SIMULTANEOUS QUANTILE INFERENCE FOR NON-STATIONARY LONG MEMORY TIME SERIES

WEICHI WU[1] AND ZHOU ZHOU

University of Toronto

September 11, 2014

Abstract

We consider simultaneous or functional inference of time-varying quantile curves for a class of non-stationary and long memory time series. New uniform Bahadur representations and Gaussian approximation schemes are established for a wide class of non-stationary and long memory linear processes. Furthermore, an asymptotic distributional theory is developed for the maxima of a class of non-stationary long memory Gaussian processes. With the latter theoretical results, simultaneous confidence bands for the above mentioned quantile curves with asymptotically correct coverage probabilities are constructed.

1 Introduction

The analysis of non-stationary time series has attracted considerable attention in statistics in the last three decades. See for instance Priestley (1988), Dahlhaus (1997), Neumann and von Sachs (1997), Mallat et al. (1998), Nason et al. (2000), Ombao et al. (2005) and Zhou and Wu (2009) among others. It seems that most of the established models, methodologies and theory are for short memory non-stationary processes. On the other hand, however, there have been relatively much fewer results on long range dependent processes with time varying or heterogeneous stochastic behaviors.

There is an increasing need for non-stationary and long-memory time series analysis in various applied fields, such as hydrology, geophysics, climate change, econometrics and

[1]Corresponding author. Department of Statistics, 100 St. George Street, Toronto, Ontario, M5S 3G3 Canada.
E-mail: *weichi.wu@mail.utoronto.ca*

Key words and phrases. Heterogeneity, long memory, local linear quantile estimation, bootstrap, simultaneous confidence bands

quantitative finance. On one hand, in the literature of econometrics and quantitative finance, long memory has been empirically identified as one of the stylized facts for many financial time series data. Among others, Ding et al.(1993) discussed the long memory property of power transformations of absolute stock returns. Cheung (1993) found that various exchange-rate data exhibited long memory using the Geweke-Porter-Hudak test. Recently, Rossi and de Magistris (2013) found that long-memory exists in some trading volume data. We also refer to Baillie (1994) and Henry and Zaffaroni (2003) for comprehensive reviews of the literature of long memory processes in finance and econometrics. In the literature of hydrology, Hurst (1951) found the well known Hurst effect phenomena in the geophysics record of water storage. In the literature of geophysics, Haslett and Raftery (1989) assess Ireland's wind power via space-time model with long memory. In the literature of climate change, numerous researches, such as Smith (1993), Eichner et al.(2003), Mills (2007), Mann (2011) study the long memory in surface temperature record.

On the other hand, it has long been recognized that the data generating mechanisms do not stay unchanged for many financial, geophysics, temperature time series which span for at least moderately long period of time. See for instance Cooley and Prescott (1976), Harvey (1989), Bekaert and Harvey (1995), Stock and Watson (1996), Orbe et al. (2006) and Ravn et al. (2008) for some representative papers in the finance and economics literature, Clarke(2007), Rea et al. (2011), Kärner (2002) for some representative papers in hydrology, geophysics and climate change. In the statistics literature, among others, Mercurio and Spokoiny (2004) proposed an approach for the estimation and forecasting of the time-varying volatility. Dahlhaus and Subba Rao (2006) and Fryzlewicz et al. (2008) analyzed a non-stationary version of the autoregressive and conditional heteroscedastic (ARCH) model to accommodate the time-varying nature of the return processes.

The purpose of the paper is to perform functional inference of the time varying quantile curves for a class of non-stationary long memory processes of the form

$$X_{in} = \sum_{j=0}^{\infty} a_j(t_i)\varepsilon_{i-j,n} + \mu(t_i), \quad i = 1, 2, \cdots, n, \tag{1}$$

where n is the time series length, $t_i = i/n$ and $\varepsilon_{i,n}$ are centered random variables satisfying

