isid/ms/2005/08
July 25, 2005
http://www.isid.ac.in/~statmath/eprints

Parameter Estimation of Chirp Signals in Presence of Stationary Noise

Debasis Kundu
Swagata Nandi

Indian Statistical Institute, Delhi Centre
7, SJSS Marg, New Delhi–110 016, India

Parameter Estimation of Chirp Signals in Presence of Stationary Noise

Debasis Kundu[1] Swagata Nandi[2]

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 it is observed that the least squares estimators 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 for small sample sizes.

Key Words and Phrases: Asymptotic distributions; Chirp signals; Least squares estimators; Multiple chirp signals.

Short Running Title: Chirp signals in stationary noise.

MS Classifications: 62J02; 62E25.

Address of correspondence: Swagata Nandi, e-mail: nandi@isid.ac.in, Phone: 91-11-51493931, Fax: 91-11-51493981.

[1] Department of Mathematics and Statistics, Indian Institute of Technology Kanpur, Pin 208016, India.

[2] Stat-Math Unit, Indian Statistical Institute, 7, S.J.S. Sansanwal Marg, New Delhi - 110016.

1 Introduction

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

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

Here $y(n)$ is the real valued signal observed at $n = 1, \ldots, N$. A^0 and B^0 are real-valued amplitudes and α^0 and β^0 are the frequency and frequency rate respectively. The error random variable $\{X(n)\}$ is a sequence of random variables with mean zero and finite fourth moment. The error random variable $X(n)$ satisfies the following assumption:

Assumption 1: The error random variable $\{X(n)\}$ can be written in the following form;

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

Here $\{e(n)\}$ is a sequence of independent and identically distributed (i.i.d.) random variables with mean zero and finite fourth moment. The coefficients $a(j)$s satisfy the following condition;

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

The signals as described in (1) are known as the chirp signals in the statistical signal processing literature (Djurić and Kay; 1990). Chirp signals are quite common in various areas of science and engineering, specifically in sonar, radar, communications, etc. Several authors considered the chirp signals model (1) when $X(n)$s are i.i.d. random variables. See for example, works of Abatzoglon (1986), Kumaresn and Verma (1987), Djurić and Kay (1990), Gini, Montanari and Verrazzani (2000), Nandi and Kundu (2004) etc. It is well known that in most practical situations the errors may not be independent. We assume stationarity through assumption 1 to make the model more realistic.

In this paper, we discuss the chirp signal model in presence of stationary noise. We consider the least squares estimators and study their properties, when the errors satisfy assumption 1. It is well known that the simple sum of sinusoidal model does not satisfy the sufficient conditions of Jennrich (1969) or Wu (1981) for the least squares estimators to be consistent. So the model (1) as a generalization of sinusoidal model also does so and it is not clear how the least squares estimators will behave in this more complicated situation.

It is observed that the least squares estimators are strongly consistent and 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 also be constructed.

The rest of the paper is organized as follows. In the section 2, we provide the consistency results of the least squares estimators. The asymptotic distributions of the least squares estimators are derived in section 3. The case of multiple chirp model is discussed in section 4. Some numerical results are presented in section 5 and finally we conclude the paper in section 6.

2 Consistency of the LSEs

Let us use the following notation: $\boldsymbol{\theta} = (A, B, \alpha, \beta)$, $\boldsymbol{\theta}^0 = (A^0, B^0, \alpha^0, \beta^0)$. Then, the least squares estimator (LSE) of $\boldsymbol{\theta}^0$, say $\hat{\boldsymbol{\theta}} = (\hat{A}, \hat{B}, \hat{\alpha}, \hat{\beta})$, can be obtained by minimizing

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

with respect to A, B, α and β. Now, we have the following result.

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 $\boldsymbol{\Theta} = (-\infty, \infty) \times (-\infty, \infty) \times [0, \pi] \times [0, \pi]$ and $A^{0^2} + B^{0^2} > 0$. If the error random variables $X(n)$ satisfy assumption 1, then $\hat{\boldsymbol{\theta}}$, the LSE of $\boldsymbol{\theta}^0$, is a strongly consistent estimator of $\boldsymbol{\theta}^0$.

We need the following lemmas to prove theorem 1.

Lemma 1: Let us denote

$$S_{C,M} = \left\{ \boldsymbol{\theta}; \boldsymbol{\theta} = (A_R, A_I, \alpha, \beta), |\boldsymbol{\theta} - \boldsymbol{\theta}^0| \geq 4C, |A_R| \leq M, |A_I| \leq M \right\}.$$

If for any $C > 0$ and for some $M < \infty$,

$$\liminf_{N \to \infty} \inf_{\boldsymbol{\theta} \in S_{C,M}} \frac{1}{N} \left[Q(\boldsymbol{\theta}) - Q(\boldsymbol{\theta}^0) \right] > 0 \quad a.s.$$

then $\hat{\boldsymbol{\theta}}$ is a strongly consistent estimator of $\boldsymbol{\theta}^0$.

Proof of Lemma 1: It is simple and therefore, it is omitted.

Lemma 2: As $N \to \infty$,

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

Proof of Lemma 2: See in the Appendix A.

Proof of Theorem 1: In this proof, we denote $\hat{\boldsymbol{\theta}}$ by $\hat{\boldsymbol{\theta}}_N = (\hat{A}_N, \hat{B}_N, \hat{\alpha}_N, \hat{\beta}_N)$ to emphasize that $\hat{\boldsymbol{\theta}}$ depends on the sample size. If $\hat{\boldsymbol{\theta}}_N$ is not consistent for $\boldsymbol{\theta}^0$, then either:

Case I: For all subsequences $\{N_k\}$ of $\{N\}$, $|\hat{A}_{N_k}| + |\hat{B}_{N_k}| \to \infty$. This implies

$$\frac{1}{N_k} \left[Q(\hat{\boldsymbol{\theta}}_{N_k}) - Q(\boldsymbol{\theta}^0) \right] \to \infty.$$

But as $\hat{\boldsymbol{\theta}}_{N_k}$ is the LSE of $\boldsymbol{\theta}^0$, therefore,

$$Q(\hat{\boldsymbol{\theta}}_{N_k}) - Q(\boldsymbol{\theta}^0) < 0,$$

which leads to a contradiction. So $\hat{\boldsymbol{\theta}}_N$ is a strongly consistent estimator of $\boldsymbol{\theta}^0$.

