On Testing Parameters of Chirp Signal Model

Subhra Sankar Dhar, Debasis Kundu and Ujjwal Das

***Abstract*— The estimation of the parameters involved in the chirp signal model is well-studied in the literature. However, to the best of our knowledge, the issues related to other component of Statistical inference, namely, the testing of hypothesis on those parameters has not been studied yet. In this article, the testing of hypothesis problem on the unknown parameters involved in one-dimensional chirp signal model is explored. Precisely speaking, we here theoretically investigate whether the vector of unknown parameters is same as the specified values of the parameter or not. For that purpose, we propose four tests based on the least squares and the least absolute deviation estimators of the unknown parameters using L_1 and L_2 distances. It is shown that the proposed tests are consistent (i.e., the asymptotic power of the tests tend to one as the sample size tends to infinity), and moreover, the asymptotic local power of the tests using contiguous (local) alternatives is also thoroughly studied. An extensive simulation study shows the satisfactory performance of the new tests, and the usefulness of the proposed tests is exhibited on a few benchmark real data sets that are closely associated with various chirp signal models.**

***Index Terms*—asymptotic power, consistency of a test, contiguous alternatives, mixture distribution.**

I. Introduction

A chirp signal sweeps linearly from low to high frequency over time. Chirps may be found in natural events as well as from man-made systems. For example, the chirp signals are encountered in many different engineering applications particularly in radar, laser, active and passive sonar systems etc. The problem of parameter estimation of chirp signals has received a considerable attention in the signal processing literature, see for example [1] – [4], and see the references cited therein. A real valued one-dimensional chirp signal model in additive noise can be written mathematically as follows:

$$y_n = A\cos(\alpha n+\beta n^2)+B\sin(\alpha n+\beta n^2)+X_n; n = 1,..,N. \quad (1)$$

Here y_n is a real valued signal observed at $n = 1,\dots,N$, n denotes the time point, A, B are amplitudes, α and β are frequency and frequency rate, respectively. The additive error X_n is a sequence of independent identically distributed (i.i.d.) random variables with mean zero and finite variance. This kind of signal arises in many applications of signal processing, one of the most important being the radar problem. For instance, consider a radar illuminating a target. Then, the transmitted signal will undergo a phase shift induced by the distance and relative motion between the target and the receiver. Assuming the motion is continuous and differentiable, the phase shift can be adequately modelled as $\phi(t) = a_0 + a_1 t + a_2 t^2$, where the parameters are either related to speed and accelerator or range and speed, depending on what the radar is intended for, and on the kind of waveforms transmitted, see for example [5] (page 56 – 65).

Apart from the radar application, the chirp signals have noteworthy application on sonar (i.e., SOund Navigation And Ranging). Briefly speaking, different materials propagate frequencies differently on or under the surface of the water. It often happens that the energy of some frequencies is absorbed faster by a specific material and hence, yields less information of return. For this and other reasons, an experimenter should have a wide range of survey signals to cover a band of frequencies. Those signals can be modelled by chirp signal model. Moreover, having the sonar signals, one may estimate the parameters involved in the chirp signal model and check that whether estimated parameters are associated with the signals emitted from a particular material or not. For detailed discussion, one may see, [6], [7] along with the references given there.

Besides the aforesaid applications, Biological or Environmental science has also various examples of chirp signals. For instance, there are plenty number of bird species around the universe, and consequently, identifying a particular species of bird may not be effortless only by watching for an untrained person. In order to avoid such kind of identification problem, one may compare the unknown bird chirp with the a priori knowledge of bird chirp to detect the particular species of bird, which is eventually a testing of hypothesis or classification problem of chirp signal model. For such kind of study, we refer the readers to [8], [9] and the references mentioned therein.

S. S. Dhar and D. Kundu are with Department of Mathematics and Statistics, IIT Kanpur, India. e-mail of corresponding author: subhra@iitk.ac.in.

U. Das is with Operations Management, Quantitative Methods and Information Systems, IIM Udaipur, India.

In addition to sonar and acoustic signals, chirp can also be found in other areas like wave physics, music to animal vocalization, speech, turbulence, complex biology, medicine, stock price modeling etc. Examples from medical science may include epileptic seizures (EEG), pregnancy contractions (uterine EMG) etc. A detailed discussion on the application of chirp signals to various scientific domains can be found in [10] as well as in its bibliography.

Although, an extensive amount of work has been done in estimating the unknown parameters and deriving their properties of the chirp signal model, in presence of additive noise as given in (1), as far as we comprehend, not much attention has been paid to the associated testing of hypothesis problem. As pointed out in the discussion of the real applications, the testing of hypothesis of the parameter vector seems to be an important issue particularly in identifying a target, see for example [11] and references mentioned in the earlier discussion. Strictly speaking, it basically tests whether the signal is coming from a specific known source or not. The main aim of this paper is to address formally the most general testing of hypothesis problem of the model (1), which has not been addressed in the statistical signal processing literature.

In connection with the earlier paragraph, first we would like to discuss how one can formulate the test statistics to test $H_0 : \boldsymbol{\theta} = \boldsymbol{\theta}_0$ against $H_1 : \boldsymbol{\theta} \neq \boldsymbol{\theta}_0$, where $\boldsymbol{\theta} = (A, B, \alpha, \beta)$, and $\boldsymbol{\theta}_0 = (A_0, B_0, \alpha_0, \beta_0)$ is known to us. In this context, it should be mentioned that at the outset one needs to estimate the unknown parameter $\boldsymbol{\theta}$ based on the given observations to construct a legitimate test statistic. Regarding the estimation of unknown parameters involved in model (1), a reasonable number of attempts have been made in developing different estimation procedures, see for example [12] – [14], and the references cited therein. It is known that among the different estimators, the least squares estimators (i.e., LSE) and the least absolute deviation (i.e., LAD) estimators have certain optimality properties, see for example [15] and [16] in this respect. Due to this reason, we will mainly consider the MLEs (i.e., Maximum Likelihood Estimator) and LAD estimators. Afterwards, once estimate of the unknown parameters are obtained, one may conceive the idea of using an appropriate notion of discrepancy between the estimator and the specified parameter to formulate the test statistic. In this article, we propose four test statistics based on the LSE and the LAD estimators of the specified parameters, where the deviation or the discrepancy between the vector of the LSE or the LAD estimates and the vector of the specified value is measured by well-known L_2 and L_1 distances. The L_2 and L_1 distances are chosen by the motivation that L_2-distance based estimator perform well for light tailed distribution whereas L_1-distance based estimator is usually more efficient when data are obtained from any heavy tailed distribution. However, strictly speaking, one may consider any distance function other than L_2 and L_1 distances in principle.

In this context, note that one needs to study the theoretical properties of the tests in understanding the performance of the tests for various types of data. We here investigate the consistency property of the tests and the sophisticated asymptotic theory of the test statistics under contiguous (local) alternatives (see, e.g. [17], Chapter 6), which has not been studied before. Roughly speaking, contiguous alternatives are local alternatives asymptotically converging to the null hypothesis of interest. It is now an appropriate place to emphasize that the LSE and the LAD estimators depend on the non-identical random variables, and the central limit theorem and the weak law of large number for non-identical random variables allow us to derive the asymptotic distributions of the test statistics, which is itself a challenging theoretical problem by its own worth. Moreover, we also compare their performances for large samples as well as for small sample. A precise definition of contiguity is provided at the beginning of Section 4.

The rest of the article is organized as follows. Section 2 describes the formulation of the test statistics and the consistency of property of the proposed tests. An extensive finite sample study is carried out in Section 3, and Section 4 contains the asymptotic distribution of the test statistics under contiguous (i.e., local) alternatives and asymptotic power study of the tests. A few benchmark data sets are analysed in Section 5, and Section 6 consists of a few concluding remarks. At the end, all technical details are provided in the Appendix.

II. Proposed Test Statistics and Properties of the Tests

Recall the model from the Introduction that

$$Y_n = A\cos(\alpha n+\beta n^2)+B\sin(\alpha n+\beta n^2)+X_n; n = 1, \ldots, N,$$

where Y_n is the real-valued signal at $n = 1, \ldots, N$, A and B are real-valued amplitudes, α and β are the frequency and the frequency rate, respectively, and X_n is a sequence of i.i.d. error random variables having variance $= \sigma^2 < \infty$. For notational convenience, we here denote $\boldsymbol{\theta} = (A, B, \alpha, \beta)$, and we now want to test $H_0 : \boldsymbol{\theta} = \boldsymbol{\theta}_0$ against $H_1 : \boldsymbol{\theta} \neq \boldsymbol{\theta}_0$, where $\boldsymbol{\theta}_0 = (A_0, B_0, \alpha_0, \beta_0)$ is known to us. It should be emphasized that under H_0,

" $=$ " indicates the componentwise equality, and under H_1, " $\neq$" indicates that at least one of the components of $\boldsymbol{\theta}$ and $\boldsymbol{\theta}_0$ are unequal. We next consider the data $\mathcal{Y} = \{y_1, \ldots, y_N\}$, and based on $\mathcal{Y}$, one may formulate the test statistics $T_{N,1} = ||D^{-1}(\hat{\boldsymbol{\theta}}_{N,LSE} - \boldsymbol{\theta}_0)||_2^2$, $T_{N,2} = ||D^{-1}(\hat{\boldsymbol{\theta}}_{N,LAD} - \boldsymbol{\theta}_0)||_2^2$, $T_{N,3} = ||D^{-1}(\hat{\boldsymbol{\theta}}_{N,LSE} - \boldsymbol{\theta}_0)||_1^2$, and $T_{N,4} = ||D^{-1}(\hat{\boldsymbol{\theta}}_{N,LAD} - \boldsymbol{\theta}_0)||_1^2$, where $D = diag\{N^{-1/2}, N^{-1/2}, N^{-3/2}, N^{-5/2}\}$, and $||.||_2$ and $||.||_1$ denote the usual Euclidean and L_1 distances, respectively. Here

$$\hat{\boldsymbol{\theta}}_{N,LSE} = arg\min_{\boldsymbol{\theta}} \sum_{n=1}^{N} \{y_n - A\cos(\alpha n + \beta n^2) - B\sin(\alpha n + \beta n^2)\}^2$$

and

$$\hat{\boldsymbol{\theta}}_{N,LAD} = arg\min_{\boldsymbol{\theta}} \sum_{n=1}^{N} |y_n - A\cos(\alpha n + \beta n^2) - B\sin(\alpha n + \beta n^2)|.$$

As it is indicated in the introduction, we now want to discuss the motivations of formulating the test statistics $T_{N,1}$, $T_{N,2}$, $T_{N,3}$ and $T_{N,4}$ to test H_0 against H_1. Note that the testing of hypothesis problem described in H_0 against H_1 states that whether the unknown parameter $\boldsymbol{\theta}$ equals with the specified $\boldsymbol{\theta}_0$ or not. In order to address this problem, one needs to measure the discrepancy between $\boldsymbol{\theta}$ and $\boldsymbol{\theta}_0$; however, $\boldsymbol{\theta}$ is unknown in the problem. The unknown $\boldsymbol{\theta}$ can be estimated by the well-known least squares and least absolute deviations procedures based on a given data, which are denoted as $\hat{\boldsymbol{\theta}}_{N,LSE}$ and $\hat{\boldsymbol{\theta}}_{N,LAD}$, respectively. Finally, the test statistics $T_{N,1}$, $T_{N,2}$, $T_{N,3}$ and $T_{N,4}$ are formulated based on well-known Euclidean and L_1 norm distances between estimated $\boldsymbol{\theta}$ (i.e., $\hat{\boldsymbol{\theta}}_{N,LSE}$ or $\hat{\boldsymbol{\theta}}_{N,LAD}$) and $\boldsymbol{\theta}_0$. As we mentioned in the introduction, there will be no harm of considering any other norm mathematically. It should be clarified that here L_2 and L_1 norms are defined on four dimensional vector, which is entirely different from the usual absolute norm for scaler valued variable. In fact, since the components of the estimated vector (i.e., the LSE and the LAD estimators) can be Statistically dependent, the properties of the tests based on $T_{N,1}$, $T_{N,2}$, $T_{N,3}$ and $T_{N,4}$ can be different from the that of the tests based on componentwise differences.

For sake of technicalities, one needs to assume the following conditions.