$$\varepsilon_{i,n} = G(t_i, \eta_i) \tag{2}$$

with η_i's independent and identically distributed. In (1), long memory is introduced by allowing coefficient functions $a_j(t)$ to decay slowly with j. And $\{X_{in}\}$ is non-stationary since functions $a_j(t)$ and $G(t, \cdot)$ vary with time t. In the sequel we shall omit the subscript n in X_{in} if no confusions arise. Let $Q_\alpha(t)$ be the αth quantile of $\{X_i\}$ at time t, $0 \leq t \leq 1$. For a fixed $\beta \in (0, 1)$, we shall construct a $100(1-\beta)\%$ asymptotic simultaneous confidence band (SCB) for $Q_\alpha(t)$; i.e., we shall find random quantities $L_{n,\alpha}(t)$ and $U_{n,\alpha}(t)$ such that

$$\lim_{n\to\infty} \mathbb{P}(L_{n,\alpha}(t) \leq Q_\alpha(t) \leq U_{n,\alpha}(t), 0 < t < 1) = 1 - \beta. \tag{3}$$

Monitoring and inference of the quantile curves is very important for risk measure and control in quantitative finance and econometrics. In particular, the high or low quantiles, depending on the context, are called value at risk (VaR) in finance. The VaR has become a widely used measure of market risk in risk management. We refer to Chapter 7 of Tsay (2005) and the monographs of Jorion (2006) and Holton (2003) for a comprehensive account of VaR in financial risk management. For non-stationary financial time series, simultaneous inference of $Q_\alpha(t)$ is a very important task as it allows one to monitor the time-varying pattern of the market risk with statistical guide and confidence.

However, constructing quantile simultaneous confidence bands for non-stationary and long memory time series is a difficult problem. To our knowledge, there are no corresponding results in the literature to date. Generally speaking, the latter problem can be solved if the following three tasks can be achieved. (i) Construct a uniform Bahadur representation for the quantile curves and establish that estimates of $Q_\alpha(t)$ can be uniformly well approximately by linear combinations of $\{X_i\}$ on $(0, 1)$. (ii) Establish that the partial sum process of the non-stationary long memory process $\{X_i\}$ can be well approximated by that of a corresponding non-stationary and long memory Gaussian process. (iii) Establish an asymptotic distribution theory for the maxima of non-stationary long memory Gaussian processes.

Task (i) relies on investigating the uniform oscillation rate of the empirical processes of $\{X_i\}$. Note that due to long memory, empirical process theory established for short memory or independent data (see for instance Zhou 2010 and Pollard 1994) cannot be applied here. For functions of stationary long memory data, Ho and Hsing (1997) proposed a deep theoretical method for an asymptotic theory. In this paper, we generalize this method to the empirical process of non-stationary long memory time series and prove a uniform Bahadur representation for the local linear quantile estimators of $Q_\alpha(t)$. The empirical process theory established here can further facilitate the asymptotic theory for a wide class of nonparametric M-estimates of non-stationary long memory processes.

Task (ii) belongs to a class of problems called Gaussian approximations or invariance principles. Invariance principles have very wide applications in statistics and probability and have received a great deal of attention in the literature. See for instance Komlös, Major and Tusnády (1975,1976), Einmahl (1987a, 1987b, 1989) and Zaitsev (2001,2002a,2002b) for some deep results for independent data and Wu and Zhou (2011) for a result on non-stationary short memory time series. To date, however, there are no results on Gaussian approximations for non-stationary long memory time series. In this paper, we utilize a representation of partial sums of (1) and establish an invariance principle with sufficiently sharp approximation rates; see Theorem 2 in Section 3.2. The established invariance principle can be of separate interest and can be useful for a large class of problems in the analysis of non-stationary and long memory data.

In the literature, the classic result to address issue (iii) is the asymptotic extreme value theory established in Bickel and Rosenblatt (1973). See for instance Härdle (1989) and Xia (1998). However, the results in Bickel and Rosenblatt (1973) are for short memory and approximately stationary Gaussian processes and cannot be directly used under the current setting. In the literature, Sun (1993) and Sun and Loader (1994) established an asymptotic extreme value theory for Gaussian random fields. In this paper we utilize the latter results and establish an extreme value theory for a class of Gaussian non-stationary and long memory processes. With the theoretical progresses on issues (i)-(iii), we construct in this paper SCB's for $Q_\alpha(t)$ with asymptotically correct coverage probabilities. The SCB's facilitate one to monitor and test the pattern and magnitude of the time-varying quantile curves which, for instance, provides useful tools for risk management of non-stationary and long memory financial time series.

