ESTIMATION OF PARAMETERS OF MULTIPLE CHIRP SIGNAL IN PRESENCE OF HEAVY TAILED ERRORS

SWAGATA NANDI¹ AND DEBASIS KUNDU²

ABSTRACT. In this paper, we consider the estimation of the unknown parameters of the multiple chirp signal model in presence of additive error. The chirp signals are quite common in many areas of science and engineering, specially sonar, radar, audio signals etc. The observed signals are usually corrupted by noise. In different signal processing applications it is observed that the errors may be heavy tailed. In this paper it is assumed that the additive errors have mean zero but may not have finite variance and are independent and identically distributed. We consider the least squares estimators and the approximate least squares estimators which maximize a periodogram like function. It has been observed that both the estimators are strongly consistent. The asymptotic distribution of the least squares estimators is obtained under the assumption that the additive errors are from a symmetric stable distribution. The approximate least squares estimators have the same asymptotic distribution as the least squares estimators. We perform some numerical simulations to see how the proposed estimators work. It is observed that the least squares estimators perform slightly better than the approximately least squares estimators in terms of the biases and mean absolute deviation.

1. INTRODUCTION

This paper is concerned with the estimation of the unknown parameters of the following multicomponent chirp signal model;

\[ y(t) = \sum_{k=1}^{p} \left( A_k^0 \cos(\theta_{1k}^0 t + \theta_{2k}^0 t^2) + B_k^0 \sin(\theta_{1k}^0 t + \theta_{2k}^0 t^2) \right) + e(t); \quad t = 1, \ldots, n. \]  

(1)

Here \( y(t) \) is a real-valued signal observed at \( n \) equi-distant points, \( t = 1, \ldots, n \). For \( k = 1, \ldots, p \), \( A_k^0 \) and \( B_k^0 \) are real-valued amplitudes and \( \theta_{1k}^0 \) and \( \theta_{2k}^0 \) are frequency and frequency rate, respectively. The number of chirp components, \( p \geq 1 \), is assumed to be known in advance. The additive error \( \{e(t)\} \) is a sequence of independent and identically distributed

\[ \text{2000 Mathematics Subject Classification.} \ 62J02; 62E20; 62C05. \]

\[ \text{Key words and phrases.} \ \text{Multiple Chirp Signal; Least Squares Estimator; Approximate Least Squares Estimators; Symmetric Stable; Asymptotic Distribution;} \]
(i.i.d.) random variables with mean zero but the variance may not exist. This includes the case when the variance is not finite. The specific structure of \( \{e(t)\} \) will be stated in Assumptions 1 and 2.

Model (1) is a generalization of fixed-frequencies sinusoidal model. In chirp signal model, frequencies change with time and the rate of change is governed by the frequency rate parameter in case of each chirp component. This model is quite useful in various areas of science and engineering, specially, sonar, radar, communications etc. For example, chirp signal model is used in various forms of trajectories of moving objects with respect to fixed receivers. Several authors have considered this or some related models when \( e(t) \)'s are i.i.d. with finite variance, see for example Abatzoglou [1], Djuric and Kay [3], Giannakis and Zhou [5], Shamsunder et al. [18], Ikram et al. [8], Gini et al. [6], Saha and Kay [16], Nandi and Kundu [14], Farquharson et al. [4], Mazumder [12], Jensen et al. [9] and see the references cited therein.

In several signal processing applications it is observed that the errors are heavy tailed, see for example Zoubir and Breich [20] and Barkat and Stankovic [2]. The main aim of this paper is to consider the case when the error random variables have a heavy tailed distribution. In a heavy tailed distribution, extreme probabilities approach zero relatively slowly. An important criterion for heavy tailedness, as noted by Mandelbrot [11], is the non-existence of the variance. We are using the same definition as of Mandelbrot [11] that a distribution is heavy tailed if and only if the variance is not finite. It is assumed that \( E|e(t)|^{1+\delta} < \infty \), for \( 0 < \delta < 1 \). It may be mentioned that the LS method is a reasonable choice for estimating the unknown parameters in any linear and nonlinear models in presence of additive errors. But when the second moment of the error random variable does not exist, the LS method may not be the ideal estimation procedure. In this paper, quite counter intuitively, it is observed that both the least squares (LS) and approximate LS methods provide consistent estimators of the unknown parameters of the model (1). Moreover, if the errors are from a symmetric stable distribution, then asymptotically LS estimators (LSEs) and approximate LS estimators (ALSEs) follow a multivariate symmetric stable distribution.

Let us recall that a symmetric (around 0) random variable \( X \) is said to have a symmetric \( \alpha \)-stable (\( S\alpha S \)) distribution with the scale parameter \( \sigma \) and the stability index \( \alpha \), if the characteristic function of the random variable \( X \) is \( E[e^{itX}] = e^{-\sigma |t|^\alpha} \). The \( S\alpha S \) distribution is a special case of the general Stable distribution with a non-zero shift and a skewness parameters. For different properties of the Stable and \( S\alpha S \) distributions, see Samorodnitsky
CHIRP SIGNAL WITH HEAVY TAILED ERROR

and Taqqu \[17\]. We need the following assumptions on the error process and the true values of the unknown parameters.

**Assumption 1.** The error random variables \( \{e(t)\} \) is a sequence of i.i.d. random variables with mean zero and \( E|e(t)|^{1+\delta} < 0 \) for some \( 0 < \delta < 1 \).

**Assumption 2.** The error random variables \( \{e(t)\} \) is a sequence of i.i.d. random variables with mean zero and distributed as symmetric \( \alpha \)-stable distribution with scale parameter \( \sigma \) where \( 1 + \delta < \alpha < 2 \), \( 0 < \delta < 1 \).

**Assumption 3.** The frequencies \( \theta^0_{1k} \)'s and the chirp rate \( \theta^0_{2k} \)'s are distinct and \((A^0_k, B^0_k, \theta^0_{1k}, \theta^0_{2k})\) is an interior point of the parameter space \( \Theta \), \( k = 1, \ldots, p \).

**Assumption 4.** The amplitudes \( A^0_k \) and \( B^0_k \), \( k = 1, \ldots, p \) satisfy the following order restrictions;

\[
\infty > M^2 \geq A^0_1 + B^0_1 > \cdots > A^0_p + B^0_p
\]

When \( \delta \geq 1 \) in Assumption \[1\] the second moment exists. That is \( \{e(t)\} \) is a sequence of i.i.d. random variables with mean zero and finite variance. Under Assumption \[2\] \( \{e(t)\} \) is distributed as symmetric \( \alpha \)-stable so that \( \alpha \)-th moment does not exist and under Assumption \[1\] \((1 + \delta)\)-th moment of \( \{e(t)\} \) exists. Therefore, the condition \( 1 + \delta < \alpha < 2 \), \( 0 < \delta < 1 \) is required.

The rest of the paper is organized as follows. In section \[2\] we discuss the consistency properties of the LSEs and ALSEs for single chirp signal. Asymptotic distribution of LSEs for single chirp model is established in section \[3\]. Theoretical properties for LSEs for multiple chirp signal model is derived in section \[4\]. Numerical experiments are presented in section \[5\] and the paper is concluded in section \[6\].

## 2. Estimation of Parameters for Single Chirp Model

In this section, we study the asymptotic properties of LSEs and approximate LSEs for unknown parameters for chirp signal model. We consider model \[1\] with \( p = 1 \) and for notational simplicity, we write the model as

\[
y(t) = A^0 \cos(\theta^0_1 t + \theta^0_2 t^2) + B^0 \sin(\theta^0_1 t + \theta^0_2 t^2) + e(t); \quad t = 1, \ldots, N.
\]
Write $\xi = (A, B, \theta_1, \theta_2)$ and $\xi^0 = (A^0, B^0, \theta_1^0, \theta_2^0)$ as the true value of $\xi$, then the LSE of $\xi^0$, say $\hat{\xi}$ is obtained by minimizing the following residual sum of squares;

$$Q(\xi) = \sum_{t=1}^{n} \left( y(t) - A \cos(\theta_1 t + \theta_2 t^2) - B \sin(\theta_1 t + \theta_2 t^2) \right)^2$$

with respect to $A$, $B$, $\theta_1$ and $\theta_2$. We assume that the true parameter vector $\xi^0$ is an interior point of the parameter space $\Theta$. Minimizing (3) with respect to $\xi$ requires a four dimensional optimization, but using separable regression technique of Richards [15], it can be obtained by solving a two dimensional minimization problem in the following way.

Consider the following $n \times 2$ matrix

$$X(\theta) = \begin{bmatrix}
\cos(\theta_1 + \theta_2) & \sin(\theta_1 + \theta_2) \\
\cos(2\theta_1 + 4\theta_2) & \sin(2\theta_1 + 4\theta_2) \\
\vdots & \vdots \\
\cos(n\theta_1 + n^2\theta_2) & \sin(n\theta_1 + n^2\theta_2)
\end{bmatrix}$$

and use $a = (A, B)^T$ and $Y = (y(1), y(2), \ldots, y(n))^T$. Then,

$$Q(\xi) = (Y - X(\theta)a)^T(Y - X(\theta)a),$$

and for a given $\theta = (\theta_1, \theta_2)$, $Q(\xi)$ is minimized at

$$\hat{a}(\theta) = (X^T(\theta)X(\theta))^{-1}X^T(\theta)Y.$$ 

Therefore, replacing $a$ by $\hat{a}(\theta)$ in $Q(\xi)$ and minimizing $Q(\xi)$ with respect to $\xi$ boils down to minimizing

$$Q(\hat{a}(\theta), \theta) = Y^T(I - P_{X(\theta)})Y,$$

with respect to $\theta$, where

$$P_{X(\theta)} = X(\theta)(X^T(\theta)X(\theta))^{-1}X^T(\theta).$$

Thus, a two step method can be implemented to obtain the LSE of $\xi$. First minimize $Q(\hat{a}(\theta), \theta)$ with respect to $\theta$, denote it as $\hat{\theta}$ and then estimate the linear parameters $A$ and $B$ using $\hat{a}(\hat{\theta})$.

We propose the approximate LSEs (ALSE) of $(\theta_1, \theta_2)$ in the same line as in case of sinusoidal model, by maximizing the following periodogram like function

$$I(\theta) = \frac{2}{n} \left| \sum_{t=1}^{n} y(t) e^{-i(\theta_1 t + \theta_2 t^2)} \right|^2$$

(5)
with respect to $\theta_1$ and $\theta_2$. If $\tilde{\theta} = (\tilde{\theta}_1, \tilde{\theta}_2)$ maximizes $I(\theta)$, then we call $\tilde{\theta}$ as ALSE of $\theta$. Following the approach of sinusoidal model, we propose the ALSEs of $A$ and $B$, say $\tilde{A}$ and $\tilde{B}$ as

$$
\tilde{A} = \frac{2}{n} \sum_{t=1}^{n} y(t) \cos(\tilde{\theta}_1 t + \tilde{\theta}_2 t^2), \quad \tilde{B} = \frac{2}{n} \sum_{t=1}^{n} y(t) \sin(\tilde{\theta}_1 t + \tilde{\theta}_2 t^2).
$$

(6)

Then $\tilde{\xi} = (\tilde{A}, \tilde{B}, \tilde{\theta}_1, \tilde{\theta}_2)$ is the ALSE $\xi^0$. In case of single chirp, the equivalent assumption of Assumption 4 is the following.

