right|^2 \right]^{\frac{1}{2}} \\
&= \left[E \sup_{\theta} \left(\sum_{n=1}^N e(n)e(n+k)e^{im\theta n} \right) \left(\sum_{n=1}^N e(n)e(n+k)e^{-im\theta n} \right) \right]^{\frac{1}{2}} \\
&\leq \left[E \sum_{n=1}^N e(n)^2 e(n+k)^2 + 2E \left| \sum_{n=1}^{N-1} e(n)e(n+1)e(n+k)e(n+k+1) \right| + \dots \right. \\
&\quad \left. + 2E \left| e(1)e(1+k)e(N)e(N+k) \right| \right]^{\frac{1}{2}} \\
&= O(N + N \cdot N^{\frac{1}{2}})^{\frac{1}{2}} \quad (\text{using Lemma A.1}) = O(N^{\frac{3}{4}}).
\end{aligned}$$

Lemma A.3. $E \sup_{\alpha, \beta} \left| \sum_{n=1}^N e(n)e^{i(\alpha n + \beta n^2)} \right|^2 = O(N^{7/4})$.

Proof of Lemma A.3.

$$\begin{aligned}
E \sup_{\alpha, \beta} \left| \sum_{n=1}^N e(n)e^{i(\alpha n + \beta n^2)} \right|^2 &= E \sup_{\alpha, \beta} \left[\sum_{n=1}^N e(n)e^{i(\alpha n + \beta n^2)} \right] \left[\sum_{n=1}^N e(n)e^{-i(\alpha n + \beta n^2)} \right] \\
&\leq O(N + NN^{\frac{3}{4}}) \quad (\text{using Lemma A.2}) = O(N^{\frac{7}{4}}).
\end{aligned}$$

Lemma A.4. $E \sup_{\alpha, \beta} \left| \sum_{n=1}^N e(n)e^{i(\alpha n + \beta n^2)} / N \right| \leq O(N^{-1/8})$.

Proof of Lemma A.4.

$$\begin{aligned}
E \sup_{\alpha, \beta} \left| \frac{1}{N} \sum_{n=1}^N e(n)e^{i(\alpha n + \beta n^2)} \right| &\leq \left[E \sup_{\alpha, \beta} \left| \frac{1}{N} \sum_{n=1}^N e(n)e^{i(\alpha n + \beta n^2)} \right|^2 \right]^{\frac{1}{2}} \\
&= O(N^{-\frac{1}{8}}) \quad (\text{using Lemma A-3}).
\end{aligned}$$

Lemma A.5. $E \sup_{\alpha, \beta} \left| \sum_{n=1}^N X(n)e^{i(\alpha n + \beta n^2)} / N \right| \leq O(N^{-1/8})$.

Proof of Lemma A.5.

$$\begin{aligned}
E \sup_{\alpha, \beta} \left| \frac{1}{N} \sum_{n=1}^N X(n)e^{i(\alpha n + \beta n^2)} \right| &= E \sup_{\alpha, \beta} \left| \frac{1}{N} \sum_{n=1}^N \sum_{k=-\infty}^{\infty} a(k)e(n-k)e^{i(\alpha n + \beta n^2)} \right| \\
&\leq \sum_{k=-\infty}^{\infty} |a(k)| \left[E \sup_{\alpha, \beta} \frac{1}{N} \left| \sum_{n=1}^N e(n-k)e^{i(\alpha n + \beta n^2)} \right| \right].
\end{aligned}$$

Note that $E \sup_{\alpha, \beta} |\sum_{n=1}^N e(n-k)e^{i(\alpha n + \beta n^2)}|/N$ is independent of k , and therefore the result follows using Lemma A.4.

