Quantile Process with Applications : Chirp Signal Model

Subhra Sankar Dhar

IIT Kanpur
Department of Mathematics & Statistics
Kanpur 208016, India
email: subhra@iitk.ac.in

November 24, 2021

Abstract

The parameters in the well-known chirp signal model controls the frequency fluctuations of the signals, and consequently, the estimation of the parameters has received considerable attention in the literature of statistical signal processing. In the same spirit with a broader view, this article investigates the quantile estimator of parameters involved in the chirp signal model, which enables us to provide basic features of the entire distribution of the signals. In the course of this study, we establish the limiting behaviour of the associated stochastic process, which we call quantile process. As the applications of this result, we obtain the limiting distributions of various quantile based measures of descriptive statistics, which give us summarized features of the fluctuations of the signals in various senses. Finally, along with extensive simulation study, the practicability of the proposed methodology is shown on a few benchmark real datasets closely related with various chirp signal models.

Keywords : Bernstein inequality, central limit theorem; Gaussian process; process convergence; sonar; statistical signal processing
1 Introduction

Very often the scientists in various interdisciplinary and engineering sciences experience a situation, where the frequency of a signal is time dependent, i.e., the frequency of the signal varies over time, and such signal is known as chirp signal in the literature of signal processing. Exemplary we mention that the chirp signals are visible in radar, laser, active and passive sonar systems etc (see, e.g., Bender et al. (2006), Nakahira et al. (2005), Caloz et al. (2013) and a few relevant references therein). A real valued one-dimensional chirp signal model in additive noise can be written as follows:

\[ Y_n = A \cos(\alpha n + \beta n^2) + B \sin(\alpha n + \beta n^2) + X_n; \quad n = 1, \ldots, N, \quad (1.1) \]

where \( Y_n \) is a real valued signal observed at \( n = 1, \ldots, N \), \( n \) denotes the time point, \( A \) and \( B \) are amplitudes, \( \alpha \) and \( \beta \) are frequency and frequency rate, respectively. Here \( X_n \) is a sequence of independent and identically distributed (i.i.d.) random variables with certain properties. To know about this model, the readers may refer to Dhar et al. (2019) and a few references in that article.

In many applications in signal processing, the frequency of signals can be modelled by (1.1). For instance, sonar (i.e., Sound Navigation And Ranging) signals very often are treated as chirp signals, and as an illustration, one may consider the following example. The frequency generated by different materials are different on or under the surface of the water. For example, very often a specific material absorbs the energy of some frequencies faster, and consequently, emit less information of return. In order to model such phenomena, a wide range of survey signals covering a band of frequencies can be fitted by the chirp signal model described in (1.1). Other than sonar signals, there are many examples in various sciences, where the signals can be fitted by model (1.1). For example, the receiving antenna of an aircraft receives Radar (i.e., Radio Detection And Ranging) signals from a target, and at the same time, the transmitting antenna transmits Radar signals to the target. As the frequencies of the Radar received by the receiving antenna and the transmitted by transmitting antenna vary over time, the respective Radar signals can be well fitted by the model (1.1). For such type of signals, the parameters \( A, B, \alpha \) and \( \beta \) involved in the model (1.1) can be estimated by the least squares or the least absolute deviation methodology, and one may work on various statistical methodologies using those point estimators (see, e.g., Wu et al. (2008), Yetik and Nehorai (2003), Nandi and Kundu (2004)).
and a few references therein). However, those estimators only give us an insight about the mean or the median surface of the response variable given the covariates; not the whole feature of the surface of the response variable conditioning on the covariates. The concept of the quantiles (see, e.g., Serfling (1980), p. 74 and Koenker (2005)) was introduced in the literature to capture the whole distributional feature of a random variable.

As an alternative approach to the least squares estimation, Koenker and Bassett (1978) proposed the idea of quantile regression in the multiple linear regression model, which provides us the quantile surface of the response variable conditioning on the covariates. The concept of the quantile regression since then has been amply available in the literature of Statistics, Economics and other interdisciplinary sciences. The idea of quantile regression is easy to interpret and because of its good robustness property, it has become a powerful toolkit in interpreting data. Moreover, like the mean and the median, the quantile can also be obtained from a certain minimization problem, where the check function is a convex function, and this fact enables researchers to extend the concepts of quantiles in many topics in Statistics. Specifically, using the similar idea of the regression quantiles, the next paragraph will describe the concept of quantiles for the chirp signal model described in (1.1). For details about quantile regression, the readers may refer to Koenker (2005).

As mentioned earlier, Statistical signal processing literature has paid considerable attention on the point estimation of $A$, $B$, $\alpha$ and $\beta$ in model (1.1) but to the best of our knowledge, the quantiles of $A$, $B$, $\alpha$ and $\beta$ or in other words, the quantile surface of the response variable $Y_n$ in model (1.1) has not yet been studied in the statistical signal processing literature. This article explores the quantiles of the unknown parameters of model (1.1), which are obtained from a certain minimization problem and studies a number of key results related to quantiles with its application. Apart from the theoretical advantage of studying quantile estimates, it has a numerous direct advantage in applications also. For instance, as described earlier, the receiving and the transmitting antenna of an aircraft receive the Radar, and the frequencies of the Radar is likely to follow the model (1.1); see also Gini et al. (2000). The least squares or the least absolute deviation estimates of $A$, $B$, $\alpha$ and $\beta$ only provides us the idea of the central tendency of the signals, and the central tendency of the signals may be highly influenced by the extreme signals occurring due to turbulence. However, even under such extreme situations, various quantile estimates of $A$, $B$, $\alpha$ and $\beta$ enable us to see the effect of extreme signals on the whole distributional feature of the signals. Besides, one can also propose various robust measure like the trimmed mean (see, e.g., Dhar and Chaudhuri (2012) and a few references therein) or inter quartile range
(see, e.g., Warr and Erich (2013) and a few reference therein) based on quantiles to interpret the various properties of the observed signals.

The main contribution of this article is three fold. The first fold is introducing the concept of quantiles in the well-known chirp signal model described in (1.1). In the literature of statistical signal processing, there were a few attempts on $L_1$ or least absolute deviation estimates of $A$, $B$, $\alpha$ and $\beta$; however, to the best of our knowledge, the statistical signal processing literature has not studied the concept quantile estimators of the parameters in chirp signal model such as (1.1). This article introduces the idea of quantile estimates of $A$, $B$, $\alpha$ and $\beta$ in model (1.1). The second fold is the development of results on the convergence of the associated sample quantile process to a certain Gaussian process. The convergence of the quantile process associated with the estimators of the unknown parameters involved in the model have been studied in the literature (see, e.g., Jureckova et al. (2020), Volgushev (2020), Zwingmann and Holzmann (2020), Volgushev et al. (2019), Parker (2019), Hsieh and Wang (2018), Yuan et al. (2017), Tse (2005), Tse (2009), Qu and Yoon (2015), Wagener et al. (2012) and many more articles); however, any kind of process convergence has not been received any attention in the statistical signal processing literature; particularly, for the parameters involved in the chirp signal model like (1.1). Along with proposing the concepts of quantiles in the chirp signal model (1.1), this article thoroughly investigates the various properties of the sample quantile process. Exemplary, the result on the weak convergence of the quantile process to a certain Gaussian process implies the uniform convergence of the quantile estimates to their population counterpart, i.e., corresponding unknown parameters, where uniformity is respect to the quantile index. It is needless to mention that this fact has numerous statistical implications, which will be elaborated in the subsequent sections. The third fold of the contribution is related to the applications of the quantile estimators of $A$, $B$, $\alpha$ and $\beta$ in model (1.1). We here propose the quantile based measures of the location, the scatter and the skewness of the observed signals, and using the result on convergence of the sample quantile process, the asymptotic properties of the sample version of the proposed quantile based measure are derived. The performance of those measures are also investigated on various simulated and real data as well along with a thorough study to see how good the proposed measures compared to well-known classical measures associated with classical descriptive Statistics.

In the course of exploring the aforesaid points, a number of mathematical challenges arise in deriving the results. The first challenge is associated with the non-differentiabiility of the criterion function of the quantile, and to overcome this problem, in the neig-
bourhood of the point of non-differentiability of the original criterion function, a three
degree polynomial is considered, which approximates the original criterion function, i.e.,
a smoothed approximation of the original criterion function by an approximated criterion
function. Afterwards, as the smoothed approximated criterion function is infinitely times
differentiable, the asymptotic analysis is carried out on that approximated criterion func-
tion. Finally, the result is obtained since the approximated criterion function converges
to the original criterion function uniformly. To be summarized, the proof of Theorem 2.1
deals with substantial Mathematical complexity related to Analysis and Probability the-
ory. The next level mathematical difficulty is involved in deriving the limiting stochastic
process of the sample quantile process of $A, B, \alpha$ and $\beta$ indexed by quantile index, which
has not been studied in the statistical signal processing literature to the best of our knowl-
edge. In order to establish the process convergence of the sample quantile process, first
we prove the arbitrary finite dimensional weak convergence, which is itself a new result in
the literature. Afterwards, we establish the complicated tightness property of the sample
quantile process, and the process convergence follows from the tightness property of the
sample quantile process along with the weak convergence of the arbitrary finite dimensional
distributions of the sample quantiles.

The rest of the article is organized as follows. Section 2 describes the model (1.1) and
proposes the quantile estimator. The properties of the estimator are discussed in Section
2.1 and the results of the quantile estimator and the associated processes are also stated in
that section. Section 2.2 investigates various applications of the quantile based estimator.
Two well-known data sets are analysed in Section 3 and Section 4 studies the performance
of the quantile based estimators when the number of signals is small or moderately small. A
few concluding remarks are discussed in Section 5. Finally, Section 6 contains all technical
details.

2 Preliminaries and model

Recall model (1.1), i.e.,

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

where $Y_n$ is a real valued signal at $n = 1, \ldots, N$, $A$ and $B$ are real valued amplitudes, $\alpha$ and $\beta$ are frequency and frequency rate, respectively, and $X_n$ is a sequence of i.i.d. error
random variables. A few other technical conditions will be stated before the statements of the results in the subsequent sections. For notational convenience, let us denote $\theta = [A, B, \alpha, \beta]$, and for any $\gamma \in (0, 1)$, $\theta_\gamma = [A_\gamma, B_\gamma, \alpha_\gamma, \beta_\gamma]$, where $A_\gamma$, $B_\gamma$, $\alpha_\gamma$ and $\beta_\gamma$ are $\gamma$-th quantile of $A$, $B$, $\alpha$ and $\beta$, respectively, which is defined as follows. Here for the sake of technicalities, we are assuming that the parameter space of $\theta$ is $\Theta = [-L, L] \times [-L, L] \times [0, \pi] \times [0, \pi]$, where $L$ is an arbitrary large constant.

Let us now denote

$$Q(\theta) = E[|X_n| + (2\gamma - 1)(X_n)]|X_{n-1}, \ldots, X_1],$$  \hspace{1cm} (2.2)$$

where $X_n = Y_n - A \cos(\alpha n + \beta n^2) - B \sin(\alpha n + \beta n^2)$ (follows from (1.1)), and note that

$$\theta_\gamma = \arg \min_{\theta \in \Theta} Q(\theta),$$

where $\theta = [A, B, \alpha, \beta]$. Next, suppose that

$$Q_N(\theta) = \frac{1}{N} \sum_{n=1}^{N} \{|x_n| + (2\gamma - 1)x_n\},$$  \hspace{1cm} (2.3)$$

where $x_n = y_n - A \cos(\alpha n + \beta n^2) - B \sin(\alpha n + \beta n^2)$ (follows from (1.1)). The proposed estimator of $\theta_\gamma = [A_\gamma, B_\gamma, \alpha_\gamma, \beta_\gamma]$ is $\hat{\theta}_{N,\gamma} = [\hat{A}_{N,\gamma}, \hat{B}_{N,\gamma}, \hat{\alpha}_{N,\gamma}, \hat{\beta}_{N,\gamma}]$, which is the following.

$$\hat{\theta}_{N,\gamma} = \arg \min_{\theta \in \Theta} Q_N(\theta),$$

where $\gamma \in (0, 1)$.

The existences of $\theta_\gamma$ and $\hat{\theta}_{N,\gamma}$ follow from their forms in (2.2) and (2.3), respectively as $Q(\theta)$ (see (2.2)) and $Q_N(\theta)$ (see (2.3)) are convex functions of $\theta$. Moreover, the uniqueness of $\theta_\gamma$ can be ensured once the probability density function of the random variable $X_N$ is non-degenerate, and for the uniqueness of $\hat{\theta}_{N,\gamma}$, one needs to assume that all observed signals $y_1, \ldots, y_n$ are not the same. Regarding the computation of the estimator $\hat{\theta}_{N,\gamma}$, it is an appropriate place to mention that the closed form of $\hat{\theta}_{N,\gamma}$ is not tractable in this case as $|.|$ in $\mathbb{R}$ is not a differentiable function, and hence, to compute $\hat{\theta}_{N,\gamma}$ for a given observations of the signals, one needs to implement numerical methods such as Newton-Raphson method to find the solution of the minimization problem in (2.3).

Next, note that at $\gamma = \frac{1}{2}$, the estimator $\hat{\theta}_{N,\gamma}$ will be the well-known least absolute
deviation (LAD) estimator of $\theta = [A, B, \alpha, \beta]$, and a version of this estimator is studied in Lahiri et al. (2014). One can expect that the LAD estimator is expected to be more robust than the least squares estimator (LSE), and consequently, more efficient than the LSE estimator when the signals are generated from any heavy tailed distribution. In the same spirit, one can expect that the quantile based estimator will be more effective than the moment based estimators in the presence of outliers in the data. In order to study the large sample properties of the quantile based estimator, one needs to investigate the large sample properties of the the sequence of stochastic process $\hat{\theta}_{N, \gamma}$ indexed by $\gamma \in [0, 1]$, which is mathematically more challenging than deriving the limiting properties of $\hat{\theta}_{N, \gamma}$ at a fixed $\gamma \in [0, 1]$. All these challenging issues along with the results are studied in Section 2.1.

### 2.1 Properties of $\hat{\theta}_{N, \gamma}$ and Results

This section discusses various Statistical properties of the $\gamma$-th quantile $\hat{\theta}_{N, \gamma}$. Firstly, note that $\hat{\theta}_{N, \gamma}$ is not equivariant under location transformation unlike the quantiles for univariate data because of the presence of trigonometric functions in (2.3) through $x_n$. Secondly, due to the same reason, $\hat{\theta}_{N, \gamma}$ is not equivariant under scale transformation. Hence, one can conclude that $\hat{\theta}_{N, \gamma}$ is not equivariant under arbitrary affine transformation, and this fact indicates that $\hat{\theta}_{N, \gamma}$ has complex geometric structure, which motives us to study the distributional feature of $\hat{\theta}_{N, \gamma}$.

