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:

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

(1)

Here $y_n$ is a real valued signal observed at $n = 1, \ldots, N$, $n$ denotes the time point, $A, B$ are amplitudes, $\alpha$ and $\beta$ 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 acceleration 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.
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 $H_0 : \theta = \theta_0$ against $H_1 : \theta \neq \theta_0$, where $\theta = (A, B, \alpha, \beta)$, and $\theta_0 = (A_0, B_0, \alpha_0, \beta_0)$ is
known to us. In this context, it should be mentioned that at the outset one needs to estimate the unknown
parameter $\theta$ 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 exten-
sive 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, $\alpha$ and $\beta$ 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
$\theta = (A, B, \alpha, \beta)$, and we now want to test $H_0 : \theta = \theta_0$
against $H_1 : \theta \neq \theta_0$, where $\theta_0 = (A_0, B_0, \alpha_0, \beta_0)$ 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 $\theta$ and $\theta_0$ are unequal. We next consider the data $Y = \{y_1, \ldots, y_N\}$, and based on $Y$, one may formulate the test statistics $T_{N,1} = \|D^{-1}(\hat{\theta}_{N,LSE} - \theta_0))\|^2_2$, $T_{N,2} = \|D^{-1}(\hat{\theta}_{N,LAD} - \theta_0))\|^2_2$, $T_{N,3} = \|D^{-1}(\hat{\theta}_{N,LSE} - \theta_0))\|^2_2$, and $T_{N,4} = \|D^{-1}(\hat{\theta}_{N,LAD} - \theta_0))\|^2_2$, where $D = \text{diag}\{N^{-1/2}, N^{-1/2}, N^{-3/2}, N^{-5/2}\}$, and $\| \cdot \|_2$ and $\| \cdot \|_1$ denote the usual Euclidean and $L_1$ distances, respectively. Here

\[
\hat{\theta}_{N,LSE} = \arg \min_{\theta} \sum_{n=1}^{N} \left\{ y_n - A\cos(\alpha_n + \beta_n^2) - B\sin(\alpha_n + \beta_n^2) \right\}
\]

and

\[
\hat{\theta}_{N,LAD} = \arg \min_{\theta} \sum_{n=1}^{N} \left\{ y_n - A\cos(\alpha_n + \beta_n^2) - B\sin(\alpha_n + \beta_n^2) \right\}
\]

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 $\theta$ equals with the specified $\theta_0$ or not. In order to address this problem, one needs to measure the discrepancy between $\theta$ and $\theta_0$; however, $\theta$ is unknown in the problem. The unknown $\theta$ can be estimated by the well-known least squares and least absolute deviations procedures based on a given data, which are denoted as $\hat{\theta}_{N,LSE}$ and $\hat{\theta}_{N,LAD}$, 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 $\theta$ (i.e., $\hat{\theta}_{N,LSE}$ or $\hat{\theta}_{N,LAD}$) and $\theta_0$. 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) $\theta = (A,B,\alpha,\beta)$ 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 $f_n(y_n, \theta)$, which is twice continuously differentiable with respect to $\theta$. It is assumed that $E\left[\frac{\partial^2}{\partial \theta^2} f_n(y_n, \theta)\right]^{2+\delta} < \infty$ for some $\delta > 0$ and $E\left[\frac{\partial^2}{\partial \theta^2} f_n(y_n, \theta)\right]^{2+\delta} < \infty$ for all $n = 1, \ldots, N$. Here $\theta_i$ and $\theta_j$ ($1 \leq i, j \leq 4$) are the $i$ and $j$-th components of $\theta$, 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 $\theta$ 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 $A = \{A_1, A_2, A_3, A_4\}$ be the four dimensional random vector associated with Gaussian distribution having the location parameter $= 0$ and the scatter matrix $= 2\sigma^2 \Sigma$ (denote it as $\Sigma_1$), where

\[
\Sigma = \begin{pmatrix}
\frac{A_1^2 + 9B_2^2}{2} & -4A_1B_0 & -4A_0B_0 & 18B_0 & -15B_0 \\
-4A_1B_0 & \frac{9A_1^2 + B_2^2}{2} & -18A_0 & -18A_0 & 96 \\
-4A_0B_0 & -18A_0 & 18B_0 & -18A_0 & 90 \\
18B_0 & -18A_0 & 90 & -90 & 90 \\
-15B_0 & 90 & 90 & 90 & -90
\end{pmatrix}
\]