Case II: For at least one subsequence $\{N_k\}$ of $\{N\}$, $\hat{\boldsymbol{\theta}}_{N_k} \in S_{C,M}$, for some $C > 0$ and for an $0 < M < \infty$. Now let us write

$$\frac{1}{N} \left[Q(\boldsymbol{\theta}) - Q(\boldsymbol{\theta}^0) \right] = f_1(\boldsymbol{\theta}) + f_2(\boldsymbol{\theta}),$$

where

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

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

Using lemma 2, it follows that

$$\lim_{N\to\infty}\sup_{\boldsymbol{\theta}\in S_{C,M}} f_2(\boldsymbol{\theta}) = 0 \quad a.s. \tag{3}$$

Now consider the following sets;

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

Note that

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

Therefore,

$$\underline{\lim}\inf_{\boldsymbol{\theta}\in S_{C,M}} \frac{1}{N}\Big[Q(\boldsymbol{\theta})-Q(\boldsymbol{\theta}^0)\Big] \geq \underline{\lim}\inf_{\boldsymbol{\theta}\in S} \frac{1}{N}\Big[Q(\boldsymbol{\theta})-Q(\boldsymbol{\theta}^0)\Big]. \tag{4}$$

First we show that

$$\underline{\lim}\inf_{\boldsymbol{\theta}\in S_{C,M,j}} \frac{1}{N}\Big[Q(\boldsymbol{\theta})-Q(\boldsymbol{\theta}^0)\Big] > 0 \qquad a.s., \tag{5}$$

for $j=1,\ldots,4$ and because of (4), it implies

$$\underline{\lim}\inf_{\boldsymbol{\theta}\in S_{C,M}} \frac{1}{N}\Big[Q(\boldsymbol{\theta})-Q(\boldsymbol{\theta}^0)\Big] > 0 \qquad a.s.$$

Therefore, due to lemma 1, theorem 1 is proved, provided we can show (5). First consider $j=1$ to prove (5). Using (3), it follows that

$$\begin{aligned} &\underline{\lim}\inf_{\boldsymbol{\theta}\in S_{C,M,1}} \frac{1}{N}\Big[Q(\boldsymbol{\theta})-Q(\boldsymbol{\theta}^0)\Big] = \underline{\lim}\inf_{\boldsymbol{\theta}\in S_{C,M,1}} f_1(\boldsymbol{\theta}) \\ &= \underline{\lim}\inf_{|A-A^0|\geq C} \frac{1}{N}\sum_{n=1}^{N}\Big[A^0\cos(\alpha^0 n+\beta^0 n^2) - A\cos(\alpha n+\beta n^2)+ \end{aligned}$$

$$B^0\sin(\alpha^0 n+\beta^0 n^2)-B\cos(\alpha n+\beta n^2)\Big]^2$$

$$= \lim_{N\to\infty}\inf_{|A-A^0|\geq C}\frac{1}{N}\sum_{n=1}^{N}\cos^2(\alpha^0 n+\beta^0 n^2)(A-A^0)^2$$

$$\geq C^2\lim_{N\to\infty}\frac{1}{N}\sum_{n=1}^{N}\cos^2(\alpha^0 n+\beta^0 n^2)>0.$$

For other j also, it can be shown along the same line and that proves theorem 1.

3 Asymptotic Distribution of the LSE

In this section we compute the asymptotic distribution of the least squares estimators. 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. Now expanding $Q'(\hat{\boldsymbol{\theta}})$ around the true parameter value $\boldsymbol{\theta}^0$ by Taylor series, we obtain

$$Q'(\hat{\boldsymbol{\theta}})-Q'(\boldsymbol{\theta}^0)=(\hat{\boldsymbol{\theta}}-\boldsymbol{\theta}^0)Q''(\bar{\boldsymbol{\theta}}), \tag{6}$$

here $\bar{\boldsymbol{\theta}}$ is a point on the line joining the points $\hat{\boldsymbol{\theta}}$ and $\boldsymbol{\theta}^0$. Suppose $\mathbf{D}$ is a 4×4 diagonal matrix as follows;

$$\mathbf{D}=\operatorname{diag}\left\{N^{-\frac{1}{2}},N^{-\frac{1}{2}},N^{-\frac{3}{2}},N^{-\frac{5}{2}}\right\}.$$

Since $Q'(\hat{\boldsymbol{\theta}})=0$, therefore (6) can be written as

$$(\hat{\boldsymbol{\theta}}-\boldsymbol{\theta}^0)\mathbf{D}^{-1}=-\left[\mathbf{Q}'(\boldsymbol{\theta}^0)\mathbf{D}\right]\left[\mathbf{D}\mathbf{Q}''(\bar{\boldsymbol{\theta}})\mathbf{D}\right]^{-1}, \tag{7}$$

as $\left[\mathbf{D}\mathbf{Q}''(\bar{\boldsymbol{\theta}})\mathbf{D}\right]$ is an invertible matrix $a.e.$ for large N. From Theorem 1, it follows that $\hat{\boldsymbol{\theta}}$ converges $a.e.$ to $\boldsymbol{\theta}^0$ and since each element of $Q''(\bar{\boldsymbol{\theta}})$ is a continuous function of $\boldsymbol{\theta}$, therefore,

$$\lim_{N\to\infty}\left[\mathbf{D}\mathbf{Q}''(\bar{\boldsymbol{\theta}})\mathbf{D}\right]=\lim_{N\to\infty}\left[\mathbf{D}\mathbf{Q}''(\boldsymbol{\theta}^0)\mathbf{D}\right]=2\boldsymbol{\Sigma}(\boldsymbol{\theta}^0)\quad\text{(say)}.$$

Now let us look at different elements of the matrix $\boldsymbol{\Sigma}(\boldsymbol{\theta})=(\sigma_{jk}(\boldsymbol{\theta}))$. We will use the following result

$$\lim_{N\to\infty}\frac{1}{N^p}\sum_{n=1}^{N}n^{p-1}=\frac{1}{p}\qquad\text{for } p=1,2,\ldots.$$

and the following notation:

$$\lim_{N\to\infty} \frac{1}{N^{p+1}} \sum_{n=1}^{N} n^p \cos^k(\alpha n + \beta n^2) = \delta_k(p, \alpha, \beta), \tag{8}$$

$$\lim_{N\to\infty} \frac{1}{N^{p+1}} \sum_{n=1}^{N} n^p \sin^k(\alpha n + \beta n^2) = \gamma_k(p, \alpha, \beta). \tag{9}$$