In order to study the distributional properties of $\hat{\theta}_{N, \gamma}$, one may be first interested to know the exact distribution of $\hat{\theta}_{N, \gamma}$. However, because of the complex structure of $\hat{\theta}_{N, \gamma}$, the exact distribution of $\hat{\theta}_{N, \gamma}$ is intractable, and for this reason, our aim here is to study the large sample properties of $\hat{\theta}_{N, \gamma}$, i.e., strictly speaking, the distributional feature of $\hat{\theta}_{N, \gamma}$ as $N \to \infty$. Now, note that as the components of the $\theta_{N, \gamma}$, i.e., $\hat{A}_{N, \gamma}$, $\hat{B}_{N, \gamma}$, $\hat{\alpha}_{N, \gamma}$ and $\hat{\beta}_{N, \gamma}$ are random variables, the random vector $\hat{\theta}_{N, \gamma}$ has its own sampling distribution for a fixed $\gamma \in [0, 1]$, and hence, one may be interested to study distribution of the random vector $\hat{\theta}_{N, \gamma}$ for a fixed $\gamma$ as $N \to \infty$. This study is also supposed to add a significant contribution by its own worth. However, such pointwise (with respect to $\gamma$) distribution of $\hat{\theta}_{N, \gamma}$ does not lead us to derive the distribution of any functional depending on $\hat{\theta}_{N, \gamma}$. In order to resolve this issue, one may consider the sequence of random (i.e., stochastic) process $\theta_{N, \gamma}$ indexed by $\gamma \in [0, 1]$, and the large sample properties of the sequence of the random process $\theta_{N, \gamma}$ ($\gamma \in [0, 1]$) in $l^\infty([0, 1])$ enables us to study the large sample properties of any functional.
depending on $\hat{\theta}_{N,\gamma}$ as the metric space $l^\infty([0,1])$ is equipped with supremum norm. The following technical assumptions will be required to establish the limiting distribution of the sequence of the stochastic process $\hat{\theta}_{N,\gamma}$ in the metric space $l^\infty([0,1])$.

(A1) The bounded error random variables $X_n$, $n = 1, \ldots, N$ are independently and identically distributed with the cumulative distribution function $F(.)$ having the positive, bounded and differentiable probability density function $f(.)$, and $F^{-1}(\gamma) = 0$.

(A2) The vector of unknown parameters, i.e., $\theta = [A, B, \alpha, \beta]$ is an interior point in $\Theta = [-L,L] \times [-L,L] \times [0,\pi] \times [0,\pi]$, where $L$ is an arbitrary constant.

(A3) $A^2 + B^2 > 0$.

Remark 2.1 The assumptions (A1) and (A2) are realistic in nature. Note that many well-known probability density functions such as normal, Cauchy, Exponential satisfy the conditions (A1). The assumption (A2) is also common in practice. It is expected that the unknown parameters $A$ and $B$ will lie in a bounded set, i.e., they will not be infinite in practice. Moreover, for the parameters $\alpha$ and $\beta$, the parameter space as $[0,\pi]$ is justifiable as $\pi$ and $0$ are the largest and the smallest possible values of $\alpha$ (and $\beta$), respectively. The assumption (A3) indicates that both $A$ and $B$ cannot be equal with zero, which is reasonable as the deterministic part involving $A$, $B$, $\alpha$ and $\beta$ in model (1) disappears when $A = B = 0$, and consequently, model (1) contains only the random part $X_n$ that makes the model irrelevant.

Let us now consider the following $4 \times 4$ dimensional matrix for a fixed $\gamma$, which is denoted as $\Sigma_{\gamma}$.

$$
\Sigma_{\gamma} = \frac{1}{\gamma(1-\gamma)f(F^{-1}(\gamma))f(F^{-1}(1-\gamma))} \begin{pmatrix}
A_1^2 + 9B_1^2 & -4A_1B_1 & -18B_1 & 15B_1 \\
-4A_1B_1 & A_1^2 + 9B_1^2 & 18A_1 & -15A_1 \\
-18B_1 & 18A_1 & 96 & -90 \\
15B_1 & -15A_1 & -90 & 90
\end{pmatrix}
$$

and moreover, let us denote $D_N = diag\left(\frac{1}{N^{1/2}}, \frac{1}{N^{1/2}}, \frac{1}{N^{1/2}}, \frac{1}{N^{1/2}}\right)$, where $diag(\ldots, \ldots)$ denotes $4 \times 4$ diagonal matrix.

The next theorem describes the asymptotic normality of $\hat{\theta}_{N,\gamma}$ for a fixed $\gamma$.

Theorem 2.1 Under the conditions (A1) - (A3), for a fixed $\gamma \in K \subset [0,1]$, as $N \to \infty$, $D_N^{-1}(\hat{\theta}_{N,\gamma} - \theta_\gamma)$ converges weakly to a four-dimensional Gaussian distribution with mean
Remark 2.2 The assertion in Theorem 2.1 indicates that $\theta_{N,\gamma}$ has asymptotically (i.e., as $N \to \infty$) normal distribution after appropriate normalization. Moreover, it is also implied by this result that $\tilde{\theta}_{N,\gamma}$ is a consistent estimator of $\theta_{\gamma}$ as $N \to \infty$. In this context, the rate of convergence of $\theta_{N,\gamma}$ should be discussed as well. Note that the rates of convergence of $\hat{A}_{N,\gamma}$ and $\hat{B}_{N,\gamma}$ are as usual $N^{-1/2}$ but the rates of convergence of $\hat{\alpha}_{N,\gamma}$ and $\hat{\beta}_{N,\gamma}$ are $N^{-3/2}$ and $N^{-5/2}$, respectively. The reason of unusual rates of convergences of $\hat{\alpha}_{N,\gamma}$ and $\hat{\beta}_{N,\gamma}$ lies in the fact that the unknown parameters $\alpha$ and $\beta$ are associated with the multipliers $n$ and $n^2$, respectively in model (1.1). Here we want to discuss the outline of the proof of Theorem 2.1 also. Note the criterion function in (2.3) is not a differentiable function of $\theta$ as $f(x) = |x|$ is not a differentiable function at $x = 0$, and hence, one cannot adopt the approach based on derivative to tackle the problem related to $\arg\min$ of a function. In order to avoid this problem, the non-differentiable criterion function is approximated by a suitable differentiable function, and the rest of the proof follows from the long and complex derivation of the Bahadur type expansion of $\theta_{N,\gamma}$. The technical details of the proof of Theorem 2.1 is provided in Section 6.

Next, we would like to discuss the process convergence of the sequence of stochastic processes $\hat{\theta}_{N,\gamma}$ indexed by $\gamma \in [0, 1]$. Let us now consider a random element $\mathbb{W}_{\gamma}$ associated with a Gaussian process with zero mean and the covariance kernel $\Sigma_{\gamma_1,\gamma_2} = \text{Covariance}(\mathbb{W}_{\gamma_1}, \mathbb{W}_{\gamma_2})$, where $\gamma_1, \gamma_2 \in K \subset [0, 1]$ are two arbitrary constants, and $K$ is the same compact set as defined in the statement of Theorem 2.1. The explicit form of $\Sigma_{\gamma_1,\gamma_2}$ is as follows.

$$\Sigma_{\gamma_1,\gamma_2} = C(\gamma_1, \gamma_2) \begin{pmatrix}
A_{\gamma_1,\gamma_2} + 9B_{\gamma_1,\gamma_2} & -4A_{\gamma_1,\gamma_2} & -18B_{\gamma_2} & 15B_{\gamma_2} \\
2(A_{\gamma_1,\gamma_2} + B_{\gamma_1,\gamma_2}) & A_{\gamma_1,\gamma_2}+9B_{\gamma_1,\gamma_2} & A_{\gamma_1,\gamma_2}+18B_{\gamma_2} & -15A_{\gamma_1,\gamma_2} \\
-4A_{\gamma_1,\gamma_2} & A_{\gamma_1,\gamma_2}+18B_{\gamma_2} & 2(A_{\gamma_1,\gamma_2} + 18B_{\gamma_2}) & 18A_{\gamma_1,\gamma_2} \\
-18B_{\gamma_2} & 18A_{\gamma_1,\gamma_2} & 18A_{\gamma_1,\gamma_2} & -90 \\
-15B_{\gamma_2} & -15A_{\gamma_1,\gamma_2} & -90 & 90 \\
\end{pmatrix}$$

where $C(\gamma_1, \gamma_2) = \frac{1}{\gamma_1(1-\gamma_2)f(F^{-1}(\gamma_1))f(F^{-1}(1-\gamma_2))}$. We now state the result on the convergence of the sequence of the stochastic process $\hat{\theta}_{N,\gamma}$.

Theorem 2.2 Under conditions (A1) – (A3), as $N \to \infty$, the sequence of stochastic process $D_{N}(\hat{\theta}_{N,\gamma} - \theta_{\gamma})$ indexed by $\gamma \in K \subset [0, 1]$ converges weakly in $l^\infty(K)$ to $\mathbb{W}_{\gamma}$, where
$K$ is the same as defined in the statement of Theorem 2.1, and $W_{\gamma}$ is a four dimensional random element associated with a Gaussian process with zero mean and the covariance kernel $\Sigma_{\gamma_1,\gamma_2} = \text{Covariance}(W_{\gamma_1}, W_{\gamma_2})$. The explicit expression of $\Sigma_{\gamma_1,\gamma_2}$ is provided above.

**Corollary 2.1** Under conditions (A1) – (A3), $\sup_{\gamma \in K} \| \hat{\theta}_{N,\gamma} - \theta_{\gamma} \| \overset{p}{\to} 0$ as $N \to \infty$, where $K$ is the same as defined in Theorem 2.1.

**Remark 2.3** The assertion of Theorem 2.2 indicates that the sequence of stochastic process $\hat{\theta}_{N,\gamma}$ converges weakly to a certain Gaussian process in $L^\infty(.)$ space after appropriate normalization. This fact (see Corollary 2.1) implies that $\hat{\theta}_{N,\gamma}$ converges weakly to $\theta_{\gamma}$ as $N \to \infty$ over uniform choice of $\gamma \in K \subset [0,1]$. Such Donsker type of result associated with $\hat{\theta}_{N,\gamma}$ have many applications like the usual Donsker type result of the empirical distribution function (see, e.g., van der Vaart (1998)), and a few of those applications will be studied in the subsequent section.

**Remark 2.4** Here we want to discuss a few issues briefly related to the proof of Theorem 2.2. In order to prove the weak convergence of the sequence of the stochastic process in $L^\infty(.)$, there are two steps. In the first step, one needs to establish the weak convergence of the sequence of the stochastic process at arbitrary finitely many time points (i.e., here $\gamma$ is the time parameter), and in the second step, it is required to establish that the sequence of the stochastic process is uniformly bounded almost surely. The first step can be established by a tactful application of Cramer-Wold device (see, e.g., Serfling (1980), p. 18) on the Bahadur expansion of $\hat{\theta}_{N,\gamma}$ (see the proof of Theorem 2.1), and the second step requires application of inequalities like Bernstein inequality (see, e.g., van der Vaart (1998), p. 285) to find the uniform bound almost surely. The details of the proof is provided in Section 6.

### 2.2 Applications

In the literature of Statistical Signal Processing, as said before, the estimation of the parameters in model (1.1) had received considerable attention but they only studied the point estimation of the parameters based on well-known methodologies (see, e.g., Gini et al. (2000), Yetik and Nehorai (2003) etc). However, none of them investigated the quantiles of the parameter estimation in the Chirp signal model, and to the best of our knowledge, this topic is an entirely untouched part in the Statistical Signal Processing literature. Recently,
Dhar et al. (2019) extensively studied one-sample testing of hypothesis problem using the points estimators of parameters involved in (1.1) but their work also do not have any study related to quantile estimators of the parameters in model (1.1). We here study a few applications based on \( \hat{\theta}_{N,\gamma} \).

2.2.1 LAD Estimation

As said earlier, there were a few research articles on the point estimation of the parameters \( A, B, \alpha \) and \( \beta \) in model (1.1), and the least absolute deviation (LAD) estimator of \( \theta \) is one of them (see, e.g., Lahiri et al. (2014)). Note that at \( \gamma = \frac{1}{2} \), \( \hat{\theta}_{N,\gamma} \) coincides with the LAD estimator of \( \theta \), which follows from (2.3). As a consequence of this fact, when \( \gamma = \frac{1}{2} \), the asymptotic normality of \( \hat{\theta}_{N,\gamma} \) after appropriate normalization stated in Theorem 2.1 coincides with the result of the asymptotic normality of the LAD estimator of \( \theta \); see Theorem 2 in Lahiri et al. (2014). Strictly speaking, the result in Theorem 2.1 generalizes the result in Theorem 2 in Lahiri et al. (2014) through the varying choices of the index parameter \( \gamma \). The advantage of using the LAD estimator as a point estimator of \( \theta \) is that the LAD estimator is likely to be more efficient than the least squares estimator (LSE) when the signals are obtained from any heavy tailed distribution. In this context, the readers may look at any standard text book of robust Statistics (see, e.g., Huber (2004)).

2.2.2 Scatter Estimation

In order to understand the spread of the signals, one may consider the absolute difference between the estimated signal values corresponding to \((1-\gamma)\)-th and \(\gamma\)-th quantile estimates of \( \theta = [A, B, \alpha, \beta] \), where \( \gamma \in [0, 1] \). Notaionally writing, let \( S_{1,N}(\gamma) \) (\( \gamma \in [0, 1] \)) denotes the \( \gamma \)-th scatter estimator of the spread of the signals, i.e.,

\[
S_{1,N}(\gamma) = \frac{1}{N} \sum_{n=1}^{N} |\hat{y}_{n,1-\gamma} - \hat{y}_{n,\gamma}|,
\]

where

\[
\hat{y}_{n,1-\gamma} = \hat{A}_{N,1-\gamma} \cos(\hat{\alpha}_{N,1-\gamma} n + \hat{\beta}_{N,1-\gamma} n^2) + \hat{B}_{N,1-\gamma} \sin(\hat{\alpha}_{N,1-\gamma} n + \hat{\beta}_{N,1-\gamma} n^2)
\]
and

\[ \hat{y}_{n,\gamma} = \hat{A}_{N,\gamma} \cos(\hat{\alpha}_{N,\gamma} n + \hat{\beta}_{N,\gamma} n^2) + \hat{B}_{N,\gamma} \sin(\hat{\alpha}_{N,\gamma} n + \hat{\beta}_{N,\gamma} n^2). \]

We now state the asymptotic distribution of \( S_{1,N}(\gamma) \) as \( N \to \infty \).

**Theorem 2.3** Suppose that \( S_{1}(\gamma) = \frac{1}{N} \sum_{n=1}^{N} \{ A_{1-\gamma} \cos(\alpha_{1-\gamma} n + \beta_{1-\gamma} n^2) + B_{1-\gamma} \sin(\alpha_{1-\gamma} n + \beta_{1-\gamma} n^2) \} \), then under conditions (A1) – (A3), as \( N \to \infty \), \( \sigma_{N}^{-1}(S_{1,N}(\gamma) - S_{1}(\gamma)) \) converges weakly to a Gaussian distribution with zero mean and Variance \( \sigma_{N}^{-1} \). Here \( \nabla h(\theta) = \left( \frac{\partial h(\theta)}{\partial A}, \frac{\partial h(\theta)}{\partial B}, \frac{\partial h(\theta)}{\partial \alpha}, \frac{\partial h(\theta)}{\partial \beta} \right) \), \( h(\theta) = A \cos(\alpha n + \beta n^2) + B \sin(\alpha n + \beta n^2) \) and \( \theta = [A, B, \alpha, \beta] \).

**Remark 2.5** The assertion in Theorem 2.3 indicates that the proposed scatter measure \( S_{1,N}(\gamma) \) has asymptotically normal distribution after appropriate normalization. Using this asymptotic normality result, one can construct a confidence interval of \( S_{1}(\gamma) \) for a fixed \( \gamma \) or carry out a certain testing of hypothesis problem based on \( S_{1,N}(\gamma) \). In this context, we would also like to mention that one may be tried to derive the weak convergence of the sequence of stochastic process \( S_{1,N}(\gamma) \) indexed by \( \gamma \in [0, 1] \) for theoretical interest; however, statistically speaking, it is of more interest to see the performance of \( S_{1,N}(\gamma) \) for a fixed \( \gamma \), and for that reason, we here skip the investigation on the process convergence result of \( S_{1,N}(\gamma) \).