Next, $B = \{B_1, B_2, B_3, B_4\}$ is another four dimensional random vector associated with Gaussian distribution having the location parameter $= 0$ and the scatter matrix $= \frac{1}{M}\Sigma$ (denote it as $\Sigma_2$), 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 $\sum_{i=1}^{4} \lambda_iZ_{i}^{2}$, where $\lambda_i$s are the eigen values of $\Sigma_1$ (defined above), and $Z_i$s are independent standard normal random variables. Under (A1)-(A3), the test based on $T_{N,1}$ will have asymptotic size $= \eta$ when $T_{N,1} \geq c_{1\eta}$. Moreover, under the same assumption, $P_{H_0}[T_{N,1} \geq c_{1\eta}] \to 1$ 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 $\sum_{i=1}^{4} \lambda^{*}_iZ_{i}^{2}$, where $\lambda^{*}_i$s are the eigen values of $\Sigma_2$ (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 \( T_{N, 2} \geq c_{2, 0} \). Moreover, under the same assumption, \( P_{H_1} [T_{N, 2} \geq c_{2, 0}] \rightarrow 1 \) as \( N \rightarrow \infty \), 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 \( \left\{ \frac{1}{\sum_{i=1}^{N} |A_i|} \right\}^2 \), 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, \( P_{H_1}[T_{N, 3} \geq c_{3, \eta}] \rightarrow 1 \) as \( N \rightarrow \infty \), 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 \( \left\{ \frac{1}{\sum_{i=1}^{N} |B_i|} \right\}^2 \), 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, \( P_{H_1}[T_{N, 4} \geq c_{4, \eta}] \rightarrow 1 \) as \( N \rightarrow \infty \), 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: \( \theta_0 \) and \( \theta_1 \), and we consider alternatives \( \theta^* = (1 - \psi)\theta_0 + \psi\theta_1 \) 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 \( \theta_0 = (1, 1, \pi/4, \pi/4) \) and \( \theta_1 = (4, 4, \pi/2, \pi/2) \) are chosen. In the course of investigation, a sample of size = 100 is drawn repeatedly 1000 times when \( \theta = \theta_0 \) (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 \( c_{0.05} \)). Afterwards, to estimate the power, we follow the same procedure when \( \theta = \theta^* \) (i.e., \( H_1 \) is true), and the proportion of times the value of the test statistic exceeding \( c_{0.05} \) is defined as the estimated power. The results are summarized in Figure 1 (given in supplementary file).

Note that at \( \psi = 0, \theta^* \) coincides with \( \theta_0 \), 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 motivated 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 \( P_n(\mathcal{A}_n) \rightarrow 0 \) implies that \( Q_n(\mathcal{A}_n) \rightarrow 0 \) for every sequence of measurable sets \( \mathcal{A}_n \), 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 \( \mathcal{A}_n \) is changing over \( n \) along with the corresponding \( \sigma \)-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 hypothesis parameter \( \theta_0 \). Suppose that we want to test \( H_0 : \theta = \theta_0 \) against \( H_{1, N} : \theta = \theta_N = \theta_0 + \delta_N \), where \( \delta_N = (\delta_0, \delta_1, \delta_2) \in \mathbb{R}^3 \). 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 \( \lambda_i \)'s are eigenvalues of \( \Sigma_2 \) (defined after Remark 2.1), and \( M = (M_1, M_2, M_3, M_4) \) is a 4-dimensional random vector associated with a Gaussian distribution having the location parameter \( \mu = (\mu_1, \mu_2, \mu_3, \mu_4) \) and the scatter parameter \( \Sigma \).