The rest of the paper is organized as follows. In Section 2 we introduce some notation and assumptions that are used throughout the paper. The main theoretical results on the Bahadur representations, Gaussian approximations and asymptotic distribution for Gaussian process extreme values are established in Section 3. Issue of bandwidth selections will also be discussed in Section 3. A discussion will be provided in Section 5. Finally, the theoretical results are proved in Section 6.

2 Preliminaries

2.1 Definitions

Traditionally, for second order stationary process X_i, it possesses long memory if $\sum_{j=-\infty}^{\infty}\Gamma(j)=\infty$, where $\Gamma(j) := cov(X_1, X_{1+j})$ is the autocovariance function. For non-stationary time series, we extend the latter classic definition of long memory and define the following uniform long memory property of non-stationary time series as follows:

Definition 1. *We say a triangle array of non-stationary time series $\{X_{in}\}_{i=1}^{n}$ is uniformly long memory if for every positive integer i,*

$$\lim_{n\to\infty}\sum_{j=-\infty}^{\infty} cov(X_{in}, X_{(i+j)n}) = \infty. \tag{4}$$

In the above definition, we set $X_{in} = 0$ if $i \le 0$ or $i > n$. Elementary calculations show that the process $\{X_{in}\}$ defined in (1) is a uniformly long memory process according to the above definition if the coefficient functions $a_j(t)$ satisfy $c \le a_j(t)/j^{d(t)} \le C$ for some positive and finite constants c and C and $1/2 < d(t) < 1$.

2.2 Assumptions

Suppose we observe

$$X_{in} := \sum_{j=0}^{\infty} a_j(i/n)\varepsilon_{i-j,n} + \mu(\frac{i}{n}), 1 \le i \le n, \tag{5}$$

where the innovations $\varepsilon_{i,n} = G(i/n, \eta_i)$, $G(\cdot)$ is a measurable function, $\{\eta_i\}_{i=-\infty}^{\infty}$ are i.i.d. random variables, and $\mathbb{E}(G(t, \eta_i)) = 0$ for all $i \in \mathbb{Z}$, $t \in (0, 1)$. Observe that $X_{in} =$

$X_{in}(i/n)$ with

$$X_{in}(t) = \sum_{j=0}^{\infty} a_j(t) G(t - j/n, \eta_{i-j}) + \mu(t), 0 \leq t \leq 1.$$

Remark 1. *Note that in* (5)*, the innovations* $\varepsilon_{i,n} = G(i/n, \eta_i)$ *are independent but non-identically distributed. We argue that allowing the innovations of the process to be non-stationary is very important for quantile analysis of non-stationary time series. The reason is that this makes our model flexible and allows the marginal distributions of* X_{in} *to change arbitrarily over time. To see this, just compare two simple models:*

$$X_{in}(t) = \sum_{j=0}^{\infty} \frac{a(t)}{j^{\beta}} G(t - j/n, \eta_{i-j}), \tag{6}$$

$$Y_{in}(t) = \sum_{j=0}^{\infty} \frac{a(t)}{j^{\beta}} \zeta_{i-j}, \tag{7}$$

where ζ_i *are i.i.d. random variables,* $\beta > 1/2$*. Let* $Q_{Y,\alpha}(t)$ *represents the* α_{th} *quantile curve of* $Y_{in}(t)$*. Define* $Z = \sum_{j=0}^{\infty} \frac{1}{j^{\beta}} \zeta_{i-j}$*, and* Z_α *be* Z*'s* α *quantile. It's obvious that* $Q_{Y,\alpha}(t) = a(t) Z_\alpha$*. Let* $0 < a < b < c < 1$ *be real numbers. Then* $Q_{Y,a}(t) - Q_{Y,b}(t) \propto a(t)$*, and* $Q_{Y,b}(t) - Q_{Y,c}(t) \propto a(t)$*, where notation* $\propto$ *means "be proportional to". Then we have*