Lemma A.6.

$$\sup_{\alpha, \beta} \left| \frac{1}{N} \sum_{n=1}^N X(n)e^{i(\alpha n + \beta n^2)} \right| \longrightarrow 0, \quad a.s..$$

Proof of Lemma A.6.. Consider the sequence N^9 . Using Lemma A.5 we obtain

$$E \sup_{\alpha, \beta} \frac{1}{N^9} \left| \sum_{n=1}^{N^9} X(n)e^{i(\alpha n + \beta n^2)} \right| \leq O(N^{-\frac{9}{8}}).$$

Therefore, using the Borel Cantelli Lemma, it follows that

$$\sup_{\alpha, \beta} \frac{1}{N^9} \left| \sum_{n=1}^{N^9} X(n)e^{i(\alpha n + \beta n^2)} \right| \longrightarrow 0, \quad a.s..$$

Now for $J, N^9 < J \leq (N+1)^9$,

$$\begin{aligned} & \sup_{\alpha, \beta} \sup_{N^9 < J \leq (N+1)^9} \left| \frac{1}{N^9} \sum_{n=1}^{N^9} X(n)e^{i(\alpha n + \beta n^2)} - \frac{1}{J} \sum_{n=1}^J X(n)e^{i(\alpha n + \beta n^2)} \right| \\ &= \sup_{\alpha, \beta} \sup_{N^9 < J \leq (N+1)^9} \left| \frac{1}{N^9} \sum_{n=1}^{N^9} X(n)e^{i(\alpha n + \beta n^2)} - \frac{1}{N^9} \sum_{n=1}^J X(n)e^{i(\alpha n + \beta n^2)} \right. \\ & \quad \left. + \frac{1}{N^9} \sum_{n=1}^J X(n)e^{i(\alpha n + \beta n^2)} - \frac{1}{J} \sum_{n=1}^J X(n)e^{i(\alpha n + \beta n^2)} \right| \\ & \leq \frac{1}{N^9} \sum_{n=N^9+1}^{(N+1)^9} |X(n)| + \sum_{n=1}^{(N+1)^9} |X(n)| \left(\frac{1}{N^9} - \frac{1}{(N+1)^9} \right). \end{aligned}$$

Note that the mean squared error of the first term is of the order $O(N^{-18} \times ((N+1)^9 - N^9)^2) = O(N^{-2})$. Similarly, the mean squared error of the second term is of the order $O(N^{18} \times [((N+1)^9 - N^9)/N^{18}]^2) = O(N^{-2})$. Therefore, both terms converge to zero almost surely, and that proves the Lemma.

Proof of Theorem 1. In this proof, we write $\hat{\theta}$ as $\hat{\theta}_N = (\hat{A}_N, \hat{B}_N, \hat{\alpha}_N, \hat{\beta}_N)$ to emphasize that $\hat{\theta}$ depends on the sample size. If $\hat{\theta}_N$ is not consistent for θ^0 , then there exists a sub-sequence $\{N_k\}$ of $\{N\}$ such that $\hat{\theta}_{N_k}$ does not converge to θ^0 .

Case I. Suppose $|\hat{A}_{N_k}| + |\hat{B}_{N_k}|$ is not bounded, at least one of $|\hat{A}_{N_k}|$ or $|\hat{B}_{N_k}|$ tends to ∞ . This implies $Q(\hat{\theta}_{N_k})/N_k \rightarrow \infty$. Since $\lim Q(\theta^0)/N_k < \infty$, $[Q(\hat{\theta}_{N_k}) -$

$Q(\hat{\boldsymbol{\theta}}_{N_k})/N_k \rightarrow \infty$. But, as $\hat{\boldsymbol{\theta}}_{N_k}$ is the LSE of $\boldsymbol{\theta}^0$, $Q(\hat{\boldsymbol{\theta}}_{N_k}) - Q(\boldsymbol{\theta}^0) < 0$, which leads to a contradiction.