\[
\mu_1 = \left( \lim_{N \to \infty} - \frac{1}{N} \sum_{n=1}^{N} E_{F_n, u_n} \{Y_n - A_0 \cos(\alpha_0 n + \beta_0 n^2) - B_0 \sin(\alpha_0 n + \beta_0 n^2)\} \{\cos(\alpha_0 n + \beta_0 n^2)\} \right) \frac{\partial f(Y_n, \theta)}{\partial A} \{f(Y_n, \theta)\}^{-1}\theta_0 + \frac{\partial f(Y_n, \theta)}{\partial B} \{f(Y_n, \theta)\}^{-1} \theta_0.
\]

(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 \( \lambda_i^* \)'s are eigenvalues of \( \Sigma_2 \) (defined after Remark 2.1), and \( M^* = (M_1^*, M_2^*, M_3^*, M_4^*) \) is a 4-dimensional random vector associated with a Gaussian distribution having the location parameter \( \mu^* = (\mu_1^*, \mu_2^*, \mu_3^*, \mu_4^*) \) and the scatter parameter \( \Sigma^* \).

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 \( \sum_{i=1}^{4} | \mu_i^* | \), where \( \mu^* = \{A_1^*, A_2^*, A_3^*, A_4^*\} \) is a four dimensional random vector associated with a Gaussian distribution with the location parameter \( \mu = \mu^* \) and the scatter matrix \( \Sigma \).

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 \( \sum_{i=1}^{4} | B_i^* | \), where \( B^* = \{B_1^*, B_2^*, B_3^*, B_4^*\} \) is a four dimensional random vector associated with a Gaussian distribution with the location parameter \( \mu = \mu^* \) and the scatter matrix \( \Sigma \).

The assertions in Theorems 4.1–4.4 enable us to compute the asymptotic power of the tests under contiguity.
ous alternatives $H_{1,N}$. To carry out the tests, one first needs to compute $c_{1i}$, $c_{2i}$, $c_{3i}$ and $c_{4i}$ described in Theorems 2.1, 2.2, 2.3 and 2.4, respectively. For sake of concise presentation, only the computation of $c_{1i}$ is explained here. To compute the critical value $c_{1i}$, we generate the data with sample size $= 1000$ (denoted as $\{x_{1i}, \ldots , x_{1000}\}$) from the Gaussian distribution associated with the random vector 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 $c_{1i}$. The computation of $c_{2i}$, $c_{3i}$ and $c_{4i}$ are done by the similar procedure as described for estimating $c_{1i}$. Next, to compute the asymptotic power of the test based on $T_{N,1}$, a data with sample size $= 1000$ (denoted as $Y = \{y_1, \ldots , y_{1000}\}$) is obtained from the Gaussian distribution associated with the random vector M defined in Theorem 4.1, and the estimated asymptotic power of the test based on $T_{N,1}$ will be $\{1 - F_Y(c_{1i})\}$, where $F_Y$ is the empirical distribution function of $\sum_{i=1}^{4} \lambda_i M_i^2$ (defined in Theorem 4.1) based on the data $Y$. 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 $\delta$, where $\delta = \delta_1 = \delta_2 = \delta_3 = \delta_4$. The equality among $\delta_1$, $\delta_2$, $\delta_3$ and $\delta_4$ is considered because of simple pictorial representation. In this study, the error distribution is assumed to be standard normal distribution and 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 $Y_{1/3}$, and $Y_{2/3}$ denotes the remaining two-third of data). We apply absolute error and squared error based methods on $Y_{1/3}$ and calculate the estimates of $(A, B, \alpha, \beta)$, which is fixed as $(A_0, B_0, \alpha_0, \beta_0)$. Afterwards, based on the remaining data, i.e., $Y_{2/3}$, we want to test whether the unknown parameter $(A, B, \alpha, \beta)$ is same as $(A_0, B_0, \alpha_0, \beta_0)$ 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, $Y_{1/3}$ can be though as the training data whereas $Y_{2/3}$ 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 $Y_{1/3}$.

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 $Y = \{y_1, \ldots , y_n\}$, 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
TABLE I

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

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


VII. APPENDIX: PROOFS

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

**Lemma 1:** Under \( H_0 \) and (A1)-(A3), \( T_{N,1} \) converges weakly to \( \sum_{i=1}^{4} \lambda_i Z_i^2 \), where \( \lambda_i \)s are the eigenvalues of \( \Sigma_1 \), and \( Z_i \)s are independent standard normal random variables. Here, \( \sum_{i=1}^{4} \) is same as defined after Remark 2.1.