We now want to propose an alternative estimator of the scatter based on the spread of distribution of \( \hat{\theta}_{N,\gamma} \). In order to measure the spread of the distribution of \( \hat{\theta}_{N,\gamma} \) or any suitable function of \( \hat{\theta}_{N,\gamma} \), one may consider the volume of the support associated with the distribution of \( \hat{\theta}_{N,\gamma} \) or any suitable function of \( \hat{\theta}_{N,\gamma} \). Let \( S_{2,N} \) denotes the proposed measure, i.e., Mathematically speaking, one can write

\[ S_{2,N} = \int_{0}^{1} \| D_{N}^{-1} \hat{\theta}_{N,\gamma} \|^2 d\gamma. \]

Here we should point out that instead of the Euclidean norm of \( \hat{\theta}_{N,\gamma} \) in the expression of \( S_{2,N} \), one may consider any other norm to accumulate the information, and the multiplier \( D_{N}^{-1} \) appears only because of sake of technicalities. We now state the asymptotic distribution of \( S_{2,N} \) as \( N \to \infty \).
Theorem 2.4 Suppose that \( S_2 = \int_0^1 ||D^{-1}_N \hat{\theta}_n||^2 d\gamma \). Then under conditions (A1) – (A3), as \( N \to \infty \), \( (S_{2,N} - S_2) \) converges weakly to the distribution of \( \int_0^1 ||\mathbb{W}_\gamma||^2 d\gamma \), where \( \mathbb{W}_\gamma \) is the same as defined in Theorem 2.2, and ||.-|| denotes the usual Euclidean norm.

Remark 2.6 The construction of \( S_{2,N} \) indicates that \( S_{2,N} \) is essentially the normalized sum of squares of the components of \( \hat{\theta}_{N,\gamma} \), i.e., in other words, it is equivalent to the variance of \( \hat{\theta}_{N,\gamma} \). Theorem 2.4 studies the distributional behaviour of \( \hat{\theta}_{N,\gamma} \) when the number of signals is large. Precisely speaking, it follows from the statement of Theorem 2.4 that the proposed measure \( S_{2,N} \) converges weakly to the distribution of a certain random variable involved integration over a bounded set \([0, 1]\). It is to be noted that in practice, the exact computation of this integral is intractable but some Riemann type approximation can be done. This issue is discussed in Section 3.1.2, where the performance of the proposed measure \( S_{2,N} \) is carried out.

2.2.3 Interquartile Range

As we proposed measures of scatter in Section 2.2.2, one can propose alternative measures of scatter similar to the concept of interquartile range. In this context, one may consider \( S_{1,N}(0.25) \) as the interquartile range. The following Corollary states the asymptotic distribution of \( S_{1,N}(0.25) \).

Corollary 2.2 Under conditions (A1) – (A3), as \( N \to \infty \), \( \sigma_{N,0.25}^{-1}(S_{1,N}(0.25) - S_1(0.25)) \) converges weakly to a Gaussian distribution with zero mean and Variance = 1, where \( \sigma_{N,0.25} = \{\nabla h(\theta)D_N\}(\Sigma_{0.75} + \Sigma_{0.25} - 2\Sigma_{0.75}\Sigma_{0.25})\{\nabla h(\theta)D_N\}^T \). Here \( \nabla h(\theta) = \left( \frac{\partial h(\theta)}{\partial A}, \frac{\partial h(\theta)}{\partial B}, \frac{\partial h(\theta)}{\partial \alpha}, \frac{\partial h(\theta)}{\partial \beta} \right) \), \( h(\theta) = A \cos(\alpha n + \beta n^2) + B \sin(\alpha n + \beta n^2) \) and \( \theta = [A, B, \alpha, \beta] \).

Using the result stated in Corollary 2.2, one can carry out the test based on \( S_{N,1}(0.25) \) to capture the spread of the data. If the data has some outliers/influential observations, \( S_{N,1}(0.25) \) may work well.

An alternative measure of previously discussed interquartile measure, one may also consider

\[
S_{2,N,0.25} = \int_{0.25}^{0.75} ||D^{-1}_N \hat{\theta}_{N,\gamma}||^2 d\gamma
\]

13
as an interquartile based spread measure. The following corollary states the asymptotic distribution of $S_{2,N,0.25}$.

**Corollary 2.3** Suppose that $S_{2,0.25} = \int_{0.25}^{0.75} \|D_N^{-1}\theta_0\|^2 d\gamma$. Then under conditions (A1) – (A3), as $N \to \infty$, $(S_{2,N,0.25} - S_{2,0.25})$ converges weakly to the distribution of $\int_{0.25}^{0.75} \|W_\gamma\|^2 d\gamma$, where $W_\gamma$ is the same as defined in Theorem 2.2, and $\|\cdot\|$ denotes the usual Euclidean norm.

Corollary 2.3 asserts that $S_{2,N,0.25}$ converges weakly to the distribution of a certain random variable, which involves integration. As we discussed in Section 2.2.2, the exact computation of the integration may not be tractable in practice, and to resolve this issue, one may adopt the idea of Riemann approximation of the integral. This measure is also expected to perform well in the presence of outliers/influential observations.

**2.2.4 Skewness**

In order to measure the skewness of the data, i.e., to check how much the data is deviated from the symmetry, one may propose a measure of skewness as

$$SK_{N,1} = \sup_{\gamma \in [0,1]} \|\hat{\theta}_{N,1-\gamma} + \hat{\theta}_{N,\gamma} - 2\hat{\theta}_{N,0}\|$$

or

$$SK_{N,2} = \int_{0}^{1} \|\hat{\theta}_{N,1-\gamma} + \hat{\theta}_{N,\gamma} - 2\hat{\theta}_{N,0}\| d\gamma.$$

The formulations of $SK_{N,1}$ and $SK_{N,2}$ are motivated by the fact that $\|\theta_{1-\gamma} + \theta_{\gamma} - 2\theta_{0}\| = 0$ for all $\gamma \in [0,1]$ if the distribution of the data is symmetric. The next theorem describes the distributional feature of $SK_{N,1}$ and $SK_{N,2}$ as $N \to \infty$.

**Theorem 2.5** Suppose that $SK_1 = \sup_{\gamma \in [0,1]} \|\theta_{1-\gamma} + \theta_{\gamma} - 2\theta_{0}\|$, $SK_2 = \int_{0}^{1} \|\theta_{1-\gamma} + \theta_{\gamma} - 2\theta_{0}\| d\gamma$. Then, under the conditions (A1) – (A3), $D_N^{-1}(SK_{N,1} - SK_1)$ and $D_N^{-1}(SK_{N,2} - SK_2)$ converge weakly to the distributions of $\sup_{\gamma \in [0,1]} \|W_{1-\gamma} + W_{\gamma} - 2W_0\|$ and $\int_{0}^{1} \|W_{1-\gamma} + W_{\gamma} - 2W_0\| d\gamma$, respectively.
Remark 2.7 Theorem 2.5 asserts that the proposed measures of the skewness, namely, $SK_{N,1}$ and $SK_{N,2}$ converges weakly to the distribution of certain random variables, which involves sup and integration, respectively. In practice, to compute $SK_{N,1}$, consider a partition $T = [\gamma_0, \gamma_1, \ldots, \gamma_k]$, where $0 \leq \gamma_0 < \gamma_\kappa \leq 1$, where $\kappa$ is a sufficiently large number, and afterwards, compute the maximum value of $SK_{N,1}$ over the partition $T$. Next, in order to compute $SK_{N,2}$, one can approximate the integration by the Riemann approximation over the partition $T$. The performance of the proposed measures $SK_{N,1}$ and $SK_{N,2}$ is studied in Sections 3.1.3 and 3.2.3.

3 Data Analysis

As we discussed in Section 1 there are many practical situations, where the frequency of signals changes over time, and in those cases, the chirp signal model described in (1.1) can be well fitted to the given signals. In this section, we analyse a few real data, which are well-known in Statistical signal processing literature and investigate the performance of various quantile based measures described in Section 2.2.

3.1 Connectionist Bench (Sonar, Mines vs. Rocks) Data Set

This data set has two hundred eight many signals observed at various 208 angles; among them, one hundred eleven signals observed of a metal cylinder at various angles, and ninety seven signals are obtained from rocks at various angles, and this well-known data set is available at https://archive.ics.uci.edu/ml/496datasets/Connectionist+Bench+(Sonar,+Mines+vs.+Rocks) This data set was earlier analysed in Gorman and Sejnowski (1988), and Figure 1 illustrates the observed signals. In this figure, the red coloured observations are the signals from the metal, and the blue coloured observations are the signals from the rock. For both type of signals, we first try to check whether the model described in (1.1) fits the data or not. In order to investigate it, we plot the residuals of both type of signals in Figure 2 where the $i$-th residuals is defined as

$$\hat{x}_n = y_n - \hat{A}_{N,0.5} \cos(\hat{\alpha}_{N,0.5} n + \hat{\beta}_{N,0.5} n^2) - \hat{B}_{N,0.5} \sin(\hat{\alpha}_{N,0.5} n + \hat{\beta}_{N,0.5} n^2).$$

The diagrams in Figures 2 indicate that the residuals are randomly distributed, and hence, one can conclude that the data is well fitted by the model (1.1). Here we would like to
briefly discuss about the formula of $\hat{x}_n$ provided above. Note that $\hat{A}_{N,0.5}$, $\hat{B}_{N,0.5}$, $\hat{\alpha}_{N,0.5}$ and $\hat{\beta}_{N,0.5}$ are the LAD estimators of $A$, $B$, $\alpha$ and $\beta$, respectively (see Section 2.2.1). However, in principle, one may use any other appropriate estimators of $A$, $B$, $\alpha$ and $\beta$ such as the least squares estimators of $A$, $B$, $\alpha$ and $\beta$ as well. Here, we consider the LAD estimators as they are consistent estimators, which follows from the assertion in Theorem 2.1.

3.1.1 Analysis associated with application in Section 2.2.1

Now, suppose that we want to check whether the parameters involved in the model for signals from the metal and the parameters involved in the model for signals from the rocks are the same or not. With notations, one can describe the problem in the following way. Let $\theta_1 = [A_1, B_1, \alpha_1, \beta_1]$ be the parameters involved in the model for signals from the metal, and $\theta_2 = [A_2, B_2, \alpha_2, \beta_2]$ is the parameters involved in the model for signals from
the rock. We now want to test

\[ H_0 : \theta_1 = \theta_2 \]

against

\[ H_1 : \theta_1 \neq \theta_2, \]

where “=” indicates the componentwise equality, and “\( \neq \)” indicates that at least one of the components are unequal. To test \( H_0 \) against \( H_1 \), one may consider the test statistic as a suitable difference between \( \hat{\theta}_{1,N_1}^{1/2} \) and \( \hat{\theta}_{2,N_2}^{1/2} \), where \( \hat{\theta}_{1,N_1}^{1/2} \) and \( \hat{\theta}_{2,N_2}^{1/2} \) are the LAD estimators (see Section 2.2.1 for detail on the LAD estimator) of \( \theta_1 \) and \( \theta_2 \), respectively, and here \( N_1 = 111 \) and \( N_2 = 97 \). Let the test statistic be

\[ T_{N_1,N_2} = \left\| \hat{\theta}_{1,N_1}^{1/2} - \hat{\theta}_{2,N_2}^{1/2} \right\|, \]

where \( \| . \| \) denotes the usual Euclidean norm. For this data, we obtain

\[ \hat{\theta}_{1,N_1}^{1/2} = [0.908, 0.691, 1.459, 0.496] \]
and
\[ \hat{\theta}_{1,N_2,\frac{1}{2}} = [0.663, 0.293, 0.440, 1.170]. \]

Therefore, for this data, \( T_{N_1,N_2} = \left\| \hat{\theta}_{1,N_1,\frac{1}{2}} - \hat{\theta}_{2,N_2,\frac{1}{2}} \right\| = 1.308. \) Next, in order to validate the null hypothesis \( H_0 \), we compute the \( p \)-value using the technique of residual Bootstrap for this data set in the following way, which can also be used as a generic algorithm.

**Algorithm**

**Input:** Signals \( \{y_{1,1}, \ldots, y_{1,N_1}\} \) and \( \{y_{2,1}, \ldots, y_{2,N_2}\} \)

**Output:** \( p \)-value

1: Compute the value of the test statistic \( T_{N_1,N_2} \). Let us denote it by \( t_0 \).

2: Obtain \( \hat{\theta}_{1,N_1,\frac{1}{2}} = [\hat{A}_{1,N_1,\frac{1}{2}}, \hat{B}_{1,N_1,\frac{1}{2}}, \hat{\alpha}_{1,N_1,\frac{1}{2}}, \hat{\beta}_{1,N_1,\frac{1}{2}}] \) and \( \hat{\theta}_{2,N_2,\frac{1}{2}} = [\hat{A}_{2,N_2,\frac{1}{2}}, \hat{B}_{2,N_2,\frac{1}{2}}, \hat{\alpha}_{2,N_2,\frac{1}{2}}, \hat{\beta}_{2,N_2,\frac{1}{2}}] \) from \( \{y_{1,1}, \ldots, y_{1,N_1}\} \) and \( \{y_{2,1}, \ldots, y_{2,N_2}\} \), respectively.

3: Estimate the residuals for \( n = 1, \ldots, N_1 \) and \( m = 1, \ldots, N_2 \):
\[
\hat{X}_{1,n} = y_{1,n} - \hat{A}_{1,N_1,\frac{1}{2}} \cos(\hat{\alpha}_{1,N_1,\frac{1}{2}} n + \hat{\beta}_{1,N_1,\frac{1}{2}} n^2) - \hat{B}_{1,N_1,\frac{1}{2}} \sin(\hat{\alpha}_{1,N_1,\frac{1}{2}} n + \hat{\beta}_{1,N_1,\frac{1}{2}} n^2)
\]
and
\[
\hat{X}_{2,m} = y_{2,m} - \hat{A}_{2,N_2,\frac{1}{2}} \cos(\hat{\alpha}_{2,N_2,\frac{1}{2}} m + \hat{\beta}_{2,N_2,\frac{1}{2}} m^2) - \hat{B}_{2,N_2,\frac{1}{2}} \sin(\hat{\alpha}_{2,N_2,\frac{1}{2}} m + \hat{\beta}_{2,N_2,\frac{1}{2}} m^2).
\]

4: Combine the data \( \{y_{1,1}, \ldots, y_{1,N_1}, y_{2,1}, \ldots, y_{2,N_2}\} \) and obtain the LAD estimates \( \hat{A}, \hat{B}, \hat{\alpha} \) and \( \hat{\beta} \) of \( A, B, \alpha \) and \( \beta \) based on this combined sample.

5: Permute the estimated error \( \hat{X}_{1,n} \) for \( n = 1, \ldots, N_1 \). Using these \( N_1! \) permuted errors \( \hat{X}_{1,n}, \hat{A}, \hat{B}, \hat{\alpha} \) and \( \hat{\beta} \) in model (1.1), obtain \( N_1! \) many signals over \( N_1 \) many time points. The \( i \)-th \( (i = 1, \ldots, N_1!) \) re-sample of the signals is denoted as \( \{y_{i1}, \ldots, y_{iN_1}\} \). Similarly, permute the estimated error \( \hat{X}_{2,m} \) for \( m = 1, \ldots, N_2 \). Using these \( N_2! \) permuted errors \( \hat{X}_{2,m}, \hat{A}, \hat{B}, \hat{\alpha} \) and \( \hat{\beta} \) in model (1.1), obtain \( N_2! \) many signals observed at \( N_2 \) many time points. The \( j \)-th \( (j = 1, \ldots, N_2!) \) re-sample of the signals is denoted as \( \{y_{j1}, \ldots, y_{jN_2}\} \).
6: Denote $M = \min(N_1!, N_2!)$ and compute the value of the test statistic for $M$ many re-samples like $\{y_{i1}, \ldots, y_{iN_1}\}$ and $\{y_{j1}, \ldots, y_{jN_2}\}$. Denote the values of the test statistic $T_{N_1,N_2}$ for those $M$ many re-samples as $t_1, \ldots, t_M$.