**Assumption 5.** $A^0$ and $B^0$ are not identically equal to zero and $A_0^2 + B_0^2 < M^2 < \infty$

Under assumption 1 on the error process $\{e(t)\}$, we prove the strong consistency of the LSEs and ALSEs and the results are stated in the following theorems. The proofs of theorem 2.1 and 2.2 are provided in appendices A and B, respectively. We derive the asymptotic distribution of the LSEs and ALSEs of the unknown parameters under assumption 2 in next section.

**Theorem 2.1.** Under assumptions 1, 3 and 2, $\hat{\xi}$, the LSE of $\xi^0$, is a strongly consistent estimator of $\xi^0$.

**Theorem 2.2.** If $A^0$ and $B^0$ are not identically equal to zero, then under assumptions 1 and 4, $\tilde{\xi}$, the ALSE of $\xi^0$, defined in (5) and (6), is a strongly consistent estimator of $\xi^0$.

### 3. Asymptotic Distribution of LSEs and ALSEs for single chirp model

In this section, we first develop the asymptotic distribution of the LSEs of the unknown parameters for single component chirp model under Assumption 2. Then we discuss the equivalence of LSEs and ALSEs in their asymptotic distributions. Under Assumption 2, $\{e(t)\}$ is a i.i.d. sequence of symmetric $\alpha$-stable random variables, where $1 + \delta < \alpha < 2$, so that $\alpha$-th moment does not exist whereas $(1 + \delta)$-th moment does. As defined earlier, $Q(\xi)$ is the residual sum of squares for the single component chirp model (2). Denote $Q'(\xi)$ and $Q''(\xi)$ as the vector of first order derivatives of order $1 \times 4$ and matrix of second order derivatives of order $4 \times 4$ of $Q(\xi)$, respectively. Expand $Q'(\tilde{\xi})$ around $\xi^0$, using multivariate Taylor series expansion

$$
Q'(\tilde{\xi}) - Q'(\xi^0) = (\tilde{\xi} - \xi^0)Q''(\tilde{\xi}),
$$

(7)
where $\tilde{\xi}$ is a point on the line joining $\hat{\xi}$ and $\xi^0$. Suppose $D_1$ and $D_2$ are two diagonal matrices defined as follows:

\[
D_1 = \text{diag}\{n^{-\frac{1}{\alpha}}, n^{-\frac{1}{\alpha}}, n^{-\frac{14\alpha}{5}}, n^{-\frac{14\alpha}{5}}\}, \\
D_2 = \text{diag}\{n^{-\frac{\alpha-1}{\alpha}}, n^{-\frac{\alpha-1}{\alpha}}, n^{-\frac{2\alpha-1}{\alpha}}, n^{-\frac{3\alpha-1}{\alpha}}\}.
\]

Since $\hat{\xi}$ minimizes $Q(\xi)$, $Q'(\hat{\xi}) = 0$, (7) can be written as

\[
(\hat{\xi} - \xi^0) D_2^{-1} = -[Q'(\xi^0)D_1] [D_2 Q''(\xi)D_1]^{-1},
\]

provided $D_2 Q''(\xi)D_1$ is an invertible matrix for large $n$. It follows that $\hat{\xi}$ converges to $\xi^0$ a.s. from theorem 2.1 and $Q''(\xi)$ is a continuous function of $\xi$. Therefore, using continuous mapping theorem

\[
\lim_{n \to \infty} [D_2 Q''(\xi)D_1] = \lim_{n \to \infty} [D_2 Q''(\xi^0)D_1] = \Gamma \quad \text{(say)}
\]

where

\[
\Gamma = \begin{pmatrix}
1 & 0 & \frac{B_0^0}{2} & \frac{B_0^0}{3} \\
0 & 1 & -\frac{A_0^0}{2} & -\frac{A_0^0}{3} \\
\frac{B_0^0}{2} & -\frac{A_0^0}{2} & \frac{1}{3}(A_0^0 + B_0^0) & \frac{1}{5}(A_0^0 + B_0^0) \\
\frac{B_0^0}{3} & -\frac{A_0^0}{3} & \frac{1}{4}(A_0^0 + B_0^0) & \frac{1}{5}(A_0^0 + B_0^0)
\end{pmatrix}
\]

and

\[
\Gamma^{-1} = ((\gamma^{ij})) = \frac{1}{A_0^2 + B_0^2} \begin{pmatrix}
A_0^2 + 9B_0^0 & -8A_0^0B_0^0 & 36B_0^0 & -30B_0^0 \\
-8A_0^0B_0^0 & 9A_0^2 + B_0^2 & -36A_0^0 & 30A_0^0 \\
36B_0^0 & -36A_0^0 & 192 & 180 \\
-30B_0^0 & 30A_0^0 & 180 & 180
\end{pmatrix}.
\]

In order to show that $Q'(\xi^0)D_1$ converges to a multivariate stable distribution, we write $Q'(\xi^0)D_1 = (Z_{1n}, Z_{2n}, Z_{3n}, Z_{4n})$ where

\[
Z_{1n} = -2 \frac{1}{n^2} \sum_{t=1}^{n} e(t) \cos(\theta_1^0 t + \theta_2^0 t^2), \quad Z_{2n} = -2 \frac{1}{n^2} \sum_{t=1}^{n} e(t) \sin(\theta_1^0 t + \theta_2^0 t^2),
\]

\[
Z_{3n} = \frac{2}{n^2} \sum_{t=1}^{n} e(t)tg(\xi^0; t), \quad Z_{4n} = \frac{2}{n^{3/2}} \sum_{t=1}^{n} e(t)t^2g(\xi^0; t),
\]

and $g(\xi^0; t) = A^0 \sin(\theta_1^0 t + \theta_2^0 t^2) - B^0 \cos(\theta_1^0 t + \theta_2^0 t^2)$. If $t = (t_1, t_2, t_3, t_4)$, the joint characteristic function of $Q'(\xi^0)D_1$ is

\[
\phi_n(t) = E\left[\exp\left\{i(t_1 Z_{1n} + t_2 Z_{2n} + t_3 Z_{3n} + t_4 Z_{4n})\right\}\right] = E\left[\exp\left\{i \frac{2}{n^2} \sum_{r=1}^{n} e(r)K_t(r)\right\}\right],
\]
where
\[ K_t(r; \xi^0) = -t_1 \cos(\theta_1 r + \theta_2 r^2) - t_2 \sin(\theta_1 r + \theta_2 r^2) + \frac{r t_3}{n} g(\xi^0; r) + \frac{r^2 t_4}{n^2} g(\xi^0; r). \]

Because \( \{e(t)\} \) is a sequence of independent random variables,
\[
\phi_n(t) = \prod_{r=1}^{n} E \exp \left\{ i \frac{2}{n \pi} e(r) K_t(r; \xi^0) \right\} = \prod_{r=1}^{n} \exp \left\{ -2^\alpha \sigma^\alpha \frac{1}{n} |K_t(r; \xi^0)|^\alpha \right\} = \exp \left\{ -2^\alpha \sigma^\alpha \frac{1}{n} \sum_{r=1}^{n} |K_t(r; \xi^0)|^\alpha \right\}.
\]

It is noted similarly as in Nandi et al. \cite{13} that \( \frac{1}{n} \sum_{r=1}^{n} |K_t(r; \xi^0)|^\alpha \) converges as \( n \to \infty \).

Now the following arguments show that it converges to a non-zero limit for \( t \neq 0 \). First, observe that for all \( r \) and \( n; 1 \leq r \leq n, n = 1, 2, \ldots \)

\[ |K_t(r; \xi^0)| \leq |t_1| + |t_2| + (|t_3| + |t_4|)(A^0 + B^0) = S; \text{ (say)}. \]

Therefore, \( |K_t(r; \xi^0)/S| \leq 1 \). Hence, for \( 0 < \alpha \leq 2 \) and \( r = 1, 2, \ldots \)

\[ |K_t(r; \xi^0)/S|^\alpha \geq |K_t(r; \xi^0)/S|^2 \Rightarrow |K_t(r; \xi^0)|^\alpha \geq \frac{S^\alpha}{S^2} |K_t(r; \xi^0)|^2. \]

Using Lemma \[1\] (stated in Appendix A), it follows that
\[
\lim_{n \to \infty} \frac{1}{n} \sum_{r=1}^{n} |K_t(r; \xi^0)|^\alpha = \tau_t(A^0, B^0, \theta_1, \theta_2, \alpha) = \tau_t(\xi^0; \alpha) > 0.
\]

This implies that
\[
\lim_{n \to \infty} \phi_n(t) = \exp \left\{ -2^\alpha \sigma^\alpha \tau_t(\xi^0; \alpha) \right\}. \quad (12)
\]

This limiting characteristic function \[12\] implies that any linear combination of \( Z_{1n}, Z_{2n}, Z_{3n} \) and \( Z_{4n} \), even for large samples, follows a symmetric stable distribution with stability index \( \alpha \).

Now, consider the following for large \( n, \)
\[
[Q'(\xi^0)D_t] [D_t Q''(\xi) D_t]^{-1} = \begin{bmatrix}
\sum_{k=1}^{4} Z_{kn} \gamma_{k1} \\
\sum_{k=1}^{4} Z_{kn} \gamma_{k2} \\
\sum_{k=1}^{4} Z_{kn} \gamma_{k3} \\
\sum_{k=1}^{4} Z_{kn} \gamma_{k4}
\end{bmatrix}^{-1} = -\frac{2}{n^{\frac{1}{2}}} \begin{bmatrix}
\sum_{r=1}^{n} e(r) U_A(r; \xi^0) \\
\sum_{r=1}^{n} e(r) U_B(r; \xi^0) \\
\sum_{r=1}^{n} e(r) U_{A_1}(r; \xi^0) \\
\sum_{r=1}^{n} e(r) U_{A_2}(r; \xi^0)
\end{bmatrix}. \quad (13)
\]
where
\[
U_A(r; \xi) = \frac{1}{A^2 + B^2} \left( (A^2 + 9B^2) \cos(\theta_1 r + \theta_2 r^2) - 8AB \sin(\theta_1 r + \theta_2 r^2) - \frac{36Br}{n} g(\xi; r) + \frac{30B r^2}{n^2} g(\xi; r) \right),
\]
\[
U_B(r; \xi) = \frac{1}{A^2 + B^2} \left( -8AB \cos(\theta_1 r + \theta_2 r^2) + (9A^2 + B^2) \sin(\theta_1 r + \theta_2 r^2) + \frac{36Ar}{n} g(\xi; r) - \frac{30Ar^2}{n^2} g(\xi; r) \right),
\]
\[
U_{\theta_1}(r; \xi) = \frac{1}{A^2 + B^2} \left( 36B \cos(\theta_1 r + \theta_2 r^2) - 36A \sin(\theta_1 r + \theta_2 r^2) - \frac{192r}{n} g(\xi; r) - \frac{180r^2}{n^2} g(\xi; r) \right),
\]
\[
U_{\theta_2}(r; \xi) = \frac{1}{A^2 + B^2} \left( -30B \cos(\theta_1 r + \theta_2 r^2) + 30A \sin(\theta_1 r + \theta_2 r^2) - \frac{180r}{n} g(\xi; r) - \frac{180r^2}{n^2} g(\xi; r) \right).
\]