**Proof of Lemma 1:** Under \( H_0 \) and assumptions (A1)-(A3), it follows from [15] that \( D^{-1}(\hat{\theta}_{N,LSE}-\theta_0) \) converges weakly to 4-dimensional random vector \( \bf{R} \) 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 \( 4 \times 4 \) orthogonal matrix (i.e., \( A' A = A A' = I_4 \)) such that \( A \Sigma A' = \text{Diag} \{ \lambda_1, \lambda_2, \lambda_3, \lambda_4 \} \). Here \( A' \) denotes the transpose of the matrix \( A \). The aforementioned fact implies that \( AD^{-1}(\hat{\theta}_N-\theta_0) \) 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 \( D \) since \( \Sigma \) is the same as defined after Remark 2.1.

**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 \( \eta \) 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 \( \sum_{i=1}^{4} \lambda_i Z_i^2 \) that described in the statement of this theorem.

We now consider

\[
P_{H_1}[T_{N,1} > c_{1\eta}]
= P_{H_1}[||D^{-1}(\hat{\theta}_{N,LSE}-\theta_0)||^2 > c_{1\eta}]
= P_{H_1}[||D^{-1}(\hat{\theta}_{N,LSE}-\theta_0) - D^{-1}(\hat{\theta}_1-\theta_0)||^2 > c_{1\eta} + ||D^{-1}(\hat{\theta}_1-\theta_0)||^2 - 2 < D^{-1}(\hat{\theta}_{N,LSE}-\theta_0), D^{-1}(\hat{\theta}_1-\theta_0) >]
\]

since under \( H_1 \), \( \theta = \theta_1 \), and \(<, , > \) denotes the inner product.

\[
P_{H_1}[||D^{-1}(\hat{\theta}_{N,LSE}-\theta_1)||^2 > c_{1\eta} + ||D^{-1}(\hat{\theta}_1-\theta_0)||^2 - 2 < D^{-1}(\hat{\theta}_{N,LSE}-\theta_0), D^{-1}(\hat{\theta}_1-\theta_0) >]
= P_{H_1}[T_{N,1} > c_{1\eta} + ||D^{-1}(\theta_1-\theta_0)||^2 - 2 < D^{-1}(\hat{\theta}_{N}-\theta_0), D^{-1}(\hat{\theta}_1-\theta_0) >]
\]

since under \( H_1 \), \( T_{N,1} = ||D^{-1}(\theta_{N,LSE}-\theta_1)||^2 \),

\[
P_{H_1}[T_{N,1} > c_{1\eta} + ||D^{-1}(\theta_1-\theta_0)||^2 - 2 < D^{-1}(\hat{\theta}_{N}-\theta_0), D^{-1}(\hat{\theta}_1-\theta_0) >
+ 2 < D^{-1}(\theta_{N}-\theta_1), D^{-1}(\theta_1-\theta_0) >]
\]

\[
P_{H_1}[T_{N,1} > c_{1\eta} + ||D^{-1}(\theta_1-\theta_0)||^2 - 2||D^{-1}(\theta_1-\theta_0)||^2 - 2 < D^{-1}(\hat{\theta}_{N,LSE}-\theta_1), D^{-1}(\theta_1-\theta_0) >]
= P_{H_1}[T_{N,1} > c_{1\eta} - ||D^{-1}(\theta_1-\theta_0)||^2 + 2 < D^{-1}(\hat{\theta}_{N,LSE}-\theta_1), D^{-1}(\theta_1-\theta_0) >]
\]

\[
\to 1 \text{ as } N \to \infty.
\]

The last step follows from the facts that \( ||D^{-1}(\theta_1-\theta_0)||^2 \to \infty \) as \( N \to \infty \) and \( ||D^{-1}(\hat{\theta}_{N,LSE}-\theta_1)|| \to 0 \) as \( N \to \infty \) in probability under \( H_1 \) and (A1)-(A3). Hence, it completes the proof. \( \square \)

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 \( \sum_{i=1}^{4} \lambda_i^* Z_i^* \), where \( \lambda_i^* \)s are the eigenvalues of \( \Sigma_2 \), and \( Z_i^* \)s are independent standard normal random variables. Here, \( \Sigma_2 \) is same as defined after Remark 2.1.