7: The $p$-value of the test is defined as $\frac{1}{m} \sum_{i=1}^{M} 1_{(t_i \geq t_0)}$, where $1_A = 1$ if $A$ is true, and $1_A = 0$ if $A$ is not true. Alternatively stated, the $p$-value is defined as the proportion of times $t_1, \ldots, t_M$ exceeding $t_0$.

\begin{verbatim}
final
return p-value
\end{verbatim}

Using this algorithm, i.e., residual bootstrap (see, e.g., Efron and Tibshirani (1994)), we obtain the $p$-value = 0.044 for this data set. This small $p$-value implies that the data does not favour the null hypothesis, i.e., precisely speaking, $\theta_1 = [A_1, B_1, \alpha_1, \beta_1]$, the parameters involved in the model for signals from the metal and $\theta_2 = [A_2, B_2, \alpha_2, \beta_2]$, the parameters involved in the model for signals from the rock are different for this data, which is expected as the signals are obtained from two entirely different objects.

3.1.2 Analysis associated with application in Section 2.2.2

In Section 3.1.1 we have seen that the parameters involved in the model for the signals obtained from the metal and the rock are different. Now, one may be interested to know whether the spread of the signals obtained from the metal and the rock are the same or not. Notationally speaking, we want to check

$$H_0 : \int_0^1 ||\theta_{1,\gamma}||^2 d\gamma = \int_0^1 ||\theta_{2,\gamma}||^2 d\gamma$$

against

$$H_1 : \int_0^1 ||\theta_{1,\gamma}||^2 d\gamma \neq \int_0^1 ||\theta_{2,\gamma}||^2 d\gamma,$$

where $\theta_{1,\gamma}$ and $\theta_{2,\gamma}$ are $\gamma$-th quantiles of $\theta_1$ and $\theta_2$, respectively. The definitions of $\theta_1$ and $\theta_2$ are provided at the beginning of Section 3.1.1
In order to test $H_0$ against $H_1$, we consider the test statistic $S_{N_1,N_2}$ as

$$S_{N_1,N_2} = |S_{2,N_1} - S_{2,N_2}|,$$

where $S_{2,N_1} = \int_0^1 ||D_{N_1}^{-1}\hat{\theta}_{N_1,\gamma}||^2 d\gamma$ and $S_{2,N_2} = \int_0^1 ||D_{N_2}^{-1}\hat{\theta}_{N_2,\gamma}||^2 d\gamma$. For detailed interpretation of the construction of $S_{2,N_1}$ and $S_{2,N_2}$, we refer the readers to look at the construction of $S_{2,N}$ in Section 2.2.2. Here we want to discuss the computation of $S_{2,N_1}$ and $S_{2,N_2}$. Note that

$$S_{2,N_1} = E_\gamma \left[ ||D_{N_1}^{-1}\hat{\theta}_{N_1,\gamma}||^2 \right],$$

where the random variable $\gamma$ follows uniform distribution over $[0, 1]$, and by law of large number, one can claim that

$$\frac{1}{R} \sum_{i=1}^R ||D_{N_1}^{-1}\theta_{N_1,\gamma_i}||^2 \overset{p/a.s.}{\to} E_\gamma \left[ ||D_{N_1}^{-1}\hat{\theta}_{N_1,\gamma}||^2 \right] = S_{2,N_1}$$

as $R \to \infty$, where $\gamma_1, \ldots, \gamma_L$ are i.i.d. sequence of random variables having the same distribution of the random variable $\gamma$, i.e., the uniform distribution over $[0, 1]$. Therefore, in order to compute $S_{2,N_1}$, we generate $\gamma_1, \ldots, \gamma_{1000}$ from uniform distribution over $[0, 1]$, and

$$\frac{1}{1000} \sum_{i=1}^{1000} ||D_{N_1}^{-1}\theta_{N_1,\gamma_i}||^2$$

is considered as an estimate of $S_{2,N_1}$ for this data. The similar approach is adopted to compute $S_{2,N_2}$ as well for this data.

For this data, we obtain $S_{N_1,N_2} = 0.091$, and as described in Section 3.1.1, we also compute the p-value, which equals with 0.537. This large value p-value indicates that this data favours the null hypothesis $H_0$, i.e., the signals obtained from the metal and the rock do not have any significant difference in terms of spread for this data. In this context, we should mention that one may formulate the test statistic based on the measure $S_{1,N}(\gamma)$ (studied in Section 2.2.2) for various choices of $\gamma$ but for the sake of concise presentation, we here skip this study.

3.1.3 Analysis associated with application in Section 2.2.4

In Section 3.1.2 we have observed that the spread of the signals generated by the metal and the rock do not have any significant difference. This fact motivates us to study whether the signals from the metal and the rock are generated from the symmetric distribution or
from skewed distribution. In order to check it, for both types of signals, we consider the following two testing of hypothesis problems:

\[ H_0^* : SK_2^* = 0 \]

against

\[ H_1^* : SK_2^* \neq 0 \]

AND

\[ H_0 : SK_1 = 0 \]

against

\[ H_1 : SK_1 \neq 0 \]

The expressions of \( SK_1 \) and \( SK_2 \) are provided in the statement of Theorem 2.5. To test \( H_0 \) against \( H_1 \), \( SK_{N,1} \) (see Section 2.2.4) is considered as the test statistic, and for testing \( H_0^* \) against \( H_1^* \), \( SK_{N,2} \) (see Section 2.2.4) is considered as the test statistic. Note that \( N = N_1 = 111 \) for the signals obtained from the metal, and \( N = N_2 = 97 \) for the signals obtained from the rock. In order to compute \( SK_{N,2} \), the same procedure described for computing \( S_{2,N_1} \) (see Section 3.1.2) is adopted here.

To test \( H_0 \) against \( H_1 \), we obtain \( SK_{N_1,1} = 0.784 \) for the signals obtained from the metal, and the \( p \)-value equals with 0.033, where the \( p \)-value is computed following the procedure described in Section 3.1.1. At the same time, the same test is carried out for the signals obtained from the rock as well. In this case, we obtain \( SK_{N_2,1} = 0.725 \), and the corresponding \( p \)-value = 0.084. These small \( p \)-values indicate that the signals obtained from the metal and the rock do not favour the null hypothesis \( H_0 \), i.e., the distribution of the signals from both the metal and the rock have the feature of skewness. The same phenomenon we observe for the testing of hypothesis problem \( H_0^* \) against \( H_1^* \). For the signals obtained from the metal, we obtain \( SK_{N_1,2} = 0.923 \) and the \( p \)-value = 0.065, and for the signals obtained from the rock, we obtain \( SK_{N_2,2} = 0.901 \) and the associated \( p \)-value = 0.083. Overall, all these investigations indicate that the distributional feature of the signals obtained from both the metal and the rock is skewed; not symmetric.
3.2 Sonar Data

This data set contains 477 sonar signals obtained from an experiment conducted in the Department of Electrical Engineering at the IIT Kanpur, India. The data set is submitted to Journal’s repository and illustrated in Figure 3. As we analysed various aspects in Section 3.1 here also we investigate all these issues for this data set. This data set is well fitted by the model (1.1) as Dhar et al. (2019) empirically argued that the estimated residuals follow some light tailed distribution using the technique of QQ plot (see, e.g., Wilk and Gnanadesikan (1968) and Doksum and Sievers (1976)), i.e., the residuals do not have any fixed pattern, and hence, one can carry out the analysis on this data assuming that the data is associated with the model (1.1).
3.2.1 Analysis associated with application in Section 2.2.1

We here want to see the central tendency of the signals using the approach of LAD techniques discussed in Section 2.2.1. Suppose that we want to test

\[ H_0 : \theta = \left[ 1, 1, \frac{\pi}{4}, \frac{\pi}{4} \right] \]

against

\[ H_1 : \theta \neq \left[ 1, 1, \frac{\pi}{4}, \frac{\pi}{4} \right], \]

where " = " denotes the componentwise equality, and " \neq " denotes at least one inequality among the components. Now, to test \( H_0 \) against \( H_1 \), we consider the test statistic as

\[ T_N = \left| \hat{\theta}_{N,1/2} - \left[ 1, 1, \frac{\pi}{4}, \frac{\pi}{4} \right] \right|, \]

where \( \hat{\theta}_{N,1/2} \) is the LAD estimator of \( \theta \), and here \( N = 477 \). For this data, we obtain

\[ \hat{\theta}_{N,1/2} = [1.443, 1.019, 1.223, 1.002], \]

and the \( p \)-value equals 0.037, where \( p \)-value is computed according to the procedure described in Section 3.1.1. This small \( p \)-value indicates that the observed signals do not follow the null hypothesis \( H_0 \), i.e., the signals are not associated with the model described in (1.1) when \( A = B = 1 \) and \( \alpha = \beta = \frac{\pi}{4} \).

3.2.2 Analysis associated with application in Section 2.2.2

In Section 3.1.2, we studied the testing of hypothesis problem regarding the equality of scatters of two different types of signals; however, this data has only one kind of sonar signal, and we want to know whether the spread equals some specified value or not. Formally speaking, suppose that we want to test

\[ H_0 : \int_0^1 \| \theta_\gamma \|^2 d\gamma = 1 \]
against

\[ H_1 : \int_0^1 ||\hat{\theta}_\gamma||^2 d\gamma \neq 1. \]

To test \( H_0 \) against \( H_1 \), we consider the test statistic as

\[ S_N = \int_0^1 ||\hat{\theta}_{N,\gamma}||^2 d\gamma - 1 \]

motivated by the fact that \( \hat{\theta}_{N,\gamma} \) is a consistent estimator of \( \theta_{\gamma} \), which follows from Theorem 2.2. The integral involved in \( S_N \) is calculated using the same approach described in Section 3.1.2. For this data, we obtain \( S_N = 1.233 \) and \( p\)-value = 0.038, where \( p\)-value is computed using the algorithm described in Section 3.1.1. This small \( p\)-value indicates that the scatter of the signals is significantly different than one, i.e., the spread of the signals is not standardized to one in some sense.

### 3.2.3 Analysis associated with application in Section 2.2.4

In Section 3.2.2, we have seen that the spread of the signals is different from unity. Afterwards, we here would like to know whether the distributional feature of the signals is symmetric in nature or not. In order to study it, as we did in Section 3.1.3, we want to test

\[ H_0 : SK_1 = 0 \]

against

\[ H_1 : SK_1 \neq 0 \]

AND

\[ H_0^* : SK_2 = 0 \]

against

\[ H_1^* : SK_2 \neq 0. \]

The expressions of \( SK_1 \) and \( SK_2 \) are provided in the statement of Theorem 2.5. In order to test \( H_0 \) against \( H_1 \), we consider the test statistic as \( SK_{N,1} \), and for testing \( H_0^* \) against \( H_1^* \), \( SK_{N,2} \) is considered. For the expressions of \( SK_{N,1} \) and \( SK_{N,2} \), the readers are advised to see Section 2.2.4. For this data, we obtain \( SK_{N,1} = 0.862 \) and the corresponding \( p\)-value.
value = 0.037, and we have $SK_{N,2} = 0.936$ and the corresponding $p$-value = 0.057. The $p$-values are computed following the algorithm provided in Section 3.1.1. These small $p$-values indicate that the signals do not follow the null hypotheses, and hence, the signals are likely to be generated from a skewed distribution.

4 Finite Sample Study

In Section 3 we analysed two real data, which are associated with the chirp signal model characterized by the regression equation (1.1), and as usual, the sizes of both data sets are large enough. In this section, we now want to see the performance of various quantile based estimators when the sample size is small or moderately small.

4.1 Comparison : LAD and LSE estimators

As we have already seen in Section 2.2.1 that the LAD estimator of $\theta = [A, B, \alpha, \beta]$ is a special case of $\hat{\theta}_{N,\gamma}$ when $\gamma = \frac{1}{2}$ while the least squares estimator (LSE) $\hat{\theta}_{N,LSE}$ is defined as

$$\hat{\theta}_{N,LSE} = \arg \min_{\theta \in \Theta} \frac{1}{N} \sum_{n=1}^{N} x_n^2,$$

where $x_n = y_n - A \cos(\alpha n + \beta n^2) - B \sin(\alpha n + \beta n^2)$. In order to study the performance of $\hat{\theta}_{N,\frac{1}{2}}$ and $\hat{\theta}_{N,LSE}$, we compute the finite sample efficiency of $\hat{\theta}_{N,\frac{1}{2}}$ relative to $\hat{\theta}_{N,LSE}$ and carry out the study in the following way.

Suppose that in the model (1.1), $A = B = 1$ and $\alpha = \beta = \frac{\pi}{4}$, and to generate the signals $y_n$, the observations of the error random variable $X_n$ are generated from various distributions like normal distribution, $t$-distribution with 4 degrees of freedom and Cauchy distribution. Here $N = 10, 50$ and 100. In order to compute the efficiency of $\hat{\theta}_{N,\frac{1}{2}}$ relative to $\hat{\theta}_{N,LSE}$, we replicate this experiment $M$ times, and let $\hat{\theta}_{N,\frac{1}{2},i}$ and $\hat{\theta}_{N,LSE,i}$ be the estimate of $\hat{\theta}_{N,\frac{1}{2}}$ and $\hat{\theta}_{N,LSE}$ for $i$-th replicate, respectively, where $i = 1, \ldots, M$. Then the empirical mean squared error (EMSE) of $\hat{\theta}_{N,\frac{1}{2}}$ is defined as

$$EMSE(\hat{\theta}_{N,\frac{1}{2}}) = \frac{1}{M} \sum_{i=1}^{M} \left\| \hat{\theta}_{N,\frac{1}{2},i} - \left[1, 1, \frac{\pi}{4}, \frac{\pi}{4}\right] \right\|^2.$$
and the empirical mean squared error (EMSE) of $\hat{\theta}_{N,LSE}$ is defined as
\[
EMSE(\hat{\theta}_{N,LSE}) = \frac{1}{M} \sum_{i=1}^{M} \left\| \hat{\theta}_{N,LSE,i} - \left[ 1, 1, \frac{\pi}{4}, \frac{\pi}{4} \right] \right\|^2.
\]
The finite sample efficiency of $\hat{\theta}_{N,\frac{1}{2}}$ relative to $\hat{\theta}_{N,LSE}$ is defined as $\frac{EMSE(\hat{\theta}_{N,LSE})}{EMSE(\hat{\theta}_{N,\frac{1}{2}})}$. In our study, we choose $M = 1000$, and the summarized result is as follows.

(i) Let $X_n$ follows standard normal distribution.

(A) When $N = 10$, we have $EMSE(\hat{\theta}_{N,\frac{1}{2}}) = 1.347$ and $EMSE(\hat{\theta}_{N,LSE}) = 1.231$. Hence, the finite sample efficiency of $\hat{\theta}_{N,\frac{1}{2}}$ relative to $\hat{\theta}_{N,LSE}$ is 0.914.

(B) When $N = 50$, we have $EMSE(\hat{\theta}_{N,\frac{1}{2}}) = 1.212$ and $EMSE(\hat{\theta}_{N,LSE}) = 1.114$. Hence, the finite sample efficiency of $\hat{\theta}_{N,\frac{1}{2}}$ relative to $\hat{\theta}_{N,LSE}$ is 0.919.

(C) When $N = 100$, we have $EMSE(\hat{\theta}_{N,\frac{1}{2}}) = 1.117$ and $EMSE(\hat{\theta}_{N,LSE}) = 1.136$. Hence, the finite sample efficiency of $\hat{\theta}_{N,\frac{1}{2}}$ relative to $\hat{\theta}_{N,LSE}$ is 0.983.

Overall, these finite sample efficiency values indicate that $\hat{\theta}_{N,LSE}$ is marginally more efficient than $\hat{\theta}_{N,\frac{1}{2}}$.