Here k takes values 1 and 2. Using these notation for limits, we compute the elements of $\mathbf{\Sigma}(\boldsymbol{\theta})$ by routine calculations and are as follows:

$$\sigma_{11}(\boldsymbol{\theta}) = \delta_2(0, \alpha, \beta), \quad \sigma_{12}(\boldsymbol{\theta}) = \frac{1}{2}\gamma_1(0, 2\alpha, 2\beta), \quad \sigma_{13}(\boldsymbol{\theta}) = -\frac{1}{2}A\gamma_1(1, 2\alpha, 2\beta) + B\delta_2(1, \alpha, \beta),$$

$$\sigma_{14}(\boldsymbol{\theta}) = -\frac{1}{2}A\gamma_1(2, 2\alpha, 2\beta) + B\delta_2(2, \alpha, \beta), \quad \sigma_{22}(\boldsymbol{\theta}) = \gamma_2(0, \alpha, \beta),$$

$$\sigma_{23}(\boldsymbol{\theta}) = -A\gamma_2(1, \alpha, \beta) + \frac{1}{2}B\gamma_1(1, 2\alpha, 2\beta), \quad \sigma_{24}(\boldsymbol{\theta}) = -A\gamma_2(2, \alpha, \beta) + \frac{1}{2}B\gamma_1(2, 2\alpha, 2\beta),$$

$$\sigma_{33}(\boldsymbol{\theta}) = A^2\gamma_2(2, \alpha, \beta) + B^2\delta_2(2, \alpha, \beta) - AB\gamma_1(2, 2\alpha, 2\beta),$$

$$\sigma_{34}(\boldsymbol{\theta}) = A^2\gamma_2(3, \alpha, \beta) + B^2\delta_2(3, \alpha, \beta) - AB\gamma_1(3, 2\alpha, 2\beta),$$

$$\sigma_{44}(\boldsymbol{\theta}) = A^2\gamma_2(4, \alpha, \beta) + B^2\delta_2(4, \alpha, \beta) - AB\gamma_1(4, 2\alpha, 2\beta).$$

The 4×1 random vector $\left[\mathbf{Q}'(\boldsymbol{\theta}^0)\mathbf{D}\right]$ takes the form;

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

Now using the central limit theorem of stochastic processes (see Fuller; 1976, page 251), it follows that $\left[\mathbf{Q}'(\boldsymbol{\theta}^0)\mathbf{D}\right]$ tends to a 4-variate normal distribution as given below;

$$\left[\mathbf{Q}'(\boldsymbol{\theta}^0)\mathbf{D}\right] \xrightarrow{d} \mathcal{N}_4(\mathbf{0}, \mathbf{G}(\boldsymbol{\theta}^0)), \tag{10}$$

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

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

$$g_{13}(\boldsymbol{\theta}) = \lim_{N\to\infty} \frac{4}{N^2} E[S_1 S_3], \qquad g_{14}(\boldsymbol{\theta}) = \lim_{N\to\infty} \frac{4}{N^3} E[S_1 S_4], \tag{12}$$

$$g_{22}(\boldsymbol{\theta}) = \lim_{N\to\infty} \frac{4}{N} E[S_2]^2, \qquad g_{23}(\boldsymbol{\theta}) = \lim_{N\to\infty} \frac{4}{N^2} E[S_2 S_3], \tag{13}$$

$$g_{24}(\boldsymbol{\theta}) = \lim_{N\to\infty} \frac{4}{N^3} E[S_2 S_4], \qquad g_{33}(\boldsymbol{\theta}) = \lim_{N\to\infty} \frac{4}{N^3} E[S_3]^2, \tag{14}$$

$$g_{34}(\boldsymbol{\theta}) = \lim_{N\to\infty} \frac{4}{N^4} E[S_3 S_4], \qquad g_{44}(\boldsymbol{\theta}) = \lim_{N\to\infty} \frac{4}{N^5} E[S_4]^2, \tag{15}$$

where

$$\begin{aligned}
S_1 &= \sum_{n=1}^{N} X(n)\cos(\alpha n + \beta n^2), \qquad S_2 = \sum_{n=1}^{N} X(n)\sin(\alpha n + \beta n^2),\\
S_3 &= \sum_{n=1}^{N} nX(n)\left[A\sin(\alpha n + \beta n^2) - B\cos(\alpha n + \beta n^2)\right],\\
S_4 &= \sum_{n=1}^{N} n^2 X(n)\left[A\sin(\alpha n + \beta n^2) - B\cos(\alpha n + \beta n^2)\right].
\end{aligned}$$

For $k < j$, $g_{jk}(\boldsymbol{\theta}) = g_{kj}(\boldsymbol{\theta})$. These limits given in (11) to (15) exist for fixed values of $\boldsymbol{\theta}$ because of (8) and (9). Therefore, from (7) the following theorem follows.

THEOREM 2: Under the same assumptions as in Theorem 1,

$$(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}^0)\mathbf{D}^{-1} \xrightarrow{d} \mathcal{N}_4\left[\mathbf{0}, \boldsymbol{\Sigma}^{-1}(\boldsymbol{\theta}^0)\mathbf{G}(\boldsymbol{\theta}^0)\boldsymbol{\Sigma}^{-1}(\boldsymbol{\theta}^0)\right]. \tag{16}$$

REMARK 1: When $X(n)$s are *i.i.d.* random variables, then the covariance matrix takes the simplified form

$$\boldsymbol{\Sigma}^{-1}(\boldsymbol{\theta}^0)\mathbf{G}(\boldsymbol{\theta}^0)\boldsymbol{\Sigma}^{-1}(\boldsymbol{\theta}^0) = 4\sigma^2\boldsymbol{\Sigma}^{-1}(\boldsymbol{\theta}^0).$$

REMARK 2: Although we could not prove it theoretically, but it is observed by extensive numerical computations that the right hand side limits of (8) and (9) for $k = 1, 2$ do not depend on α. So we assume that these quantities are independent of their second argument and we write them as

$$\delta_k(p;\beta) = \delta_k(p,\alpha,\beta), \qquad \gamma_k(p;\beta) = \gamma_k(p,\alpha,\beta).$$

Let us denote

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

c_c and c_s are functions of α^0 and β^0, but we do not make it explicit here to keep the notation simple.