$$Q_{Y,a}(t) - Q_{Y,b}(t) \propto Q_{Y,b}(t) - Q_{Y,c}(t). \tag{8}$$

The above restriction on the shapes of the quantile curves makes model (7) *not useful for quantile analysis in many cases. In particular, under model* (7)*, if the* a_{th} *and* b_{th} *quantile curves stay unchanged across time for some* $a < b$*, then (8) implies that the* c_{th} *quantile curve should also be a constant function over time for any* $c \in (0, 1)$*. However, in many practical situations, it is possible that some quantile curves stay constant while others exhibit interesting patterns of changes over time. On the other hand, note that the set up for* X_{in} *in* (5) *does not post any restrictions on the shapes of the quantile curves.*

Our target is to estimate the α_{th} quantile $Q_{\alpha,n}(t)$ of $X_{in}(t)$. We have several assumptions as follows:

(A0) For fixed α, $Q_{\alpha,n}(t)$ and $Q'_{\alpha,n}(t) := \frac{\partial Q_{\alpha,n}(t)}{\partial t}$ are bounded on [0,1], and have bounded derivative.

(A1) There exists postive constant C such that $\|G(t,\eta_i)\|_p \leq C$, and for any $i \in \mathbb{Z}$, $t,s \in (0,1)$, $\|G(t,\eta_i)-G(s,\eta_i)\|_p \leq C|t-s|$ for some $p \geq 2$, where $\|\cdot\|_p = \{\mathbb{E}[|\cdot|^p]\}^{1/p}$. $\mathbb{E}(G(t,\eta_i)) = 0$, $Var(G(t,\eta_i)) = \sigma(t)$, with $|\sigma'(t)|$ bounded for all $i \in \mathbb{Z}$, $t \in (0,1)$.

(A2) Let $g_{i,n}(x)$ be the density of $\varepsilon_{i,n}, g_{i,n} \in \mathcal{C}^r(-\infty,\infty)$,where $\mathcal{C}^r(-\infty,\infty)$ represents r times differentiability in $\mathbb{R}$. Let$\frac{\partial^r g_{i,n}(x)}{\partial x^r} := g_{i,n}^{(r)}(x)$. We require that $|g_{i,n}^{(r)}(x)|$ is bounded and integrable for $r = 0,1,...,l$, $l \geq 3$.

(A3) $|a_j(t)| = O(\frac{1}{(j+1)^\gamma})\ \forall t \in [0,1], j = \{0,1,2,..\infty\}$, where $1/2 < \gamma < 1$. We also assume that $a_j(t)$ has derivative $\dot{a}_j(t) := \frac{\partial a_j(t)}{\partial t}$, such that $\dot{a}_j(t) = O(\frac{1}{(j+1)^\gamma})$. Without loss of generality, let $a_0(t) \equiv 1$ for all $t \in [0,1]$.

(A4) $f_n(t,x) \in \mathcal{C}([0,1]\times\mathbb{R})$, where $f_n(t,\cdot)$ is the density of $X_{in}(t)$. Also $\frac{\partial}{\partial t}f_n(t,x)$ and $\frac{\partial}{\partial x}f_n(t,x)$ are bounded.

(A5) $K(\cdot) \in \mathbf{K}$,where $\mathbf{K}$ is the collection of density functions which is symmetric with support [-1,1] and is $\mathcal{C}^1[-1,1]$. We write $K_{b_n}(\cdot) = K(\cdot/b_n)$ for short.

Note that if we assume (A5), then we have

$$\begin{aligned}\Sigma_n(t) :&= \sum_{i=1}^n (1,(i/n-t)/b_n)^T(1,(i/n-t)/b_n)K_{b_n}(i/n-t) \\ &= nb_n\mu_k + O(1) \text{ uniformly on any closed interval of (0,1).} \qquad (9)\\ &\text{where } \mu_k = diag(1, 2\int_0^1 u^2K(u)du) \text{ which is a } 2\times 2 \text{ diagonal matrix.}\end{aligned}$$

Also we define $\mu_2 = 2\int_0^1 u^2K(u)du$.