(ii) Let $X_n$ follows standard t distribution with 4 degrees of freedom.

(A) When $N = 10$, we have $EMSE(\hat{\theta}_{N,\frac{1}{2}}) = 2.448$ and $EMSE(\hat{\theta}_{N,LSE}) = 2.675$. Hence, the finite sample efficiency of $\hat{\theta}_{N,\frac{1}{2}}$ relative to $\hat{\theta}_{N,LSE}$ is 1.093.

(B) When $N = 50$, we have $EMSE(\hat{\theta}_{N,\frac{1}{2}}) = 2.221$ and $EMSE(\hat{\theta}_{N,LSE}) = 2.335$. Hence, the finite sample efficiency of $\hat{\theta}_{N,\frac{1}{2}}$ relative to $\hat{\theta}_{N,LSE}$ is 1.051.

(C) When $N = 100$, we have $EMSE(\hat{\theta}_{N,\frac{1}{2}}) = 2.178$ and $EMSE(\hat{\theta}_{N,LSE}) = 2.310$. Hence, the finite sample efficiency of $\hat{\theta}_{N,\frac{1}{2}}$ relative to $\hat{\theta}_{N,LSE}$ is 1.061.

Overall, these finite sample efficiency values indicate that $\hat{\theta}_{N,\frac{1}{2}}$ is marginally more efficient than $\hat{\theta}_{N,LSE}$.

(iii) Let $X_n$ follows standard Cauchy distribution.
(A) When $N = 10$, we have $EMSE(\hat{\theta}_{N,1/2}) = 3.175$ and $EMSE(\hat{\theta}_{N,LSE}) = 12.469$. Hence, the finite sample efficiency of $\hat{\theta}_{N,1/2}$ relative to $\hat{\theta}_{N,LSE}$ is 3.927.

(B) When $N = 50$, we have $EMSE(\hat{\theta}_{N,1/2}) = 2.992$ and $EMSE(\hat{\theta}_{N,LSE}) = 31.783$. Hence, the finite sample efficiency of $\hat{\theta}_{N,1/2}$ relative to $\hat{\theta}_{N,LSE}$ is 10.623.

(C) When $N = 100$, we have $EMSE(\hat{\theta}_{N,1/2}) = 2.867$ and $EMSE(\hat{\theta}_{N,LSE}) = 74.849$. Hence, the finite sample efficiency of $\hat{\theta}_{N,1/2}$ relative to $\hat{\theta}_{N,LSE}$ is 26.107.

Overall, these finite sample efficiency values indicate that $\hat{\theta}_{N,1/2}$ outperform $\hat{\theta}_{N,LSE}$.

All these studies in (i), (ii) and (iii) clearly indicate that the LAD estimator is more efficient compared to the LSE estimator when the signals are obtained from the heavy tailed distribution in view of the fact that Cauchy distribution has heavier tail than $t$-distribution with 4 degrees of freedom, and $t$-distribution with 4 degrees of freedom has heavier tail than normal distribution. Strictly speaking, as the tail of the distribution becomes heavier, the efficiency of the LAD estimator increases. This also gives us an impression of having advantage in using a quantile based estimator when the data is generated from any heavy tailed distribution.

4.2 Comparison: $S_{2,N}$ and Variance-Covariance matrix

In Section 4.1, we have studied the performance of the quantile based location estimator, namely, the LAD estimator and the moment based location estimator, namely, the LSE estimator, and we have seen that for the signals generated from the heavy tailed distributions, the quantile based LAD estimator performs better, i.e., in other words, the quantile based estimator is more robust against the outliers/influential observations than the moment based estimator. One may be now interested to see whether the same phenomena is true for the scatter estimator or not. In order to investigate it, we here consider the same setting of simulation as we did in Section 4.1 which is briefly described in the following for the sake of completeness.

Consider the model (1.1), and choose $A = B = 1$ and $\alpha = \beta = \frac{\pi}{4}$. Here also, $N = 10, 50$ and 100, and three cases are investigated when the error random variable $X_n$ follows normal distribution, $t$-distribution with 4 degrees of freedom and Cauchy distribution. As
a quantile based estimator of scatter, we consider

\[
SK_{2,N} = \int_0^1 ||D_N^{-1}\hat{\theta}_{N,\gamma}||^2 d\gamma
\]
as we proposed in Section 2.2.2 and as a moment based estimator of scatter, we here consider

\[
S_N = \text{det}(D_N^{-1}(\hat{\theta}_{N,LSE} - \theta_0))\text{det}(D_N^{-1}(\hat{\theta}_{N,LSE} - \theta_0)),
\]
where \(\theta_0 = [1, 1, \frac{\pi}{4}, \frac{\pi}{4}]\) is specified, \(\hat{\theta}_{N,LSE}\) is defined in Section 4.1 and \(\text{det}(.)\) denotes the determinant of the matrix \((.)\). As earlier, the experiment is replicated \(M = 1000\) times, and \(SK_{2,N,i}\) denotes the value of \(SK_{2,N}\) for \(i\)-th replication, and \(S_{N,i}\) denotes the value of \(S_N\) for \(i\)-th replication. The finite sample efficiency of \(SK_{2,N}\) relative to \(S_N\) is defined as

\[
\frac{1}{M} \sum_{i=1}^{M} S_{N,i} \quad \frac{1}{M} \sum_{i=1}^{M} S_{2,N,i}
\]

Now the results are summarized in the following.

(i) Let \(X_n\) follows standard normal distribution.

(A) When \(N = 10\), we have \(\frac{1}{M} \sum_{i=1}^{M} S_{2,N,i} = 2.796\) and \(\frac{1}{M} \sum_{i=1}^{M} S_{N,i} = 2.365\). Hence, the finite sample efficiency of \(S_{2,N}\) relative to \(S_N\) is 0.846.

(B) When \(N = 50\), we have \(\frac{1}{M} \sum_{i=1}^{M} S_{2,N,i} = 2.668\) and \(\frac{1}{M} \sum_{i=1}^{M} S_{N,i} = 2.299\). Hence, the finite sample efficiency of \(S_{2,N}\) relative to \(S_N\) is 0.862.

(C) When \(N = 100\), we have \(\frac{1}{M} \sum_{i=1}^{M} S_{2,N,i} = 2.558\) and \(\frac{1}{M} \sum_{i=1}^{M} S_{N,i} = 2.172\). Hence, the finite sample efficiency of \(S_{2,N}\) relative to \(S_N\) is 0.849.

Overall, these finite sample efficiency values indicate that \(S_N\) is marginally more efficient than \(S_{2,N}\).

(ii) Let \(X_n\) follows \(t\) distribution with 4 degrees of freedom.
(A) When \( N = 10 \), we have \( \frac{1}{M} \sum_{i=1}^{M} S_{2,N,i} = 3.112 \) and \( \frac{1}{M} \sum_{i=1}^{M} S_{N,i} = 3.257 \). Hence, the finite sample efficiency of \( S_{2,N} \) relative to \( S_{N} \) is 1.046.

(B) When \( N = 50 \), we have \( \frac{1}{M} \sum_{i=1}^{M} S_{2,N,i} = 3.114 \) and \( \frac{1}{M} \sum_{i=1}^{M} S_{N,i} = 3.201 \). Hence, the finite sample efficiency of \( S_{2,N} \) relative to \( S_{N} \) is 1.028.

(C) When \( N = 100 \), we have \( \frac{1}{M} \sum_{i=1}^{M} S_{2,N,i} = 3.198 \) and \( \frac{1}{M} \sum_{i=1}^{M} S_{N,i} = 3.068 \). Hence, the finite sample efficiency of \( S_{2,N} \) relative to \( S_{N} \) is 1.042.

Overall, these finite sample efficiency values indicate that \( S_{2,N} \) is marginally more efficient than \( S_{N} \).

(iii) Let \( X_n \) follows Cauchy distribution.

(A) When \( N = 10 \), we have \( \frac{1}{M} \sum_{i=1}^{M} S_{2,N,i} = 4.259 \) and \( \frac{1}{M} \sum_{i=1}^{M} S_{N,i} = 13.781 \). Hence, the finite sample efficiency of \( S_{2,N} \) relative to \( S_{N} \) is 3.236.

(B) When \( N = 50 \), we have \( \frac{1}{M} \sum_{i=1}^{M} S_{2,N,i} = 3.997 \) and \( \frac{1}{M} \sum_{i=1}^{M} S_{N,i} = 26.759 \). Hence, the finite sample efficiency of \( S_{2,N} \) relative to \( S_{N} \) is 6.695.

(C) When \( N = 100 \), we have \( \frac{1}{M} \sum_{i=1}^{M} S_{2,N,i} = 3.839 \) and \( \frac{1}{M} \sum_{i=1}^{M} S_{N,i} = 34.736 \). Hence, the finite sample efficiency of \( S_{2,N} \) relative to \( S_{N} \) is 9.048.

Overall, these finite sample efficiency values indicate that \( S_{2,N} \) is substantially more efficient than \( S_{N} \).

All inclusive, as we have observed in Section 4.1, one can conclude that the quantile based estimator \( S_{2,N} \) performs better than the moment based estimator \( S_{N} \) when the signals are obtained from the heavy tailed distribution, i.e., to be summarized, the quantile based scatter estimator is also more robust against the presence of outliers/influential observations.

5 Concluding Remarks

This article studies the quantile estimators of the unknown amplitude parameters \( A \) and \( B \), frequency parameter \( \alpha \) and frequency rate parameter \( \beta \). At first, we here establish the limiting distribution of those quantile estimators at a fixed value of quantile index. Next,
the major contribution of this article is deriving the Donsker-type result of the quantile process indexed by the quantile index parameter, which is formally stated in Theorem 2.2. Strictly speaking, this theorem enables us to derive the limiting distribution of any continuous functional of the quantiles based estimator, which has wide applications, and a few of them are studied in Section 2.2. Moreover, we implement those applications on real and simulated data as well.

We assume the condition \( A^2 + B^2 > 0 \) to establish the results in Theorems 2.1 and 2.2. This means that both \( A \) and \( B \) cannot be equal to zero in order to satisfy the aforementioned condition on \( A \) and \( B \). However, note that model 1.1 reduces to a complete random model when \( A = B = 0 \), and hence, the results in Theorems 2.1 and 2.2 are incapable to test whether the signals are generated from any chirp signal model or a random model. The derivation of the similar results asserted in Theorems 2.1 and 2.2 may be of interest for future research when \( A = B = 0 \).

The choice of model (1.1) may also be of interest to many practitioners. For instance, one may be interested to consider the model as

\[
Y_n = \sum_{i=1}^{p} \left\{ A_i \cos(\alpha_i n + \beta_i n^2) + B_i \sin(\alpha_i n + \beta_i n^2) \right\} + X_n, n = 1, \ldots, N,
\]

where \((A_i, B_i, \alpha_i, \beta_i), i = 1, \ldots, p\) is a 4p-dimensional unknown parameter, and \(X_n\) is a sequence of i.i.d. random errors. For such models, one of the major issues is the appropriate choice of \(p\) and to resolve it, one may adopt or propose variable selection methodology to choose appropriate \(p\). Afterwards, for such models also, one may consider the quantile estimators of \((A_i, B_i, \alpha_i, \beta_i), i = 1, \ldots, p\) as we have done here for \(p = 1\). However, for a large \(p\), the curse of dimensionality may arise, and it is also of interest to see whether the limiting distribution of the quantile based estimators is possible or not when \(p\) equals with some order of \(N\).

As we consider the LAD estimator as a location estimator in Section 2.2.1, one can use the trimmed mean also a location estimator (see, e.g., Dhar and Chaudhuri (2012) and a few references therein), and the concept of the trimmed mean is extended for various regression models (see, e.g., Welsh (1987) and Dhar (2016)). To the best of our knowledge, the approach of trimmed mean to estimate the unknown parameters in the chirp signal model has not been paid any attention in the literature. To study this research problem, one may use the results presented in Theorems 2.1 and 2.2 in view of the fact that the
trimmed is a certain linear combination of the order statistic, and hence, there is a one-to-one correspondence with the quantiles as well. In the same spirit, one may derive the limiting distribution of the $L$-estimator (see, e.g., Serfling (1980), Chapter 8 and Jureckova and Welsh (1990) in the context of the linear model) also using the results in Theorems 2.1 and 2.2. Moreover, one can conceive the idea of rank as well using the concept of quantiles studied here.

Acknowledgement: The author is supported by the MATRICS (MTR/2019/000039), a research grant from the SERB, Government of India.

6 Supplementary Material : Technical Details

Proof of Theorem 2.1: Let us first concentrate on the criterion function $Q_N(\theta) = \frac{1}{N} \sum_{i=1}^{N} g(x_n)$ (see (2.3)), where $x_n = y_n - A \cos(\alpha n + \beta n^2) - B \sin(\alpha n + \beta n^2)$, $g : \mathbb{R} \rightarrow \mathbb{R}$, and $g(x) = |x| + (2\gamma - 1)x$ for any $x \in \mathbb{R}$, and $\gamma \in (0,1)$. Let us now consider a new sequence of functions $g_N(.)$, which is the following.

$$g_N(x) = \left( \frac{x^3}{3N^6} - \frac{x^2}{N^3} - \frac{N^3}{3} \right) 1_{\left(\frac{-1}{N^{\frac{3}{2}}} \leq x \leq 0\right)} - x 1_{\left(x < -\frac{1}{N^{\frac{3}{2}}}\right)} + \left( - \frac{x^3}{3N^6} + \frac{x^2}{N^3} + \frac{N^3}{3} \right) 1_{\left(0 \leq x \leq \frac{1}{N^{\frac{3}{2}}}\right)} + x 1_{\left(x > \frac{1}{N^{\frac{3}{2}}}\right)} + (2\gamma - 1)x.$$

Suppose now that $Q_N^*(\theta) = \frac{1}{N} \sum_{i=1}^{N} g_N(x_n)$, and

$$\hat{\theta}_{N,\gamma}^* = \arg \min_{\theta \in \Theta} Q_N^*(\theta)$$

whereas recall that

$$\hat{\theta}_{N,\gamma} = \arg \min_{\theta \in \Theta} Q_N(\theta).$$

Let us now try to show that

$$\sup_{\theta \in \Theta} |Q_N^*(\theta) - Q_N(\theta)| = o_{a.s}(1) \text{ as } N \rightarrow \infty,$$ (6.4)
where \(a.s.\) denotes the almost sure convergence. Note that for an arbitrary \(\epsilon > 0\), we have

\[
P[|Q_N^*(\theta) - Q_N(\theta)| > \epsilon] \leq \frac{E[|Q_N^*(\theta) - Q_N(\theta)|]}{\epsilon} \text{ using Markov's inequality}
\]

\[
\leq \frac{K_1}{N^4} \sum_{n=1}^{N} E\left[1_{(0 \leq |X_n| \leq \frac{1}{N^3})}\right] \text{ for some } K_1
\]

\[
= \frac{K_1}{N^4} \sum_{n=1}^{N} P\left[-\frac{1}{N^3} \leq X_n \leq \frac{1}{N^3}\right]
\]

\[
= \frac{K_1}{N^4} \sum_{n=1}^{N} \left\{F\left(\frac{1}{N^3}\right) - F\left(-\frac{1}{N^3}\right)\right\}
\]

\[
= \frac{K_2}{N^7} \sum_{n=1}^{N} f(\xi) \text{ for some } \xi \in \left(0, \frac{1}{N^3}\right)
\]

\[
\leq \frac{K_3}{N^6} \text{ for some } K_3.
\]

Hence, we firstly have

\[
|Q_N^*(\theta) - Q_N(\theta)| = o_p(N^{-6+\delta}) \text{ for any } \delta > 0
\]

\[
\Rightarrow \sup_{\theta \in \Theta} |Q_N^*(\theta) - Q_N(\theta)| = o_p(N^{-6+\delta}) \text{ as } \Theta \text{ is a compact set,}
\]

(6.5)

where \(o_p\) denotes the convergence in probability.