Now according to the above assumption, δs and γs are independent of α and based on it, we can explicitly compute the elements of $\mathbf{G}$ matrix as follows:

$$
\begin{aligned}
g_{11}(\boldsymbol{\theta}^0) &= 4\sigma^2\left[c_c^2\delta_2(0;\beta^0) + c_s^2\gamma_2(0;\beta^0) - c_c c_s\gamma_1(0;2\beta^0)\right], \\
g_{12}(\boldsymbol{\theta}^0) &= 4\sigma^2\left[\frac{1}{2}(c_c^2 - c_s^2)\gamma_1(0;2\beta^0) + c_c c_s(\delta_2(0;\beta^0) - \gamma_2(0;\beta^0))\right], \\
g_{13}(\boldsymbol{\theta}^0) &= 4\sigma^2\Big[\gamma_2(1;\beta^0)(Ac_c + Bc_s)c_s - \delta_2(1;\beta^0)(Ac_s - Bc_c)c_c \\
&\qquad\qquad\qquad\qquad -\frac{1}{2}\gamma_1(1;2\beta^0)(Ac_c^2 - Ac_s^2 - Bc_c c_s)\Big], \\
g_{14}(\boldsymbol{\theta}^0) &= 4\sigma^2\Big[\gamma_2(2;\beta^0)(Ac_c + Bc_s)c_s - \delta_2(2;\beta^0)(Ac_s - Bc_c)c_c \\
&\qquad\qquad\qquad\qquad -\frac{1}{2}\gamma_1(2;2\beta^0)(Ac_c^2 - Ac_s^2 - Bc_c c_s)\Big], \\
g_{22}(\boldsymbol{\theta}^0) &= 4\sigma^2\left[c_c^2\gamma_2(0;\beta^0) + c_s^2\delta_2(0;\beta^0) + c_c c_s\gamma_1(0;2\beta^0)\right], \\
g_{23}(\boldsymbol{\theta}^0) &= -4\sigma^2\Big[\gamma_2(1;\beta^0)(Ac_c + Bc_s)c_c + \delta_2(1;\beta^0)(Ac_s - Bc_c)c_s \\
&\qquad\qquad\qquad\qquad -\frac{1}{2}\gamma_1(1:2\beta^0)(Bc_c^2 - Bc_s^2 - Ac_c c_s)\Big], \\
g_{24}(\boldsymbol{\theta}^0) &= -4\sigma^2\Big[\gamma_2(2;\beta^0)(Ac_c + Bc_s)c_c + \delta_2(2;\beta^0)(Ac_s - Bc_c)c_s \\
&\qquad\qquad\qquad\qquad -\frac{1}{2}\gamma_1(2;2\beta^0)(Bc_c^2 - Bc_s^2 - Ac_c c_s)\Big], \\
g_{33}(\boldsymbol{\theta}^0) &= 4\sigma^2\Big[\gamma_2(2;\beta^0)(Ac_c + Bc_s)^2 + \delta_2(2;\beta^0)(Ac_s - Bc_c)^2 \\
&\qquad\qquad\qquad\qquad +\gamma_1(2;2\beta^0)\{(A^2 - B^2)c_c c_s - AB(c_c^2 - c_s^2)\}\Big], \\
g_{34}(\boldsymbol{\theta}^0) &= 4\sigma^2\Big[\gamma_2(3;\beta^0)(Ac_c + Bc_s)^2 + \delta_2(3;\beta^0)(Ac_s - Bc_c)^2 \\
&\qquad\qquad\qquad\qquad +\gamma_1(3;2\beta^0)\{(A^2 - B^2)c_c c_s - AB(c_c^2 - c_s^2)\}\Big], \\
g_{44}(\boldsymbol{\theta}^0) &= 4\sigma^2\Big[\gamma_2(4;\beta^0)(Ac_c + Bc_s)^2 + \delta_2(4;\beta^0)(Ac_s - Bc_c)^2 \\
&\qquad\qquad\qquad\qquad +\gamma_1(4;2\beta^0)\{(A^2 - B^2)c_c c_s - AB(c_c^2 - c_s^2)\}\Big].
\end{aligned}
$$

So obtaining the explicit expressions of entries of variance-covariance matrix of $(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}^0)\mathbf{D}^{-1}$ is possible by inverting the matrix $\boldsymbol{\Sigma}(\boldsymbol{\theta}^0)$. But they are not provided here due to the complex (notational) structure of matrices $\boldsymbol{\Sigma}(\boldsymbol{\theta}^0)$ and $\mathbf{G}(\boldsymbol{\theta}^0)$. If the true value of β is zero (i.e. frequency does not change over time) and if that information is used in the model, then the model (1) is nothing but the usual sinusoidal model. In that case amplitudes are asymptotically independent of the frequency which has not been observed in case of chirp signal model.

4 Multiple Chirp Signal

In this section, we introduce the multiple chirp signal model in stationary noise. The complex-valued single chirp model was generalized as superimposed chirps by Saha and Kay (2002). The following model is a similar generalization of model (1). We assume that the observed data $y(n)$ have the following representation.

$$y(n) = \sum_{k=1}^{p} \Big[A_k^0 \cos(\alpha_k^0 n + \beta_k^0 n^2) + B_k^0 \sin(\alpha_k^0 n + \beta_k^0 n^2)\Big] + X(n); \quad n = 1, \ldots, N. \tag{17}$$

Similarly as single chirp, the parameters $\alpha_k^0, \beta_k^0 \in (0, \pi)$ are frequency and frequency rate respectively. A_k^0s and B_k^0s are real-valued amplitudes. Again the aim is to estimate the parameters and study their properties. We assume that the number of components, p is known and $X(n)$s satisfy assumption 1. Estimation of p is an important problem and will be addressed elsewhere. Now let us define, $\boldsymbol{\theta}_k = (A_k, B_k, \alpha_k, \beta_k)$ and $\boldsymbol{\nu} = (\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_p)$ be the parameter vector. The least squares estimators of the parameters are obtained by minimizing the objective function, say $R(\boldsymbol{\nu})$ (defined similarly as $Q(\boldsymbol{\theta})$; see eq. (2), sec. 2). Let $\hat{\boldsymbol{\nu}}$ and $\boldsymbol{\nu}^0$ denote the least squares estimator and the true value of $\boldsymbol{\nu}$. The consistency of $\hat{\boldsymbol{\nu}}$ follows similarly as the consistency of $\hat{\boldsymbol{\theta}}$, considering the parameter vector as $\boldsymbol{\nu}$. We will state the asymptotic distribution of $\hat{\boldsymbol{\nu}}$ here. The proof involves routine calculations and use of multiple Taylor series expansion and central limit theorem for stochastic processes. For asymptotic distribution of $\hat{\boldsymbol{\nu}}$, we introduce the following notation; $\boldsymbol{\psi}_k^N = (\hat{\boldsymbol{\theta}}_k - \boldsymbol{\theta}_k^0)\mathbf{D}^{-1} = \Big(N^{1/2}(\hat{A}_k - A_k^0), N^{1/2}(\hat{B}_k - B_k^0), N^{3/2}(\hat{\alpha}_k - \alpha_k^0), N^{5/2}(\hat{\beta}_k - \beta_k^0)\Big)$, moreover c_c^k and c_s^k are obtained from c_c and c_s by replacing α^0 and β^0 by α_k^0 and β_k^0 respectively. Let us denote $\beta_j + \beta_k = \beta_{jk}^+$, $\beta_j - \beta_k = \beta_{jk}^-$, $d_1 = c_c^1 c_c^2 + c_s^1 c_s^2$, $d_2 = c_c^1 c_s^2 + c_s^1 c_c^2$, $d_3 = c_c^1 c_c^2 - c_s^1 c_s^2$ and $d_4 = c_c^1 c_s^2 - c_s^1 c_c^2$. Then the asymptotic distribution of $(\boldsymbol{\psi}_1^N, \ldots, \boldsymbol{\psi}_p^N)$ is as follows.