Each element of \( [Q'(\xi^0)D_1] [D_2 Q''(\xi)D_1]^{-1} \) is a linear combination of \( Z_{1n}, Z_{2n}, Z_{3n} \) and \( Z_{4n} \) and therefore, distributed as symmetric \( \alpha \)-stable distribution. Using Theorem 2.1.5 of Samorodnitsky and Taqqu \[17\] that a random vector is symmetric \( \alpha \)-stable in \( \mathbb{R}^d \), if and only if any linear combination is symmetric stable in \( \mathbb{R}^1 \), where \( d \) is the order of the vector, it immediately follows that

\[
[Q'(\xi^0)D_1] [D_2 Q''(\xi)D_1]^{-1}
\]

converges to a symmetric \( \alpha \)-stable random vector in \( \mathbb{R}^d \) as \( n \to \infty \) and has the following limiting characteristic function

\[
\phi(t) = \exp\left\{-2^\alpha \sigma^\alpha \tau_v (\xi^0; \alpha) \right\}.
\]  

(14)

The vector \( v \) is defined as in (12), replacing \( t \) by \( v \) where

\[
v_1(t; A^0, B^0) = \frac{1}{A^0 + B^0} \left( (A^0 + 9B^0) t_1 - 8A^0 B^0 t_2 + 36B^0 t_3 - 30B^0 t_4 \right),
\]
\[
v_2(t; A^0, B^0) = \frac{1}{A^0 + B^0} \left( -8A^0 B^0 t_1 + (9A^0 + B^0) t_2 - 36B^0 t_3 + 30A^0 t_4 \right),
\]
\[
v_3(t; A^0, B^0) = \frac{1}{A^0 + B^0} \left( 36B^0 t_1 - 36A^0 t_2 + 192 t_3 + 180 t_4 \right),
\]
\[
v_4(t; A^0, B^0) = \frac{1}{A^0 + B^0} \left( -30B^0 t_1 + 30A^0 t_2 + 180 t_3 + 180 t_4 \right).
\]

So, we have the following theorem.
Theorem 3.1. In case of model (2), under Assumptions 2, 3 and 5,
\[
(\hat{\xi} - \xi^0)D_2^{-1} = \left( n^{\frac{\alpha-1}{\alpha}} (\hat{A} - A^0), n^{\frac{\alpha-1}{\alpha}} (\hat{B} - B^0), n^{\frac{2\alpha-1}{\alpha}} (\hat{\theta}_1 - \theta_1^0), n^{\frac{3\alpha-1}{\alpha}} (\hat{\theta}_2 - \theta_2^0) \right)
\]
converges to a multivariate S\(\alpha\)S distribution in \(\mathbb{R}^4\) having the characteristic function given in (14).

It has been observed that ALSEs have the same asymptotic distribution as the LSEs in case of model 2 and the result is stated in the following theorem.

Theorem 3.2. If \(A^0\) and \(B^0\) are not identically equal to zero, then under Assumptions 2 and 3, the limiting distribution of
\[
(\hat{\xi} - \xi^0)D_2^{-1} = \left( n^{\frac{\alpha-1}{\alpha}} (\hat{A} - A^0), n^{\frac{\alpha-1}{\alpha}} (\hat{B} - B^0), n^{\frac{2\alpha-1}{\alpha}} (\hat{\theta}_1 - \theta_1^0), n^{\frac{3\alpha-1}{\alpha}} (\hat{\theta}_2 - \theta_2^0) \right)
\]
is same as the limiting distribution of \((\hat{\xi} - \xi^0)D_2^{-1}\).

Proof of Theorem 3.2: See in Appendix C.

4. Consistency and asymptotic distribution for multiple chirp model

In this section, we discuss the asymptotic results of the LSEs of the unknown parameters for multiple chirp model (1). The parameter vector is denoted by \(\eta = (\eta_1, \ldots, \eta_p)\), \(\eta_k = (A_k, B_k, \theta_{1k}, \theta_{2k})\), \(k = 1, \ldots, p\) and \(\eta^0\) denote the true value of \(\eta\). The LSE of \(\eta\), say \(\hat{\eta}\) is obtained by minimizing the residual sum of squares for model (1), defined similarly as \(Q(\xi)\) in (3). We write \(R(\eta)\) as the residual sum of squares in this case and defined as follows;
\[
R(\eta) = \sum_{t=1}^{n} \left[ y(t) - \sum_{k=1}^{p} \left( A_k \cos(\theta_{1k} t + \theta_{2k} t^2) + B_k \sin(\theta_{1k} t + \theta_{2k} t^2) \right) \right]^2.
\] (15)

In matrix notation,
\[
R(\eta) = (Y - X_p(\theta)\psi)^T (Y - X_p(\theta)\psi)
= \left( Y - \sum_{k=1}^{p} X(\theta_k)a_k \right)^T \left( Y - \sum_{k=1}^{p} X(\theta_k)a_k \right),
\]
where \(X_p(\theta) = [X(\theta_1), \ldots, X(\theta_p)]\) and \(X(\theta_k)\) is same as \(X(\theta)\), \(\theta\) replaced by \(\theta_k = (\theta_{1k}, \theta_{2k})\); \(\psi^T = (a_1, \ldots, a_p)\), \(\psi = (A_k, B_k)^T\). Then, for a given \(\theta = (\theta_1, \ldots, \theta_p)\), the vector of linear parameters \(\psi\) can be estimated as
\[
\hat{\psi}(\theta_1, \ldots, \theta_p) = (X_p(\theta)X_p(\theta))^{-1}X_p(\theta)Y.
\]
Replacing $\hat{\psi}(\theta_1, \ldots, \theta_p)$ in $R(\eta)$, similarly as single chirp model, $\theta_{1k}$ and $\theta_{2k}$, $k = 1, \ldots, p$ can be estimated. Finally, $\psi$ is estimated as $\hat{\psi}(\hat{\theta}_1, \ldots, \hat{\theta}_p)$.

The ALSEs of the unknown parameters in case of model (1) are obtained similarly as the single chirp model. We maximize the periodogram like function

$$I(\theta) = \frac{2}{n} \left| \sum_{t=1}^{n} y(t) e^{-i(\theta_1 t + \theta_2 t^2)} \right|^2$$

with respect to $\theta_1$ and $\theta_2$, locally and sequentially. Under assumption 4, at the first step, $(\theta_1, \theta_2)$ that maximizes $I(\theta)$ is the estimate of $(\theta_{11}, \theta_{21})$; write it as $(\tilde{\theta}_{11}, \tilde{\theta}_{21})$. At the second step, we obtain the estimate of $(\theta_{11}, \theta_{21})$, say $(\tilde{\theta}_{11}, \tilde{\theta}_{22})$ and at the $p$-th step we the estimate of $(\theta_{p1}, \theta_{p1})$, say $(\tilde{\theta}_{p1}, \tilde{\theta}_{p2})$. The linear parameters $A_k$ and $B_k$ are estimated as

$$\tilde{A}_k = \frac{2}{n} \sum_{t=1}^{n} y(t) \cos(\tilde{\theta}_{1k} t + \tilde{\theta}_{2k} t^2), \quad \tilde{B}_k = \frac{2}{n} \sum_{t=1}^{n} y(t) \sin(\tilde{\theta}_{1k} t + \tilde{\theta}_{2k} t^2), \quad k = 1, \ldots, p.$$  

We write $\tilde{\eta}_k$ as the ALSE of $\eta_k$. We also note that the boundedness condition on linear parameters (stated in Assumption 4) is not required in case of ALSEs. The LSEs and ALSEs of the unknown parameters in case of multiple chirp model (1), similarly as single chirp model (2), provide strongly consistent estimators. We state the results in the following theorems.

**Theorem 4.1.** In model (1), if the error process $\{e(t)\}$ satisfies Assumption 7, frequencies satisfy Assumption 3 and amplitudes $A^0_j$ and $B^0_j$, $j = 1, \ldots, p$ satisfy Assumption 4, then $\hat{\eta}_j$ is a strongly consistent estimator of $\eta^0_j$, $j = 1, \ldots, p$.

**Theorem 4.2.** In model (1), if the error process $\{e(t)\}$ satisfies Assumption 7, frequencies satisfy Assumption 3 and amplitudes $A^0_j$ and $B^0_j$, $j = 1, \ldots, p$ satisfy

$$\infty > A^0_1 + B^0_1 > \cdots > A^0_p + B^0_p$$

then $\tilde{\eta}_j$ is a strongly consistent estimator of $\eta^0_j$, $j = 1, \ldots, p$.

Theorems 4.1 and 4.2 can be proved along the same lines as theorems 2.1 and 2.2 so the proof are not provided.

We discuss the asymptotic distribution of the LSEs of the unknown parameters, $\hat{\eta}$ under Assumptions 2, 3, and 4. Similarly as previous section, let $R'(\eta)$ and $R''(\eta)$ denote the vector of first derivatives and matrix of second derivatives of $R(\eta)$ of orders $1 \times 4p$ and $4p \times 4p$, respectively.
respectively. Define two diagonal matrices $\Delta_1$ and $\Delta_2$ of order $4p \times 4p$ as follows:

$$
\Delta_1 = \begin{bmatrix}
D_1 & 0 & \cdots & 0 \\
0 & D_1 & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & D_1
\end{bmatrix}, \quad \Delta_2 = \begin{bmatrix}
D_2 & 0 & \cdots & 0 \\
0 & D_2 & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & D_2
\end{bmatrix},
$$

where $D_1$ and $D_2$ are same as defined in section 3. Using multivariate Taylor series expansion along the same line as in section 3, we have

$$
(\hat{\eta} - \eta^0) \Delta_2^{-1} = [R'(\eta^0) \Delta_1] \Delta_2 R''(\eta) \Delta_1^{-1},
$$

because $R'(\hat{\eta}) = 0$ and $\Delta_2 R''(\eta) \Delta_1$ is an invertible matrix for large $n$. Here, $\hat{\eta}$ is a point on the line joining $\eta$ and $\eta^0$. Similarly, as in case of model (2) for $D_2 Q''(\xi) D_1$, we can show that

$$
\lim_{n \to \infty} [\Delta_2 R''(\eta) \Delta_1] = \lim_{n \to \infty} [\Delta_2 R''(\eta^0) \Delta_1] = \begin{bmatrix}
\Gamma_1 & 0 & \cdots & 0 \\
0 & \Gamma_1 & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & \Gamma_1
\end{bmatrix} = \Gamma_G, \quad \text{(say)},
$$

where $\Gamma_k$ is a $4 \times 4$ matrix obtained from $\Gamma$ by replacing $A^0$ and $B^0$ by $A^0_k$ and $B^0_k$, respectively. Consider $t = (t_1, \ldots, t_p)$, $t_j = (t_{ij}, t_{2j}, t_{3j}, t_{4j})$ and write $R'(\eta) \Delta_2 = (Z_1^j, \ldots, Z_p^j)$, $Z_i^j = (Z_i^{j1}, Z_i^{j2}, Z_i^{j3}, Z_i^{j4})$. The elements of $Z_i^j$, $Z_k^{ij}$, $k = 1, \ldots, 4$ are defined similarly as $Z_{kn}$, $k = 1, \ldots, 4$; $A^0$, $B^0$, $\theta_1^0$ and $\theta_2^0$ replaced by $A^0_k$, $B^0_k$, $\theta_1^0_k$ and $\theta_2^0_k$, respectively. With these notation, the joint characteristic function of the elements of $R'(\eta) \Delta_2$ is

$$
\phi_n^p(t) = E \exp \left\{ i \frac{2}{n^\alpha} \sum_{r=1}^n e(r) K_t^M(r; \eta^0) \right\}, \quad K_t^M(r; \eta^0) = \sum_{k=1}^p K_{tk}(r; \eta_k^0).
$$