**Proof of Lemma 2:** Under \( H_0 \) and assumptions (A1)-(A2), [3] showed that \( D^{-1}(\hat{\theta}_{N,LAD}-\theta_0) \) converges weakly to 4-dimensional random vector associated with a Gaussian distribution with the location parameter \( (0,0,0,0) \) and the scatter matrix \( \Sigma_2 \). Define a new random vector \( L^* = A^* R^* \), where \( A^* \) is a \( 4 \times 4 \) orthogonal matrix (i.e., \( A^* A^* = A^* A^* = I_4 \)) such that \( A^* \Sigma_2 A^* = \text{Diag} \{ \lambda_1^*, \lambda_2^*, \lambda_3^*, \lambda_4^* \} \). Here \( A^* \) denotes the transpose of the matrix \( A^* \). The aforementioned fact implies that \( A^* D^{-1}(\theta_{N,LAD}-\theta_0) \) 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 \(= \text{Diag}\{\lambda_1^2, \lambda_2^2, \lambda_3^2, \lambda_4^2\}\) since \(A^* \Sigma_2 A^{T} = \text{Diag}\{\lambda_1^*, \lambda_2^*, \lambda_3^*, \lambda_4^*\}\). Therefore \(\{A^* D^{-1}(\theta_{N,LAD} - \theta_0)\}^T \{A^* D^{-1}(\theta_{N,LAD} - \theta_0)\} = \{D^{-1}(\theta_{N,LAD} - \theta_0)\}^T A^T A \{D^{-1}(\theta_{N,LAD} - \theta_0)\} = ||D^{-1}(\theta_{N,LAD} - \theta_0)||^2 := T_{N,2}\) converges weakly to \(\sum_{i=1}^{4} \lambda_i^* Z_i^2\), where \(\lambda_i^*\)’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 \(P_{H_1}[T_{N,2} > c_{2n}]\) → 1 as \(N \to \infty\), where \(c_{2n}\) is the critical value of the level of significance \(\eta\) to test \(H_0\) against \(H_1\), which follows from Lemma 2. In other words, \(c_{1n}\) 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

\[
P_{H_1}[T_{N,2} > c_{2n}] = P_{H_1}[||D^{-1}(\theta_{N,LAD} - \theta_0)||^2 > c_{2n}] = P_{H_1}[||D^{-1}(\theta_{N,LAD} - \theta_0) - D^{-1}(\theta_1 - \theta_0)||^2 > c_{2n} + ||D^{-1}(\theta_1 - \theta_0)||^2 - 2 < D^{-1}(\theta_{N,LAD} - \theta_0), D^{-1}(\theta_1 - \theta_0) >]  
\]

since under \(H_1\), \(\theta = \theta_1\), and \(<, >\) denotes the inner product.

\[
P_{H_1}[T_{N,2} > c_{2n}] = P_{H_1}[||D^{-1}(\theta_{N,LAD} - \theta_0)||^2 - 2 < D^{-1}(\theta_{N,LAD} - \theta_0), D^{-1}(\theta_1 - \theta_0) >]  
\]

since under \(H_1\), \(T_{N,2} = ||D^{-1}(\theta_{N,LAD} - \theta_1)||^2\).

\[
P_{H_1}[T_{N,2} > c_{2n}] = P_{H_1}[||D^{-1}(\theta_1 - \theta_0)||^2 - 2 < D^{-1}(\theta_1 - \theta_0), D^{-1}(\theta_1 - \theta_0) >]  
\]

→ 1 as \(N \to \infty\).