$$(\boldsymbol{\psi}_1^N, \ldots, \boldsymbol{\psi}_p^N) \xrightarrow{d} \mathcal{N}_{4p}\left(\mathbf{0}, 2\sigma^2 \Lambda^{-1}(\boldsymbol{\nu}^0)\mathbf{H}(\boldsymbol{\nu}^0)\Lambda^{-1}(\boldsymbol{\nu}^0)\right), \tag{18}$$

$$\Lambda(\boldsymbol{\nu}) = \begin{pmatrix} \boldsymbol{\Lambda}_{11} & \boldsymbol{\Lambda}_{12} & \cdots & \boldsymbol{\Lambda}_{1p} \\ \boldsymbol{\Lambda}_{21} & \boldsymbol{\Lambda}_{22} & \cdots & \boldsymbol{\Lambda}_{2p} \\ \vdots & \vdots & \vdots & \vdots \\ \boldsymbol{\Lambda}_{p1} & \boldsymbol{\Lambda}_{p2} & \cdots & \boldsymbol{\Lambda}_{pp} \end{pmatrix}, \qquad \mathbf{H}(\boldsymbol{\nu}) = \begin{pmatrix} \mathbf{H}_{11} & \mathbf{H}_{12} & \cdots & \mathbf{H}_{1p} \\ \mathbf{H}_{21} & \mathbf{H}_{22} & \cdots & \mathbf{H}_{2p} \\ \vdots & \vdots & \vdots & \vdots \\ \mathbf{H}_{p1} & \mathbf{H}_{p2} & \cdots & \mathbf{H}_{pp} \end{pmatrix}. \tag{19}$$

The sub-matrices $\boldsymbol{\Lambda}_{jk}$ and $\mathbf{H}_{jk}$ are of the 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 $\boldsymbol{\theta} = \boldsymbol{\theta}_j$. The entries of off-diagonal sub-matrices $\boldsymbol{\Lambda}_{jk} = ((\lambda_{rs}))$ and $\mathbf{H}_{jk} = ((h_{rs}))$ are given in appendix B. The elements of matrices $\boldsymbol{\Lambda}_{jk}$ and $\mathbf{H}_{jk}$ are non-zero. So 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. frequency rates βs vanish, the model (17) is equivalent to multiple frequency model. In such case, the off-diagonal matrices in $\mathbf{H}$ and $\boldsymbol{\Lambda}$ are zero matrices and the estimators of parameters in different components are independent. This is due to the reason that $\delta_1(p, \alpha, 0) = 0 = \gamma_1(p, \alpha, 0)$ for all $p \geq 0$ and $\alpha \in (0, \pi)$.

5 An Example

In this section, we present an example of the estimation of parameters. For this purpose, we consider a single chirp model with $A = 2.93$, $B = 1.91$, $\alpha = 2.5$ and $\beta = .10$. The sample size used here is 50. Though, $\alpha, \beta \in (0, \pi)$, we have considered the true value of β, much less than the initial frequency α, as β, being the frequency rate is comparatively small in general. We consider different stationary processes as error random variables for our simulations. The errors are generated from $(a)\quad X(t) = \rho e(t+1) + e(t)$, $(b)\quad X(t) = \rho X(t-1) + e(t)$ and $(c)\quad X(t) = \rho_1 X(t-1) + \rho_2 X(t-2) + e(t)$. The processes (a), (b) and (c) are stationary $MA(1)$, $AR(1)$ and $AR(2)$ processes. Here $MA(q)$ and $AR(p)$ are usual notation for moving average process of order q and autoregressive process of order p respectively. For simulations, $\rho = .5$, $\rho_1 = 1.4$ and $\rho_2 = -.48$ have been used. We consider different values of σ_x^2, the error variance of $X(t)$. Accordingly σ^2, the variance of $e(t)$ is used for different processes to generate the error vector. The LSEs of the parameters are obtained by minimizing the residual sum of squares. The starting estimates of frequency and frequency rate are obtained by maximizing the following 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 (2-d) grid of $(0, \pi) \times (0, \pi)$. The linear parameters A and B are expressible in terms of α and β. So minimization of $Q(\boldsymbol{\theta})$ with respect to $\boldsymbol{\theta}$ involves a 2-d search. The LSEs of all parameters are reported in Table 1 for different values of σ_x^2. In section 3, we have obtained the asymptotic distribution of a single chirp signal model under quite general assumptions. So, it is possible to obtain confidence intervals of the unknown parameters for fixed finite length data. But due to the complexity involved in the distribution, it is extremely complicated to implement it in practice. For this reason we have used the percentile bootstrap (Boot-p) method to obtain the confidence intervals of the parameters as a simple alternative method as suggested by Nandi, Iyer and Kundu (2002). These bootstrap confidence intervals for all the parameters are also reported in Table 1. We have seen in simulations, that the maximizer of the periodogram like function defined above over a fine grid provides a reasonably good initial estimates of the non-linear parameters, α and β.

The lengths of the intervals are reasonably small in almost all the cases. The lengths of the intervals are in decreasing order of linear parameters, α and β. The asymptotic distribution also suggests accordingly as the rates of convergence are $N^{-1/2}$, $N^{-3/2}$, $N^{-5/2}$ respectively. This has been reflected in the bootstrap intervals to some extent. The length of the interval of each parameter increases as the error variance increases. We have considered different types of errors which are stationary and we observe that the performances of least squares estimator and bootstrap method in obtaining confidence interval are quite good. Now to see how the fitted signal looks like we generated a realization with error (b) and $\sigma_x^2 = .1$, the fitted signal is plotted in Fig. 1 along with the original one.