$K_{tk}(r; \eta_k^0)$ is same as $K_t(r; \xi^0)$ with $A^0$, $B^0$, $\theta_1^0$ and $\theta_2^0$ replaced by $A^0_k$, $B^0_k$, $\theta_1^0_k$ and $\theta_2^0_k$, respectively for $k = 1, \ldots, p$. This particular form of $K_t^M(r; \eta^0)$ and independence of $\{e(r)\}$ enable us to write

$$
\phi_n^p(t) = \prod_{j=1}^p \prod_{r=1}^n \exp \left\{ i \frac{2}{n^\alpha} e(r) K_{tk}(r; \eta_k^0) \right\} = \prod_{j=1}^p \prod_{r=1}^n \exp \left\{ -\frac{2\alpha}{n} |K_{tk}(r; \eta_k^0)|^\alpha \right\} = \prod_{j=1}^p \exp \left\{ -\frac{2\alpha}{n} \sum_{r=1}^n |K_{tk}(r; \eta_k^0)|^\alpha \right\}.
$$
Take limit as $n \to \infty$,

$$
\lim_{n \to \infty} \phi_n^p(t) = \prod_{j=1}^{p} \exp \left\{ -2^\alpha \sigma^\alpha \tau_j(\eta^0_j, \alpha) \right\}
$$

$$
= \prod_{j=1}^{p} \{ \text{Joint characteristic function of}(Z^j_{1n}, Z^j_{2n}, Z^j_{3n}, Z^j_{4n}) \}.
$$

This implies that $(Z^j_{1n}, Z^j_{2n}, Z^j_{3n}, Z^j_{4n})$ and $(Z^k_{1n}, Z^k_{2n}, Z^k_{3n}, Z^k_{4n})$ for $j \neq k, 1, \ldots, p$ are asymptotically independently distributed.

Now considering linear combinations similarly as in section 3, it is observed that as $n \to \infty$,

\[(\hat{\eta} - \eta^0)\Delta_2^{-1} = \left((\hat{\eta}_1 - \eta^0_1)D_2^{-1}, \ldots, (\hat{\eta}_p - \eta^0_p)D_2^{-1}\right)\]

converges to a $S\alpha S$ random vector in $\mathbb{R}^{4p}$ having the following characteristic function

$$
\phi_t^p = \exp \left\{ -2^\alpha \sigma^\alpha \sum_{j=1}^{p} \tau_{w_j}(\eta^0_j, \alpha) \right\}, \quad (19)
$$

\[w_j = (w_{1j}(t_j), w_{2j}(t_j), w_{3j}(t_j), w_{4j}(t_j)), \quad w_{kj}(t_j) = v_k(t_j; A_j^0, B_j^0), \quad k = 1, \ldots, 4; \quad j = 1, \ldots, p, \quad (20)\]

where $v_k, k = 1, \ldots, 4$ are defined in section 3. Therefore, we state the asymptotic distribution of the LSEs of the unknown parameters in model (1) in the following theorem.

**Theorem 4.3.** In model (1), if the error process $\{e(t)\}$ satisfies Assumption 2, frequencies satisfy Assumption 3 and amplitudes $A_j^0$ and $B_j^0, j = 1, \ldots, p$ satisfy Assumption 4, then $(\hat{\eta}_j - \eta^0_j)D_2^{-1}$ converges to a multivariate $\alpha$-stable distribution in $\mathbb{R}^4$ with limiting characteristic function $\exp \left\{ -2^\alpha \sigma^\alpha \tau_{w_j}(\eta^0_j, \alpha) \right\}$ where $w_j$ is defined in (20). Also $(\hat{\eta}_j - \eta^0_j)D_2^{-1}$ and $(\hat{\eta}_k - \eta^0_k)D_2^{-1}$ for $j \neq k$ are asymptotically independently distributed.

In case of multiple chirp model, ALSEs are asymptotically equivalent to LSEs in distribution and it can be shown similarly as Theorem 3.2. The joint asymptotic distribution of the ALSEs, $\tilde{\eta}_1, \ldots, \tilde{\eta}_p$ is same as that of the $\hat{\eta}_1, \ldots, \hat{\eta}_p$ and stated in the following theorem.

**Theorem 4.4.** Under Assumptions 2, if model (1) satisfies

$$
\infty > A_1^0 + B_1^0 > \cdots > A_p^0 + B_p^0,
$$

then the limiting distribution of $(\hat{\eta}_j - \eta^0_j)D_2^{-1}$ is same as the limiting distribution of $(\hat{\eta}_j - \eta^0_j)D_2^{-1}$ and $(\hat{\eta}_j - \eta^0_j)D_2^{-1}$ and $(\hat{\eta}_k - \eta^0_k)D_2^{-1}$ for $j \neq k$ are asymptotically independently distributed.
5. Numerical Experiments

In this section, we present results of numerical experiments. We consider the following two models, first one is a single chirp signal model and second one is a two component multiple chirp model.

Model 1: \[ y(t) = A \cos(\theta_1^0 t + \theta_2^0 t^2) + B \sin(\theta_1^0 t + \theta_2^0 t^2) + e(t), t = 1, \ldots, n \] (21)

Model 2: \[ y(t) = A_1 \cos(\theta_{11}^0 t + \theta_{12}^0 t^2) + B_1 \sin(\theta_{11}^0 t + \theta_{12}^0 t^2) + A_2 \cos(\theta_{21}^0 t + \theta_{22}^0 t^2) + B_2 \sin(\theta_{21}^0 t + \theta_{22}^0 t^2) + e(t), t = 1, \ldots, n \] (22)

\[ A_0^0 = B_0^0 = 2.5, \quad \theta_1^0 = 1.5 \text{ and } \theta_2^0 = 0.1, \]
\[ A_1^0 = B_1^0 = 4.0, \quad \theta_{11}^0 = 1.5, \quad \theta_{12}^0 = 0.1 \quad A_2^0 = B_2^0 = 3.0, \quad \theta_{11}^0 = 2.5, \quad \theta_{12}^0 = 0.2. \]

The sequence of random variables \( \{e(t)\} \) is from a symmetric stable distribution with mean zero, scale parameters \( \sigma > 0 \) and stability index \( \alpha < 2 \). We have taken the value of the frequency rate to be much less than the initial frequency, as the frequency rate represents the rate of change and is comparatively small in general. We would like to see how the LSEs and ALSEs of the unknown parameters, specially the nonlinear parameters, behave for different values of \( \alpha, \sigma \) and \( n \), the sample size. That is, for single chirp model (21), the frequency \( \theta_1 \) and the frequency rate \( \theta_2 \) and for model (22), the frequencies \( \theta_{11} \) and \( \theta_{21} \) and frequency rates \( \theta_{12} \) and \( \theta_{22} \). For simulation, we consider \( \alpha = 1.5, 1.7, 1.9; \sigma = 0.1, 1.0 \) and \( n = 250, 500, 1000 \).

We generate a sample of size \( n \) for each combination of \( \alpha, \sigma \) and \( n \) and compute the LSEs and ALSEs of the unknown parameters. We replicate the process of data generation and estimation 5000 times and calculate the average estimate (AVE) and mean absolute deviation (MAD) for each parameter estimator over these replications. We have not considered mean squared error as a criterion for evaluating the estimators because in the present set-up, estimators are asymptotically multivariate stable such that the variance does not exist for \( \alpha < 2 \). The results for single chirp model (21) are reported in Tables 1-3 for sample sizes 250, 500 and 1000, respectively. The results for two component model (22) are presented in Tables 4-6. It is worth to be noted that for obtaining LSEs, we need to implement a four dimensional minimization whereas for ALSEs a two dimensional maximization is required for single chirp model. For model (22), as the number of chirp component \( p = 2 \), the LSEs are obtained by solving a \( 4p = 8 \) dimensional minimization problem and ALSEs are obtained by solving \( p = 2 \), two dimensional maximization problem. Therefore, as \( p \) increases, the complexity
in finding LSEs is much more as compared to the ALSEs. We have used downhill simplex method to carry out these optimizations.

Some of the salient features of the numerical experiments based on Tables 1-6 are given below.

(1) Performances of both LSEs and ALSEs improve as sample size $n$ increases, scale parameters $\sigma$ decreases and the stability index $\alpha$ increases.

(2) The estimates are close to true values in all the cases considered as biases are quite small in both LSEs and ALSEs in general. Though, biases are small, ALSEs have slightly larger bias in case of frequency estimates.

(3) As expected, mean absolute deviation decreases as sample size increases. The rate of convergence $O\left(\frac{1}{n^{3/2}}\right)$ for frequencies and $O\left(\frac{1}{n^{5/2}}\right)$ for frequency rates are in general observed in each mean absolute deviation values in all the cases for LSEs and in most of the cases for ALSEs.

(4) Mean absolute deviation increases as the scale parameter, $\sigma$ increases.

(5) Mean absolute deviation decreases as the stability index $\alpha$ increases. It indicates that for heavier tail, it is more difficult to estimate the unknown parameters.

(6) Comparing LSE and ALSE, it is observed that although they are asymptotically equivalent, LSE behaves marginally better than the ALSE, in terms of minimum mean absolute deviation for most of the cases considered here.

(7) Computationally, ALSEs are much easier to compute than the LSEs, at least for large $p$. Therefore, if $p$ is large, ALSE is preferable, but when $p$ is small, LSE is recommended.