The last step follows from the facts that \(||D^{-1}(\theta_1 - \theta_0)||^2 \to \infty\) as \(N \to \infty\) and \(||D^{-1}(\theta_{N,LAD} - \theta_1)||^2 \to 0\) 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 \(A_i\)’s 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 \(D^{-1}(\theta_{N,LSE} - \theta_0)\) 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 \(||D^{-1}(\theta_{N,LSE} - \theta_0)||^2 := T_{N,3}\) 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 \(P_{H_0}[T_{N,3} > c_{3n}]\) → 1 as \(N \to \infty\), where \(c_{3n}\) is the critical value of the level of significance \(\eta\) to test \(H_0\) against \(H_1\), which follows from Lemma 3. In other words, \(c_{3n}\) 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
\[
P_{H_1}[T_{N,3} > c_{3\eta}] = P_{H_1}[\|D^{-1}(\hat{\theta}_{N,LSE} - \theta_0)\|_1 > c_{3\eta}]
= P_{H_1}[\|D^{-1}(\hat{\theta}_{N,LSE} - \theta_0)\|_1 > k] \quad (k = \sqrt{c_{3\eta}} \text{ is an arbitrary constant})
= P_{H_1}[\|D^{-1}(\hat{\theta}_{N,LSE} - \theta_1 + \theta_1 - \theta_0)\|_1 > k]
= P_{H_1}[\|D^{-1}(\hat{\theta}_{N,LSE} - \theta_1) - (\theta_0 - \theta_1)\|_1 > k]
\geq P_{H_1}[\|D^{-1}(\hat{\theta}_{N,LSE} - \theta_1)\|_1 > k - D^{-1}\|\theta_0 - \theta_1\|_1] \quad (\text{since } |a - b| \geq |a| - |b|)
\rightarrow 1 \text{ as } N \rightarrow \infty.
\]

The last step follows from the fact that \(\|D^{-1}(\hat{\theta}_{N,LSE} - \theta_1)\|_1\) converges weakly to \(\sum_{i=1}^{4} |A_i|\), i.e., by Prohorov’s theorem (see, e.g., [17]), \(\|D^{-1}(\hat{\theta}_{N,LSE} - \theta_1)\|_1 = O_p(1)\), and \(D^{-1}\|\theta_0 - \theta_1\|_1 \rightarrow \infty\) as \(N \rightarrow \infty\). Hence, it completes the proof. \(\square\)

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 \(B_i\)s are same as defined after Remark 2.1.

**Proof of Lemma 4:** As it is established in Lemma 2, we have that \(D^{-1}(\hat{\theta}_{N,LAD} - \theta_0)\) 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 \((0, 0, 0, 0)^T\) 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 \(\|D^{-1}(\hat{\theta}_{N,LAD} - \theta_0)\|_1^2 = T_{N,4}\) converges weakly to \(\left\{\sum_{i=1}^{4} |B_i|\right\}^2\), and hence, the proof is complete. \(\square\)

**Proof of Theorem 2.4:** To prove this theorem, one needs to show that \(P_{H_1}[T_{N,4} > c_{4\eta}] \rightarrow 1\) as \(N \rightarrow \infty\), where \(c_{4\eta}\) is the critical value of the level of significance \(\eta\) 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

\[
P_{H_1}[T_{N,4} > c_{4\eta}] = P_{H_1}[\|D^{-1}(\hat{\theta}_{N,LAD} - \theta_0)\|_1^2 > c_{4\eta}]
= P_{H_1}[\|D^{-1}(\hat{\theta}_{N,LAD} - \theta_0)\|_1 > k^*] \quad (k^* = \sqrt{c_{4\eta}} \text{ is an arbitrary constant})
= P_{H_1}[\|D^{-1}(\hat{\theta}_{N,LAD} - \theta_1 + \theta_1 - \theta_0)\|_1 > k^*]
= P_{H_1}[\|D^{-1}(\hat{\theta}_{N,LAD} - \theta_1) - (\theta_0 - \theta_1)\|_1 > k^*]
\geq P_{H_1}[\|D^{-1}(\hat{\theta}_{N,LAD} - \theta_1)\|_1 > k^* - D^{-1}\|\theta_0 - \theta_1\|_1] \quad (\text{since } |a - b| \geq |a| - |b|)
\rightarrow 1 \text{ as } N \rightarrow \infty.
\]