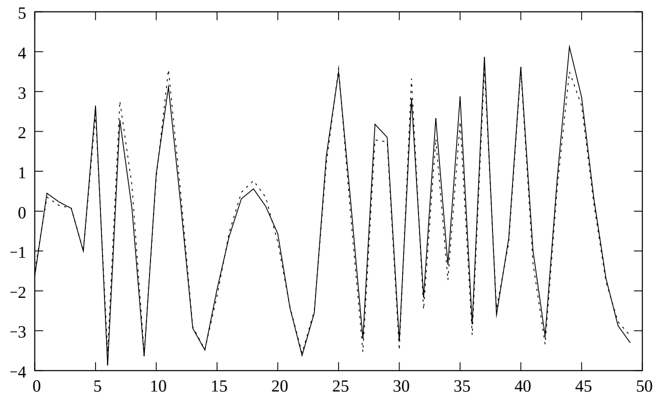

Figure 1: Plot of original signal and estimated signals.

Table 1: Parameter estimates and their bootstrap confidence intervals.

			Parameters			
Error	σ_x^2		A	B	α	β
(a)	0.1	Est.	2.92817	1.98915	2.50219	.099995
		Boot	(2.7511, 3.1521)	(1.7753, 2.2809)	(2.4956, 2.5105)	(.09986, .10012)
	0.5	Est.	2.93097	2.09466	2.50484	.099990
		Boot	(2.5088, 3.4279)	(1.6117, 2.7059)	(2.4906, 2.5239)	(.09968, .10026)
	1.0	Est.	2.93473	2.18057	2.50687	.099984
		Boot	(2.3219, 3.6316)	(1.4948, 3.0022)	(2.4874, 2.5330)	(.09955, .10037)
(b)	0.1	Est.	2.97191	2.00987	2.50321	.099920
		Boot	(2.8207, 3.1810)	(1.7411, 2.2841)	(2.4947, 2.5105)	(.09984, .10009)
	0.5	Est.	3.02630	2.13125	2.50684	.099829
		Boot	(2.6689, 3.4733)	(1.5003, 2.7194)	(2.4881, 2.5236)	(.09963, .10020)
	1.0	Est.	3.07107	2.21792	2.50924	.099770
		Boot	(2.5514, 3.7006)	(1.2960, 3.0176)	(2.4829, 2.5334)	(.09947, .10030)
(c)	0.1	Est.	2.88462	1.95833	2.49750	.100088
		Boot	(2.6189, 3.1625)	(1.6520, 2.2731)	(2.4890, 2.5089)	(.09987, .10026)
	0.5	Est.	2.82969	2.02281	2.49446	.10020
		Boot	(2.2206, 3.4412)	(1.3232, 2.6779)	(2.4759, 2.5197)	(.09971, .10058)
	1.0	Est.	2.78774	2.07561	2.49234	.100271
		Boot	(1.8830, 3.6420)	(-1.2886, 2.9606)	(.67524, 2.5295)	(.09959, 3.0401)

6 Conclusions

In this paper, we study the problem of estimation of parameters of the real single chirp signal model as well as multiple chirp signal model in stationary noise. It is a generalization of multiple frequency model the way the complex-valued chirp model is a generalization of the exponential model. We propose the least squares estimator to estimate the parameters. It has been observed that the estimators are consistent and asymptotically normally distributed. The estimates are obtained for a simulated data. But for confidence intervals, a bootstrap procedure has been applied as the asymptotic dispersion matrix turns out to be quite complicated for practical implementation purposes. Initial estimates of the frequency and 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 5. Also generalization of some existing iterative and non-iterative methods for frequency model to chirp signal model is another problem which need to be addressed as well as the estimation of the number of chirp components for multiple chirp model.

Appendix A

To prove Lemma 2, we need the following lemmas.

Lemma A-1: Let $e(n)$ be a sequence of i.i.d. random variables with mean zero and finite fourth moment, then

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

$$E\left|\sum_{n=1}^{N-k-1} e(n)e(n+1)e(n+k)e(n+k+1)\right| = O(N^{\frac{1}{2}}), \tag{21}$$

for $k = 2, 3, \ldots, N-2$.

Proof of Lemma A-1: We prove (20), (21) follows similarly. Note that

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

Lemma A-2: For an arbitrary integer m,

$$E\sup_\theta\left|\sum_{n=1}^{N} e(n)e(n+k)e^{im\theta n}\right| = O(N^{\frac{3}{4}}).$$

Proof of Lemma A-2:

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

Lemma A-3:

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

Proof of Lemma A-3:

$$E\sup_{\alpha,\beta}\left|\sum_{n=1}^{N} e(n)e^{i(\alpha n+\beta n^2)}\right|^2 = E\sup_{\alpha,\beta}\left[\sum_{n=1}^{N} e(n)e^{i(\alpha n+\beta n^2)}\right]\left[\sum_{n=1}^{N} e(n)e^{-i(\alpha n+\beta n^2)}\right]$$
$$= E\sup_{\alpha,\beta}\left[\sum_{n=1}^{N} e(n)^2 + \left(e^{i(\alpha+\beta)}\sum_{n=1}^{N-1} e(n)e(n+1)e^{2i\beta n} + e^{-i(\alpha+\beta)}\sum_{n=1}^{N-1} e(n)e(n+1)e^{-2i\beta n}\right)\right.$$
$$+\left(e^{i(2\alpha+4\beta)}\sum_{n=1}^{N-2} e(n)e(n+2)e^{4i\beta n} + e^{-i(2\alpha+4\beta)}\sum_{n=1}^{N-2} e(n)e(n+2)e^{-4i\beta n}\right) +$$

$$\vdots$$

$$+\left(e^{i((N-1)\alpha+(N-1)^2\beta)}e(1)e(N)e^{2i(N-1)\beta}+e^{-i((N-1)\alpha+(N-1)^2\beta)}e(1)e(N)e^{-2i(N-1)\beta}\right)\Big]$$

$$\leq O(N+NN^{\frac{3}{4}}) \text{ (using Lemma A-2) } = O(N^{\frac{7}{4}}).$$

Lemma A-4:

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

Proof of Lemma A-4:

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

Lemma A-5:

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

Proof of Lemma A-5:

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

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

Lemma A-6:

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

Proof of Lemma A-6:

Consider the sequence N^9, then 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 Borel Cantelli lemma it follows that