Remark 2. : *Condition A0) puts some requirements on the smoothness of $Q_{\alpha,n}(t)$ in order to perform the local polynomial quantile regression. A1) and A2) make some assumptions on the tail behavior of the innovations $\{\varepsilon_{i,n}\}_{i=-\infty}^\infty$ for techniqical convenience. Furthermore, by A1), the existence of $\sigma'(t)$ implies the boundness of $|\sigma'(t)|$. (Hence we could also make assumptions on the existence of $\sigma'(t)$). Condition A3) characterizes the long memory structure in this paper. The differentiability of time varying $a(t)$ actually* *makes the non-stationary time series locally stationary. In particular, if we consider the subseries of $\{X_{in}\}$ for which the time t is near some $t_0 \in [0,1]$, then the subseries are approximately stationary. Observe that our setting admits the FARIMA$(0,d,0)$ model. A3) implies that $d \leq 1-\gamma$. Condition A4) makes some smoothness assumptions on the density of series $X_{in}(t)$. This type of relationship has been investigated for stationary long range dependence. In the literature of non-stationary process, a similar but slightly weaker assumption is made in the non-stationary short-memory settings in Zhou and Wu (2009). We shall provide proposition 1 to state the relationship between condition A1), A3) and A4). Condition A5) makes mild assumptions on the kernel function $K(\cdot)$, and we also give out the consequent convergence of $\Sigma_n(t)$ in equation (9).*

We also provide the following proposition, which shows that under certain conditions, A1), A3) imply A4).

Proposition 1. *Suppose conditions A1), A3) hold. Let $f_{G(t,\eta_i)}(x)$ be the density of $G(t,\eta_i)$. Assume $\frac{\partial^2}{\partial x \partial t} f_{G(t,\eta_i)}(x)$ exists and bounded. Suppose i) $\frac{\partial}{\partial x} f_{G(t,\eta_i)}(x)$ is bounded for $t \in [0,1]$, $x \in \mathbb{R}$. Denote $A(x,t) = \lim_{t\to s} \frac{\partial}{\partial x}[\frac{\mathbb{P}\{G(t,\eta_i)\leq x\} - \mathbb{P}\{G(s,\eta_i)\leq x\}}{t-s}]$. Moreover, assume that ii) $A(x,t)$ exists and is bounded for $x \in \mathbb{R}$, $t \in [0,1]$. Then A4) holds.*

3 Main Results

Let $F_n(t,x) = \mathbb{P}(X_{in}(t) \leq x)$. Define the quantiles $Q_{\alpha,n}(t) = \inf_x\{F_n(t,x) \geq \alpha\}$. We estimate $Q_{\alpha,n}(t)$ and $Q'_{\alpha,n}(t)$ by

$$(\hat{Q}_{\alpha,n,b_n}(t), \hat{Q}'_{\alpha,n,b_n}(t)) = \arg\min_{\beta_0,\beta_1} \sum_{i=1}^{n} \rho_\alpha(X_{in} - \beta_0 - \beta_1(i/n - t))K_{b_n}(i/n - t), \tag{10}$$

where $\rho_\alpha(x) = \alpha x^+ + (1-\alpha)(-x)^+$ is the check function (Koenker (2005)). Equation (10) defines the local linear quantile estimators. Let $\Psi_\alpha(x) := \alpha - I(x \leq 0)$ be the left derivative of $\rho_\alpha(x)$. Define $\hat{\theta}_{\alpha,n}(t) = (\hat{\theta}_{\alpha,n,1}(t), \hat{\theta}_{\alpha,n,2}(t))^T := (\hat{Q}_{\alpha,n}(t) - Q_{\alpha,n}(t), b_n(\hat{Q}'_{\alpha,n}(t) -$

$Q'_{\alpha,n}(t)))^T$. Let $\mathbf{z}_{i,n}(t) = (1, (i/n - t)/b_n)^T$. For $\theta = (\theta_1, \theta_2)^T$, let

$$S_{\alpha,n}(t,\theta) = \sum_{i=1}^{n} \Psi_\alpha(X_{in} - Q_{\alpha,n}(t) - (i/n - t)Q'_{\alpha,n}(t) - \theta^T \mathbf{z}_{i,n}(t)) K_{b_n}(i/n - t)\mathbf{z}_{i,n}(t). \tag{11}$$