Next, note that

\[
\sum_{N=1}^{\infty} P[|Q_N^*(\theta) - Q_N(\theta)| > \epsilon] \leq \sum_{N=1}^{\infty} \frac{K_3}{N^6} < \infty.
\]

As it is true for any \(\epsilon > 0\), by first Borel-Cantelli Lemma (see, e.g., [Serfling (1980), p. 351]) and in view of compactness of \(\Theta\), the claim in (6.4) is established.

Finally, in view of (6.4), we have

\[
|\hat{\theta}_{N,\gamma}^* - \hat{\theta}_{N,\gamma}| = o_{as}(1) \text{ as } N \to \infty.
\]

(6.6)

The analogous convergence in probability result follows from (6.5). Moreover, arguing in
a similar way as [Kim et al. (2000)](#) studied and in view of (6.6), we further have

\[ |\hat{\theta}_{N,\gamma}^* - \theta_\gamma| = o_{as}(1) \text{ as } N \to \infty. \] (6.7)

Now, we want to establish that

\[ D_N H(Q_N^*(\theta_\gamma)) D_N \overset{p}{\to} \Sigma_\gamma \text{ as } N \to \infty, \] (6.8)

where \( H(Q_N^*(\theta)) = \left( \frac{\partial^2 Q_N^*(\theta)}{\partial \theta_i \partial \theta_j} \right)_{1 \leq i, j \leq 4} \), \( \theta := [\theta_1, \theta_2, \theta_3, \theta_4] = [A, B, \alpha, \beta] \) is a generic variable, and \( D_N \) and \( \Sigma_\gamma \) are the same as defined in Section 2.1.

In order to establish it, recall that

\[
g_N(x) = \left( \frac{x^3}{3N^6} - \frac{x^2}{N^3} - \frac{N^3}{3} \right) 1(-\frac{1}{N^3} \leq x \leq 0) - x1(x< -\frac{1}{N^3}) \\
+ \left( -\frac{x^3}{3N^6} + \frac{x^2}{N^3} + \frac{N^3}{3} \right) 1(0 \leq x \leq \frac{1}{N^3}) + x1(x> \frac{1}{N^3}) + (2\gamma - 1)x.
\]

Hence, we have

\[
\frac{d}{dx}\{g_N(x)\} = \left( \frac{x^2}{N^6} - \frac{2x}{N^3} \right) 1(-\frac{1}{N^3} \leq x \leq 0) - 1(x< -\frac{1}{N^3}) \\
+ \left( -\frac{x^2}{N^6} + \frac{2x}{N^3} \right) 1(0 \leq x \leq \frac{1}{N^3}) + 1(x> \frac{1}{N^3}) + (2\gamma - 1)
\]

and

\[
\frac{d^2}{dx^2}\{g_N(x)\} = \left( \frac{2x}{N^6} - \frac{x}{N^3} \right) 1(-\frac{1}{N^3} \leq x \leq 0) + \left( -\frac{2x}{N^6} + \frac{2}{N^3} \right) 1(0 \leq x \leq \frac{1}{N^3}).
\]

Now, recall that \( Q_N^*(\theta) = \frac{1}{N} \sum_{n=1}^{N} g_N(X_n) \), and note that

\[
E[H(Q_N^*(\theta))] = -E[\{\nabla Q_N^*(\theta)\}^T \{\nabla Q_N^*(\theta)\}],
\]

where \( \nabla Q_N^*(\theta) = \left( \frac{\partial Q_N^*(\theta)}{\partial A}, \frac{\partial Q_N^*(\theta)}{\partial B}, \frac{\partial Q_N^*(\theta)}{\partial \alpha}, \frac{\partial Q_N^*(\theta)}{\partial \beta} \right) \). Therefore, the key term involves in analyzing \( H(Q_N^*(\theta)) \) is \( \frac{d}{dx}\{g_N(x)\}_{x=X_n} \). Let us now try to investigate the behaviour of the moments of \( \frac{d}{dx}\{g_N(x)\}_{x=X_n} \).
Now, note that
\[
E \left[ \frac{d}{dx} \{g_N(x)\} \big|_{x=x_n} \right] \\
= \int_{-\frac{1}{N^3}}^{0} \left( \frac{x^2}{N^6} - \frac{2x}{N^3} \right) f(x) dx - \int_{0}^{\frac{1}{N^3}} f(x) dx + \int_{0}^{\frac{1}{N^3}} \left( -\frac{x^2}{N^6} + \frac{2x}{N^3} \right) f(x) dx \\
+ \int_{\frac{1}{N^3}}^{\frac{1}{N^3}} f(x) dx + (2\gamma - 1) \int_{\frac{1}{N^3}}^{\infty} f(x) dx \\
= \int_{-\frac{1}{N^3}}^{0} \left( \frac{x^2}{N^6} - \frac{2x}{N^3} \right) f(x) dx - \left[ F \left( -\frac{1}{N^3} \right) - F(0) \right] + \int_{0}^{\frac{1}{N^3}} \left( -\frac{x^2}{N^6} + \frac{2x}{N^3} \right) f(x) dx \\
+ \left[ 1 - F \left( \frac{1}{N^3} \right) \right] + (2\gamma - 1).
\]

In the above expression, observe that \( \int_{0}^{\frac{1}{N^3}} \left( -\frac{x^2}{N^6} + \frac{2x}{N^3} \right) f(x) dx \to 0 \) as \( N \to \infty \) since \( F \) is a continuous distribution function, and due to the same reason along with the fact of (A1), \( [1 - F \left( \frac{1}{N^3} \right)] \to [1 - F(0)] = (1 - \gamma) \) as \( N \to \infty \). Hence, we have
\[
- \left[ F \left( -\frac{1}{N^3} \right) - F(0) \right] + \left[ 1 - F \left( \frac{1}{N^3} \right) \right] + (2\gamma - 1) \to \gamma \text{ as } N \to \infty. \tag{6.9}
\]

Next, consider
\[
\int_{0}^{\frac{1}{N^3}} \frac{2x}{N^3} f(x) dx = \frac{2}{N^3} \int_{0}^{\frac{1}{N^3}} x f(x) dx \\
= \frac{2}{N^3} \left[ \frac{1}{N^3} \left\{ F \left( \frac{1}{N^3} \right) - F(0) \right\} - \int_{0}^{\frac{1}{N^3}} F(x) dx \right] \\
= \frac{2}{N^3} \left[ \frac{1}{N^3} \left\{ F \left( \frac{1}{N^3} \right) - F(0) \right\} - \frac{1}{N^3} F(\xi) dx \right] \quad (\xi \in [0, \frac{1}{N^3}]) \\
\to 0 \text{ as } N \to \infty. \tag{6.10}
\]
Afterwards, we consider
\[
\left| \frac{1}{N^3} \int_0^{x^2} f(x) \, dx \right|^2 = \left| \frac{1}{N^6} \int_0^{x^2} f(x) \, dx \right|^2 = \frac{1}{N^{12}} \left| \int_0^{x} x f(x) \, dx \right|^2 \leq \frac{1}{N^{12}} \left| \int_0^{x} x^2 f(x) \, dx \right| \to 0
\]
(6.11)
as \(N \to \infty\).

Hence, using (6.9), (6.10) and (6.11), we have
\[
E \left[ \frac{d}{dx} \{ g_N(x) \} \right] \to \lambda \text{ as } N \to \infty.
\]
(6.12)

We now want to work on \( E \left[ \frac{d}{dx} \{ g_N(x) \} \right] \). Note that
\[
E \left[ \frac{d}{dx} \{ g_N(x) \} \right] = \int_{-\frac{1}{N^3}}^{\frac{1}{N^3}} \frac{x^2}{N^6} \left( \frac{2x}{N^3} \right)^2 f(x) \, dx + \int_{\frac{1}{N^3}}^{\infty} f(x) \, dx + (2\gamma - 1)^2
\]
\[
+ 2(2\gamma - 1) \int_{-\frac{1}{N^3}}^{\frac{1}{N^3}} \left( \frac{x^2}{N^6} - \frac{2x}{N^3} \right) f(x) \, dx + 2(2\gamma - 1) \int_{-\infty}^{-\frac{1}{N^3}} f(x) \, dx
\]
\[
+ 2(2\gamma - 1) \int_{0}^{\frac{1}{N^3}} \left( -\frac{x^2}{N^6} + \frac{2x}{N^3} \right) f(x) \, dx + 2(2\gamma - 1) \int_{\frac{1}{N^3}}^{\infty} f(x) \, dx.
\]

Observe now the followings.
\[
\int_{0}^{\frac{1}{N^3}} f(x) \, dx = F \left( -\frac{1}{N^3} \right) - F(0) \to 0 \text{ as } N \to \infty.
\]
(6.13)
\[
\int_{\frac{1}{N^3}}^{\infty} f(x) \, dx = F(\infty) - F \left( \frac{1}{N^3} \right) \to (1 - \gamma) \text{ as } N \to \infty.
\]
(6.14)
\[
\int_{-\infty}^{\infty} f(x) \, dx = F \left( -\frac{1}{N^3} \right) - F (-\infty) \to \gamma \text{ as } N \to \infty. \tag{6.15}
\]

Arguing exactly in a similar way as in (6.10) and (6.11), we have
\[
\int_{-\frac{1}{N^3}}^{0} \left( \frac{x^2}{N^6} - \frac{2x}{N^3} \right) f(x) \, dx \to 0 \text{ as } N \to \infty,
\]
and also,
\[
\int_{-\frac{1}{N^3}}^{0} \left( -\frac{x^2}{N^6} + \frac{2x}{N^3} \right) f(x) \, dx \to 0 \text{ as } N \to \infty.
\]

Next, consider
\[
\int_{-\frac{1}{N^3}}^{\frac{1}{N^3}} \left( \frac{x^2}{N^6} - \frac{2x}{N^3} \right)^2 f(x) \, dx
\]
\[
= \int_{-\frac{1}{N^3}}^{\frac{1}{N^3}} x^4 f(x) \, dx + 4 \int_{-\frac{1}{N^3}}^{\frac{1}{N^3}} x^2 f(x) \, dx - 4 \int_{-\frac{1}{N^3}}^{\frac{1}{N^3}} x^3 f(x) \, dx.
\]
First observe that
\[
\int_{-\frac{1}{N^3}}^{\frac{1}{N^3}} x^4 f(x) \, dx \leq G \int_{-\frac{1}{N^3}}^{\frac{1}{N^3}} x^4 \, dx \to 0 \text{ as } N \to \infty \text{ using the fact that } f(x)
\]
is bounded by some constant \(G\). The exact arguments also leads to
\[
\int_{-\frac{1}{N^3}}^{\frac{1}{N^3}} x^2 f(x) \, dx \to 0 \text{ as } N \to \infty,
\]
and
\[
\int_{-\frac{1}{N^3}}^{\frac{1}{N^3}} x^3 f(x) \, dx \to 0 \text{ as } N \to \infty.
\]
Hence, all these above facts imply that
\[
E \left[ \left. \frac{d}{dx} \{g_N(x)\} \right|_{x=X_n} \right]^2 \to (1-\gamma) + (\gamma - 1) \gamma = \gamma^2,
\]
and hence,
\[
\text{Variance} \left[ \left. \frac{d}{dx} \{g_N(x)\} \right|_{x=X_n} \right] \to 0 \text{ as } N \to \infty. \tag{6.16}
\]

Therefore, (6.12) and (6.16) imply that
\[
\left. \frac{d}{dx} \{g_N(x)\} \right|_{x=X_n} \overset{P}{\to} \gamma \text{ as } N \to \infty, \tag{6.17}
\]
where \( \overset{P}{\to} \) denotes the convergence in probability.
Further, observe that the (1, 1)-th element of $D_N H(Q_N^*(\theta_\gamma)) D_N$ equals with

$$\frac{1}{N} \sum_{i=1}^{N} \left\{ \frac{d}{dx} \{ g(x) \} \big|_{x=x_n} \right\}^2 \cos^2(\alpha_\gamma n + \beta_\gamma n^2),$$

and for all $t \in \mathbb{R}$, it follows from Vinogradov and Karatsuba (1984) that

$$\lim_{N \to \infty} \frac{1}{N^t} \sum_{i=1}^{N} n^t \cos^2(\alpha_\gamma n + \beta_\gamma n^2) = \frac{1}{2(t+1)}, \quad (6.18)$$

$$\lim_{N \to \infty} \frac{1}{N^t} \sum_{i=1}^{N} n^t \sin^2(\alpha_\gamma n + \beta_\gamma n^2) = \frac{1}{2(t+1)}, \quad (6.19)$$

and

$$\lim_{N \to \infty} \frac{1}{N^t} \sum_{i=1}^{N} n^t \sin(\alpha_\gamma n + \beta_\gamma n^2) \cos(\alpha_\gamma n + \beta_\gamma n^2) = \frac{1}{2(t+1)}. \quad (6.20)$$

Hence, using (6.17), (6.18), (6.19) and (6.20), we have (6.8).

We now want to work on the distribution of $\{D_N \nabla Q_N^*(\theta_\gamma)\}$, where

$$\nabla Q_N^*(\theta_\gamma) = \left( \frac{\partial Q_N^*(\theta)}{\partial A}, \frac{\partial Q_N^*(\theta)}{\partial B}, \frac{\partial Q_N^*(\theta)}{\partial \alpha}, \frac{\partial Q_N^*(\theta)}{\partial \beta} \right) \bigg|_{\theta = \theta_\gamma}.$$ 

Let us denote

$$\{D_N \nabla Q_N^*(\theta_\gamma)\} := (\{D_N \nabla Q_N^*(\theta_\gamma)\}_1, \{D_N \nabla Q_N^*(\theta_\gamma)\}_2, \{D_N \nabla Q_N^*(\theta_\gamma)\}_3, \{D_N \nabla Q_N^*(\theta_\gamma)\}_4),$$

and observe that

$$\{D_N \nabla Q_N^*(\theta_\gamma)\}_1 = - \frac{1}{\sqrt{N}} \sum_{n=1}^{N} \left\{ \frac{d}{dx} g_N(x) \right\} \bigg|_{x=X_n} \cos(\alpha_\gamma n + \beta_\gamma n^2), \quad (6.21)$$

$$\{D_N \nabla Q_N^*(\theta_\gamma)\}_2 = - \frac{1}{\sqrt{N}} \sum_{n=1}^{N} \left\{ \frac{d}{dx} g_N(x) \right\} \bigg|_{x=X_n} \sin(\alpha_\gamma n + \beta_\gamma n^2), \quad (6.22)$$
\begin{align}
\{D_N \nabla Q_N^*(\theta_\gamma)\}_3 &= \frac{1}{N^{3/2}} \sum_{n=1}^{N} n \left\{ \frac{d}{dx} g_N(x) \right\}_{x=X_n} [A_\gamma \sin(\alpha_\gamma n + \beta_\gamma n^2) - B_\gamma \cos(\alpha_\gamma n + \beta_\gamma n^2)], \\
\text{and} \\
\{D_N \nabla Q_N^*(\theta_\gamma)\}_4 &= \frac{1}{N^{5/2}} \sum_{n=1}^{N} n^2 \left\{ \frac{d}{dx} g_N(x) \right\}_{x=X_n} [A_\gamma \sin(\alpha_\gamma n + \beta_\gamma n^2) - B_\gamma \cos(\alpha_\gamma n + \beta_\gamma n^2)].
\end{align}