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

Now consider for J, such that $N^9 < J \leq (N+1)^9$, then

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

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

Appendix B

Here we provide entries of the matrix $\mathbf{\Lambda}_{jk} = ((\lambda_{rs}))$ and $\mathbf{H}_{jk} = ((h_{rs}))$

$$\lambda_{11} = \delta_1(0;\beta^+_{j,k}) + \delta_1(0;\beta^-_{j,k}), \quad \lambda_{12} = \gamma_1(0;\beta^+_{j,k}) - \gamma_1(0;\beta^-_{j,k}),$$

$$\lambda_{13} = -A_k\gamma_1(1;\beta^+_{j,k}) + A_k\gamma_1(1;\beta^-_{j,k}) + B_k\delta_1(1;\beta^+_{j,k}) + B_k\delta_1(1;\beta^-_{j,k}),$$

$$\lambda_{14} = -A_k\gamma_1(2;\beta^+_{j,k}) + A_k\gamma_1(2;\beta^-_{j,k}) + B_k\delta_1(2;\beta^+_{j,k}) + B_k\delta_1(2;\beta^-_{j,k}),$$

$$\lambda_{21} = \gamma_1(0;\beta^+_{j,k}) + \gamma_1(0;\beta^-_{j,k}), \quad \lambda_{22} = \delta_1(0;\beta^-_{j,k}) - \delta_1(0;\beta^+_{j,k}),$$

$$\lambda_{23} = -A_k\delta_1(1;\beta^-_{j,k}) + A_k\delta_1(1;\beta^+_{j,k}) + B_k\gamma_1(1;\beta^+_{j,k}) + B_k\gamma_1(1;\beta^-_{j,k}),$$

$$\lambda_{24} = -A_k\delta_1(2;\beta^-_{j,k}) + A_k\delta_1(2;\beta^+_{j,k}) + B_k\gamma_1(2;\beta^+_{j,k}) + B_k\gamma_1(2;\beta^-_{j,k}),$$

$$\lambda_{31} = -A_j\gamma_1(1;\beta^+_{j,k}) - A_j\gamma_1(1;\beta^-_{j,k}) + B_j\delta_1(1;\beta^+_{j,k}) + B_j\delta_1(1;\beta^-_{j,k}),$$

$$\lambda_{32} = -A_j\delta_1(1;\beta^-_{j,k}) + A_j\delta_1(1;\beta^+_{j,k}) + B_j\gamma_1(1;\beta^+_{j,k}) - B_j\gamma_1(1;\beta^-_{j,k}),$$

$$\lambda_{33} = (A_jA_k + B_jB_k)\delta_1(2;\beta^-_{j,k}) - (A_jA_k - B_jB_k)\delta_1(2;\beta^+_{j,k})$$
$$-(A_jB_k + A_kB_j)\gamma_1(2;\beta^+_{j,k}) - (A_jB_k - A_kB_j)\gamma_1(2;\beta^-_{j,k}),$$

$$\lambda_{34} = (A_jA_k + B_jB_k)\delta_1(3;\beta^-_{j,k}) - (A_jA_k - B_jB_k)\delta_1(3;\beta^+_{j,k})$$
$$-(A_jB_k + A_kB_j)\gamma_1(3;\beta^+_{j,k}) - (A_jB_k - A_kB_j)\gamma_1(3;\beta^-_{j,k}),$$

$$\lambda_{41} = -A_j\gamma_1(2;\beta^+_{j,k}) - A_j\gamma_1(2;\beta^-_{j,k}) + B_j\delta_1(2;\beta^+_{j,k}) + B_j\delta_1(2;\beta^-_{j,k}),$$

$$\lambda_{42} = -A_j\delta_1(2;\beta^-_{j,k}) + A_j\delta_1(2;\beta^+_{j,k}) + B_j\gamma_1(2;\beta^+_{j,k}) - B_j\gamma_1(2;\beta^-_{j,k}),$$

$$\lambda_{43} = (A_jA_k + B_jB_k)\delta_1(3;\beta^-_{j,k}) - (A_jA_k - B_jB_k)\delta_1(3;\beta^+_{j,k})$$
$$-(A_jB_k + A_kB_j)\gamma_1(3;\beta^+_{j,k}) - (A_jB_k - A_kB_j)\gamma_1(3;\beta^-_{j,k}),$$

$$\lambda_{44} = (A_jA_k + B_jB_k)\delta_1(4;\beta^-_{j,k}) - (A_jA_k - B_jB_k)\delta_1(4;\beta^+_{j,k})$$
$$-(A_jB_k + A_kB_j)\gamma_1(4;\beta^+_{j,k}) - (A_jB_k - A_kB_j)\gamma_1(4;\beta^-_{j,k}).$$

$$h_{11} = d_3\delta_1(0;\beta^+_{j,k}) + d_1\delta_1(0;\beta^-_{j,k}) - d_2\gamma_1(0;\beta^+_{j,k}) + d_4\gamma_1(0;\beta^-_{j,k}),$$

$$h_{12} = d_4\delta_1(0;\beta^-_{j,k}) + d_2\delta_1(0;\beta^+_{j,k}) - d_1\gamma_1(0;\beta^-_{j,k}) + d_3\gamma_1(0;\beta^+_{j,k}),$$

$$h_{13} = \delta_1(1;\beta^+_{j,k})(B_kd_3 - A_kd_2) - \delta_1(1;\beta^-_{j,k})(B_kd_1 - A_kd_4)$$
$$-\gamma_1(1;\beta^+_{j,k})(B_kd_2 + A_kd_3) + \gamma_1(1;\beta^-_{j,k})(A_kd_1 + B_kd_4),$$

$$h_{14} = \delta_1(2;\beta^+_{j,k})(B_kd_3 - A_kd_2) - \delta_1(2;\beta^-_{j,k})(B_kd_1 - A_kd_4)$$
$$-\gamma_1(2;\beta^+_{j,k})(B_kd_2 + A_kd_3) + \gamma_1(2;\beta^-_{j,k})(A_kd_1 + B_kd_4),$$

$$h_{21} = d_2\delta_1(0;\beta^+_{j,k}) - d_4\delta_1(0;\beta^-_{j,k}) + d_3\gamma_1(0;\beta^+_{j,k}) + d_1\gamma_1(0;\beta^-_{j,k}),$$

$$h_{22} = -d_3\delta_1(0;\beta^+_{j,k}) + d_1\delta_1(0;\beta^-_{j,k}) + d_2\gamma_1(0;\beta^+_{j,k}) + d_4\gamma_1(0;\beta^-_{j,k}),$$

$$h_{23} = \delta_1(1;\beta^+_{j,k})(B_kd_2 + A_kd_3) - \delta_1(1;\beta^-_{j,k})(A_kd_1 + B_kd_4)$$
$$+\gamma_1(1;\beta^+_{j,k})(B_kd_3 - A_kd_2) + \gamma_1(1;\beta^-_{j,k})(B_kd_1 - A_kd_4),$$