6. Concluding Remarks

In this paper, we consider the multiple chirp signal model observed with additive i.i.d. errors which are heavy tailed. We propose LS and approximate LS methods to estimate the unknown parameters and prove the strong consistency of the proposed estimators. No distributional assumption is required to prove the consistency. We obtain the asymptotic distribution as multivariate symmetric stable distribution when errors are from a symmetric stable distribution. Although, we have addressed the problem under the assumption of i.i.d. errors, the results can be extended to finite order moving average type. Another point, we have not considered in the paper, is the estimation of $p$, the number of chirp component.
Table 1. Average estimates and mean absolute deviations of the estimators of frequency and frequency rate for Model 1 when the sample size $n = 250$.

| $\sigma$ | $\alpha$ | LSE  | ALSE  |
|---------|---------|------|------|
|         |         | $\theta_1$ | $\theta_2$ | $\theta_1$ | $\theta_2$ |
| .1      | 1.5     | 1.50001  | 9.9997e-2       | 1.49579  | .10002  |
|         |         | 1.6988e-4 | 8.5680e-7       | 4.2101e-3 | 1.3754e-5 |
| 1.7     |         | 1.50001  | 9.9996e-2       | 1.49576  | .10002  |
|         |         | 1.0927e-4 | 5.5196e-7       | 4.2383e-3 | 1.3843e-5 |
| 1.9     |         | 1.50002  | 9.9996e-2       | 1.49574  | .10002  |
|         |         | 7.9631e-5 | 4.0493e-7       | 4.2544e-3 | 1.3894e-5 |
| 1.0     | 1.5     | 1.50001  | 9.9999e-2       | 1.49605  | .10001  |
|         |         | 1.7665e-3 | 8.9748e-6       | 5.5351e-3 | 2.0923e-5 |
| 1.7     |         | 1.50001  | 9.9998e-2       | 1.49596  | .10001  |
|         |         | 1.0913e-3 | 5.5570e-6       | 4.4925e-3 | 1.5973e-5 |
| 1.9     |         | 1.50000  | 9.9997e-2       | 1.49592  | .10001  |
|         |         | 7.8374e-4 | 4.0271e-6       | 4.1747e-3 | 1.4240e-5 |

Table 2. Average estimates and mean absolute deviations of the estimators of frequency and frequency rate for Model 1 when the sample size $n = 500$.

| $\sigma$ | $\alpha$ | LSE  | ALSE  |
|---------|---------|------|------|
|         |         | $\theta_1$ | $\theta_2$ | $\theta_1$ | $\theta_2$ |
| .1      | 1.5     | 1.50002  | 9.9996e-2       | 1.49894  | 9.9998e-2 |
|         |         | 5.4276e-5 | 1.4192e-7       | 1.0683e-3 | 2.4361e-6 |
| 1.7     |         | 1.50004  | 9.9996e-2       | 1.49893  | 9.9998e-2 |
|         |         | 3.5417e-5 | 9.2149e-8       | 1.0902e-3 | 2.4783e-6 |
| 1.9     |         | 1.50005  | 9.9996e-2       | 1.49893  | 9.9998e-2 |
|         |         | 2.6116e-5 | 6.7640e-8       | 1.1004e-3 | 2.4978e-6 |
| 1.0     | 1.5     | 1.50000  | 9.9997e-2       | 1.49915  | .10000  |
|         |         | 5.3122e-4 | 1.4430e-6       | 1.3131e-3 | 2.9331e-6 |
| 1.7     |         | 1.50000  | 9.9997e-2       | 1.49909  | 9.9999e-2 |
|         |         | 3.2529e-4 | 8.8386e-7       | 1.0788e-3 | 2.4303e-6 |
| 1.9     |         | 1.50000  | 9.9997e-2       | 1.49905  | 9.9998e-2 |
|         |         | 2.3134e-4 | 6.3097e-7       | 1.0094e-3 | 2.3002e-6 |
Table 3. Average estimates and mean absolute deviations of the estimators of frequency and frequency rate for Model 1 when the sample size \( n = 1000 \).

| \( \sigma \) | \( \alpha \) | \( \theta_1 \) | \( \theta_2 \) | \( \theta_1 \) | \( \theta_2 \) |
|---|---|---|---|---|---|
| .1 | 1.5 | AVE 1.50005 9.9996e-2 | 1.50019 9.9996e-2 | MAD 1.6431e-5 2.2745e-8 | 1.7077e-4 2.0331e-7 |
| .17 | AVE 1.50006 9.9996e-2 | 1.50020 9.9996e-2 | MAD 1.0062e-5 1.3919e-8 | 1.8534e-4 2.1688e-7 |
| .19 | AVE 1.50006 9.9996e-2 | 1.50020 9.9996e-2 | MAD 6.8519e-6 9.4219e-9 | 1.8407e-4 2.1479e-7 |
| 1.0 | 1.5 | AVE 1.50002 9.9996e-2 | 1.50013 9.9996e-2 | MAD 9.7189e-05 1.6477e-7 | 1.7441e-4 2.3960e-7 |
| .17 | AVE 1.50002 9.9996e-2 | 1.50014 9.9996e-2 | MAD 6.6935e-5 1.0701e-7 | 1.4911e-4 1.8752e-7 |
| .19 | AVE 1.50003 9.9996e-2 | 1.50014 9.9996e-2 | MAD 5.2772e-5 8.0917e-8 | 1.7226e-4 3.5407e-5 |

Table 4. Average estimates and mean absolute deviations of the estimators of frequencies and frequency rates for Model 2 when the sample size \( n = 250 \).

| \( \sigma \) | \( \alpha \) | \( \theta_{11} \) | \( \theta_{12} \) | \( \theta_{21} \) | \( \theta_{22} \) | \( \theta_{11} \) | \( \theta_{12} \) | \( \theta_{21} \) | \( \theta_{22} \) |
|---|---|---|---|---|---|---|---|---|---|
| .1 | 1.5 | AVE 1.50004 9.9996e-2 | 2.49990 .19999 | 1.50022 9.9996e-2 | 2.50739 .19997 |
| MAD 3.8164e-5 2.4028e-7 | 6.9776e-5 4.0953e-7 | 2.4268e-4 1.2877e-6 | 7.4010e-3 3.5307e-5 |
| 1.7 | AVE 1.50005 9.9996e-2 | 2.49989 .19999 | 1.50018 9.9996e-2 | 2.50740 .19997 |
| MAD 2.6862e-5 1.6240e-7 | 5.2056e-5 2.9084e-7 | 1.8204e-4 1.0235e-6 | 7.4186e-3 3.5370e-5 |
| .19 | AVE 1.50005 9.9996e-2 | 2.49989 .19999 | 1.50015 9.9996e-2 | 2.50741 .19997 |
| MAD 2.0193e-5 1.1984e-7 | 4.1530e-5 2.2659e-7 | 1.5003e-4 8.9151e-7 | 7.4291e-3 3.5407e-5 |
| 1.0 | 1.5 | AVE 1.49999 9.9997e-2 | 2.49991 .19999 | 1.50049 9.9997e-2 | 2.50701 .19997 |
| MAD 3.5143e-4 2.6203e-6 | 5.9594e-4 4.1564e-6 | 2.5948e-3 1.1264e-5 | 7.3139e-3 3.4540e-5 |
| 1.7 | AVE 1.50000 9.9997e-2 | 2.49992 .19999 | 1.50041 9.9996e-2 | 2.50713 .19997 |
| MAD 2.2141e-4 1.6372e-6 | 3.6566e-4 2.5262e-6 | 1.6068e-3 7.0855e-6 | 7.1544e-3 3.4295e-5 |
| .19 | AVE 1.50000 9.9997e-2 | 2.49991 .19999 | 1.50040e 9.9996e-2 | 2.50726e .19997 |
| MAD 1.6048e-4 1.1801e-6 | 2.6970e-4 1.8251e-6 | 1.1269e-3 5.0824e-6 | 7.2557e-3 3.4733e-5 |

Some information theoretic criterion or cross validation may be required. Further work is needed in that direction.
Table 5. Average estimates and mean absolute deviations of the estimators of frequencies and frequency rates for Model 2 when the sample size $n = 500$.

| $\sigma$ | $\alpha$ | LSE | ALSE |
|---------|----------|-----|------|
|        |          | $\theta_{11}$ | $\theta_{12}$ | $\theta_{21}$ | $\theta_{22}$ | $\theta_{11}$ | $\theta_{12}$ | $\theta_{21}$ | $\theta_{22}$ |
| .1     | 1.5      | 1.50006  | 9.9996e-2 | 2.49991 | .19999 | 1.49987 | 9.9996e-2 | 2.49989 | .19999 |
|        |          | 9.1188e-6 | 3.5758e-8 | 1.1666e-5 | 4.7214e-8 | 1.2204e-4 | 3.1807e-7 | 7.8160e-5 | 2.6177e-7 |
| 1.7    |          | 1.50006  | 9.9996e-2 | 2.49991 | .19999 | 1.49988 | 9.9996e-2 | 2.49989 | .19999 |
|        |          | 7.1482e-6 | 2.4736e-8 | 1.3089e-5 | 3.9897e-8 | 1.0069e-4 | 6.0609e-5 | 2.3821e-7 |
| 1.9    |          | 1.50006  | 9.9996e-2 | 2.49991 | .19999 | 1.49990 | 9.9996e-2 | 2.49989 | .19999 |
|        |          | 5.8370e-6 | 1.8806e-8 | 1.3971e-5 | 4.7214e-8 | 9.1188e-6 | 2.6399e-7 | 2.3137e-7 |

Table 6. Average estimates and mean absolute deviations of the estimators of frequencies and frequency rates for Model 2 when the sample size $n = 1000$.

| $\sigma$ | $\alpha$ | LSE | ALSE |
|---------|----------|-----|------|
|        |          | $\theta_{11}$ | $\theta_{12}$ | $\theta_{21}$ | $\theta_{22}$ | $\theta_{11}$ | $\theta_{12}$ | $\theta_{21}$ | $\theta_{22}$ |
| .1     | 1.5      | 1.50006  | 9.9996e-2 | 2.49991 | .19999 | 1.50010 | 9.9996e-2 | 2.49983 | .19999 |
|        |          | 2.1993e-5 | 9.9996e-2 | 2.49990 | .19999 | 1.49991 | 9.9997e-2 | 2.49993 | .19999 |
| 1.7    |          | 1.50004  | 9.9996e-2 | 2.49990 | .19999 | 1.50008 | 9.9996e-2 | 2.49984 | .19999 |
|        |          | 2.4137e-5 | 9.9996e-2 | 2.49990 | .19999 | 1.49990 | 9.9997e-2 | 2.49994 | .19999 |
| 1.9    |          | 1.50004  | 9.9996e-2 | 2.49990 | .19999 | 1.50008 | 9.9996e-2 | 2.49993 | .19999 |
|        |          | 2.0650e-5 | 9.9996e-2 | 2.49990 | .19999 | 1.50012 | 9.9996e-2 | 2.49993 | .19999 |

Appendix A

In this Appendix, we prove Theorem 2.1 and following lemmas are required.
Lemma 1. (Lahiri et al. [10]) If \((\theta_1, \theta_2) \in (0, \pi) \times (0, \pi)\), then except for a countable number of points, the following results are true.

\[
\lim_{n \to \infty} \frac{1}{n} \sum_{t=1}^{n} \cos(\theta_1 t + \theta_2 t^2) = \lim_{n \to \infty} \frac{1}{n} \sum_{t=1}^{n} \sin(\theta_1 t + \theta_2 t^2) = 0,
\]

\[
\lim_{n \to \infty} \frac{1}{n^{k+1}} \sum_{t=1}^{n} t^k \cos^2(\theta_1 t + \theta_2 t^2) = \lim_{n \to \infty} \frac{1}{n^{k+1}} \sum_{t=1}^{n} t^k \sin^2(\theta_1 t + \theta_2 t^2) = \frac{1}{2(k+1)},
\]

\[
\lim_{n \to \infty} \frac{1}{n^{k+1}} \sum_{t=1}^{n} t^k \cos(\theta_1 t + \theta_2 t^2) \sin(\theta_1 t + \theta_2 t^2) = 0, \quad k = 0, 1, 2.
\]

Lemma 2. If a sequence of random variables \(\{\epsilon(t)\}\) satisfies Assumption 1, then as \(n \to \infty\) and \(k \geq 0\)

\[
\sup_{\alpha, \beta} \left| \frac{1}{n^{k+1}} \sum_{t=1}^{n} t^k \epsilon(t) \cos(\alpha t) \cos(\beta t^2) \right| \to 0, \quad \text{a.s.} \quad (23)
\]

The result is true for all combinations of cosine and sine functions.