The last step follows from the fact that \(\|D^{-1}(\hat{\theta}_{N,LAD} - \theta_1)\|_1\) converges weakly to \(\sum_{i=1}^{4} |B_i|\), i.e., \(\|D^{-1}(\hat{\theta}_{N,LAD} - \theta_1)\|_1 = O_p(1)\), and \(D^{-1}\|\theta_0 - \theta_1\|_1 \rightarrow \infty\) as \(N \rightarrow \infty\). Hence, the proof is complete. \(\square\)

**Proof of Theorem 4.1:** In order to establish the contiguity of the sequence \(F_n(y_n, \theta_N)\) (or \(f_n(y_n, \theta_N)\)) relative to \(F_n(y_n, \theta_0)\) (or \(f_n(y_n, \theta_0)\)), 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 \(\sigma^2\) (see [20] (p. 253, Corollary 2 to Lecam’s first Lemma)), where \(\sigma\) is a positive constant. We now have
\[
L_N = \log \prod_{n=1}^{N} \frac{f_n(y_n, \theta_N)}{f_n(y_n, \theta_0)} = \sum_{n=1}^{N} \log \frac{f_n(y_n, \theta_0 + \delta_N)}{f_n(y_n, \theta_0)} \text{ since } f_n \text{s are independent.}
\]

\[
= \sum_{n=1}^{N} \{ \log f_n(y_n, \theta_0 + \delta_N) - \log f_n(y_n, \theta_0) \}
\]

\[
= \sum_{n=1}^{N} \{ g_n(y_n, \theta_0 + \delta_N) - g_n(y_n, \theta_0) \} \text{ (here we denote } g_n(.) = \log f_n(.) \)
\]

\[
= \sum_{n=1}^{N} \{ g_n(y_n, \theta_0) + < \delta_N, \nabla g_n(y_n, \theta_0) > + \frac{1}{2} < \delta_N, H_{\theta}(g_n(y_n, \theta_*)) \delta_N > - g_n(y_n, \theta_0) \}
\]

\[
= \sum_{n=1}^{N} \left\{ < \delta_N, \nabla g_n(y_n, \theta_0) > + \frac{1}{2} < \delta_N, H_{\theta}(g_n(y_n, \theta_*)) \delta_N > \right\},
\]

where \( \nabla g_n(y_n, \theta_0) = \left( \frac{\partial g_n(y_n, \theta)}{\partial \theta_1}, \frac{\partial g_n(y_n, \theta)}{\partial \theta_2}, \frac{\partial g_n(y_n, \theta)}{\partial \theta_3}, \frac{\partial g_n(y_n, \theta)}{\partial \theta_4} \right) \) \( i, j \leq 4 \), \( \theta_1, \theta_2, \theta_3, \theta_4 = (A, B, \alpha, \beta) \), \( \theta_* \) is lying on the straight line connecting \( \theta_0 \) and \( \theta_* \). The expression of \( L_N \) follows from the Taylor series expansion of \( g_n(y_n, \theta_0 + \delta_N) \) up to the second term. Further, note that both of \( < \delta_N, \nabla g_n(y_n, \theta_0) > \) and \( < \delta_N, H_{\theta}(g_n(y_n, \theta_*)) \delta_N > \) are sequence of independent random variables, and \( \theta_* \rightarrow \theta_0 \) as \( N \rightarrow \infty \). 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 \( < \delta_N, \nabla g_n(y_n, \theta_0) > \) converges weakly to a normal random variable with mean zero and variance \( = E_{F_{n, \eta_{n}}}(\nabla g_n(y_n, \theta_0))_1 \) and \( \frac{1}{2} < \delta_N, H_{\theta}(g_n(y_n, \theta_*)) \delta_N > \) converges in probability to \( \frac{1}{2} E_{F_{n, \eta_{n}}} < 1, H_{\theta}(g_n(y_n, \theta_*)) > \) where \( 1 = (1,1,0,0) \). Further, we have \( E_{F_{n, \eta_{n}}} < 1, H_{\theta}(g_n(y_n, \theta_*)) > = - E_{F_{n, \eta_{n}}}(\nabla g_n(y_n, \theta_0))_1 \). Hence, the sequence \( F_n(y_n, \theta_N) \) (or \( f_n(y_n, \theta_N) \)) is contiguous relative to \( F_n(y_n, \theta_0) \) (or \( f_n(y_n, \theta_0) \)).