(A1) $\boldsymbol{\theta} = (A, B, \alpha, \beta)$ is an interior point of the parameter space $\Theta = (-\infty, \infty) \times (-\infty, \infty) \times (0, \pi) \times (0, \pi)$.

(A2) $A^2 + B^2 > 0$.

(A3) The i.i.d. error random variables have the positive density function $f(.)$ with finite second moment.

(A4) Let F_n be the distribution function of Y_n with the density function $f_n(y_n, \boldsymbol{\theta})$, which is twice continuously differentiable with respect to $\boldsymbol{\theta}$. It is assumed that $E\left[\frac{\partial}{\partial\theta_i} f_n(y_n, \boldsymbol{\theta})\right]^{2+\delta}_{\boldsymbol{\theta}=\boldsymbol{\theta}_0} < \infty$ for some $\delta > 0$ and $E\left[\frac{\partial^2}{\partial\theta_i\partial\theta_j} f_n(y_n, \boldsymbol{\theta})\right]^2_{\boldsymbol{\theta}=\boldsymbol{\theta}_0} < \infty$ for all $n = 1, \ldots, N$. Here θ_i and θ_j $(1 \le i, j \le 4)$ are the i and j-th components of $\boldsymbol{\theta}$, respectively.

Remark 2.1: *Note that the conditions assumed in* **(A1)–(A4)** *are realistic in practice. Condition* **(A1)** *gives an idea that the results investigated in this article can be applicable when $\boldsymbol{\theta}$ belongs to its natural parameter space. Condition* **(A2)** *ensures that both A and B cannot be equal to zero. The assumption in* **(A3)** *holds for most of the well-known probability density functions, e.g., normal, Laplace and Cauchy probability density functions. The smoothness assumptions in* **(A4)** *is required to prove the asymptotic normality of the test statistics under contiguous alternatives. Such assumptions are common across the asymptotic statistics.*

We now consider two 4-dimensional Gaussian distributions: Let $\mathbf{A} = \{A_1, A_2, A_3, A_4\}$ be the four dimensional random vector associated with Gaussian distribution having the location parameter $= \mathbf{0}$ and the scatter matrix $= 2\sigma^2\Sigma$ (denote it as Σ_1), where

$$\Sigma = \frac{1}{A_0^2 + B_0^2}\begin{pmatrix} \frac{A_0^2+9B_0^2}{2} & -4A_0B_0 & 18B_0 & -15B_0 \\ -4A_0B_0 & \frac{9A_0^2+B_0^2}{2} & -18A_0 & 15A_0 \\ 18B_0 & -18A_0 & 96 & -90 \\ -15B_0 & 15A_0 & -90 & 90 \end{pmatrix}$$

Next, $\mathbf{B} = \{B_1, B_2, B_3, B_4\}$ is another four dimensional random vector associated with Gaussian distribution having the location parameter $= \mathbf{0}$ and the scatter matrix $= \frac{1}{\{f(M)\}^2}\Sigma$ (denote it as Σ_2), where M is the population median of the distribution associated with the density function f. The asymptotic property of the tests based on $T_{N,1}$, $T_{N,2}$, $T_{N,3}$ and $T_{N,4}$ are described below.

Theorem 2.1: *Let $c_{1\eta}$ be $(1-\eta)$-th quantile of the distribution of $\sum_{i=1}^{4} \lambda_i Z_i^2$, where λ_is are the eigen values of Σ_1 (defined above), and Z_is are independent standard normal random variables. Under* **(A1)-(A3)**, *the test based on $T_{N,1}$ will have asymptotic size $= \eta$ when $T_{N,1} \ge c_{1\eta}$. Moreover, under the same assumption, $P_{H_1}[T_{N,1} \ge c_{1\eta}] \to 1$ as $N \to \infty$, i.e., the test based on $T_{N,1}$ will be a consistent test.*

Theorem 2.2: *Let $c_{2\eta}$ be $(1-\eta)$-th quantile of the distribution of $\sum_{i=1}^{4} \lambda_i^* Z_i^{*2}$, where λ_i^*s are the eigen values of Σ_2 (defined before Theorem 2.1), and Z_i^*s are independent standard normal random variables. Under* **(A1)-(A3)**, *the test based on $T_{N,2}$ will have asymptotic* *size* $= \eta$ *when* $T_{N,2} \geq c_{2\eta}$. *Moreover, under the same assumption,* $P_{H_1}[T_{N,2} \geq c_{2\eta}] \to 1$ *as* $N \to \infty$*, i.e., the test based on* $T_{N,2}$ *will be a consistent test.*

Theorem 2.3: *Let* $c_{3\eta}$ *be* $(1-\eta)$*-th quantile of the distribution of* $\left\{\sum_{i=1}^{4} |A_i|\right\}^2$, *where* A_i*s are same as defined after Remark 2.1. Under* **(A1)-(A3)**, *the test based on* $T_{N,3}$ *will have asymptotic size* $= \eta$ *when* $T_{N,3} \geq c_{3\eta}$. *Moreover, under the same assumption,* $P_{H_1}[T_{N,3} \geq c_{3\eta}] \to 1$ *as* $N \to \infty$*, i.e., the test based on* $T_{N,3}$ *will be a consistent test.*

Theorem 2.4: *Let* $c_{4\eta}$ *be* $(1-\eta)$*-th quantile of the distribution of* $\left\{\sum_{i=1}^{4} |B_i|\right\}^2$, *where* B_i*s are same as defined after Remark 2.1. Under* **(A1)-(A3)**, *the test based on* $T_{N,4}$ *will have asymptotic size* $= \eta$ *when* $T_{N,4} \geq c_{4\eta}$. *Moreover, under the same assumption,* $P_{H_1}[T_{N,4} \geq c_{4\eta}] \to 1$ *as* $N \to \infty$*, i.e., the test based on* $T_{N,4}$ *will be a consistent test.*

The assertions in Theorems 2.1-2.4 imply that the tests based on $T_{N,1}$, $T_{N,2}$, $T_{N,3}$ and $T_{N,4}$ are consistent, i.e., the power of the tests converge to one as the sample sizes tend to infinity. In other words, this property indicates that these tests will pose substantial power for large sample size.

III. Finite Sample Study

In Section II, we studied the tests based on $T_{N,1}$, $T_{N,2}$, $T_{N,3}$ and $T_{N,4}$ when the sample size was infinite, and in that case, the power of all of them converge to one. In other words, they are comparable with each other for infinite sample size, and hence, at the next level, we are interested to see the performance of the tests when the sample size is finite. In this section, we investigate the finite sample performance of those tests at 5% level of significance. We study their behavior assuming different error distributions for X_n as well as for different values of N in the model described in Section II. We first fix the numerical values of the parameters under null and alternative hypotheses: $\boldsymbol{\theta}_0$ and $\boldsymbol{\theta}_1$, and we consider alternatives $\boldsymbol{\theta}^* = (1-\psi)\boldsymbol{\theta}_0 + \psi\boldsymbol{\theta}_1$ for a wide range of $\psi \in [0, 1]$. As the error distribution, three different distributions are considered, namely, standard normal, standard Laplace and t-distribution with 5 degrees of freedom. In this study, we choose $N = 30$, 50 and 100, and $\boldsymbol{\theta}_0 = (1, 1, \pi/4, \pi/4)$ and $\boldsymbol{\theta}_1 = (4, 4, \pi/2, \pi/2)$ are chosen. In the course of investigation, a sample of size $= 100$ is drawn repeatedly 1000 times when $\boldsymbol{\theta} = \boldsymbol{\theta}_0$ (i.e., H_0 is true), and 0.95-th quantile of the empirical distribution of the test statistic is considered as the estimated critical value (denote it as $\hat{c}_{0.05}$). Afterwards, to estimate the power, we follow the same procedure when $\boldsymbol{\theta} = \boldsymbol{\theta}^*$ (i.e, H_1 is true), and the proportion of times the value of the test statistic exceeding $\hat{c}_{0.05}$ is defined as the estimated power. The results are summarized in Figure 1 (given in supplementary file).

Note that at $\psi = 0$, $\boldsymbol{\theta}^*$ coincides with $\boldsymbol{\theta}_0$, and for that reason, the estimated powers are close to the estimated level (i.e., 0.05) in all cases, i.e., in other words, the tests are able to retain the level of significance. In terms of power, it is indicated from the power curves (see Figure 1) that the tests based on $T_{N,1}$ and $T_{N,3}$ (i.e., based on least square methodology) performs well when the data is obtained from the light tailed distribution like normal distribution. On the other hand, the tests based on $T_{N,2}$ and $T_{N,4}$ (i.e., based on least absolute deviation methodology) perform better than that of $T_{N,1}$ and $T_{N,3}$ for heavy tailed distribution like Laplace and t distribution with 5 degrees of freedom. Overall, one can prefer the least absolute deviation based methodologies when that data are likely to have influential observations/ outliers.

IV. Asymptotic Power Study : Local Alternatives

In Section 2, we have already seen that the tests based on $T_{N,1}$, $T_{N,2}$, $T_{N,3}$ and $T_{N,4}$ are consistent, i.e., in other words, the powers of those tests tend to one as the sample sizes tend to infinity for a fixed alternative. This fact motived us to investigate how well those tests perform when the alternatives are local (or contiguous) in nature. Recently, [19] described the concept of local alternatives in details; we, however, here briefly describe this concept for sake of completeness. The sequence of probability measures Q_n will be contiguous relative to another sequence of probability measures P_n if $P_n(A_n) \to 0$ implies that $Q_n(A_n) \to 0$ for every sequence of measurable sets A_n, where $(\Omega_n, \mathcal{A}_n)$ is the sequence of measurable spaces, and P_n and Q_n are two probability measures defined on $(\Omega_n, \mathcal{A}_n)$. Since the sequence of sets A_n is changing over n along with the corresponding σ-field $\mathcal{A}_n$, the verification of contiguity from its definition is not easily doable. To overcome this issue, Le Cam' s first lemma (e.g., see [20]) based on the log likelihood ratio is used to establish the contiguity in the proofs (see the proof of Theorem 4.1).

Consider now a sequence of local alternatives relative to the null hypothesized parameter $\boldsymbol{\theta}_0$. Suppose that we want to test $H_0 : \boldsymbol{\theta} = \boldsymbol{\theta}_0$ against $H_{1,N} : \boldsymbol{\theta} = \boldsymbol{\theta}_N = \boldsymbol{\theta}_0 + \boldsymbol{\delta}_N$, where $\boldsymbol{\delta}_N = \left(\frac{\delta_1}{N^{1/2}}, \frac{\delta_2}{N^{1/2}}, \frac{\delta_3}{N^{3/2}}, \frac{\delta_4}{N^{5/2}}\right)$. The asymptotic distributions of $T_{N,1}$, $T_{N,2}$, $T_{N,3}$ and $T_{N,4}$ under $H_{1,N}$ are described in the following theorems.