$$h_{24} = \delta_1(2;\beta^+_{j,k})(B_kd_2 + A_kd_3) - \delta_1(2;\beta^-_{j,k})(A_kd_1 + B_kd_4)$$
$$+\gamma_1(2;\beta^+_{j,k})(B_kd_3 - A_kd_2) + \gamma_1(2;\beta^-_{j,k})(B_kd_1 - A_kd_4),$$

$$h_{31} = \delta_1(1;\beta^+_{j,k})(B_jd_3 - A_jd_2) + \delta_1(1;\beta^-_{j,k})(A_jd_4 + B_jd_1)$$

$$-\gamma_1(1;\beta^+_{j,k})(B_jd_2+A_jd_3)-\gamma_1(1;\beta^-_{j,k})(B_jd_4-A_jd_1),$$

$$h_{32}=\delta_1(1;\beta^+_{j,k})(B_jd_2+A_jd_3)+\delta_1(1;\beta^-_{j,k})(B_jd_4-A_jd_1)$$

$$+\gamma_1(1;\beta^+_{j,k})(B_jd_3-A_jd_2)-\gamma_1(1;\beta^-_{j,k})(B_jd_1+A_jd_4),$$

$$h_{33}=\delta_1(2;\beta^+_{j,k})\left\{-A_j(A_kd_3+B_kd_2)-B_j(A_kd_2-B_kd_3)\right\}$$

$$+\delta_1(2;\beta^-_{j,k})\left\{A_j(A_kd_1+B_kd_4)-B_j(A_kd_4-B_kd_1)\right\}$$

$$+\gamma_1(2;\beta^+_{j,k})\left\{A_j(A_kd_2-B_kd_3)-B_j(A_kd_3+B_kd_2)\right\}$$

$$+\gamma_1(2;\beta^-_{j,k})\left\{A_j(A_kd_4-B_kd_1)+B_j(A_kd_1+B_kd_4)\right\},$$

$$h_{34}=\delta_1(3;\beta^+_{j,k})\left\{-A_j(A_kd_3+B_kd_2)-B_j(A_kd_2-B_kd_3)\right\}$$

$$+\delta_1(3;\beta^-_{j,k})\left\{A_j(A_kd_1+B_kd_4)-B_j(A_kd_4-B_kd_1)\right\}$$

$$+\gamma_1(3;\beta^+_{j,k})\left\{A_j(A_kd_2-B_kd_3)-B_j(A_kd_3+B_kd_2)\right\}$$

$$+\gamma_1(3;\beta^-_{j,k})\left\{A_j(A_kd_4-B_kd_1)+B_j(A_kd_1+B_kd_4)\right\},$$

$$h_{41}=\delta_1(2;\beta^+_{j,k})(B_jd_3-A_jd_2)+\delta_1(2;\beta^-_{j,k})(A_jd_4+B_jd_1)$$

$$-\gamma_1(2;\beta^+_{j,k})(B_jd_2+A_jd_3)+\gamma_1(2;\beta^-_{j,k})(B_jd_4-A_jd_1),$$

$$h_{42}=\delta_1(2;\beta^+_{j,k})(B_jd_2+A_jd_3)+\delta_1(2;\beta^-_{j,k})(B_jd_4-A_jd_1)$$

$$+\gamma_1(2;\beta^+_{j,k})(B_jd_3-A_jd_2)-\gamma_1(2;\beta^-_{j,k})(B_jd_1+A_jd_4),$$

$$h_{43}=\delta_1(3;\beta^+_{j,k})\left\{-A_j(A_kd_3+B_kd_2)-B_j(A_kd_2-B_kd_3)\right\}$$

$$+\delta_1(3;\beta^-_{j,k})\left\{A_j(A_kd_1+B_kd_4)-B_j(A_kd_4-B_kd_1)\right\}$$

$$+\gamma_1(3;\beta^+_{j,k})\left\{A_j(A_kd_2-B_kd_3)-B_j(A_kd_3+B_kd_2)\right\}$$

$$+\gamma_1(3;\beta^-_{j,k})\left\{A_j(A_kd_4-B_kd_1)+B_j(A_kd_1+B_kd_4)\right\},$$

$$h_{44}=\delta_1(4;\beta^+_{j,k})\left\{-A_j(A_kd_3+B_kd_2)-B_j(A_kd_2-B_kd_3)\right\}$$

$$+\delta_1(4;\beta^-_{j,k})\left\{A_j(A_kd_1+B_kd_4)-B_j(A_kd_4-B_kd_1)\right\}$$

$$+\gamma_1(4;\beta^+_{j,k})\left\{A_j(A_kd_2-B_kd_3)-B_j(A_kd_3+B_kd_2)\right\}$$

$$+\gamma_1(4;\beta^-_{j,k})\left\{A_j(A_kd_4-B_kd_1)+B_j(A_kd_1+B_kd_4)\right\}.$$

References

Abatzoglou, T. (1986). Fast maximum likelihood joint estimation of frequency and frequency rate. *IEEE Trans. on Aerosp. Electron. Syst.* **22**, 708-714.

Djuric, P.M. and Kay, S.M. (1990). Parameter estimation of chirp signals. *IEEE Trans. on Acoustics, Speech and Signal Process.* **38**, No. 12, 2118-2126.

Fuller, W.A. (1976), *Introduction to Statistical Time Series*, John Wiley and Sons, New York.

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**, No. 4, 1029-1039.

Jennrich, R.I. (1969). Asymptotic properties of the nonlinear least squares estimators. *Annals of Mathematical Statistics.* **40**, 633-643.

Kumaresan, R. and Verma, S. (1987). On estimating the parameters of chirp signals using rank reduction techniques. *Proc. 21st. Asilomar Conf.*, Pacific Grove, CA, 555-558.

Nandi, S., Iyer, S.K. and Kundu, D. (2002). Estimation of frequencies in presence of heavy tail errors. *Statistics and Probability Letters* **58**, No. 3, 265-282.

Nandi, S. and Kundu, D. (2004). Asymptotic properties of the least squares estimators of the parameters of the chirp signals. *Annals of the Institute of Statistical Mathematics.* **56**, 529-544.

Rao, C.R. (1973), *Linear Statistical Inferences and Its Applications*, Wiley and Sons, New York, 2nd. Edition.

Rihaczek, A.W. (1969) *Principles of High Resolution Radar*, New York McGraw Hill.

Saha, S. and Kay, S.M. (2002). Maximum likelihood parameter estimation of superimposed chirps using Monte Carlo importance sampling. *IEEE Trans. Signal Processing* **50**, No. 2, 224-230.

Wu, C.F.J. (1981). Asymptotic theory of the nonlinear least squares estimation. *Annals of Statistics* **9**, 501-513.