Case II. Suppose $|\hat{A}_{N_k}| + |\hat{B}_{N_k}|$ is bounded. Then there exists a set $S_{C,M}$ (as defined in Lemma 1) such that $\hat{\boldsymbol{\theta}}_{N_k} \in S_{C,M}$ for some $C > 0$ and $0 < M < \infty$. Write $[Q(\boldsymbol{\theta}) - Q(\boldsymbol{\theta}^0)]/N = f_1(\boldsymbol{\theta}) + f_2(\boldsymbol{\theta})$, where

$$\begin{aligned} f_1(\boldsymbol{\theta}) &= \frac{1}{N} \sum_{n=1}^N \left[A^0 \cos(\alpha^0 n + \beta^0 n^2) - A \cos(\alpha n + \beta n^2) \right. \\ &\quad \left. + B^0 \sin(\alpha^0 n + \beta^0 n^2) - B \sin(\alpha n + \beta n^2) \right]^2, \\ f_2(\boldsymbol{\theta}) &= \frac{2}{N} \sum_{n=1}^N X(n) \left[A^0 \cos(\alpha^0 n + \beta^0 n^2) - A \cos(\alpha n + \beta n^2) \right. \\ &\quad \left. + B^0 \sin(\alpha^0 n + \beta^0 n^2) - B \sin(\alpha n + \beta n^2) \right]. \end{aligned}$$

Using Lemma 2, it follows that

$$\lim_{N \rightarrow \infty} \sup_{\boldsymbol{\theta} \in S_{C,M}} f_2(\boldsymbol{\theta}) = 0 \quad a.s.. \quad (21)$$

Consider the following sets:

$$\begin{aligned} S_{C,M,1} &= \{\boldsymbol{\theta} : \boldsymbol{\theta} = (A, B, \alpha, \beta), |A - A^0| \geq C, |A| \leq M, |B| \leq M\}, \\ S_{C,M,2} &= \{\boldsymbol{\theta} : \boldsymbol{\theta} = (A, B, \alpha, \beta), |B - B^0| \geq C, |A| \leq M, |B| \leq M\}, \\ S_{C,M,3} &= \{\boldsymbol{\theta} : \boldsymbol{\theta} = (A, B, \alpha, \beta), |\alpha - \alpha^0| \geq C, |A| \leq M, |B| \leq M\}, \\ S_{C,M,4} &= \{\boldsymbol{\theta} : \boldsymbol{\theta} = (A, B, \alpha, \beta), |\beta - \beta^0| \geq C, |A| \leq M, |B| \leq M\}. \end{aligned}$$

Note that $S_{C,M} \subset S_{C,M,1} \cup S_{C,M,2} \cup S_{C,M,3} \cup S_{C,M,4} = S$ (say). Therefore,

$$\underline{\lim}_{\boldsymbol{\theta} \in S_{C,M}} \inf \frac{1}{N} [Q(\boldsymbol{\theta}) - Q(\boldsymbol{\theta}^0)] \geq \underline{\lim}_{\boldsymbol{\theta} \in S} \inf \frac{1}{N} [Q(\boldsymbol{\theta}) - Q(\boldsymbol{\theta}^0)]. \quad (22)$$

First we show that

$$\underline{\lim}_{\boldsymbol{\theta} \in S_{C,M,j}} \inf \frac{1}{N} [Q(\boldsymbol{\theta}) - Q(\boldsymbol{\theta}^0)] > 0 \quad a.s., \quad (23)$$

for $j = 1, \dots, 4$, and then, because of (22), one has $\underline{\lim}_{\boldsymbol{\theta} \in S_{C,M}} \inf [Q(\boldsymbol{\theta}) - Q(\boldsymbol{\theta}^0)]/N > 0$ a.s.. Because of Lemma 1, Theorem 1 is proved provided we can show (23). First consider $j = 1$. Using (21), it follows that

$$\underline{\lim}_{\boldsymbol{\theta} \in S_{C,M,1}} \inf \frac{1}{N} [Q(\boldsymbol{\theta}) - Q(\boldsymbol{\theta}^0)] = \underline{\lim}_{\boldsymbol{\theta} \in S_{C,M,1}} \inf f_1(\boldsymbol{\theta})$$

$$\begin{aligned}
&= \underline{\lim}_{|A-A^0| \geq C} \inf \frac{1}{N} \sum_{n=1}^N [A^0 \cos(\alpha^0 n + \beta^0 n^2) - A \cos(\alpha n + \beta n^2) \\
&\quad + B^0 \sin(\alpha^0 n + \beta^0 n^2) - B \sin(\alpha n + \beta n^2)]^2 \\
&= \lim_{N \rightarrow \infty} \inf_{|A-A^0| \geq C} \frac{1}{N} \sum_{n=1}^N \cos^2(\alpha^0 n + \beta^0 n^2) (A - A^0)^2 \\
&\geq C^2 \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1}^N \cos^2(\alpha^0 n + \beta^0 n^2) > 0.
\end{aligned}$$

For other j one proceeds along the same lines and that proves Theorem 1.