Theorem 4.1: *Under* **(A1)**-**(A4)**, *the sequence of alternatives* $\{H_{1N}\}$ *will be contiguous alternatives relative to the null hypothesis* H_0. *Further, under* H_{1N}, $T_{N,1}$ *converges weakly to* $\sum_{i=1}^{4}\lambda_i M_i^2$, *where* λ_i*s are eigen values of* Σ_1 *(defined after Remark 2.1), and* $\mathbf{M} = (M_1, M_2, M_3, M_4)$ *is a* 4*-dimensional random vector associated with a Gaussian distribution having the location parameter* $\boldsymbol{\mu} = (\mu_1, \mu_2, \mu_3, \mu_4)$ *and the scatter parameter* $= I_4$. *Here* $\mu_1 = (\lim_{N\to\infty} -\frac{1}{N}\sum_{n=1}^{N} E_{F_{n,H_0}}[\{Y_n - A_0\cos(\alpha_0 n + \beta_0 n^2) - B_0\sin(\alpha_0 n + \beta_0 n^2)\}\{cos(\alpha_0 n + \beta_0 n^2)\}\{\delta_1 \frac{\partial f(Y_n,\boldsymbol{\theta})}{\partial A}\{f(Y_n,\boldsymbol{\theta})\}^{-1}|_{\boldsymbol{\theta}=\boldsymbol{\theta}_0} + \delta_2 \frac{\partial f(Y_n,\boldsymbol{\theta})}{\partial B}\{f(Y_n,\boldsymbol{\theta})\}^{-1}|_{\boldsymbol{\theta}=\boldsymbol{\theta}_0}\}])\Sigma_1^{-1}$,

$\mu_2 = (\lim_{N\to\infty} -\frac{1}{N}\sum_{n=1}^{N} E_{F_{n,H_0}}[\{Y_n - A_0\cos(\alpha_0 n + \beta_0 n^2) - B_0\sin(\alpha_0 n + \beta_0 n^2)\}\{sin(\alpha_0 n + \beta_0 n^2)\}\{\delta_1 \frac{\partial f(Y_n,\boldsymbol{\theta})}{\partial A}\{f(Y_n,\boldsymbol{\theta})\}^{-1}|_{\boldsymbol{\theta}=\boldsymbol{\theta}_0} + \delta_2 \frac{\partial f(Y_n,\boldsymbol{\theta})}{\partial B}\{f(Y_n,\boldsymbol{\theta})\}^{-1}|_{\boldsymbol{\theta}=\boldsymbol{\theta}_0}\}])\Sigma_1^{-1}$,

$\mu_3 = (\lim_{N\to\infty} -\frac{1}{N^2}\sum_{n=1}^{N} E_{F_{n,H_0}}[n\{Y_n - A_0\cos(\alpha_0 n + \beta_0 n^2) - B_0\sin(\alpha_0 n + \beta_0 n^2)\}\{A_0 sin(\alpha_0 n + \beta_0 n^2) - B_0 cos(\alpha_0 n + \beta_0 n^2)\}\{\delta_1 \frac{\partial f(Y_n,\boldsymbol{\theta})}{\partial A}\{f(Y_n,\boldsymbol{\theta})\}^{-1}|_{\boldsymbol{\theta}=\boldsymbol{\theta}_0} + \delta_2 \frac{\partial f(Y_n,\boldsymbol{\theta})}{\partial B}\{f(Y_n,\boldsymbol{\theta})\}^{-1}|_{\boldsymbol{\theta}=\boldsymbol{\theta}_0}\}])\Sigma_1^{-1}$, *and*

$\mu_4 = (\lim_{N\to\infty} -\frac{1}{N^3}\sum_{n=1}^{N} E_{F_{n,H_0}}[n^2\{Y_n - A_0\cos(\alpha_0 n + \beta_0 n^2) - B_0\sin(\alpha_0 n + \beta_0 n^2)\}\{A_0 sin(\alpha_0 n + \beta_0 n^2) - B_0 cos(\alpha_0 n + \beta_0 n^2)\}\{\delta_1 \frac{\partial f(Y_n,\boldsymbol{\theta})}{\partial A}\{f(Y_n,\boldsymbol{\theta})\}^{-1}|_{\boldsymbol{\theta}=\boldsymbol{\theta}_0} + \delta_2 \frac{\partial f(Y_n,\boldsymbol{\theta})}{\partial B}\{f(Y_n,\boldsymbol{\theta})\}^{-1}|_{\boldsymbol{\theta}=\boldsymbol{\theta}_0}\}])\Sigma_1^{-1}$.

Theorem 4.2: *Under* **(A1)**-**(A4)**, *the sequence of alternatives* $\{H_{1N}\}$ *will be contiguous alternatives relative to the null hypothesis* H_0. *Further, under* H_{1N}, $T_{N,2}$ *converges weakly to* $\sum_{i=1}^{4}\lambda_i^* M_i^{*2}$, *where* λ_i^**s are eigen values of* Σ_2 *(defined after Remark 2.1), and* $\mathbf{M}^* = (M_1^*, M_2^*, M_3^*, M_4^*)$ *is a* 4*-dimensional random vector associated with a Gaussian distribution having the location parameter* $\boldsymbol{\mu}^* = (\mu_1^*, \mu_2^*, \mu_3^*, \mu_4^*)$ *and the scatter parameter* $= I_4$. *Here*

$\mu_1^* = (\lim_{N\to\infty} -\frac{1}{N}\sum_{n=1}^{N} E_{F_{n,H_0}}[\{\rho_N^{'}(X_n)\}\{cos(\alpha_0 n + \beta_0 n^2)\}\{\delta_1 \frac{\partial f(Y_n,\boldsymbol{\theta})}{\partial A}\{f(Y_n,\boldsymbol{\theta})\}^{-1}|_{\boldsymbol{\theta}=\boldsymbol{\theta}_0} + \delta_2 \frac{\partial f(Y_n,\boldsymbol{\theta})}{\partial B}\{f(Y_n,\boldsymbol{\theta})\}^{-1}|_{\boldsymbol{\theta}=\boldsymbol{\theta}_0}\}])\Sigma_2^{-1}$,

$\mu_2^* = (\lim_{N\to\infty} -\frac{1}{N}\sum_{n=1}^{N} E_{F_{n,H_0}}[\{\rho_N^{'}(X_n)\}\{sin(\alpha_0 n + \beta_0 n^2)\}\{\delta_1 \frac{\partial f(Y_n,\boldsymbol{\theta})}{\partial A}\{f(Y_n,\boldsymbol{\theta})\}^{-1}|_{\boldsymbol{\theta}=\boldsymbol{\theta}_0} + \delta_2 \frac{\partial f(Y_n,\boldsymbol{\theta})}{\partial B}\{f(Y_n,\boldsymbol{\theta})\}^{-1}|_{\boldsymbol{\theta}=\boldsymbol{\theta}_0}\}])\Sigma_2^{-1}$,

$\mu_3^* = (\lim_{N\to\infty} -\frac{1}{N^2}\sum_{n=1}^{N} E_{F_{n,H_0}}[n\{\rho_N^{'}(X_n)\}\{A_0 sin(\alpha_0 n + \beta_0 n^2) - B_0 cos(\alpha_0 n + \beta_0 n^2)\}\{\delta_1 \frac{\partial f(Y_n,\boldsymbol{\theta})}{\partial A}\{f(Y_n,\boldsymbol{\theta})\}^{-1}|_{\boldsymbol{\theta}=\boldsymbol{\theta}_0} + \delta_2 \frac{\partial f(Y_n,\boldsymbol{\theta})}{\partial B}\{f(Y_n,\boldsymbol{\theta})\}^{-1}|_{\boldsymbol{\theta}=\boldsymbol{\theta}_0}\}])\Sigma_2^{-1}$, *and*

$\mu_4^* = (\lim_{N\to\infty} -\frac{1}{N^3}\sum_{n=1}^{N} E_{F_{n,H_0}}[n^2\{\rho_N^{'}(X_n)\}\{A_0 sin(\alpha_0 n + \beta_0 n^2) - B_0 cos(\alpha_0 n + \beta_0 n^2)\}\{\delta_1 \frac{\partial f(Y_n,\boldsymbol{\theta})}{\partial A}\{f(Y_n,\boldsymbol{\theta})\}^{-1}|_{\boldsymbol{\theta}=\boldsymbol{\theta}_0} + \delta_2 \frac{\partial f(Y_n,\boldsymbol{\theta})}{\partial B}\{f(Y_n,\boldsymbol{\theta})\}^{-1}|_{\boldsymbol{\theta}=\boldsymbol{\theta}_0}\}])\Sigma_2^{-1}$, *where* $\rho_N^{'}(x) = [-\gamma_N^2 x^2 + 2\gamma_N x]I_{(0<x\leq\frac{1}{\gamma_N})} + [\gamma_N^2 x^2 + 2\gamma_N x]I_{(-\frac{1}{\gamma_N}<x\leq 0\frac{1}{\gamma_N})} + 1I_{(x>\frac{1}{\gamma_N})} - 1I_{(x<-\frac{1}{\gamma_N})}$ *when* γ_N *is an increasing function of* N, $N^2 = o(\gamma_N^3)$, $\gamma_N = o(N)$ *and* $\sum_{N=1}^{\infty}\frac{1}{\gamma_N^2} < \infty$.

Theorem 4.3: *Under* **(A1)**-**(A4)**, *the sequence of alternatives* $\{H_{1N}\}$ *will be contiguous alternatives relative to the null hypothesis* H_0. *Further, under* H_{1N}, $T_{N,3}$ *converges weakly to the distribution of* $\left\{\sum_{i=1}^{4}|A_i^*|\right\}^2$, *where* $\mathbf{A}^* = \{A_1^*, A_2^*, A_3^*, A_4^*\}$ *is a four dimensional random vector associated with a Gaussian distribution with the location parameter* $= \boldsymbol{\mu}$ *and the scatter matrix* $= \Sigma_1$. *Here* $\boldsymbol{\mu}$ *is same as defined in the statement of Theorem 4.1, and* Σ_1 *is same as defined after Remark 2.1.*

Theorem 4.4: *Under* **(A1)**-**(A4)**, *the sequence of alternatives* $\{H_{1N}\}$ *will be contiguous alternatives relative to the null hypothesis* H_0. *Further, under* H_{1N}, $T_{N,4}$ *converges weakly to the distribution of* $\left\{\sum_{i=1}^{4}|B_i^*|\right\}^2$, *where* $\mathbf{B}^* = \{B_1^*, B_2^*, B_3^*, B_4^*\}$ *is a four dimensional random vector associated with a Gaussian distribution with the location parameter* $= \boldsymbol{\mu}^*$ *and the scatter matrix* $= \Sigma_2$. *Here* $\boldsymbol{\mu}^*$ *is same as defined in the statement of Theorem 4.2, and* Σ_2 *is same as defined after Remark 2.1.*

The assertions in Theorems 41–4.4 enable us to compute the asymptotic power of the tests under contiguous alternatives $H_{1,N}$. To carry out the tests, one first needs to compute $c_{1\eta}$, $c_{2\eta}$, $c_{3\eta}$ and $c_{4\eta}$ described in Theorems 2.1, 2.2, 2.3 and 2.4, respectively. For sake of concise presentation, only the computation of $c_{1\eta}$ is explained here. To compute the critical value $c_{1\eta}$, we generate the data with sample size $= 1000$ (denoted as $\{\mathbf{x}_1, \ldots, \mathbf{x}_{1000}\}$) from the Gaussian distribution associated with the random vector $\mathbf{A}$. Afterwards, one can derive the empirical distribution function of $\sum_{i=1}^{4} \lambda_i Z_i^2$ (defined in Theorem 2.1), and $(1-\eta)$-th quantile of that empirical distribution function is considered as the estimated value of $c_{1\eta}$. The computation of $c_{2\eta}$, $c_{3\eta}$ and $c_{4\eta}$ are done by the similar procedure as described for estimating $c_{1\eta}$. Next, to compute the asymptotic power of the test based on $T_{N,1}$, a data with sample size $= 1000$ (denoted as $\mathcal{Y} = \{\mathbf{y}_1, \ldots, \mathbf{y}_{1000}\}$) is obtained from the Gaussian distribution associated with the random vector $\mathbf{M}$ defined in Theorem 4.1, and the estimated asymptotic power of the test based on $T_{N,1}$ will be $\{1 - F_{\mathcal{Y}}(c_{1\eta})\}$, where $F_{\mathcal{Y}}$ is the empirical distribution function of $\sum_{i=1}^{4} \lambda_i M_i^2$ (defined in Theorem 4.1) based on the data $\mathcal{Y}$. Similarly, one can estimate the asymptotic power of the tests based on $T_{N,2}$, $T_{N,3}$ and $T_{N,4}$ using the results stated in Theorems 4.2, 4.3 and 4.4, respectively.

The asymptotic power of the tests based on $T_{N,1}$, $T_{N,2}$, $T_{N,3}$ and $T_{N,4}$ is shown in Figure 2 (given in supplementary file) for different values of δ, where $\delta = \delta_1 = \delta_2 = \delta_3 = \delta_4$. The equality among δ_1, δ_2, δ_3 and δ_4 is considered because of simple pictorial representation. In this study, the error distribution is assumed to be standard normal distribution and t distribution with 5 degrees of freedom. We see that $T_{N,1}$ based test seems to have higher power than the other three tests when the data came from a standard normal distribution, whereas $T_{N,2}$ based test perform better than the other three tests when the error distribution is t distribution with 5 degrees of freedom.