Proof of Lemma 2: We prove (23) for \(k = 0\) and it follows similarly for \(k = 1, 2, \ldots\). Define \(z(t) = \epsilon(t) I[|\epsilon(t)| \leq t^{1+\delta}]\). Then

\[
\sum_{t=1}^{\infty} P[z(t) \neq \epsilon(t)] = \sum_{t=1}^{\infty} P[|\epsilon(t)| > t^{1+\delta}] = \sum_{t=1}^{\infty} \sum_{2^{t-1} \leq n < 2^t} P[|\epsilon(1)| > n^{1+\delta}]
\]

\[
\leq \sum_{t=1}^{\infty} \sum_{2^{t-1} \leq n < 2^t} P[|\epsilon(1)| > 2^{(t-1)/(1+\delta)}]
\]

\[
\leq \sum_{t=1}^{\infty} 2^t P[|\epsilon(1)| > 2^{(t-1)/(1+\delta)}]
\]

\[
\leq \sum_{t=1}^{\infty} 2^t \sum_{j=t}^{\infty} P[2^{(j-1)/(1+\delta)} \leq \epsilon(1) < 2^{(t-1)/(1+\delta)}]
\]

\[
\leq \sum_{j=1}^{\infty} P[2^{(j-1)/(1+\delta)} \leq \epsilon(1) < 2^{(t-1)/(1+\delta)}] \sum_{t=1}^{j} 2^t
\]

\[
\leq C \sum_{j=1}^{\infty} E|\epsilon(1)|^{1+\delta} I[2^{(j-1)/(1+\delta)} \leq \epsilon(1) < 2^{(t-1)/(1+\delta)}] \leq E|\epsilon(1)|^{1+\delta} < \infty.
\]
Therefore, \( P(z(t) \neq \epsilon(t) \text{ i.o.}) = 0 \) and \( \{z(t)\} \) and \( \{\epsilon(t)\} \) are equivalent sequences. Hence,

\[
\sup_{\alpha, \beta} \left| \frac{1}{n} \sum_{t=1}^{n} \epsilon(t) \cos(\alpha t) \cos(\beta t^2) \right| \to 0, \quad \text{a.s.}
\]

\[
\Leftrightarrow \sup_{\alpha, \beta} \left| \frac{1}{n} \sum_{t=1}^{n} z(t) \cos(\alpha t) \cos(\beta t^2) \right| \to 0, \quad \text{a.s.}
\]

Write \( u(t) = z(t) - E(z(t)) \), then for large \( n \)

\[
\sup_{\alpha, \beta} \left| \frac{1}{n} \sum_{t=1}^{n} E(z(t)) \cos(\alpha t) \cos(\beta t^2) \right| \leq \frac{1}{n} \sum_{t=1}^{n} |E(z(t))| \to 0.
\] (24)

Thus, we only need to show that

\[
\sup_{\alpha, \beta} \left| \frac{1}{n} \sum_{t=1}^{n} u(t) \cos(\alpha t) \cos(\beta t^2) \right| \to 0, \quad \text{a.s.}
\] (25)

Now, for any fixed \( \alpha, \beta \in (0, \pi) \) and \( \epsilon > 0 \), let \( 0 \leq h \leq \frac{1}{2n^{1/(1+\delta)}} \). Since \( |hu(t) \cos(\alpha t) \cos(\beta t^2)| \leq \frac{1}{2}, e^x \leq 1 + x + 2|x|^{1+\delta} \) for \( x \leq \frac{1}{2} \) and \( E[|u(t)|^{1+\delta}] < C \) for some \( C > 0 \),

\[
P \left[ \left| \frac{1}{n} \sum_{t=1}^{n} u(t) \cos(\alpha t) \cos(\beta t^2) \right| \geq \epsilon \right] \leq 2e^{-h\epsilon n} \prod_{t=1}^{n} E e^{h u(t) \cos(\alpha t) \cos(\beta t^2)}
\]

\[
\leq 2e^{-h\epsilon n} \prod_{t=1}^{n} (1 + Ch^{1+\delta})
\]

\[
\leq 2e^{-h\epsilon n + 2n Ch^{1+\delta}}.
\]

Choose \( h = \frac{1}{2n^{1/(1+\delta)}} \), then for large \( n \),

\[
P \left[ \left| \frac{1}{n} \sum_{t=1}^{n} u(t) \cos(\alpha t) \cos(\beta t^2) \right| \geq \epsilon \right] \leq 2e^{-\frac{\epsilon n}{2n^{1/(1+\delta)}} + C} = 2e^{-\frac{\epsilon}{2n^{\frac{1}{1+\delta}}}}.
\] (26)

Take \( K = n^6 \). Choose \( K \) points \((\alpha_1, \beta_1), \ldots, (\alpha_K, \beta_K)\) such that for any point \((\alpha, \beta)\) in \((0, \pi) \times (0, \pi)\), we have a point \((\alpha_j, \beta_j)\) satisfying

\[
|\alpha_j - \alpha| \leq \frac{\pi}{n^3} \quad \text{and} \quad |\beta_j - \beta| \leq \frac{\pi}{n^3}.
\]
Now Taylor series expansion can be used to bound $|\cos(\beta t^2) - \cos(\beta_j t^2)| \leq t^2|\beta - \beta_j|$ and $|\cos(\alpha t) - \cos(\alpha_j t)| \leq |t||\alpha - \alpha_j|$, therefore,

\[
\bigg| \frac{1}{n} \sum_{t=1}^{n} u(t) \left\{ \cos(\alpha t) \cos(\beta t^2) - \cos(\alpha_j t) \cos(\beta_j t^2) \right\} \bigg| \\
\leq \bigg| \frac{1}{n} \sum_{t=1}^{n} u(t) \cos(\alpha t) \cos(\beta t^2) - \cos(\beta_j t^2) \bigg| + \bigg| \frac{1}{n} \sum_{t=1}^{n} u(t) \left\{ \cos(\beta_j t^2)(\cos(\alpha t) - \cos(\alpha_j t)) \right\} \bigg| \\
\leq C \left[ \frac{1}{n} \sum_{t=1}^{n} t^{1+\delta} t^2 \frac{\pi}{n^3} + \frac{1}{n} \sum_{t=1}^{n} t^{1+\delta} t \frac{\pi}{n^3} \right] \leq C \left[ \frac{\pi}{n^{1+\delta}} + \frac{\pi}{n^{1+2\delta}} \right] \rightarrow 0 \quad \text{as} \quad n \rightarrow \infty.
\]

Therefore for large $n$, we have

\[
P \left[ \sup_{\alpha,\beta} \bigg| \frac{1}{n} \sum_{t=1}^{n} u(t) \cos(\alpha t) \cos(\beta t^2) \bigg| \geq 2\varepsilon \right] \leq P \left[ \max_{j \leq n^6} \bigg| \frac{1}{n} \sum_{t=1}^{n} u(t) \cos(\alpha_j t) \cos(\beta_j t^2) \bigg| \geq \varepsilon \right] \\
\leq 2n^6 e^{-\frac{\varepsilon^2}{2n^{1+\delta}}}.
\]

Since $\sum_{n=1}^{\infty} 2n^6 e^{-\frac{\varepsilon^2}{2n^{1+\delta}}} < \infty$, therefore because of Borel-Cantelli lemma \cite{25} holds true and that proves the lemma. \hfill \qed

**Lemma 3.** Write

\[
S_c = \{ \xi \in \Omega : \xi = (A, B, \theta_1, \theta_2), |\xi - \xi^0| \geq 3c \}
\]

where $\Omega = (-M, M) \times (-M, M) \times (0, \pi) \times (0, \pi)$. If for any $c > 0$,

\[
\lim_{\theta \in S_c} \inf_{n} \frac{1}{n} \left[ \mathcal{Q}(\xi) - \mathcal{Q}(\xi^0) \right] > 0 \quad \text{a.s.,} \quad (27)
\]

then $\hat{\xi}$, which minimizes $\mathcal{Q}(\xi)$, is a strongly consistent estimator of $\xi^0$.

**Proof of Lemma 3:** The proof can be obtained along the same line as the proof of Lemma 1 of \cite{19}.

**Proof of Theorem 2.1** Write $\frac{1}{n} \left[ \mathcal{Q}(\xi) - \mathcal{Q}(\xi^0) \right] = f_1(\xi) + f_2(\xi)$, where

\[
f_1(\xi) = \frac{1}{t} \sum_{t=1}^{n} \left[ A^0 \cos(\theta_1^0 t + \theta_2^0 t^2) - A \cos(\theta_1 t + \theta_2 t^2) \right. \\
\left. + B^0 \sin(\theta_1^0 t + \theta_2^0 t^2) - B \sin(\theta_1 t + \theta_2 t^2) \right]^2,
\]

and $f_2(\xi)$ is the bounded part.
One can proceed along the same lines for Lemma 4. Let \( \tilde{\theta} = (\tilde{\theta}_1, \tilde{\theta}_2) \) be an estimate of \( \theta^0 = (\theta_1^0, \theta_2^0) \) that maximizes \( I(\theta_1, \theta_2) \) and for any \( \epsilon > 0 \), let \( S_\epsilon = \{ \theta : |\tilde{\theta} - \theta^0| > \epsilon \} \) for some fixed \( \theta^0 \in (0, \pi) \times (0, \pi) \). If for any \( \epsilon > 0 \),

\[
\lim_{n \to \infty} \sup_{S_\epsilon} \frac{1}{n} [I(\theta_1, \theta_2) - I(\tilde{\theta}_1, \tilde{\theta}_2)] < 0, \quad \text{a.s.}
\]

(29) Appendix B

In this Appendix, we provide the proofs of consistency results related to ALSEs.