Write $S_{\alpha,n}(t) = S_{\alpha,n}(t,(0,0)^T)$, $\hat{v}_{\alpha n}(t) = (\hat{v}_{\alpha n1}(t), \hat{v}_{\alpha n2}(t))^T := \Sigma_n^{1/2}(t)(\hat{\theta}_{\alpha,n,1}(t), \hat{\theta}_{\alpha,n,2}(t))^T$. We also write $\|\cdot\| := \|\cdot\|_2$. Furthermore, for series X_n, Y_n, denote $\lim_{n\to\infty} \frac{X_n}{Y_n} = 1$ by $X_n \asymp Y_n$ for short. We write $X_n = O_p(Y_n)$ if X_n is bounded by Y_n in probability, and $X_n = o_p(Y_n)$ if $X_n/Y_n \to_p 0$. To simplify notation, we define that, for vector $u = \{u_1, .., u_n\}, v = \{v_1, ..., v_n\}$, $\langle u, v\rangle = \sum_{i=1}^n u_i v_i$, and $\|u\|_2^2 = \sum_{i=1}^n u_i^2$. Define $\lfloor x \rfloor$ to be the largest integer which is smaller than or equal to x. For two random series X_n and Y_n, $X_n \sim Y_n$ if $X_n = O_p(Y_n)$ and $Y_n = O_p(X_n)$. In the proofs, if there is no confusion, write $Q_\alpha(\cdot)$ instead of $Q_{\alpha,n}(\cdot)$ for short. Without loss of generality, we just set $\alpha = 0.5$ and thus omit the subscript α of $Q_\alpha(\cdot)$ and $\Psi_\alpha(\cdot)$, etc. Let $\mathcal{B}$ be the lag operator.

3.1 Uniform Bahadur Representation

The Bahadur representation asymptotically approximates the regression estimators by certain linear forms of the data. See for instance He and Shao (1996), Koenker (2005), Wu (2007) among others. In the literature of local polynomial quantile regression, Chaudhuri (1991) provides a Bahadur representation for i.i.d. d dimensional observations. Zhou and Wu (2009) provides a Bahadur representation for non-stationary series with short range dependence. For linear models with stationary long memory and heavy tail errors, Zhou and Wu (2010) provides a Bahadur representation for regression parameter estimated by general convex checking function ρ (which includes the OLS and quantile regression). For functional of Gaussian dependent sequences, Coeurjolly (2008) obtains Bahadur representation of its sample quantiles. In the following, we shall provide a uniform Bahadur representation of the local linear quantile estimators for non-stationary, long memory processes:

Theorem 1. *Let $T_n = [\delta b_n, 1 - \delta b_n]$, where $\delta > 1$ is a constant. Assume $f_n(t,x) > 0$ for any $t \in (0,1)$, $x \in \mathbb{R}$, $b_n \to 0$, $nb_n \to \infty$, $nb_n/\log^2 n \to \infty$, $(nb_n)^{1/2-\gamma}(b_n)^{-1/p} \to 0$, then* *under conditions (A0)-(A5), we have*

$$\begin{aligned}&\sup_{t\in T_n}|f_n(t,Q_{\alpha,n}(t))\mu_{\mathbf{k}}\hat{\theta}_{\alpha,n}(t)-S_{\alpha,n}(t)/nb_n|\\&=O_p\left((\pi_n)^{1/2}\log n/\sqrt{nb_n}+(nb_n)^{1/2-\gamma}\pi_n b_n^{-1/p}+b_n\pi_n+(\pi_n)^2\right).\end{aligned}\tag{12}$$

where $\pi_n=(nb_n)^{-1/2}\left(\log n+(nb_n^5)^{1/2}+(nb_n)^{1-\gamma}b_n^{-1/p}\right)$.