V. Real Data Analysis

We here implement the proposed tests on a few real data sets to see the practicability of the tests. As earlier, we are interested to test $H_0 : (A, B, \alpha, \beta) = (A_0, B_0, \alpha_0, \beta_0)$; however, note that in case of real data, $(A_0, B_0, \alpha_0, \beta_0)$ is unknown to us. The values of the parameters under null hypothesis (i.e., $(A_0, B_0, \alpha_0, \beta_0)$) are determined from the initial one-third of each data set (denote it as $\mathcal{Y}_{1/3}$, and $\mathcal{Y}_{2/3}$ denotes the remaining two-third of data). We apply absolute error and squared error based methods on $\mathcal{Y}_{1/3}$ and calculate the estimates of (A, B, α, β), which is fixed as $(A_0, B_0, \alpha_0, \beta_0)$. Afterwards, based on the remaining data, i.e., $\mathcal{Y}_{2/3}$, we want to test whether the unknown parameter (A, B, α, β) is same as $(A, B, \alpha_0, \beta_0)$ or not. The philosophy of such partitioning is that one can check whether the feature of the data is changing over time or not. In other words, $\mathcal{Y}_{1/3}$ can be though as the training data whereas $\mathcal{Y}_{2/3}$ is the test data. In Table I, along with the p-values from test statistics, we also provide the null hypothesized value of the parameter estimated on $\mathcal{Y}_{1/3}$.

We analyse four data set, which are obtained from four experiments, and these data set are named as Exp 1, Exp 2, Exp 3 and Exp 4. These data sets are collected from the lab in the Department of Electrical Engineering, IIT Kanpur, India and available to the journal's repository. Exp 1 data consist of 512 signal values drawn randomly at $10kHz$ frequency, and Exp 2 data contain 477 observations from a sonar signal experiments. Exp 3 and Exp 4 data contain 469 and 512 signals from certain experiments. All data set were earlier analyzed in [21]. In each row in Figure 3 (given in supplementary file), the observed variable is plotted in first diagram, and the corresponding histogram is presented in second diagram.

Next in order to carry out the test, we compute the p-values of the tests based on $T_{N,1}$, $T_{N,2}$, $T_{N,3}$ and $T_{N,4}$ using the idea of well-known permutation test (see, e.g., [22]). For sake of completeness, the implementation of the tests for real data is described here. Let t_0 be the value of the test statistic T_n (a generic notation) for a given data $\mathcal{Y} = \{y_1, \ldots, y_n\}$, and $t_{0,1}, \ldots, t_{0,m}$ are the values of T_n for the m many permuted samples $\{y_{i_1}, \ldots, y_{i_n}\}$, where $\{i_1, \ldots, i_n\}$ is a permutation of $\{1, \ldots, n\}$. Finally, the p-value is defined as the proportion of times $t_{0,1}, \ldots, t_{0,m}$ exceeding t_0. The p-values obtained by the proposed tests are summarized in Table I.

The figures in Table I indicate that the null hypothesis is favoured by all four tests for each data set, which leads to the conclusion that the inherent features in the initial one-third portion of each data is not significantly different from that of the remaining two-third portion of each data .

VI. Concluding Remarks

In this article, we investigate the testing of hypothesis problem that address whether the unknown parameters involved in the one-dimensional chirp signal model equal to any specified value or not. We here propose four test statistics; two of them are based on the least square

Data	$H_0 = (A_0, B_0, \alpha_0, \beta_0)$	$T_{N,1}$	$T_{N,2}$	$T_{N,3}$	$T_{N,4}$
Exp 1	LAD: (0.4489, -0.0499, 1.3079, 1.1533) LSE:(0.1329, 0.0217, 0.9909, 0.6961)	0.695	0.694	0.697	0.696
Exp 2	LAD:(0.0179, 0.2205, 1.4192, 1.2127) LSE:(0.0283, 0.1727, 1.3312, 1.0270)	0.849	0.816	0.850	0.817
Exp 3	LAD:(0.1878, 0.0025, 1.2114, 0.9597) LSE:(-0.0843, 0.0966, 1.4507, 1.2469)	0.656	0.595	0.657	0.602
Exp 4	LAD:(0.1219, -0.0591, 1.4907, 1.2453) LSE:(0.1485, -0.1056, 1.3499, 0.6976)	0.358	0.297	0.361	0.297

TABLE I
THE SECOND COLUMN PROVIDES THE LAD VALUES ARE FOR $T_{N,2}$ AND $T_{N,4}$ WHEREAS LSE VALUES ARE FOR $T_{N,1}$ AND $T_{N,3}$. THE VALUES IN THIRD, FOURTH, FIFTH AND SIXTH COLUMNS REPORT THE p-VALUES OBTAINED BY THE TESTS BASED ON $T_{N,1}$, $T_{N,2}$, $T_{N,3}$ AND $T_{N,4}$, RESPECTIVELY.

estimator, and the remaining two of them are based on the least absolute deviation estimator. We have observed that the least square estimator based procedures perform well for light tailed distribution while for the data obtained from heavy tailed distribution, the least absolute deviation based procedures outperform the other ones.

The asymptotic results described here are valid under the condition $A^2 + B^2 > 0$, i.e., in other words, A and B cannot both be equal to zero. However, one may have interest to test $H_0^* : A = B = 0$ against $H_1^* : H_0$ is not true. To test H_0^* against H_1^*, one needs to adopt different procedure to derive the asymptotic distribution of the least square estimator or the least absolute deviation estimator, and based on those asymptotic distributions, it may be possible to derive the consistency and the asymptotic power properties of the proposed tests.

As we mentioned earlier, the least absolute deviation based procedures outperform the least square based procedure when the data are obtained from a heavy tailed distribution. In this context, a natural question will arise that how one can formulate theoretically this worthy property of the least absolute deviation based procedure. Regarding the estimation procedure, one may adopt the concepts like breakdown point or the gross error sensitivity (see, e.g., [23]) and for the testing of hypothesis problem, one may have to define the breakdown point in terms of the power function.

To prove Theorems 4.1–4.4, Le Cam's third lemma is used to obtain the asymptotic distribution of $T_{n,1}$, $T_{n,2}$, $T_{n,3}$ and $T_{n,4}$ under H_n, and Le Cam's third lemma uses the fact that log likelihood ratio converges weakly to a random variable associated with a normal distribution having certain location and scale parameters (see the proofs of Theorems 4.1–4.4). We should point out that the asymptotic normality of log likelihood ratio is a sufficient condition but not a necessary condition to establish the contiguity. Instead of Le Cam's third lemma, one can also follow [24]'s approach based on a specific truncation method for contiguity of the density functions associated with H_n with respect to the density function associated with H_0. Also, [25] investigated the asymptotic relative efficiency of some tests for independence against general contiguous alternatives of positive quadrant dependence. However, neither [25] nor [24] considered the form of contiguous alternatives as we consider here.

REFERENCES

[1] Wu, Y., So, H.C. and Liu, H. (2008) Subspace-Based algorithm for parameter estimation of polynomial phase signals. *IEEE Transactions on Signal Processing*, **56**, 4977 – 4983.

[2] Lahiri, A. (2012) Estimators of Parameters of Chirp Signals and Their Properties. *unpublished PhD thesis, Indian Institute of Technology, Kanpur*.

[3] Lahiri, A., Kundu, D. and Mitra, A. (2014) On least absolute deviation estimator of one dimensional chirp model. *Statistics*, **4**8, 405–420.

[4] Lahiri, A. and Kundu, D. (2017) On parameter estimation of two-dimensional polynomial phase signal model. *Statistica Sinica*, **2**7, 1779 – 1792.

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

[6] Lenon, M.J. and McLaughlin S. (2002) Analysing sonar data using fractional Fourier transform. *Proceedings of the* 5^{th} *Nordic Signal Processing Symposium NORSIG*. Finland.

[7] Yetik, I. S. and Nehorai A. (2003) Beamforming using fractional Fourier transform. *IEEE Transactions on Signal Processing*, **5**1, 1663 – 1668.

[8] Bardeli R., Wolff D., Kurth F., Koch M., Tauchert K.H., Frommolt K.H. (2010) Detecting bird sounds in a complex acoustic environment and application to Bioacoustic monitoring. *Pattern Recognition Letters*, **3**1, 1524 - 1534.

[9] Somervuo, P., Harma, A., & Fagerlund, S. (2006) Parametric representations of bird sounds for automatic species recognition. *IEEE Transactions on Audio, Speech, and Language Processing*, **1**4, 2252-2263.

[10] Flandrin, P. (2001) Time frequency and chirps. In Wavelet Applications VIII. *International Society for Optics and Photonics*, **4**391, 161 – 176.

[11] Gini, F., Monteanari, M. and Verrazzani, L. (2000) Estimation of Chirp Signals in compound Gaussian clutter : A cyclostationary approach. *IEEE Transaction on Acoustics, Speech and Signal Processing*, **4**8, 1029–1039.

[12] Djuric, P.M. and Kay, S.M. (1990) Parameter estimation of chirp signals. *IEEE Transactions on Acoustics, Speech and Signal Processing*, **3**8, 2118 – 2126.

[13] Francos, J.M. and Friedlander, B. (1998) Two-Dimensional Polynomial Phase Signals: Parameter Estimation and Bounds. *Multidimensional Systems and Signal Processing*, **9**, 173–205.

[14] Francos, J.M. and Friedlander, B. (1999) Parameter Estimation of 2-D Random Amplitude Polynomial-Phase Signals. *IEEE Transactions on Signal Processing*, **4**7, 1795–1810.

[15] 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*, **5**6, 529–544.

[16] Kundu, D. and Nandi, S. (2008) Parameter estimation of chirp signals in presence of stationary noise. *Statistica Sinica*, **1**8, 187–201.

[17] Van Der Vaart, A. W. (1998) *Asymptotic Statistics.* Cambridge University Press.

[18] Lehmann, E. L. and Romano, J. P. (2005) *Testing Statistical Hypotheses.* New Delhi: Springer.

[19] Dhar, S. S., Bergsma, W. and Dassios, A. (2018) Testing Independence of covariates and errors in non-parametric regression. *Scandinavian Journal of Statistics*, **4**5, 421–443.

[20] Hajek, J., Sidak, Z. and Sen, P. K. (1999) *Theory of Rank Tests.* Academic Press.

[21] Kundu, D. and Nandi, S. (2012) *Statistical Signal Processing: Frequency Estimation*. Springer.

[22] Good, P. (1994) *Permutation Tests*. Springer.

[23] Huber, P. J. (1981) *Robust Statistics.* New York: Wiley.

[24] Behnen, K. and Neuhaus, G. (1975) A central limit theorem under contiguous alternatives. *The Annals of Statistics*, **3**, 1349–1353.

[25] Behnen, K. (1971) Asymptotic optimality and ARE of certain rank-order tests under contiguity. *The Annals of Mathematical Statistics*, **4**2, 325–329.

[26] Shao, J. (2003) *Mathematical Statistics*. Springer.

[27] Rao, C. R. (1973) *Linear Statistical Inference and its Applications*. John Wiley.

VII. Appendix: Proofs

In order to proof Theorem 2.1, one needs to use the following Lemma.

Lemma 1: *Under H_0 and* **(A1)-(A3)***, $T_{N,1}$ converges weakly to $\sum\limits_{i=1}^{4} \lambda_i Z_i^2$, where λ_is are the eigen values of Σ_1, and Z_is are independent standard normal random variables. Here, Σ_1 is same as defined after Remark 2.1.*

Proof of Lemma 1: Under H_0 and assumptions **(A1)-(A3)**, it follows from [15] that $D^{-1}(\hat{\boldsymbol{\theta}}_{N,LSE} - \boldsymbol{\theta}_0)$ converges weakly to $4-$dimensional random vector $\mathbf{R}$ associated with a Gaussian distribution with the location parameter $= (0,0,0,0)'$ and the scatter matrix $= \Sigma_1$. Define a new random vector $L = AR$, where A is a 4×4 orthogonal matrix (i.e., $A^T A = AA^T = I_4$) such that $A \Sigma A^T = Diag\{\lambda_1, \lambda_2, \lambda_3, \lambda_4\}$. Here A^T denotes the transpose of the matrix A. The aforementioned fact implies that $AD^{-1}(\hat{\boldsymbol{\theta}}_N - \boldsymbol{\theta}_0)$ converges weakly to the 4-dimensional random vector AR, which is associated with the Gaussian distribution with the location parameter $= (0,0,0,0)$ and the scatter parameter $= Diag\{\lambda_1, \lambda_2, \lambda_3, \lambda_4\}$ since $A\Sigma A^T = Diag\{\lambda_1, \lambda_2, \lambda_3, \lambda_4\}$. Therefore $\{AD^{-1}(\hat{\boldsymbol{\theta}}_N - \boldsymbol{\theta}_0)\}^T \{AD^{-1}(\hat{\boldsymbol{\theta}}_N - \boldsymbol{\theta}_0)\} = \{D^{-1}(\hat{\boldsymbol{\theta}}_N - \boldsymbol{\theta}_0)\}^T A^T A \{D^{-1}(\hat{\boldsymbol{\theta}}_N - \boldsymbol{\theta}_0)\} = ||D^{-1}(\hat{\boldsymbol{\theta}}_N - \boldsymbol{\theta}_0)||^2 := T_{N,1}$ converges weakly to $\sum\limits_{i=1}^{4} \lambda_i Z_i^2$, where λ_is and Z_is are same as defined in the statement of the theorem. Hence, the proof is completed. □

Proof of Theorem 2.1: To prove this theorem, one needs to show that $P_{H_1}[T_{N,1} > c_{1\eta}] \to 1$ as $N \to \infty$, where $c_{1\eta}$ is the critical value of the level of significance η to test H_0 against H_1, which follows from Lemma 1. In other words, $c_{1\eta}$ is $(1-\eta)$-th quantile of the distribution of $\sum\limits_{i=1}^{4} \lambda_i Z_i^2$ that described in the statement of this theorem.