Acknowledgement

The authors would like to thank the referee and an associate editor for some constructive suggestions, and the Editor Professor Jane-Ling Wang for encouragement.

References

- Abatzoglou, T. (1986). Fast maximum likelihood joint estimation of frequency and frequency rate. *IEEE Trans. on Aerosp. Electron. Syst.* **22**, 708-714.
- Djurić, P. M. and Kay, S. M. (1990). Parameter estimation of chirp signals. *IEEE Trans. on Acoustics, Speech and Signal Process.* **38**, 2118-2126.
- Fuller, W. A. (1976), *Introduction to Statistical Time Series*. John Wiley and Sons, New York.
- Giannakis, G. B. and Zhou, G. (1995), Harmonics in multiplicative and additive noise: Parameter estimation using cyclic statistics. *IEEE Trans. Signal Process.* **43**, 2217-2221.
- Gini, F., Montanari, M. and Verrazzani, L. (2000). Estimation of chirp signals in compound Gaussian clutter: a cyclostationary approach. *IEEE Trans. on Acoustics, Speech and Signal Process.* **48**, 1029-1039.
- Jennrich, R. I. (1969). Asymptotic properties of the nonlinear least squares estimators. *Ann. Math. Statist.* **40**, 633-643.
- Kumaresan, R. and Verma, S. (1987). On estimating the parameters of chirp signals using rank reduction techniques. *Proc. 21st. Asilomar Conf.*, 555-558, Pacific Grove, CA,
- Kundu, D. (1997), Asymptotic properties of the least squares estimators of sinusoidal signals. *Statistics* **30**, 221-238.
- Nandi, S., Iyer, S. K. and Kundu, D. (2002). Estimation of frequencies in presence of heavy tail errors. *Statist. Probab. Lett.* **58**, 265-282.
- Nandi, S. and Kundu, D. (2004). Asymptotic properties of the least squares estimators of the parameters of the chirp signals. *Ann. Inst. Statist. Math.* **56**, 529-544.
- Saha, S. and Kay, S. M. (2002). Maximum likelihood parameter estimation of superimposed chirps using Monte Carlo importance sampling. *IEEE Trans. Signal Process.* **50**, 224-230.
- Shamsunder, S., Giannakis, G. B. and Friedlander, B. (1995). Estimation random amplitude polynomial phase signals: a cyclostationary approach. *IEEE Trans. Signal Process.* **43**, 492-505.

- Swami, A. (1996). Cramér-Rao bounds for deterministic signals in additive and multiplicative noise. *Signal Process.* **53**, 231-244.
- Wu, C. F. J. (1981). Asymptotic theory of the nonlinear least squares estimation. *Ann. Statist.* **9**, 501-513.
- Zhou, G. and Giannakis, G. B. (1995). Harmonics in multiplicative and additive noise: Performance analysis of cyclic estimators. *IEEE Trans. Signal Process.* **43**, 1445-1460.
- Zhou, G., Giannakis, G. B. and Swami, A. (1996). On polynomial phase signals with time-varying amplitudes. *IEEE Trans. Signal Process.* **44**, 848-861.

Department of Mathematics and Statistics, Indian Institute of Technology, Kanpur, Kanpur-208016, India.

E-mail: kundu@iitk.ac.in

Theoretical Statistics and Mathematics Unit, Indian Statistical Institute, Delhi Centre 7, S.J.S. Sansanwal Marg, New Delhi - 110016, India.

E-mail: nandi@isid.ac.in

(Received July 2005; accepted April 2006)