Note that \( E[\{D_N \nabla Q_N^*(\theta_\gamma)\}] = 0 := (0, 0, 0, 0) \), which follows from the definition of \( Q_N^*(\theta_\gamma) \), and in view of (6.16), we have \( \text{Variance}[\{D_N \nabla Q_N^*(\theta_\gamma)\}] \to 0 \) as \( N \to \infty \). Hence, by the central limit theorem stated in [Fuller (1996)], and in view of (6.13), (6.14), (6.15), (6.21), (6.22), (6.23) and (6.24), we have

\begin{equation}
\{D_N \nabla Q_N^*(\theta_\gamma)\} \overset{d}{\rightarrow} \mathbf{Z},
\end{equation}

where \( \mathbf{Z} \) is a \( 4 \times 1 \)-dimensional random vector associated with four dimensional Gaussian distribution with mean = 0 and the variance-covariance matrix = \( \Sigma_\gamma \). Moreover, using (6.4), (6.8), (6.6) and (6.7), we have

\begin{equation}
D_N \{H(Q_N^*(\theta_\gamma)) - H(Q_N^*(\hat{\theta}_{N,\gamma}^*)))\} D_N \overset{P}{\rightarrow} 0 \text{ as } N \to \infty,
\end{equation}

and

\begin{equation}
D_N \{H(Q_N^*(\theta_\gamma)) - H(Q_N^*(\hat{\theta}_{N,\gamma}^*)))\} D_N \overset{P}{\rightarrow} 0 \text{ as } N \to \infty,
\end{equation}

Finally, note that

\begin{align}
\nabla Q_N^*(\hat{\theta}_{N,\gamma}^*) &= 0 \text{ (follows from the definition of } Q_N^*(\cdot) \text{)} \\
\Rightarrow \nabla Q_N^*(\theta_\gamma) + H(Q_N^*(\theta_\gamma))(\hat{\theta}_{N,\gamma}^* - \theta_\gamma) &= 0 \\
(\theta_\gamma^* \text{ is some point on the straight line joining } \theta_\gamma \text{ and } \hat{\theta}_{N,\gamma}^*) \\
\Leftrightarrow D_N \nabla Q_N^*(\theta_\gamma) + D_N H(Q_N^*(\theta_\gamma))(\hat{\theta}_{N,\gamma}^* - \theta_\gamma) &= 0 \\
\Leftrightarrow D_N \nabla Q_N^*(\theta_\gamma) + D_N H(Q_N^*(\theta_\gamma)) D_N D_N^{-1}(\hat{\theta}_{N,\gamma}^* - \theta_\gamma) &= 0.
\end{align}
Using (6.28), (6.25) and (6.8) along with (6.26) and (6.27) lead to the proof of the theorem.

**Proof of Theorem 2.2:** In order to prove this theorem, let us consider arbitrary \( L \)-dimensional \((L \geq 1)\) vector \((\gamma_1, \ldots, \gamma_L)\) and corresponding \( L \)-dimensional random vector \( (D_N^{-1}(\hat{\theta}_{N,\gamma_1} - \theta_{\gamma_1}), \ldots, D_N^{-1}(\hat{\theta}_{N,\gamma_L} - \theta_{\gamma_L}))\). Suppose that \( a_1, \ldots, a_L \), where \( a_i \in \mathbb{R}^4 \) \((i = 1, \ldots, L)\), are arbitrary constants and try to derive the asymptotic distribution of the following random variable:

\[
a_1^T D_N^{-1}(\hat{\theta}_{N,\gamma_1} - \theta_{\gamma_1}) + \ldots + a_L^T D_N^{-1}(\hat{\theta}_{N,\gamma_L} - \theta_{\gamma_L}).
\]

Note that

\[
D_N^{-1}(\hat{\theta}_{N,\gamma_1} - \theta_{\gamma_1}) = -\{D_N H(Q_N^*(\theta_{\gamma_1}))D_N\}^{-1} D_N \nabla Q_N^*(\theta_{\gamma_1}) + o_p(1),
\]

which follows from (6.28), (6.26) and (6.27), and \( o_p(1) \) denotes the convergence in probability to 0 as \( N \to \infty \), and hence, we have

\[
a_1^T D_N^{-1}(\hat{\theta}_{N,\gamma_1} - \theta_{\gamma_1}) = -a_1^T \{D_N H(Q_N(\theta_{\gamma_1}))D_N\}^{-1} D_N \nabla Q_N(\theta_{\gamma_1}) + o_p(1) \text{ as } a_1^T o_p(1) = o_p(1).
\]

This implies that

\[
a_1^T D_N^{-1}(\hat{\theta}_{N,\gamma_1} - \theta_{\gamma_1}) \overset{d}{\to} Z_1,
\]

where \( Z_1 \) is a random variable associated with normal distribution with mean = 0 and variance = \( a_1^T \Sigma_\gamma a_1 \). The fact follows from (6.25), (6.8), (6.26) and (6.27). Therefore,

\[
a_1^T D_N^{-1}(\hat{\theta}_{N,\gamma_1} - \theta_{\gamma_1}) + \ldots + a_L^T D_N^{-1}(\hat{\theta}_{N,\gamma_L} - \theta_{\gamma_L}) \overset{d}{\to} Z^*,
\]

where \( Z^* \) is a some random variable associated with certain zero mean one-dimensional Gaussian distribution. Now, by Cramer-Wald device (see, e.g., Serfling (1980), p. 18), one can conclude that

\[
(D_N^{-1}(\hat{\theta}_{N,\gamma_1} - \theta_{\gamma_1}), \ldots, D_N^{-1}(\hat{\theta}_{N,\gamma_L} - \theta_{\gamma_L})) \overset{d}{\to} Z^{**},
\]

where \( Z^{**} \) is a random vector associated with a certain \( L \)-dimensional Gaussian distribution.
We next want to establish that $\sup_{\gamma} \sup_{N} \| D_{N}^{-1}(\hat{\theta}_{N,\gamma} - \theta_{\gamma}) \|$ is bounded almost surely, where $K$ is a compact subset of $[0, 1]$. Note that as $K$ is a compact set, it is enough to establish that $\sup_{N} \| D_{N}^{-1}(\hat{\theta}_{N,\gamma} - \theta_{\gamma}) \|$ is bounded almost surely for a fixed $\gamma$ since $K$ can be covered by finitely many sub-covers. Next, observe that

$$
\| D_{N}^{-1}(\hat{\theta}_{N,\gamma} - \theta_{\gamma}) \| \leq \sqrt{\bar{N}} \| \hat{A}_{N,\gamma} - A_{\gamma} \| + \sqrt{\bar{N}} \| \hat{B}_{N,\gamma} - B_{\gamma} \| + N^{3/2} \| (\hat{\alpha}_{N,\gamma} - \alpha_{\gamma}) \| + N^{5/2} \| (\hat{\beta}_{N,\gamma} - \beta_{\gamma}) \|,
$$

where $\hat{\theta}_{N,\gamma} = (\hat{A}_{N,\gamma}, \hat{B}_{N,\gamma}, \hat{\alpha}_{N,\gamma}, \hat{\beta}_{N,\gamma})$ and $\theta_{\gamma} = (A_{\gamma}, B_{\gamma}, \alpha_{\gamma}, \beta_{\gamma})$. Hence, it is enough to show that $\sup_{N} \sqrt{\bar{N}} \| \hat{A}_{N,\gamma} - A_{\gamma} \|, \sup_{N} \sqrt{\bar{N}} \| \hat{B}_{N,\gamma} - B_{\gamma} \|, \sup_{N} N^{3/2} \| (\hat{\alpha}_{N,\gamma} - \alpha_{\gamma}) \| \text{ and } \sup_{N} N^{5/2} \| (\hat{\beta}_{N,\gamma} - \beta_{\gamma}) \|$ are bounded almost surely for a fixed $\gamma$, which follows from (6.34).

Let us first try to show that $\sup_{N} \sqrt{\bar{N}} \| \hat{A}_{N,\gamma} - A_{\gamma} \|$ is bounded almost surely for a fixed $\gamma$. Now, take an arbitrary $\epsilon > 0$ and observe that

$$
\sum_{N=1}^{\infty} P \left( \sqrt{\bar{N}} \| \hat{A}_{N,\gamma} - A_{\gamma} \| > \epsilon \right) = \sum_{N=1}^{\infty} P \left( \sqrt{\bar{N}} \| \nabla Q_{N}^{*}(\theta_{\gamma})_{1} \| / NH(Q_{N}^{*}(\theta_{\gamma}))_{1,1} > \epsilon \right)
$$

(Follows from (6.29)). Here $\nabla Q_{N}^{*}(\theta_{\gamma})_{1}$ is the first component of $\nabla Q_{N}^{*}(\theta_{\gamma})$, and $NH(Q_{N}^{*}(\theta_{\gamma}))_{1,1}$ is the $(1, 1)$-th element of $D_{N}H(Q_{N}^{*}(\theta_{\gamma}))D_{N}$.)

$$
= \sum_{N=1}^{\infty} P \left( \sqrt{\bar{N}} \| \nabla Q_{N}^{*}(\theta_{\gamma})_{1} \| > NH(Q_{N}^{*}(\theta_{\gamma}))_{1,1} \epsilon \right)
$$

$$
= \sum_{N=1}^{\infty} P \left( \sqrt{\bar{N}} \| \nabla Q_{N}^{*}(\theta_{\gamma})_{1} \| > \Sigma_{\gamma, (1, 1)} \epsilon \right) \quad \text{(Here } \Sigma_{\gamma, (1, 1)} \text{ is the } (1, 1)\text{-th element of } \Sigma_{\gamma} \text{, and the fact follows from (6.8))}
$$

$$
= \sum_{N=1}^{\infty} P \left( \sqrt{N} \| \nabla \left( \frac{1}{N} \sum_{n=1}^{N} g_{N}(y_{n} - A_{\gamma} \cos(\alpha_{n} n + \beta_{n} n^{2}) - B_{\gamma} \sin(\alpha_{n} n + \beta_{n} n^{2})) \right)_{1} \| > \Sigma_{\gamma, (1, 1)} \epsilon \right)
$$

$$
= \sum_{N=1}^{\infty} P \left( \sqrt{N} \| \nabla \left( \sum_{n=1}^{N} g_{N}(y_{n} - A_{\gamma} \cos(\alpha_{n} n + \beta_{n} n^{2}) - B_{\gamma} \sin(\alpha_{n} n + \beta_{n} n^{2})) \right)_{1} \| > \sqrt{N} \Sigma_{\gamma, (1, 1)} \epsilon \right)
$$

$$
\leq \sum_{N=1}^{\infty} P \left( \sum_{n=1}^{N} | T_{i} | > \sqrt{N} \Sigma_{\gamma, (1, 1)} \epsilon \right) \quad \text{(here } T_{i} = \{ \frac{d}{dx} g_{N}(x) \} \}_{x=y_{n} - A_{\gamma} \cos(\alpha_{n} n + \beta_{n} n^{2}) - B_{\gamma} \sin(\alpha_{n} n + \beta_{n} n^{2})} \cos(\alpha_{n} n + \beta_{n} n^{2}) \}
$$
Now, using Bernstein inequality (see, e.g., van der Vaart (1998), p. 285), we have

\[ P \left[ \sum_{n=1}^{N} |T_i| > \sqrt{N} \Sigma_{\gamma,(1,1)} \epsilon \right] \leq e^{-\frac{N\epsilon^2}{2\Sigma_{\gamma,(1,1)^2}}} \] (here \( M \) is such that \( |T_i| \leq M \) pointwise)

\[ \sum_{N=1}^{\infty} P \left[ \sum_{n=1}^{N} |T_i| > \sqrt{N} \Sigma_{\gamma,(1,1)} \epsilon \right] \leq \sum_{N=1}^{\infty} e^{-\frac{N\epsilon^2}{2\Sigma_{\gamma,(1,1)^2}}} < \infty. \]

Hence, for any \( \epsilon > 0 \), we have

\[ \sum_{N=1}^{\infty} P \left[ \sqrt{N} \left| \hat{A}_{N,\gamma} - A_{\gamma} \right| > \epsilon \right] < \infty \Rightarrow \sup_{N} \sqrt{N} \left| \hat{A}_{N,\gamma} - A_{\gamma} \right| = O_{a.s.}(1), \quad (6.35) \]

which follows from Borel-Cantelli Lemma 1 (see, e.g., Serfling (1980), p. 12), and \( O_{a.s.}(1) \) denotes bounded almost surely.

Next, we try to show that \( \sup_{N} \sqrt{N} |\hat{B}_{N,\gamma} - B_{\gamma}| \) is bounded almost surely for a fixed \( \gamma \).

Now, take an arbitrary \( \epsilon_1 > 0 \) and observe that

\[
\sum_{N=1}^{\infty} P \left[ \sqrt{N} \left| \hat{B}_{N,\gamma} - B_{\gamma} \right| > \epsilon_1 \right] = \sum_{N=1}^{\infty} P \left[ \sqrt{N} \left| \nabla Q_N^*(\theta_\gamma) \right| > NH(Q_N^*(\theta_\gamma))_{2,2} \right] > \epsilon_1 \]

(Follows from (6.29)). Here \( \nabla Q_N^*(\theta_\gamma) \) is the second component of \( \nabla Q_N^*(\theta_\gamma) \), and \( NH(Q_N^*(\theta_\gamma))_{2,2} \) is the \((2,2)\)-th element of \( D_N H(Q_N^*(\theta_\gamma)) D_N \)

\[ = \sum_{N=1}^{\infty} P \left[ \sqrt{N} \left| \nabla Q_N^*(\theta_\gamma) \right| > \Sigma_{\gamma,(2,2)}\epsilon_1 \right] \]

(Here \( \Sigma_{\gamma,(2,2)} \) is the \((2,2)\)-th element of \( \Sigma_\gamma \), and the fact follows from (6.8))

\[ = \sum_{N=1}^{\infty} P \left[ \sqrt{N} \left| \nabla \left( \frac{1}{N} \sum_{n=1}^{N} g_N(y_n - A_{\gamma} \cos(\alpha_n + \beta_n n^2) - B_\gamma \sin(\alpha_n + \beta_n n^2)) \right) \right| > \Sigma_{\gamma,(2,2)}\epsilon_1 \right] \]

\[ = \sum_{N=1}^{\infty} P \left[ \left| \nabla \left( \sum_{n=1}^{N} g_N(y_n - A_{\gamma} \cos(\alpha_n + \beta_n n^2) - B_\gamma \sin(\alpha_n + \beta_n n^2)) \right) \right| > \sqrt{N}\Sigma_{\gamma,(2,2)}\epsilon_1 \right] \]

41
\[
\begin{align*}
&\leq \sum_{N=1}^{\infty} P \left[ \sum_{n=1}^{N} |T^*_i| > \sqrt{N} \Sigma_{\gamma,(2,2)} \epsilon_1 \right] \\
&\quad \quad \quad \text{(here } T^*_i = \{ d\Psi g_N(x) | x = y_n - A_{\gamma} \cos(\alpha_{\gamma} n + \beta_{\gamma} n^2) - B_{\gamma} \sin(\alpha_{\gamma} n + \beta_{\gamma} n^2) \} \sin(\alpha_{\gamma} n + \beta_{\gamma} n^2))
\end{align*}
\]

Now, again using Bernstein inequality, we have
\[
P \left[ \sum_{n=1}^{N} |T^*_i| > \sqrt{N} \Sigma_{\gamma,(2,2)} \epsilon_1 \right] \leq e^{-\frac{N \epsilon_1^2 \Sigma_{\gamma,(2,2)}^2}{M^*}} \quad \text{(here } M^* \text{ is such that } |T^*_i| \leq M^* \text{ pointwise)}
\]
\[
\sum_{N=1}^{\infty} P \left[ \sum_{n=1}^{N} |T^*_i| > \sqrt{N} \Sigma_{\gamma,(2,2)} \epsilon_1 \right] \leq \sum_{N=1}^{\infty} e^{-\frac{N \epsilon_1^2 \Sigma_{\gamma,(2,2)}^2}{M^*}} \leq \infty.
\]

Hence, for any \( \epsilon_1 > 0 \), we have
\[
\sum_{N=1}^{\infty} P \left[ \sqrt{N} |(\hat{B}_{N,\gamma} - B_{\gamma})| > \epsilon_1 \right] < \infty \Rightarrow \sup_{N} \sqrt{N} |(\hat{B}_{N,\gamma} - B_{\gamma})| = O_{a.s.}(1), \quad (6.36)
\]
which again follows from Borel-Cantelli Lemma 1.