We now consider

$$
\begin{aligned}
& P_{H_1}[T_{N,1} > c_{1\eta}] \\
= \; & P_{H_1}[||D^{-1}(\hat{\boldsymbol{\theta}}_{N,LSE} - \boldsymbol{\theta}_0)||^2 > c_{1\eta}] \\
= \; & P_{H_1}[||D^{-1}(\hat{\boldsymbol{\theta}}_{N,LSE} - \boldsymbol{\theta}_0) - D^{-1}(\boldsymbol{\theta}_1 - \boldsymbol{\theta}_0)||^2 > c_{1\eta} + ||D^{-1}(\boldsymbol{\theta}_1 - \boldsymbol{\theta}_0)||^2 - 2 < D^{-1}(\hat{\boldsymbol{\theta}}_{N,LSE} - \boldsymbol{\theta}_0), D^{-1}(\boldsymbol{\theta}_1 - \boldsymbol{\theta}_0) \\
& \text{ since under } H_1,\ \boldsymbol{\theta} = \boldsymbol{\theta}_1, \text{ and } <.,.> \text{ denotes the inner product.} \\
= \; & P_{H_1}[||D^{-1}(\hat{\boldsymbol{\theta}}_{N,LSE} - \boldsymbol{\theta}_1)||^2 > c_{1\eta} + ||D^{-1}(\boldsymbol{\theta}_1 - \boldsymbol{\theta}_0)||^2 - 2 < D^{-1}(\hat{\boldsymbol{\theta}}_{N,LSE} - \boldsymbol{\theta}_0), D^{-1}(\boldsymbol{\theta}_1 - \boldsymbol{\theta}_0) >] \\
= \; & P_{H_1}[T_{N,1} > c_{1\eta} + ||D^{-1}(\boldsymbol{\theta}_1 - \boldsymbol{\theta}_0)||^2 - 2 < D^{-1}(\hat{\boldsymbol{\theta}}_N - \boldsymbol{\theta}_0), D^{-1}(\boldsymbol{\theta}_1 - \boldsymbol{\theta}_0) >] \\
& \text{ since under } H_1,\ T_{N,1} = ||D^{-1}(\hat{\boldsymbol{\theta}}_{N,LSE} - \boldsymbol{\theta}_1)||^2. \\
= \; & P_{H_1}[T_{N,1} > c_{1\eta} + ||D^{-1}(\boldsymbol{\theta}_1 - \boldsymbol{\theta}_0)||^2 - 2 < D^{-1}(\boldsymbol{\theta}_1 - \boldsymbol{\theta}_0), D^{-1}(\boldsymbol{\theta}_1 - \boldsymbol{\theta}_0) > \\
+ \; & 2 < D^{-1}(\hat{\boldsymbol{\theta}}_N - \boldsymbol{\theta}_1), D^{-1}(\boldsymbol{\theta}_1 - \boldsymbol{\theta}_0) >] \\
= \; & P_{H_1}[T_{N,1} > c_{1\eta} + ||D^{-1}(\boldsymbol{\theta}_1 - \boldsymbol{\theta}_0)||^2 - 2||D^{-1}(\boldsymbol{\theta}_1 - \boldsymbol{\theta}_0)||^2 + 2 < D^{-1}(\hat{\boldsymbol{\theta}}_{N,LSE} - \boldsymbol{\theta}_1), D^{-1}(\boldsymbol{\theta}_1 - \boldsymbol{\theta}_0) >] \\
= \; & P_{H_1}[T_{N,1} > c_{1\eta} - ||D^{-1}(\boldsymbol{\theta}_1 - \boldsymbol{\theta}_0)||^2 + 2 < D^{-1}(\hat{\boldsymbol{\theta}}_{N,LSE} - \boldsymbol{\theta}_1), D^{-1}(\boldsymbol{\theta}_1 - \boldsymbol{\theta}_0) >] \\
\to \; & 1 \text{ as } N \to \infty.
\end{aligned}
$$

The last step follows from the facts that $||D^{-1}(\boldsymbol{\theta}_1 - \boldsymbol{\theta}_0)||^2 \to \infty$ as $N \to \infty$ and $||D^{-1}(\hat{\boldsymbol{\theta}}_{N,LSE} - \boldsymbol{\theta}_1)|| \to 0$ as $N \to \infty$ in probability under H_1 and **(A1)-(A3)** . Hence, it completes the proof. □

The proof of Theorem 2.2 is based on the fact of Lemma 2.

Lemma 2: *Under H_0 and* **(A1)-(A3)***, $T_{N,2}$ converges weakly to $\sum\limits_{i=1}^{4} \lambda_i^* Z_i^{*2}$, where λ_i^*s are the eigen values of Σ_2, and Z_i^*s are independent standard normal random variables. Here, Σ_2 is same as defined after Remark 2.1.*

Proof of Lemma 2: Under H_0 and assumptions **(A1)-(A2)**, [3] showed that $D^{-1}(\hat{\boldsymbol{\theta}}_{N,LAD} - \boldsymbol{\theta}_0)$ converges weakly to $4-$dimensional random vector associated with a Gaussian distribution with the location parameter $= (0,0,0,0)'$ and the scatter matrix $= \frac{1}{\{f(0)\}^2}\Sigma$ when the probability density function of the error random variables is symmetric about zero. A straightforward modification of their proof shows that under H_0 and assumptions **(A1)-(A3)**, $D^{-1}(\hat{\boldsymbol{\theta}}_{N,LAD} - \boldsymbol{\theta}_0)$ converges weakly to $4-$dimensional random vector $\mathbf{R}^*$ associated with a Gaussian distribution with the location parameter $= (0,0,0,0)'$ and the scatter matrix $= \Sigma_2$. Define a new random vector $L^* = A^* R^*$, where A^* is a 4×4 orthogonal matrix (i.e., $A^{*T} A^* = A^* A^{*T} = I_4$) such that $A^* \Sigma_2 A^{*T} = Diag\{\lambda_1^*, \lambda_2^*, \lambda_3^*, \lambda_4^*\}$. Here A^{*T} denotes the transpose of the matrix A^*. The aforementioned fact implies that $A^* D^{-1}(\hat{\boldsymbol{\theta}}_{N,LAD} - \boldsymbol{\theta}_0)$ converges weakly to the 4-dimensional random vector $A^* R^*$, which is associated with the

Gaussian distribution with the location parameter $= (0,0,0,0)$ and the scatter parameter $= Diag\{\lambda_1^*, \lambda_2^*, \lambda_3^*, \lambda_4^*\}$ since $A^*\Sigma_2 A^{*T} = Diag\{\lambda_1^*, \lambda_2^*, \lambda_3^*, \lambda_4^*\}$. Therefore $\{A^* D^{-1}(\hat{\boldsymbol{\theta}}_{N,LAD} - \boldsymbol{\theta}_0)\}^T \{A^* D^{-1}(\hat{\boldsymbol{\theta}}_{N,LAD} - \boldsymbol{\theta}_0)\} = \{D^{-1}(\hat{\boldsymbol{\theta}}_{N,LAD} - \boldsymbol{\theta}_0)\}^T A^{*T} A^* \{D^{-1}(\hat{\boldsymbol{\theta}}_{N,LAD} - \boldsymbol{\theta}_0)\} = ||D^{-1}(\hat{\boldsymbol{\theta}}_{N,LAD} - \boldsymbol{\theta}_0)||^2 := T_{N,2}$ converges weakly to $\sum_{i=1}^{4} \lambda_i^* Z_i^{*2}$, where λ_i^*s and Z_i^*s are same as defined in the statement of the theorem. Hence, the proof is completed. □

Proof of Theorem 2.2: To prove this theorem, one needs to show that $P_{H_1}[T_{N,2} > c_{2\eta}] \to 1$ as $N \to \infty$, where $c_{2\eta}$ is the critical value of the level of significance η to test H_0 against H_1, which follows from Lemma 2. In other words, $c_{1\eta}$ is $(1-\eta)$-th quantile of the distribution of $\sum_{i=1}^{4} \lambda_i^* Z_i^{*2}$ that described in the statement of this theorem.

We now consider

$$
\begin{aligned}
& P_{H_1}[T_{N,2} > c_{2\eta}] \\
= \; & P_{H_1}[||D^{-1}(\hat{\boldsymbol{\theta}}_{N,LAD} - \boldsymbol{\theta}_0)||^2 > c_{2\eta}] \\
= \; & P_{H_1}[||D^{-1}(\hat{\boldsymbol{\theta}}_{N,LAD} - \boldsymbol{\theta}_0) - D^{-1}(\boldsymbol{\theta}_1 - \boldsymbol{\theta}_0)||^2 > c_{2\eta} + ||D^{-1}(\boldsymbol{\theta}_1 - \boldsymbol{\theta}_0)||^2 - \\
& 2 < D^{-1}(\hat{\boldsymbol{\theta}}_{N,LAD} - \boldsymbol{\theta}_0), D^{-1}(\boldsymbol{\theta}_1 - \boldsymbol{\theta}_0) >] \\
& \text{ since under } H_1,\ \boldsymbol{\theta} = \boldsymbol{\theta}_1, \text{ and } < .,. > \text{ denotes the inner product.} \\
= \; & P_{H_1}[||D^{-1}(\hat{\boldsymbol{\theta}}_{N,LAD} - \boldsymbol{\theta}_1)||^2 > c_{2\eta} + ||D^{-1}(\boldsymbol{\theta}_1 - \boldsymbol{\theta}_0)||^2 - 2 < D^{-1}(\hat{\boldsymbol{\theta}}_{N,LAD} - \boldsymbol{\theta}_0), D^{-1}(\boldsymbol{\theta}_1 - \boldsymbol{\theta}_0) >] \\
= \; & P_{H_1}[T_{N,2} > c_{2\eta} + ||D^{-1}(\boldsymbol{\theta}_1 - \boldsymbol{\theta}_0)||^2 - 2 < D^{-1}(\hat{\boldsymbol{\theta}}_N - \boldsymbol{\theta}_0), D^{-1}(\boldsymbol{\theta}_1 - \boldsymbol{\theta}_0) >] \\
& \text{ since under } H_1,\ T_{N,2} = ||D^{-1}(\hat{\boldsymbol{\theta}}_{N,LAD} - \boldsymbol{\theta}_1)||^2. \\
= \; & P_{H_1}[T_{N,2} > c_{2\eta} + ||D^{-1}(\boldsymbol{\theta}_1 - \boldsymbol{\theta}_0)||^2 - 2 < D^{-1}(\boldsymbol{\theta}_1 - \boldsymbol{\theta}_0), D^{-1}(\boldsymbol{\theta}_1 - \boldsymbol{\theta}_0) > \\
+ \; & 2 < D^{-1}(\hat{\boldsymbol{\theta}}_N - \boldsymbol{\theta}_1), D^{-1}(\boldsymbol{\theta}_1 - \boldsymbol{\theta}_0) >] \\
= \; & P_{H_1}[T_{N,2} > c_{2\eta} + ||D^{-1}(\boldsymbol{\theta}_1 - \boldsymbol{\theta}_0)||^2 - 2||D^{-1}(\boldsymbol{\theta}_1 - \boldsymbol{\theta}_0)||^2 + 2 < D^{-1}(\hat{\boldsymbol{\theta}}_{N,LAD} - \boldsymbol{\theta}_1), D^{-1}(\boldsymbol{\theta}_1 - \boldsymbol{\theta}_0) >] \\
= \; & P_{H_1}[T_{N,2} > c_{2\eta} - ||D^{-1}(\boldsymbol{\theta}_1 - \boldsymbol{\theta}_0)||^2 + 2 < D^{-1}(\hat{\boldsymbol{\theta}}_{N,LAD} - \boldsymbol{\theta}_1), D^{-1}(\boldsymbol{\theta}_1 - \boldsymbol{\theta}_0) >] \\
\to \; & 1 \text{ as } N \to \infty.
\end{aligned}
$$

The last step follows from the facts that $||D^{-1}(\boldsymbol{\theta}_1 - \boldsymbol{\theta}_0)||^2 \to \infty$ as $N \to \infty$ and $||D^{-1}(\hat{\boldsymbol{\theta}}_{N,LAD} - \boldsymbol{\theta}_1)|| \to 0$ as $N \to \infty$ in probability under H_1 and **(A1)-(A3)**. Hence, it completes the proof. □

Lemma 3 is stated in the following, which will be used in the proof of Theorem 2.3.