Lemma 4. Let \( \tilde{\theta} = (\tilde{\theta}_1, \tilde{\theta}_2) \) be an estimate of \( \theta^0 = (\theta_1^0, \theta_2^0) \) that maximizes \( I(\theta_1, \theta_2) \) and for any \( \epsilon > 0 \), let \( S_\epsilon = \{ \theta : |\tilde{\theta} - \theta^0| > \epsilon \} \) for some fixed \( \theta^0 \in (0, \pi) \times (0, \pi) \). If for any \( \epsilon > 0 \),

\[
\lim_{n \to \infty} \sup_{S_\epsilon} \frac{1}{n} [I(\theta_1, \theta_2) - I(\tilde{\theta}_1, \tilde{\theta}_2)] < 0, \quad \text{a.s.}
\]
then as $n \to \infty$, $\tilde{\theta} \to \theta^0$ a.s., that is, $\tilde{\theta}_1 \to \theta^0_1$ and $\tilde{\theta}_2 \to \theta^0_2$ a.s.

**Proof of Lemma 4:** We write $\tilde{\theta}$ by $\tilde{\theta}_n$ and $I(\theta_1, \theta_2) = I(\theta)$ by $I_n(\theta)$ to emphasize that these quantities depend on $n$. Suppose (29) is true but $\tilde{\theta}_n$ does not converge to $\theta^0$ as $n \to \infty$. Then, there exists an $\epsilon > 0$ and a subsequence $\{n_k\}$ of $\{n\}$ such that $|\tilde{\theta}_n - \theta^0| > \epsilon$ for $k = 1, 2, \ldots$. Therefore, $\tilde{\theta}_{n_k} \in S_\epsilon$ for all $k = 1, 2, \ldots$. By definition, $\tilde{\theta}_{n_k}$ is the ALSE of $\theta^0$ and hence maximizes $I_{n_k}(\theta)$ when $n = n_k$. This implies that

$$I_{n_k}(\tilde{\theta}_{n_k}) \geq I_{n_k}(\theta^0) \Rightarrow \frac{1}{n_k} [I_{n_k}(\tilde{\theta}_{n_k}) - I_{n_k}(\theta^0)] \geq 0.$$  

Therefore, $\lim_{n \to \infty} \sup_{\theta_{n_k} \in S_\epsilon} \frac{1}{n_k} [I_{n_k}(\tilde{\theta}_{n_k}) - I_{n_k}(\theta^0)] \geq 0$, which contradicts the inequality (29).

Hence, the result follows.

**Lemma 5.** If the error process $\{e(t)\}$ satisfies Assumption 1 and the amplitudes $A^0$ and $B^0$ satisfy Assumption 3 then the estimator $\tilde{\theta}$ of $\theta^0$ which maximizes $I(\theta)$, as defined in (5), is a strongly consistent estimator of $\theta^0$.

**Proof of Lemma 5:** In this proof, we write $\theta_1 t + \theta_2 t^2 = h(\theta; t)$. Consider

$$\frac{1}{n} [I(\theta) - I(\theta^0)] = \frac{1}{n} \left[ \frac{2}{n} \left| \sum_{t=1}^{n} y(t)e^{-i(\theta_1 t + \theta_2 t^2)} \right|^2 - \frac{2}{n} \left| \sum_{t=1}^{n} y(t)e^{-i(\theta^0_1 t + \theta^0_2 t^2)} \right|^2 \right]$$

$$= \frac{2}{n^2} \left[ \left( \sum_{t=1}^{n} y(t) \cos(h(\theta; t)) \right)^2 + \left( \sum_{t=1}^{n} y(t) \sin(h(\theta; t)) \right)^2 - \left( \sum_{t=1}^{n} y(t) \cos(h(\theta^0; t)) \right)^2 - \left( \sum_{t=1}^{n} y(t) \sin(h(\theta^0; t)) \right)^2 \right]$$

$$= \frac{2}{n^2} \left[ \left( \sum_{t=1}^{n} \left\{ A^0 \cos(h(\theta^0; t)) + B^0 \sin(h(\theta^0; t)) + e(t) \right\} \cos(h(\theta; t)) \right)^2$$

$$+ \left( \sum_{t=1}^{n} \left\{ A^0 \cos(h(\theta^0; t)) + B^0 \sin(h(\theta^0; t)) + e(t) \right\} \sin(h(\theta; t)) \right)^2$$

$$- \left( \sum_{t=1}^{n} \left\{ A^0 \cos(h(\theta^0; t)) + B^0 \sin(h(\theta^0; t)) + e(t) \right\} \cos(h(\theta^0; t)) \right)^2$$

$$- \left( \sum_{t=1}^{n} \left\{ A^0 \cos(h(\theta^0; t)) + B^0 \sin(h(\theta^0; t)) + e(t) \right\} \sin(h(\theta^0; t)) \right)^2 \right].$$
Using identities given in Lemma 1, we have
\[
\lim_{n \to \infty} \sup_{S_\epsilon} \frac{1}{n} [I(\theta) - I(\theta^0)]
\]
\[
= 2 \lim_{n \to \infty} \sup_{S_\epsilon} \left[ \left( \frac{A_0}{n} \sum_{t=1}^{n} \cos(h(\theta^0; t)) \cos(h(\theta; t)) \right)^2 + \left( \frac{B_0}{n} \sum_{t=1}^{n} \sin(h(\theta^0; t)) \sin(h(\theta; t)) \right)^2 \right]
\]
\[
- \left( \frac{A_0}{n} \sum_{t=1}^{n} \cos^2(h(\theta^0; t)) \right)^2 - \left( \frac{A_0}{n} \sum_{t=1}^{n} \sin^2(h(\theta^0; t)) \right)^2 \right]
\]
\[
= -2A_0^2 \frac{1}{4} - 2B_0^2 \frac{1}{4} = -\frac{1}{2}(A_0^2 + B_0^2), \text{ a.s.}
\]
Therefore, using Lemma 4, the result follows.

**Lemma 6.** Under the same assumptions as in Lemma 5, \( n(\hat{\theta}_1 - \theta_1^0) \to 0 \) and \( n^2(\hat{\theta}_2 - \theta_2^0) \to 0 \), a.s. as \( n \to \infty \).

**Proof of Lemma 6:** Let \( I'(\theta) = \left( \frac{\partial I(\theta)}{\partial \theta_1}, \frac{\partial I(\theta)}{\partial \theta_2} \right) \) and \( I''(\theta) \) be the 2 \times 2 matrix of second derivatives of \( I(\theta) \) with respect to \( \theta \). Applying similar steps applied to \( Q(\xi) \) in section 3, expand \( I'(\theta) \) at \( \tilde{\theta} \) around \( \theta^0 \) using Taylor series expansion
\[
I'(\tilde{\theta}) - I'(\theta^0) = (\tilde{\theta} - \theta^0)I''(\theta),
\]
where \( \tilde{\theta} \) is a point on the line joining \( \tilde{\theta} \) and \( \theta^0 \). Define a diagonal matrix of order 2 as \( D = \text{diag}\{\frac{1}{n}, \frac{1}{n^2}\} \), then (30) can be written as
\[
(\tilde{\theta} - \theta^0)D^{-1} = \left[ \frac{1}{n} I'(\theta^0)D \right] \left[ \frac{1}{n} D I''(\theta)D \right]^{-1}
\]
because \( I'(\tilde{\theta}) = 0 (\tilde{\theta} \text{ maximizes } I(\theta)) \). Now note that
\[
\lim_{n \to \infty} \frac{1}{n} DI''(\theta)D = \lim_{n \to \infty} \frac{1}{n} DI''(\theta^0)D = -\frac{1}{A_0^2 + B_0^2} \left( \frac{12}{12} \frac{12}{45} \right)
\]
using Lemma 1 and additionally using Lemma 2, we have
\[
\left[ \frac{1}{n} I'(\theta^0)D \right] \to 0, \text{ a.s. as } n \to \infty.
\]
Using (32) and (33) in (31), we have \( (\tilde{\theta} - \theta^0)D^{-1} \to 0 \) a.s. as \( n \to \infty \) which implies that \( n(\hat{\theta}_1 - \theta_1^0) \to 0 \) and \( n^2(\hat{\theta}_2 - \theta_2^0) \to 0 \), a.s. as \( n \to \infty \).

**Lemma 7.** Under the same assumptions as in Lemma 5, \( \hat{A} \) and \( \hat{B} \) are strongly consistent estimator of \( A_0 \) and \( B_0 \).
Proof of Lemma 7. Expanding $\cos(\tilde{\theta}_1 t + \tilde{\theta}_2 t^2) = \cos(h(\tilde{\theta}; t))$ around $\theta^0$ upto the first order term using multivariate Taylor series

$$
\tilde{A} = \frac{2}{n} \sum_{t=1}^{n} y(t) \cos(\tilde{\theta}_1 t + \tilde{\theta}_2 t^2) = \frac{2}{n} \sum_{t=1}^{n} y(t) \cos(h(\tilde{\theta}; t))
$$

$$
= \frac{2}{n} \sum_{t=1}^{n} \left[ A^0 \cos(h(\theta^0; t)) + B^0 \sin(h(\theta^0; t)) + e(t) \right] \left[ \cos(h(\theta^0; t)) - t(\tilde{\theta}_1 - \theta^0_1) \sin(h(\theta; t)) - t^2(\tilde{\theta}_2 - \theta^0_2) \sin(h(\theta; t)) \right]
$$

$$
= \frac{2A^0}{n} \sum_{t=1}^{n} \cos^2(h(\theta^0; t)) + \frac{2B^0}{n} \sum_{t=1}^{n} \sin(h(\theta^0; t)) \cos(h(\theta^0; t)) + \frac{2}{n} \sum_{t=1}^{n} e(t) \cos(h(\theta^0; t))
$$

$$
-2n(\tilde{\theta}_1 - \theta^0_1) \frac{1}{n^2} \sum_{t=1}^{n} \left[ A^0 \cos(h(\theta^0; t)) \sin(h(\theta; t)) + B^0 \sin(h(\theta^0; t)) \sin(h(\theta; t)) + e(t) \sin(h(\theta; t)) \right]
$$

$$
-2n^2(\tilde{\theta}_2 - \theta^0_2) \frac{1}{n^3} \sum_{t=1}^{n} t^2 \left[ A^0 \cos(h(\theta^0; t)) \sin(h(\theta; t)) + B^0 \sin(h(\theta^0; t)) \sin(h(\theta; t)) + e(t) \sin(h(\theta; t)) \right]
$$

$$
\to A^0 \text{ a.s. as } n \to \infty,
$$

using Lemmas 1, 2 and 6. Similarly expanding $\sin(\tilde{\theta}_1 t + \tilde{\theta}_2 t^2)$, it can be shown that $\tilde{B} \to B^0$ a.s.