Now, we want to show that \( \sup_{N} N^{3/2} |(\hat{\alpha}_{N,\gamma} - \alpha_{\gamma})| \) bounded almost surely. For any \( \epsilon_2 > 0 \), let us observe that
\[
\sum_{N=1}^{\infty} P \left[ N^{3/2} |(\hat{\alpha}_{N,\gamma} - \alpha_{\gamma})| > \epsilon_2 \right] = \sum_{N=1}^{\infty} P \left[ N^{3/2} \frac{|\nabla Q_N^*(\theta_{\gamma})|_{3,3}}{N^3 H(Q_N^*(\theta_{\gamma}))_{3,3}} > \epsilon_2 \right]
\]
(Follows from (6.29)). Here \( \nabla Q_N^*(\theta_{\gamma})_3 \) is the third component of \( \nabla Q_N^*(\theta_{\gamma}) \), and \( N^3 H(Q_N^*(\theta_{\gamma}))_{3,3} \) is the (3,3)-th element of \( D_N H(Q_N^*(\theta_{\gamma})) D_N \)
\[
= \sum_{N=1}^{\infty} P \left[ N^{3/2} |\nabla Q_N^*(\theta_{\gamma})|_3 > N^3 H(Q_N^*(\theta_{\gamma}))_{3,3} \epsilon_2 \right]
\]
\[
= \sum_{N=1}^{\infty} P \left[ N^{3/2} |\nabla Q_N^*(\theta_{\gamma})|_3 > \Sigma_{\gamma,(3,3)} \epsilon_2 \right] \quad \text{(Here } \Sigma_{\gamma,(3,3)} \text{ is the (3,3)-th element of } \Sigma_{\gamma},
\]
and the fact follows from (6.8).
Now, by further application of Bernstein inequality, we have

\[
P \left[ \sum_{n=1}^{N} |T_i^{**}| > \sqrt{N(N+1)/2} \Sigma_{\gamma,(3,3)} \right] \leq e^{-\frac{N(N+1)^2/2 \Sigma_{\gamma,(3,3)}^2}{2 M^{**} \sqrt{N(N+1) \Sigma_{\gamma,(3,3)}^2}}} < \infty,
\]

(here \(M^{**}\) is such that \(|T_i^{**}| \leq M^{**}\) pointwise)

\[
\sum_{N=1}^{\infty} P \left[ \sum_{n=1}^{N} |T_i^{**}| > \sqrt{N(N+1)(1/2)} \Sigma_{\gamma,(3,3)} \right] \leq \sum_{N=1}^{\infty} e^{-\frac{N(N+1)^2/2 \Sigma_{\gamma,(3,3)}^2}{2 M^{**} \sqrt{N(N+1) \Sigma_{\gamma,(3,3)}^2}}} < \infty.
\]

Hence, we have

\[
\sum_{N=1}^{\infty} P \left[ \sum_{n=1}^{N} |T_i^{**}| > \sqrt{N(N+1)(1/2)} \Sigma_{\gamma,(3,3)} \right] < \infty \Rightarrow \sup_N N^{3/2} |(\hat{\alpha}_{N,\gamma} - \alpha_\gamma)| = O_{a.s.}(1),
\]

which follows from Borel-Cantelli Lemma 1.

Finally, we want to show that \(\sup_N N^{5/2} |(\hat{\beta}_{N,\gamma} - \beta_\gamma)|\) is bounded almost surely. For any \(\epsilon_3 > 0\), let us observe that
\[
\sum_{N=1}^{\infty} P \left[ N^{5/2} \left| \hat{\beta}_{N,\gamma} - \beta_\gamma \right| > \epsilon_3 \right] = \sum_{N=1}^{\infty} P \left[ N^{5/2} \frac{|\nabla Q_N^*(\theta_\gamma)_4|}{N^5 H(Q_N^*(\theta_\gamma))_{4,4}} > \epsilon_3 \right]
\]

(Follows from (6.29). Here \( \nabla Q_N^*(\theta_\gamma)_4 \) is the fourth component of \( \nabla Q_N^*(\theta_\gamma) \), and \( N^5 H(Q_N^*(\theta_\gamma))_{4,4} \) is the \((4,4)\)-th element of \( D_N H(Q_N^*(\theta_\gamma)) D_N \))

\[
= \sum_{N=1}^{\infty} P \left[ N^{5/2} |\nabla Q_N^*(\theta_\gamma)_4| > N^5 H(Q_N^*(\theta_\gamma))_{4,4} \epsilon_3 \right] = \sum_{N=1}^{\infty} P \left[ N^{5/2} |\nabla Q_N^*(\theta_\gamma)_4| > \Sigma_{\gamma,(4,4)} \epsilon_3 \right] \quad \text{(Here } \Sigma_{\gamma,(4,4)} \text{ is the } (4,4)\text{-th element of } \Sigma_{\gamma}, \text{ and the fact follows from (6.8)})
\]

\[
= \sum_{N=1}^{\infty} P \left[ N^{5/2} \left| \nabla \left( \frac{1}{N} \sum_{n=1}^{N} g_N(y_n - A_\gamma \cos(\alpha_\gamma n + \beta_\gamma n^2) - B_\gamma \sin(\alpha_\gamma n + \beta_\gamma n^2)) \right) \right|_4 > \frac{1}{N^{3/2}} \Sigma_{\gamma,(4,4)} \epsilon_3 \right]
\]

\[
\leq \sum_{N=1}^{\infty} \sum_{n=1}^{N} \left| T_{i}^{***} \right| > \frac{1}{N^{3/2}} \frac{n^2}{\Sigma_{\gamma,(4,4)} \epsilon_3} \left[ \text{here } T_{i}^{***} = \left\{ \frac{d}{dx} g_N(x) \right|_{x=y_n - A_\gamma \cos(\alpha_\gamma n + \beta_\gamma n^2) - B_\gamma \sin(\alpha_\gamma n + \beta_\gamma n^2))} (-A_\gamma \sin(\alpha_\gamma n + \beta_\gamma n^2)) \right] = \sum_{N=1}^{\infty} \frac{N(N+1)(2N+1)}{6 \sqrt{N}} \Sigma_{\gamma,(4,4)} \epsilon_3 \left[ \text{here } M_{\gamma} \text{ is such that } |T_{i}^{***}| \leq M_{\gamma} \text{ pointwise} \right] \leq e^{-\frac{(1/36) (N+1)^2 (2N+1)^2 \Sigma_{\gamma,(4,4)} \epsilon_3}{M_{\gamma} \sqrt{N} (N+1)(2N+1) \Sigma_{\gamma,(4,4)} \epsilon_3}}
\]

Now, by further application of Bernstein inequality, we have

\[
P \left[ \sum_{n=1}^{N} \left| T_{i}^{***} \right| > \sqrt{N} (N+1)(2N+1)(1/6) \Sigma_{\gamma,(4,4)} \epsilon_3 \right] \leq e^{-\frac{(1/36) (N+1)^2 (2N+1)^2 \Sigma_{\gamma,(4,4)} \epsilon_3}{M_{\gamma} \sqrt{N} (N+1)(2N+1) \Sigma_{\gamma,(4,4)} \epsilon_3}}.
\]
\[
\sum_{N=1}^{\infty} P \left[ \sum_{n=1}^{N} |T_{i}^{**}| > \sqrt{N}(N + 1)(2N + 1)(1/6) \Sigma_{\gamma,\epsilon,3} e_3 \right] \leq \sum_{N=1}^{\infty} e^{-\left(1/36 \right) N(N+1)^2(2N+1)^2 \gamma^2 / 3 N} < \infty.
\]

Hence, we have

\[
\sum_{N=1}^{\infty} P \left[ \sum_{n=1}^{N} |T_{i}^{**}| > \sqrt{N}(N + 1)(2N + 1)(1/6) \Sigma_{\gamma,\epsilon,3} e_3 \right] < \infty \Rightarrow \sup_{N} N^{5/2} \left| \beta_{N,\gamma} - \beta_{\gamma} \right| = O_{a.s.}(1),
\]

which again follows from Borel-Cantelli Lemma 1.

Therefore, using (6.34), (6.35), (6.36), (6.37) and (6.38) along with the fact that the range space of \( \gamma \), i.e., \( K \) is a compact set, one can conclude that

\[
\sup_{N} \sup_{\gamma \in K} \left| D^{-1}_{N}(\hat{\theta}_{N,\gamma} - \theta_{\gamma}) \right| = O_{a.s.}(1).
\]

Proof of Corollary 2.1: The assertion in Theorem 2.2 implies that

\[
\sup_{\gamma \in K} \left| D^{-1}_{N}(\hat{\theta}_{N,\gamma} - \theta_{\gamma}) \right| = O_{p}(1) \Rightarrow \sup_{\gamma \in K} \left| (\hat{\theta}_{N,\gamma} - \theta_{\gamma}) \right| = o_{p}(1) \text{ as } N \to \infty.
\]

This completes the proof.

Proof of Theorem 2.3: The proof of this theorem directly follows from the assertion in Theorem 2.1.

Proof of Theorem 2.4: The integration over \([0, 1]\) is a continuous mapping, and the direct use of the continuous mapping theorem (see, e.g., van der Vaart (1998), Theorem 2.3 in p. 7) on the assertion in Theorem 2.2 leads to the proof of this theorem.

Proof of Corollary 2.2: The result follows when \( \gamma = 0.25 \) in the result in Theorem 2.3.

Proof of Corollary 2.3: Arguing in a similar way as in the proof of Theorem 2.4 considering the interval \([0.25, 0.75]\) instead of \([0, 1]\), the proof follows.

Proof of Theorem 2.5: The supremum over \([0, 1]\) and the integration over \([0, 1]\) are continuous mappings. A straightforward application of the continuous mapping theorem

\[
\]
on the assertion of Theorem 2.2 leads to the proof of this theorem.

References

Bender, D. A., Hasselbeck, M. P., and Sheik-Bahae, M. (2006). Sensitive ultrashort pulse chirp measurement. *Optics letters*, 31(1):122–124.

Caloz, C., Gupta, S., Zhang, Q., and Nikfal, B. (2013). Analog signal processing: A possible alternative or complement to dominantly digital radio schemes. *IEEE Microwave Magazine*, 14(6):87–103.

Dhar, S. S. (2016). Trimmed mean isotonic regression. *Scandinavian Journal of Statistics*, 43(1):202–212.

Dhar, S. S. and Chaudhuri, P. (2012). On the derivatives of the trimmed mean. *Statistica Sinica*, 22(2):655–679.

Dhar, S. S., Kundu, D., and Das, U. (2019). Tests for the parameters of chirp signal model. *IEEE Transactions on Signal Processing*, 67(16):4291–4301.

Doksum, K. A. and Sievers, G. L. (1976). Plotting with confidence: Graphical comparisons of two populations. *Biometrika*, 63(3):421–434.

Efron, B. and Tibshirani, R. J. (1994). *An Introduction to the Bootstrap*. CRC Press, Chapman Hall/CRC Monographs on Statistics Applied Probability.

Fuller, W. A. (1996). *Introduction to Statistical Time Series*. John Wiley & Sons.

Gini, F., Montanari, M., and Verrazzani, L. (2000). Estimation of chirp radar signals in compound-gaussian clutter: a cyclostationary approach. *IEEE Transactions on Signal Processing*, 48:1029–1039.

Gorman, R. P. and Sejnowski, T. J. (1988). Learned classification of sonar targets using a massively parallel network. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 36(7):1135–1140.

Hsieh, J. J. and Wang, H. (2018). Quantile regression based on counting process approach under semi-competing risks data. *Annals of the Institute of Statistical Mathematics*, 70(2):395–419.
Huber, P. J. (2004). *Robust Statistics*. John Wiley & Sons.

Jureckova, J., Picek, J., and Schindler, M. (2020). Empirical regression quantile processes. *Applications of Mathematics, 65*(3):257–269.

Jureckova, J. and Welsh, A. H. (1990). Asymptotic relations between $l$- and $m$-estimators in the linear model. *Annals of the Institute of Statistical Mathematics, 42*(4):671–698.

Kim, T., Kim, H. K., and Choi, H. C. (2000). Asymptotic properties of lad estimators of a nonlinear time series regression model. *Journal of the Korean Statistical Society, 29*(2):187–199.

Koenker, R. (2005). *Quantile Regression*. Cambridge University Press.

Koenker, R. and Bassett, G. (1978). Regression quantiles. *Econometrica, 46*:33–50.

Lahiri, A., Kundu, D., and Mitra, A. (2014). On least absolute deviation estimators for one-dimensional chirp model. *Statistics, 48*(2):405–420.

Nakahira, K., Kodama, T., Furuhashi, T., and Maeda, H. (2005). Design of digital polarity correlators in a multiple-user sonar ranging system. *IEEE Transactions on Instrumentation and Measurement, 54*(1):305–310.

Nandi, S. and Kundu, D. (2004). Asymptotic properties of the least squares estimators of the parameters of the chirp signals. *Annals of the Institute of Statistical Mathematics, 56*(3):529–544.

Parker, T. (2019). Asymptotic inference for the constrained quantile regression process. *Journal of Econometrics, 213*(1):174–189.

Qu, Z. and Yoon, J. (2015). Nonparametric estimation and inference on conditional quantile processes. *Journal of Econometrics, 185*(1):1–19.

Serfling, R. (1980). *Approximation Theorems of Mathematical Statistics*. John Wiley & Sons.

Tse, S. (2005). Quantile process for left truncated and right censored data. *Annals of the Institute of Statistical Mathematics, 57*(1):61–69.

Tse, S. (2009). On the cumulative quantile regression process. *Mathematical Methods of Statistics, 18*(3):270–279.
van der Vaart, A. W. (1998). *Asymptotic Statistics*. Cambridge University Press.

Vinogradov, I. M. and Karatsuba, A. A. (1984). The method of trigonometric sums in number theory. *Trudy Matematicheskogo Instituta Imeni V. A. Steklova*, 168:4–30.

Volgushev, S. (2020). Smoothed quantile regression processes for binary response models. *Econometric Theory*, 36(2):292–330.

Volgushev, S., Chao, S., and Cheng, G. (2019). Distributed inference for quantile regression processes. *The Annals of Statistics*, 47(3):1634–1662.

Wagener, J., Volgushev, S., and Dette, H. (2012). The quantile process under random censoring. *Mathematical Methods of Statistics*, 21(2):127–141.

Warr, R. L. and Erich, R. A. (2013). Should the interquartile range divided by the standard deviation by used to assess normality? *The American Statistician*, 67(4):242–244.

Welsh, A. H. (1987). The trimmed mean in the linear model. *The Annals of Statistics*, 15(1):20–45.

Wilk, M. and Gnanadesikan, R. (1968). Probability plotting methods for the analysis of data. *Biometrika*, 55(1):1–17.

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(10):4977–4983.

Yetik, I. S. and Nehorai, A. (2003). Beamforming using fractional fourier transform. *IEEE Transactions on Signal Processing*, 51(6):1663–1668.

Yuan, Y., Chen, N., and Zhou, S. (2017). Modeling regression quantile process using monotone b-splines. *Technometrics*, 59(3):338–350.

Zwingmann, T. and Holzmann, H. (2020). Weak convergence of quantile and expectile processes under general assumptions. *Bernoulli*, 26(1):323–351.