Lemma 3: *Under H_0 and* **(A1)-(A3)***, $T_{N,3}$ converges weakly to $\left\{\sum_{i=1}^{4} |A_i|\right\}^2$, where A_is are same as defined after Remark 2.1.*

Proof of Lemma 3: As it is mentioned in the proof of Lemma 1, it follows from [15] that $D^{-1}(\hat{\boldsymbol{\theta}}_{N,LSE} - \boldsymbol{\theta}_0)$ converges weakly to 4−dimensional random vector $\mathbf{A} = \{A_1, A_2, A_3, A_4\}$ associated with a Gaussian distribution with the location parameter $= (0,0,0,0)'$ and the scatter matrix $= \Sigma_1$. Next, since the L_1 norm and the square are continuous mapping, the continuous mapping theorem (see, e.g., [17]) implies that $||D^{-1}(\hat{\boldsymbol{\theta}}_{N,LSE} - \boldsymbol{\theta}_0)||_1^2 := T_{N,3}$ converges weakly to $\left\{\sum_{i=1}^{4} |A_i|\right\}^2$, and hence, the proof is complete. □

Proof of Theorem 2.3: To prove this theorem, one needs to show that $P_{H_1}[T_{N,3} > c_{3\eta}] \to 1$ as $N \to \infty$, where $c_{3\eta}$ is the critical value of the level of significance η to test H_0 against H_1, which follows from Lemma 3. In other words, $c_{3\eta}$ is $(1-\eta)$-th quantile of the distribution of $\left\{\sum_{i=1}^{4} |A_i|\right\}^2$ that described in the statement of this theorem.

We now consider

$$\begin{aligned}
& P_{H_1}[T_{N,3} > c_{3\eta}] = P_{H_1}[||D^{-1}(\hat{\boldsymbol{\theta}}_{N,LSE} - \boldsymbol{\theta}_0)||_1^2 > c_{3\eta}] \\
= \;& P_{H_1}[||D^{-1}(\hat{\boldsymbol{\theta}}_{N,LSE} - \boldsymbol{\theta}_0)||_1 > k] \; (k = \sqrt{c_{3\eta}} \text{ is an arbitrary constant}) \\
= \;& P_{H_1}[||D^{-1}(\hat{\boldsymbol{\theta}}_{N,LSE} - \boldsymbol{\theta}_1 + \boldsymbol{\theta}_1 - \boldsymbol{\theta}_0)||_1 > k] \\
= \;& P_{H_1}[||D^{-1}\{(\hat{\boldsymbol{\theta}}_{N,LSE} - \boldsymbol{\theta}_1) - (\boldsymbol{\theta}_0 - \boldsymbol{\theta}_1)\}||_1 > k] \\
\geq \;& P_{H_1}[||D^{-1}(\hat{\boldsymbol{\theta}}_{N,LSE} - \boldsymbol{\theta}_1)||_1 > k - D^{-1}||\boldsymbol{\theta}_0 - \boldsymbol{\theta}_1||_1] \; (\text{since } |a-b| \geq |a| - |b|) \\
\to \;& 1 \text{ as } N \to \infty.
\end{aligned}$$

The last step follows from the fact that $||D^{-1}(\hat{\boldsymbol{\theta}}_{N,LSE} - \boldsymbol{\theta}_1)||_1$ converges weakly to $\sum_{i=1}^{4} |A_i|$, i.e., by Prohorov's theorem (see, e.g., [17]), $||D^{-1}(\hat{\boldsymbol{\theta}}_{N,LSE} - \boldsymbol{\theta}_1)||_1 = O_p(1)$, and $D^{-1}||\boldsymbol{\theta}_0 - \boldsymbol{\theta}_1||_1 \to \infty$ as $N \to \infty$. Hence, it completes the proof. □

Lemma 4 is stated in the following, which will be used in the proof of Theorem 2.4.

Lemma 4: *Under H_0 and* **(A1)-(A3)***, $T_{N,4}$ converges weakly to* $\left\{\sum_{i=1}^{4} |B_i|\right\}^2$*, where B_is are same as defined after Remark 2.1.*

Proof of Lemma 4: As it is established in Lemma 2, we have that $D^{-1}(\hat{\boldsymbol{\theta}}_{N,LAD} - \boldsymbol{\theta}_0)$ converges weakly to 4−dimensional random vector $\mathbf{B} = \{B_1, B_2, B_3, B_4\}$ associated with a Gaussian distribution with the location parameter $= (0,0,0,0)'$ and the scatter matrix $= \Sigma_2$. Arguing in a similar way as in the proof of Lemma 3, since the L_1 norm and the square are continuous mapping, the continuous mapping theorem implies that $||D^{-1}(\hat{\boldsymbol{\theta}}_{N,LAD} - \boldsymbol{\theta}_0)||_1^2 := T_{N,4}$ converges weakly to $\left\{\sum_{i=1}^{4} |B_i|\right\}^2$, and hence, the proof is complete. □

Proof of Theorem 2.4: To prove this theorem, one needs to show that $P_{H_1}[T_{N,4} > c_{4\eta}] \to 1$ as $N \to \infty$, where $c_{4\eta}$ is the critical value of the level of significance η to test H_0 against H_1, which follows from Lemma 4. In other words, $c_{4\eta}$ is $(1-\eta)$-th quantile of the distribution of $\left\{\sum_{i=1}^{4} |B_i|\right\}^2$ that described in the statement of this theorem.

We now consider

$$\begin{aligned}
& P_{H_1}[T_{N,4} > c_{4\eta}] = P_{H_1}[||D^{-1}(\hat{\boldsymbol{\theta}}_{N,LAD} - \boldsymbol{\theta}_0)||_1^2 > c_{4\eta}] \\
= \;& P_{H_1}[||D^{-1}(\hat{\boldsymbol{\theta}}_{N,LAD} - \boldsymbol{\theta}_0)||_1 > k^*] \; (k^* = \sqrt{c_{4\eta}} \text{ is an arbitrary constant}) \\
= \;& P_{H_1}[||D^{-1}(\hat{\boldsymbol{\theta}}_{N,LAD} - \boldsymbol{\theta}_1 + \boldsymbol{\theta}_1 - \boldsymbol{\theta}_0)||_1 > k^*] \\
= \;& P_{H_1}[||D^{-1}\{(\hat{\boldsymbol{\theta}}_{N,LAD} - \boldsymbol{\theta}_1) - (\boldsymbol{\theta}_0 - \boldsymbol{\theta}_1)\}||_1 > k^*] \\
\geq \;& P_{H_1}[||D^{-1}(\hat{\boldsymbol{\theta}}_{N,LAD} - \boldsymbol{\theta}_1)||_1 > k^* - D^{-1}||\boldsymbol{\theta}_0 - \boldsymbol{\theta}_1||_1] \; (\text{since } |a-b| \geq |a| - |b|) \\
\to \;& 1 \text{ as } N \to \infty.
\end{aligned}$$

The last step follows from the fact that $||D^{-1}(\hat{\boldsymbol{\theta}}_{N,LAD} - \boldsymbol{\theta}_1)||_1$ converges weakly to $\sum_{i=1}^{4} |B_i|$, i.e., $||D^{-1}(\hat{\boldsymbol{\theta}}_{N,LAD} - \boldsymbol{\theta}_1)||_1 = O_p(1)$, and $D^{-1}||\boldsymbol{\theta}_0 - \boldsymbol{\theta}_1||_1 \to \infty$ as $N \to \infty$. Hence, the proof is complete. □

Proof of Theorem 4.1: In order to establish the contiguity of the sequence $F_n(y_n, \boldsymbol{\theta}_N)$ (or $f_n(y_n, \boldsymbol{\theta}_N)$) relative to $F_n(y_n, \boldsymbol{\theta}_0)$ (or $f_n(y_n, \boldsymbol{\theta}_0)$), it is enough to show that L_n, the logarithm of the likelihood ratio, is asymptotically normal with mean $-\frac{1}{2}\sigma^2$ and variance σ^2 (see [20] (p. 253, Corollary 2 to Lecams first Lemma)), where σ is a positive constant. We now have

$$\begin{aligned} L_N &= \log\prod_{n=1}^{N}\frac{f_n(y_n,\boldsymbol{\theta}_N)}{f_n(y_n,\boldsymbol{\theta}_0)} = \sum_{n=1}^{N}\log\frac{f_n(y_n,\boldsymbol{\theta}_0+\boldsymbol{\delta}_N)}{f_n(y_n,\boldsymbol{\theta}_0)} \text{ since } f_n\text{s are independent.} \\ &= \sum_{n=1}^{N}\{\log f_n(y_n,\boldsymbol{\theta}_0+\boldsymbol{\delta}_N)-\log f_n(y_n,\boldsymbol{\theta}_0)\} \\ &= \sum_{n=1}^{N}\{g_n(y_n,\boldsymbol{\theta}_0+\boldsymbol{\delta}_N)-g_n(y_n,\boldsymbol{\theta}_0)\} \text{ (here we denote } g_n(.)=\log f_n(.)) \\ &= \sum_{n=1}^{N}\{g_n(y_n,\boldsymbol{\theta}_0)+<\boldsymbol{\delta}_N,\bigtriangledown g_n(y_n,\boldsymbol{\theta}_0)>+\frac{1}{2}<\boldsymbol{\delta}_N,H_{\boldsymbol{\theta}}(g_n(y_n,\boldsymbol{\theta}_*))\boldsymbol{\delta}_N>-g_n(y_n,\boldsymbol{\theta}_0)\} \\ &= \sum_{n=1}^{N}\{<\boldsymbol{\delta}_N,\bigtriangledown g_n(y_n,\boldsymbol{\theta}_0)>+\frac{1}{2}<\boldsymbol{\delta}_N,H_{\boldsymbol{\theta}}(g_n(y_n,\boldsymbol{\theta}_*))\boldsymbol{\delta}_N>\}, \end{aligned}$$