Next, under \( H_0 \) and assumptions (A1)-(A3), it follows from [15] that \( D^{-1}(\theta_{N,LAD} - \theta_0) \) 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 \( D^{-1}(\theta_{N,LSE} - \theta_0) \) under \( H_0 \), we have that the joint distribution of \( (D^{-1}(\theta_{N,LSE} - \theta_0), L_N) \) 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 \( D^{-1}(\theta_{N,LAD} - \theta_0) \) under 

---

**Proof of Theorem 4.2**: It follows from [3] that under \( H_0 \) and assumptions (A1)-(A3), \( D^{-1}(\theta_{N,LAD} - \theta_0) \) 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 \( D^{-1}(\theta_{N,LAD} - \theta_0) \) under \( H_0 \), we have that the joint distribution of \( (D^{-1}(\theta_{N,LAD} - \theta_0), L_N) \) 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 \( D^{-1}(\theta_{N,LAD} - \theta_0) \) under
Here also, note that under $H_{1,N}$, $D^{-1}(\hat{\theta}_{N,LAD} - \theta_0)$ 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 $D^{-1}(\theta_{N,LAD} - \theta_0)$ and $L_N$. [3] derived that $D^{-1}(\hat{\theta}_{N,LAD} - \theta_0) = -\Sigma_2^{-1}D\{\nabla Q'(\theta_0)\} + R_N$, where $R_N \xrightarrow{p} 0$, and $\Sigma_2$ is same as defined after Remark 2.1. The expressions of $D\{\nabla Q'(\theta_0)\}$ are provided in [3], where $Q'(\theta_0) = \sum_{n=1}^{N} |y_n - A_0 \cos(\alpha_0 n + \beta_0 n^2) - B_0 \sin(\alpha_0 n + \beta_0 n^2)|$. Based on these expressions, the component wise asymptotic covariance between $D^{-1}(\hat{\theta}_{N,LAD} - \theta_0)$ and $L_N$ as given in the expression of $\mu^*$ 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, $D^{-1}(\theta_{N,LSE} - \theta_0)$ converges weakly to a random vector $A^*$, 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, $D^{-1}(\hat{\theta}_{N,LAD} - \theta_0)$ converges weakly to a random vector $B^*$, which is defined in the statement of this theorem. Afterwards, a direct application of continuous mapping theorem leads to the proof.
Finite sample power for $t(3)-$distribution with $N=100$  

Finite sample power for $t(5)-$distribution with $N=100$  

Finite sample power for $N(0,1)-$distribution with $N=100$  

Finite sample power for $N(0,1)-$distribution with $N=100$  

Finite sample power for $DE(0,1)$ distribution with $N=20$  

Finite sample power for $DE(0,1)$ distribution with $N=12$  

Finite sample power for $DE(0,1)$ distribution with $N=6$  

Finite sample power for $t(3)-$distribution with $N=50$  

Finite sample power for $t(3)-$distribution with $N=35$  

Finite sample power for $t(3)-$distribution with $N=30$  

Finite sample power for $t(3)-$distribution with $N=20$  

Finite sample power for $DE(0,1)$ distribution with $N=6$  

Finite sample power for $DE(0,1)$ distribution with $N=12$  

Finite sample power for $DE(0,1)$ distribution with $N=20$  

Finite sample power for $N(0,1)-$distribution with $N=50$  

Finite sample power for $N(0,1)-$distribution with $N=30$  

Finite sample power for $N(0,1)-$distribution with $N=100$  

Finite sample power for $N(0,1)-$distribution with $N=100$  

Normal, $DE$, $t(5)$ 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 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.
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., $\delta$). 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 $A_0$, $B_0$, $\alpha_0$ and $\beta_0$ are estimated by LSE methodology using the data $Y_{1/3}$. 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 summarized, 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.