### Appendix C

In this Appendix, we show for model (2) that the asymptotic distribution of ALSE $\tilde{\xi}$ of $\xi^0$ is same as the asymptotic distribution of the LSE $\hat{\xi}$ under Assumption 2.

**Lemma 8.** If $(\theta_1, \theta_2) \in (0, \pi) \times (0, \pi)$, then except for a countable number of points, the following results are true.

$$
\lim_{n \to \infty} \frac{1}{\sqrt{n}} \sum_{t=1}^{n} \cos(\theta_1 t + \theta_2 t^2) = \lim_{n \to \infty} \frac{1}{\sqrt{n}} \sum_{t=1}^{n} \sin(\theta_1 t + \theta_2 t^2) = 0,
$$

$$
\lim_{n \to \infty} \frac{1}{n^2} \sum_{t=1}^{n} t \cos(\theta_1 t + \theta_2 t^2) = \lim_{n \to \infty} \frac{1}{n^2} \sum_{t=1}^{n} t \sin(\theta_1 t + \theta_2 t^2) = 0,
$$

$$
\lim_{n \to \infty} \frac{1}{n^2} \sum_{t=1}^{n} t^2 \cos(\theta_1 t + \theta_2 t^2) = \lim_{n \to \infty} \frac{1}{n^2} \sum_{t=1}^{n} t^2 \sin(\theta_1 t + \theta_2 t^2) = 0.
$$
Proof of Lemma 8: Refer to Grover et al. \[7\]

Proof of Theorem 3.2: Let $Q(\xi)$ be the residual sum of squares defined in \[3\]. Then

$$
\frac{1}{n} Q(\xi) = \frac{1}{n} \sum_{t=1}^{n} \left( y(t) - A \cos(\theta_1 t + \theta_2 t^2) - B \sin(\theta_1 t + \theta_2 t^2) \right)^2
$$

$$
= \frac{1}{n} \sum_{t=1}^{n} y(t)^2 - \frac{2}{n} \sum_{t=1}^{n} y(t) \left( A \cos(\theta_1 t + \theta_2 t^2) + B \sin(\theta_1 t + \theta_2 t^2) \right)
$$

$$
+ \frac{1}{n} \sum_{t=1}^{n} \left( A \cos(\theta_1 t + \theta_2 t^2) + B \sin(\theta_1 t + \theta_2 t^2) \right)^2
$$

$$
= \frac{1}{n} \sum_{t=1}^{n} y(t)^2 - \frac{2}{n} \sum_{t=1}^{n} y(t) \left( A \cos(\theta_1 t + \theta_2 t^2) + B \sin(\theta_1 t + \theta_2 t^2) \right) + \frac{1}{2} (A^2 + B^2) + o(1)
$$

$$
= C - \frac{1}{n} J(\xi) + o(1),
$$

where $C = \frac{1}{n} \sum_{t=1}^{n} y(t)^2$ and $J(\xi) = -\frac{2}{n} \sum_{t=1}^{n} y(t) \left( A \cos(\theta_1 t + \theta_2 t^2) + B \sin(\theta_1 t + \theta_2 t^2) \right) + \frac{1}{2} (A^2 + B^2)$. Now compute the elements of $\frac{1}{n} J'(\xi)$ at $\xi^0$ where $J'(\xi) = \left( \frac{\partial J(\xi)}{\partial A}, \frac{\partial J(\xi)}{\partial B}, \frac{\partial J(\xi)}{\partial \theta_1}, \frac{\partial J(\xi)}{\partial \theta_2} \right)$.

Using Lemmas \[1\], \[2\] and \[8\] we observe that

$$
\frac{1}{n} \frac{\partial J(\xi)}{\partial A} \bigg|_{\xi^0} = \frac{2}{n} \sum_{t=1}^{n} e(t) \cos(\theta^0_1 t + \theta^0_2 t^2) + o\left(\frac{1}{\sqrt{n}}\right),
$$

$$
\frac{1}{n} \frac{\partial J(\xi)}{\partial B} \bigg|_{\xi^0} = \frac{2}{n} \sum_{t=1}^{n} e(t) \sin(\theta^0_1 t + \theta^0_2 t^2) + o\left(\frac{1}{\sqrt{n}}\right),
$$

$$
\frac{1}{n} \frac{\partial J(\xi)}{\partial \theta_1} \bigg|_{\xi^0} = \frac{2}{n} \sum_{t=1}^{n} t e(t) \left[ A^0 \sin(\theta^0_1 t + \theta^0_2 t^2) - B^0 \cos(\theta^0_1 t + \theta^0_2 t^2) \right] + o\left(\sqrt{n}\right),
$$

$$
\frac{1}{n} \frac{\partial J(\xi)}{\partial \theta_2} \bigg|_{\xi^0} = \frac{2}{n} \sum_{t=1}^{n} t^2 e(t) \left[ A^0 \sin(\theta^0_1 t + \theta^0_2 t^2) - B^0 \cos(\theta^0_1 t + \theta^0_2 t^2) \right] + o\left(n\sqrt{n}\right).\]

Comparing with $\frac{1}{n} Q'(\xi^0)$, we have

$$
\frac{1}{n} Q'(\xi^0) D_1 = -\frac{1}{n} J'(\xi^0) D_1 + \left( o\left(\frac{1}{\sqrt{n}}\right), o\left(\frac{1}{\sqrt{n}}\right), o(\sqrt{n}), o(n\sqrt{n}) \right) D_1,
$$

where $D_1$ is same as defined in section \[3\]. This implies that

$$
\lim_{n \to \infty} Q'(\xi^0) D_1 = -\lim_{n \to \infty} J'(\xi^0) D_1,
$$

(34)

because $\lim_{n \to \infty} \left( o(\sqrt{n}), o(\sqrt{n}), o(n\sqrt{n}), o(n^2\sqrt{n}) \right) D_1 = 0$ for $1 + \delta < \alpha < 2$, $0 < \delta < 1$. \[\]
We note that ALSEs $\tilde{A}$ and $\tilde{B}$ are expressed as functions of $(\theta_1, \theta_2)$ and replacing $A$ and $B$ with $\tilde{A}$ and $\tilde{B}$ in $J(\xi)$, it is observed that

$$J(\tilde{A}, \tilde{B}, \theta_1, \theta_2) = I(\theta_1, \theta_2).$$

Therefore, the estimator of $\xi^0$ that maximizes $J(\xi)$, is equivalent to $\bar{\xi}$, the ALSE of $\xi^0$. The estimating equation for ALSE $\bar{\xi}$ in terms of $J(\xi)$ is

$$\left(\bar{\xi} - \xi^0\right) D_2^{-1} = -\left[J'(\xi^0)D_1\right] \left[D_2 J''(\bar{\xi}) D_1 \right]^{-1}. \quad (35)$$

Now proceeding similarly as in case of $Q''(\xi)$ is section 3, we have for $J''(\xi)$

$$\lim_{n \to \infty} \left[D_2 J''(\xi) D_1\right] = \lim_{n \to \infty} \left[D_2 J''(\xi^0) D_1\right] = -\Gamma = \lim_{n \to \infty} \left[D_2 Q''(\xi^0) D_1\right] \quad (36)$$

where $\Gamma$ is defined in (10). Hence, using (33) and (36) in (35), we have for large $n$,

$$\left(\bar{\xi} - \xi^0\right) D_2^{-1} = -\left[J'(\xi^0)D_1\right] \left[D_2 J''(\bar{\xi}) D_1 \right]^{-1} = -\left[Q'(\xi^0)D_1\right] \left[D_2 Q''(\bar{\xi}) D_1 \right]^{-1} = \left(\hat{\xi} - \xi^0\right) D_2^{-1}.$$ 

Therefore, it follows that the LSE, $\hat{\xi}$ and ALSE, $\bar{\xi}$ of $\xi^0$ are asymptotically equivalent in distribution and asymptotic distribution of $\bar{\xi}$ is same as $\hat{\xi}$.

References

1. Abatzoglou, T. (1986). Fast maximum likelihood joint estimation of frequency and frequency rate. IEEE Trans. on Aerosp. Electron. Syst. 22, 708-714.
2. Barkat, B. and Stankovic, L. (2004), “Analysis of polynomial FM signals corrupted by heavy-tailed noise”, Signal Processing, 84, 680 - 686.
3. Djuric, P. M. and Kay, S. M. (1990), “Parameter estimation of chirp signals”, IEEE Transactions on Acoustics, Speech, and Signal Processing, 38(12), 2118-2126.
4. Farquharson, F., O’Shea, P. and Ledwich, G. (2005), “A computationally efficient technique for estimating the parameters of polynomial-phase signals from noisy observations”, IEEE Transactions on Signal Processing, 53, 3337-3342.
5. 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.
6. 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.
7. Grover, R., Kundu, D. and Mitra, A. “On approximate least squares estimators of parameters on one dimensional chirp signal”, Statistics (to appear).
[8] Ikram, M. Z., Abed-Meraim, K. and Hua, Y. (1997), “Fast quadratic phase transform for estimating the parameters of multicomponent chirp signals”, Digital Signal Processing, 7(2), 127-135.

[9] Jensen, T.L., Nielsen, J.K., Jensen, J.R., Christensen, M.G., Jensen, S.H. (2017), “A fast algorithm for maximum likelihood estimation of harmonic charp parameters”, IEEE Transactions on Signal Processing, 65, 5137 - 5152.

[10] Lahiri, A., Kundu, D. and Mitra, A. (2015), “Estimating the parameters of multiple chirp signals”, Journal of Multivariate Analysis, 139, 189 - 205.

[11] Mandelbrot, B.B. (1963), “The variation of certain speculative prices”, Jour. Business, Vol. 36, 394-419.

[12] Mazumder, S. (2017), “Single-step and multiple-step forecasting in one-dimensional single chirp signal using MCMC-based Bayesian analysis”, Communications in Statistics-Simulation and Computation, 46(4), 2529-2547.

[13] Nandi, S., Iyer, S. K. and Kundu, D. (2002) “Estimating the Frequencies in Presence of Heavy Tail Errors ”, Statistics and Probability Letters, 58, No. 3, 265-282.

[14] 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, no. 3, 529-544.

[15] Richards, F. S. G. (1961) “A method of maximum likelihood estimation”, J. Roy. Stat. Soc., B, 469–475.

[16] 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.

[17] Samorodnitsky, G and Taqqu, M.S. (1994), Stable Non-Gaussian Random Processes; Stochastic Models with Infinite Variance, Chapman and Hall, New York.

[18] 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.

[19] Wu, C. F. (1981), “Asymptotic theory of non-linear least-squares estimation”, Annals of Statistics, 9, 501 - 513.

[20] Zoubir, A.M. and Brcich, R. F. (2002), “Multiuser detection in heavy tailed noise”, Digital Signal Processing, 12, 262 - 273.

1Theoretical Statistics and Mathematics Unit, Indian Statistical Institute, 7, S.J.S. Sansanwal Marg, New Delhi - 110016, India, nandi@isid.ac.in

2Department of Mathematics and Statistics, Indian Institute of Technology Kanpur, Pin 208016, India, kundu@iitk.ac.in