where $\bigtriangledown g_n(y_n,\boldsymbol{\theta}_0) = (\frac{\partial g_n(y_n,\boldsymbol{\theta})}{\partial\theta_1}|_{\boldsymbol{\theta}=\boldsymbol{\theta}_0}, \frac{\partial g_n(y_n,\boldsymbol{\theta})}{\partial\theta_2}|_{\boldsymbol{\theta}=\boldsymbol{\theta}_0}, \frac{\partial g_n(y_n,\boldsymbol{\theta})}{\partial\theta_3}|_{\boldsymbol{\theta}=\boldsymbol{\theta}_0}, \frac{\partial g_n(y_n,\boldsymbol{\theta})}{\partial\theta_4}|_{\boldsymbol{\theta}=\boldsymbol{\theta}_0})$, $H_{\boldsymbol{\theta}}(g_n(y_n,\boldsymbol{\theta})) = \left(\left(\frac{\partial^2 g_n(y_n,\boldsymbol{\theta})}{\partial\theta_i\partial\theta_j}\right)\right)_{i,j}$; $1\le i,j\le 4$, $(\theta_1,\theta_2,\theta_3,\theta_4)=(A,B,\alpha,\beta)$, $\boldsymbol{\theta}_*$ is lying on the straight line connecting $\boldsymbol{\theta}_0$ and $\boldsymbol{\theta}_N$. The expression of L_N follows from the Taylor series expansion of $g_n(y_n,\boldsymbol{\theta}_0+\boldsymbol{\delta}_N)$ up to the second term. Further, note that both of $<\boldsymbol{\delta}_N,\bigtriangledown g_n(y_n,\boldsymbol{\theta}_0)>$ and $<\boldsymbol{\delta}_N,H_{\boldsymbol{\theta}}(g_n(y_n,\boldsymbol{\theta}_*))\boldsymbol{\delta}_N>$ are sequence of independent random variables, and $\boldsymbol{\theta}_*\to\boldsymbol{\theta}_0$ as $N\to\infty$. Hence, using the central limit theorem and the weak law of large number of the sequence of independent random variables ((i.e., Liapounov's CLT and Tchebychevs WLLN, e.g., see [26]), under assumption **(A4)**, we have that $<\boldsymbol{\delta}_N,\bigtriangledown g_n(y_n,\boldsymbol{\theta}_0)>$ converges weakly to a normal random variable with mean zero and variance $= E_{F_{n,H_0}}\{\bigtriangledown g_n(y_n,\boldsymbol{\theta}_0)\mathbf{1}\}^2$, and $\frac{1}{2}<\boldsymbol{\delta}_N,H_{\boldsymbol{\theta}}(g_n(y_n,\boldsymbol{\theta}_*))\boldsymbol{\delta}_N>$ converges in probability to $\frac{1}{2}E_{F_{n,H_0}}<\mathbf{1},H_{\boldsymbol{\theta}}(g_n(y_n,\boldsymbol{\theta}_0))\mathbf{1}>$, where $\mathbf{1}=(1,1,0,0)$. Further, we have $E_{F_{n,H_0}}<\mathbf{1},H_{\boldsymbol{\theta}}(g_n(y_n,\boldsymbol{\theta}_0))\mathbf{1}> = -E_{F_{n,H_0}}\{\bigtriangledown g_n(y_n,\boldsymbol{\theta}_0)\mathbf{1}\}^2$, which follows from Rao-Bhattyacharya lower bound (see [27]). The above fact implies that L_N converges weakly to a normal random variable with mean $= -\frac{1}{2}E_{F_{n,H_0}}\{\bigtriangledown g_n(y_n,\boldsymbol{\theta}_0)\mathbf{1}\}^2$ and variance $= E_{F_{n,H_0}}\{\bigtriangledown g_n(y_n,\boldsymbol{\theta}_0)\mathbf{1}\}^2$. Hence, the sequence $F_n(y_n,\boldsymbol{\theta}_N)$ (or $f_n(y_n,\boldsymbol{\theta}_N)$) is contiguous relative to $F_n(y_n,\boldsymbol{\theta}_0)$ (or $f_n(y_n,\boldsymbol{\theta}_0)$).

Next, under H_0 and assumptions **(A1)-(A3)**, it follows from [15] that $D^{-1}(\hat{\boldsymbol{\theta}}_{N,LSE}-\boldsymbol{\theta}_0)$ converges weakly to 4−dimensional random vector associated with a Gaussian distribution with the location parameter $=(0,0,0,0)'$ and the scatter matrix $=\Sigma_1$ under H_0. Further, in view of expansion of L_N along with the asymptotic normality of $D^{-1}(\hat{\boldsymbol{\theta}}_{N,lSE}-\boldsymbol{\theta}_0)$ under H_0, we have that the joint distribution of $(D^{-1}(\hat{\boldsymbol{\theta}}_{N,LSE}-\boldsymbol{\theta}_0),L_N)$ is asymptotically 5-dimensional normal distribution under H_0. Hence, one can apply Le Cam's third lemma (see [20]) to establish the asymptotic distribution of $D^{-1}(\hat{\boldsymbol{\theta}}_{N,LSE}-\boldsymbol{\theta}_0)$ under H_{1N}. It is here appropriate to note that under H_{1N} also, $D^{-1}(\hat{\boldsymbol{\theta}}_{N,LSE}-\boldsymbol{\theta}_0)$ weakly converges to a Gaussian random variable as under H_0 but a location shift occurs in the expression of location parameter of the Gaussian distribution. Le Cam's third lemma implies that the location shift is essentially the component wise asymptotic covariance between $D^{-1}(\hat{\boldsymbol{\theta}}_{N,LSE}-\boldsymbol{\theta}_0)$ and L_N. [15] established that $D^{-1}(\hat{\boldsymbol{\theta}}_{N,LSE}-\boldsymbol{\theta}_0) = -\Sigma_1^{-1}D\{\bigtriangledown Q(\boldsymbol{\theta}_0)\}+R_N$, where $R_N\xrightarrow{p}0$, and Σ_1 is same as defined after Remark 2.1. The expressions of $D\{\bigtriangledown Q(\boldsymbol{\theta}_0)\}$ are provided in [16], where $Q(\boldsymbol{\theta}_0)=\sum_{n=1}^{N}\{y_n-A_0\cos(\alpha_0 n+\beta_0 n^2)-B_0\sin(\alpha_0 n+\beta_0 n^2)\}^2$. Based on these expressions, the component wise asymptotic covariance between $D^{-1}(\hat{\boldsymbol{\theta}}_{N,LSE}-\boldsymbol{\theta}_0)$ and L_N as given in the expression of $\boldsymbol{\mu}$ in the statement of the theorem can be obtained by a direct algebraic computation. Afterwards, the result follows from the same arguments related to an appropriate orthogonal transformation as in the proof of Lemma 1. This completes the proof. □

Proof of Theorem 4.2: It follows from [3] that under H_0 and assumptions **(A1)-(A3)**, $D^{-1}(\hat{\boldsymbol{\theta}}_{N,LAD}-\boldsymbol{\theta}_0)$ converges weakly to 4−dimensional random vector associated with a Gaussian distribution with the location parameter $=(0,0,0,0)'$ and the scatter matrix $=\Sigma_2$. Further, in view of expansion of L_N (see the proof of Theorem 4.1) along with the asymptotic normality of $D^{-1}(\hat{\boldsymbol{\theta}}_{N,LAD}-\boldsymbol{\theta}_0)$ under H_0, we have that the joint distribution of $(D^{-1}(\hat{\boldsymbol{\theta}}_{N,LAD}-\boldsymbol{\theta}_0),L_N)$ is asymptotically 5-dimensional normal distribution under H_0. Hence, one can again apply Le Cam's third lemma to establish the asymptotic distribution of $D^{-1}(\hat{\boldsymbol{\theta}}_{N,LAD}-\boldsymbol{\theta}_0)$ under

H_{1N}. Here also, note that under H_{1N}, $D^{-1}(\hat{\boldsymbol{\theta}}_{N,LAD} - \boldsymbol{\theta}_0)$ weakly converges to a Gaussian random variable as under H_0 but a location shift occurs in the expression of the location parameter of the Gaussian distribution. Now, Le Cam's third lemma implies that the location shift is essentially the component wise asymptotic covariance between $D^{-1}(\hat{\boldsymbol{\theta}}_{N,LAD} - \boldsymbol{\theta}_0)$ and L_N. [3] derived that $D^{-1}(\hat{\boldsymbol{\theta}}_{N,LAD} - \boldsymbol{\theta}_0) = -\Sigma_2^{-1} D\{\bigtriangledown Q^{'}(\boldsymbol{\theta}_0)\} + R_N$, where $R_N \stackrel{p}{\rightarrow} 0$, and Σ_2 is same as defined after Remark 2.1. The expressions of $D\{\bigtriangledown Q^{'}(\boldsymbol{\theta}_0)\}$ are provided in [3], where $Q^{'}(\boldsymbol{\theta}_0) = \sum\limits_{n=1}^{N} |y_n - A_0 \cos(\alpha_0 n + \beta_0 n^2) - B_0 \sin(\alpha_0 n + \beta_0 n^2)|$. Based on these expressions, the component wise asymptotic covariance between $D^{-1}(\hat{\boldsymbol{\theta}}_{N,LAD} - \boldsymbol{\theta}_0)$ and L_N as given in the expression of $\boldsymbol{\mu}^*$ in the statement of the theorem can be obtained by a direct algebraic computation. Afterwards, the result follows from the same arguments related to an appropriate orthogonal transformation as in the proof of Lemma 2. This completes the proof. □

Proof of Theorem 4.3: As we established in the proof of Theorem 4.1, $D^{-1}(\hat{\boldsymbol{\theta}}_{N,LSE} - \boldsymbol{\theta}_0)$ converges weakly to a random vector $\mathbf{A}^*$, which is defined in the statement of this theorem. Afterwards, a direct application of continuous mapping theorem (see, e.g., [17]) leads to the proof. □

Proof of Theorem 4.4: As we established in the proof of Theorem 4.2, $D^{-1}(\hat{\boldsymbol{\theta}}_{N,LAD} - \boldsymbol{\theta}_0)$ converges weakly to a random vector $\mathbf{B}^*$, which is defined in the statement of this theorem. Afterwards, a direct application of continuous mapping theorem leads to the proof. □

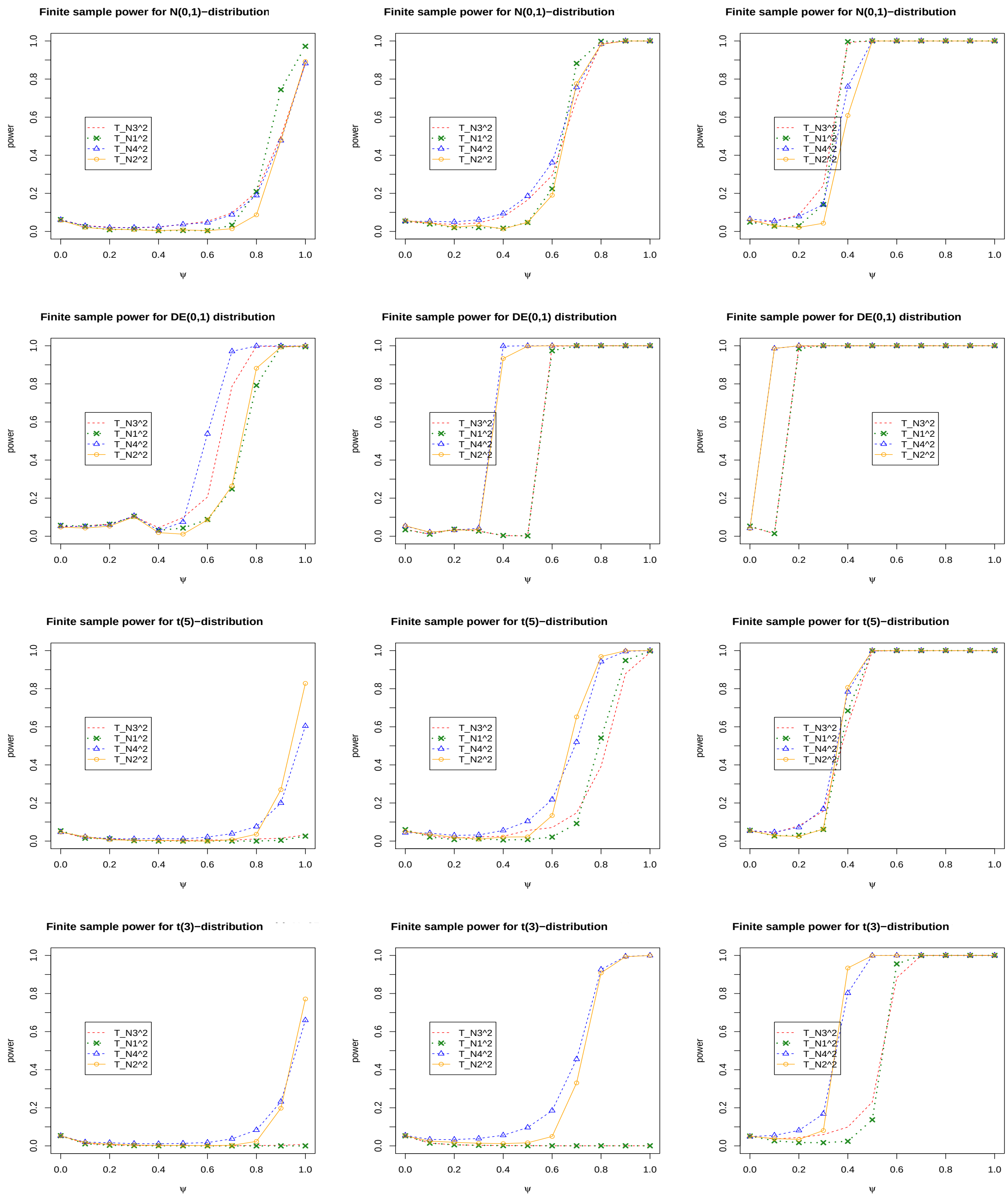

Figure 3 : The finite sample power of different tests based on four different test statistics as described in Section 2 at 5% level of significance. In the diagrams, $Normal$, DE, $t(5)$ and $t(3)$ denote the standard normal, the standard Laplace, t_5 and t_3 distributions, respectively. The first, the second and the third diagrams in each row are based on the sample sizes = 30, 50 and 100, respectively.

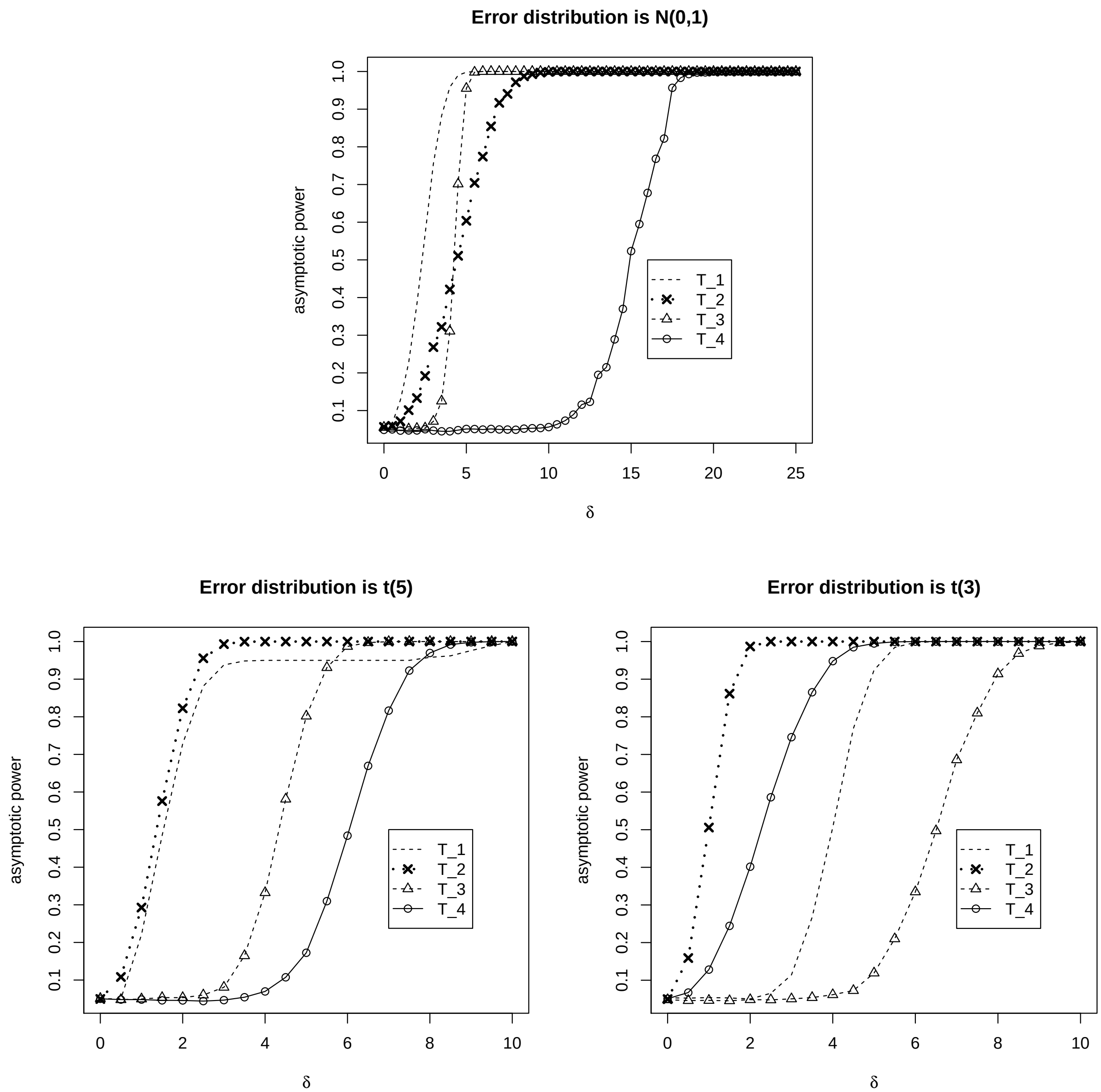

Figure 4 : The asymptotic power of different tests based on $T_{N,1}$, $T_{N,2}$, $T_{N,3}$ and $T_{N,4}$ at 5% level of significance for different values of the mixing proportion (i.e., δ). In the diagrams, $N(0,1)$, $t(3)$ and $t(5)$ denote the standard normal distribution, t-distribution with 3 degrees of freedom and t-distribution with 5 degrees of freedom.

Real Data Analysis :

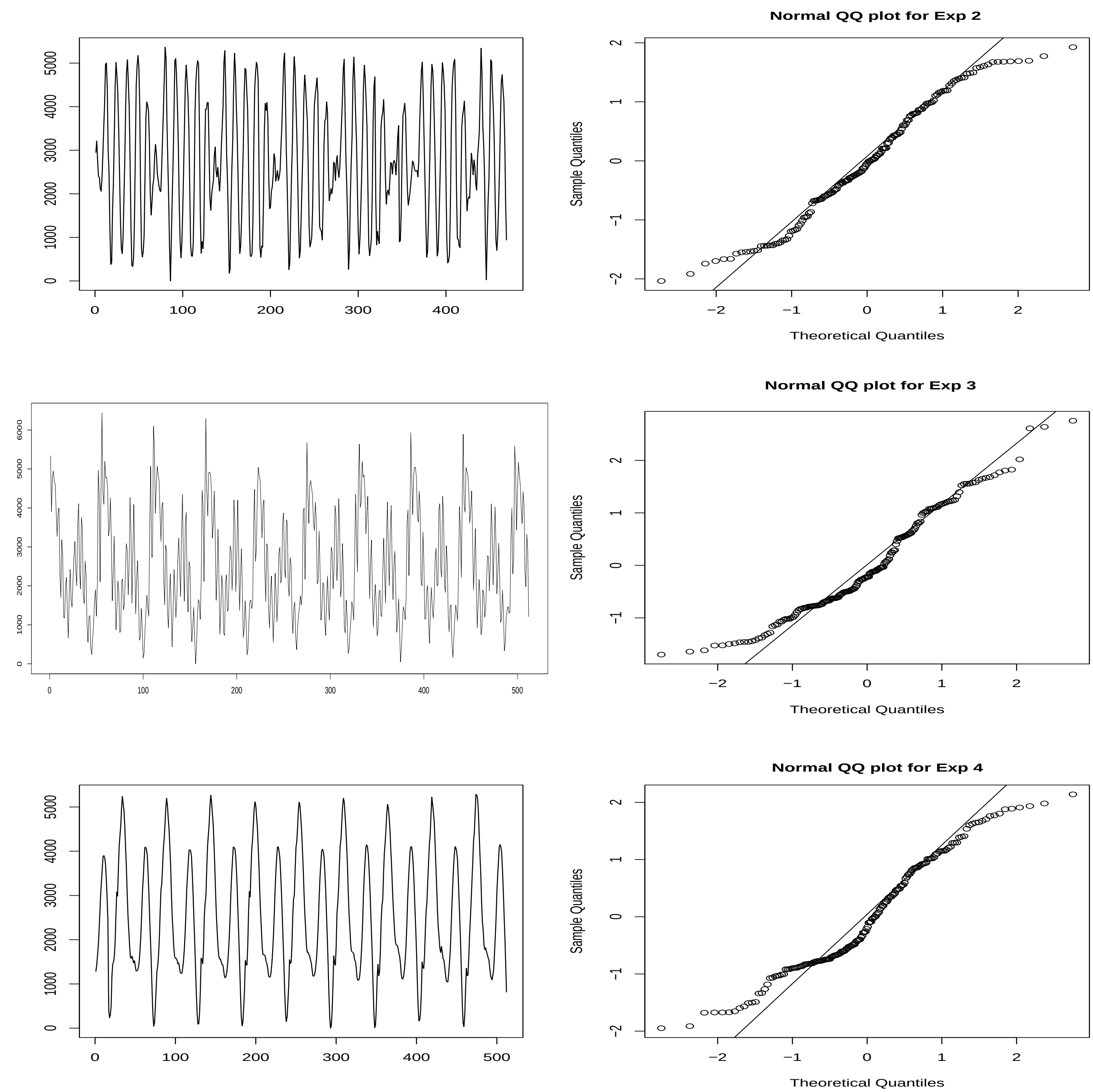

Figure 5 : First to third row: signals from Exp 2, Exp 3 and Exp 4 data. In each row, the observed value is plotted in the first diagram, and second diagram presents the quantile-quantile plot of the residuals of the corresponding data.

Here we briefly analyse three data, namely, Exp 2, Exp 3 and Exp 4. These data sets are collected from the lab in the Department of Electrical Engineering, IIT Kanpur, India and available to the journal's repository. Exp 2 data consists of 469 signals values drawn randomly at $10kHz$ frequency, and Exp 3 and Exp 4 data consist of 512 signals randomly drawn at the same frequency. It should be mentioned that Grover, Kundu and Mitra (2018) showed that these data set conform to chirp signal model, and for details, the readers may refer to Section V in that article. For each data, it is indicated from the q-q plots (in right side) in Figure 5, the residuals are obtained from light tailed distribution, and hence, the fixed parameters A_0, B_0, α_0 and β_0 are estimated by LSE methodology using the data $\mathcal{Y}_{1/3}$. The result of each data set is reported below one by one.

Exp 2 data : For this data, $A_0 = 0.1230$, $B_0 = 0.0217$, $\alpha_0 = 0.9909$ and $\beta_0 = 0.6961$, and the p values of the tests based on $T_{N,1}$, $T_{N,2}$, $T_{N,3}$ and $T_{N,4}$ are 0.659, 0.524, 0.666 and 0.525, respectively. The p-values of the tests based on $T_{N,1}$ and $T_{N,3}$ much favour the null hypothesis, and the p-values of the tests based on $T_{N,2}$ and $T_{N,4}$ are also indicating the validity of the null hypothesis. Overall, one can conclude that the feature of the data is not changing over time.

Exp 3 data : For this data, $A_0 = 0.02276$, $B_0 = 0.1727$, $\alpha_0 = 1.3312$ and $\beta_0 = 1.0270$, and the p values of the tests based on $T_{N,1}$, $T_{N,2}$, $T_{N,3}$ and $T_{N,4}$ are 0.931, 0.889, 0.934 and 0.891, respectively. The p-values of all tests, i.e., the tests based on $T_{N,1}$, $T_{N,2}$, $T_{N,3}$ and $T_{N,4}$ are indicating the validity of the null hypothesis. To be summerized, one can conclude that the feature of the data is not changing over time.

Exp 4 data : For this data, $A_0 = -0.0843$, $B_0 = 0.0966$, $\alpha_0 = 1.4507$ and $\beta_0 = 1.2469$, and the p values of the tests based on $T_{N,1}$, $T_{N,2}$, $T_{N,3}$ and $T_{N,4}$ are 0.541, 0.426, 0.549 and 0.427, respectively. The p-values of all tests, i.e., the tests based on $T_{N,1}$, $T_{N,2}$, $T_{N,3}$ and $T_{N,4}$ are reasonably large, and hence, the tests favour the null hypothesis. Overall, one can conclude that the feature of the data is not changing over time.

Reference : Grover, R., Kundu, D. and Mitra, A. (2018) On Approximate Least Squares Estimators of Parameters of One-Dimensional Chirp Signal. *Statistics*, 52, 1060–1085.