Theorem 1 asserts that the uniform probabilistic oscillations of $\hat{Q}_{\alpha,n}(t)$ can be well approximately by those of $S_{\alpha,n}(t)$ which has a much simpler mathematical form. Consequently, Theorem 1 enables us to construct the simultaneous confidence bands of $Q_\alpha(t)$ over $t\in T_n$ via a Gaussian approximations to $\{S_{\alpha,n}(t),t\in T_n\}$. The Gaussian approximation can be obtained by the following Theorem 2 and Theorem 3.

3.2 Gaussian approximation

Theorem 2. *Under conditions (A1)-(A5), on a possibly richer probability space, there exist* $Y_{kn}=\sum_{j=0}^{\infty}a_j(k/n)\sigma(\frac{k-j}{n})\varsigma_{k-j}+\mu(k/n)$ *where* ς_i*'s are i.i.d.* $N(0,1)$*, such that*

$$\max_{1\le s\le n}|\sum_{k=1}^{s}(X_{kn}-Y_{kn})|=O_p(n^{1+\nu(1/2-\gamma)}),$$

where $\nu=\frac{1}{1/2+1/p}$.

This theorem is of general interest. It provides a Gaussian approximation result for the partial sum processes of a class of non-stationary and long memory processes. The Gaussian approximation schemes or invariance principles are powerful tools and are widely applied in statistics and probability. Among others, Komlós, et al., (1975, 1976) reached the optimal rate for strong approximation of the partial sum of independent random variables. Zaitsev (2001, 2002a, 2002b) extend the previous univariate results to the multi-dimensional case. In the context of non-stationary processes, Wu and Zhou (2011) acquired a Gaussian approximation result of the partial sums with O_p bounds. For stationary and long memory process, Wang et al. (2003) proposed a strong approximation result. For more details about strong approximation, see Csörgö and Révész (1981) and the references therein. The following theorem, which is proved with the help of Theorem

2, enables us to uniformly approximate the estimated quantile curves by non-stationary and long memory Gaussian processes:

Theorem 3. *Suppose the conditions of Theorem 1 hold. Then on a possibly richer probability space, there exists a sequence of independent standard normal random variables* $\{\vartheta_i\}_{i=-\infty}^{\infty}$*, such that, for* $V_{in} = \sum_{j=0}^{\infty} a_j(\frac{i}{n})\vartheta_{i-j}$*, and* $\frac{\varsigma_n}{(nb_n)^{1/2-\gamma}} \to 0$ *as* $n \to \infty$*, then*

$$\sup_{t\in T_n} |f_n(t, Q_{\alpha,n}(t))(\mu_{\mathbf{k}}\hat{\theta}_{\alpha,n}(t) - \frac{\sigma(t)}{nb_n}\sum_{i=1}^{n} V_{in}K_{b_n}(i/n-t)\mathbf{z}_{i,n}(t) - \frac{b_n^2 Q''_{\alpha,n}(t)(\mu_2, 0)^T}{2})| = O_p(\varsigma_n), \tag{13}$$

where

$$\varsigma_n = \zeta_n + K_n^p/nb_n, K_n^p = nb_n^4 + \log n\sqrt{nb_n} + b_n^{-1/p}g_n + (b_n)^{1-\frac{1}{p}}(nb_n)^{3/2-\gamma},$$
$$g_n = (nb_n)^{2-2\gamma}(\log(nb_n)(\mathbf{1}\{\gamma = 3/4\}) + \mathbf{1}\{\gamma < 3/4\}) + (nb_n)^{1/2}\mathbf{1}\{\gamma > 3/4\}$$

and $\mathbf{1}(\cdot)$ *is indicator function ,*

$$\zeta_n = (\pi_n)^{1/2}\log n/\sqrt{nb_n} + (nb_n)^{1/2-\gamma}\pi_n b_n^{-1/p} + b_n\pi_n + (\pi_n)^2 + n^{\frac{1/2-\gamma}{1/2+1/p}}/(b_n) + n^{1/2-\gamma}b_n^{-1/p},$$
$$\pi_n = (nb_n)^{-1/2}\left(\log n + (nb_n^5)^{1/2} + (nb_n)^{1-\gamma}b_n^{-1/p}\right).$$

This theorem can be shown by the results of Theorem 8, Lemma 5 and Theorem 2. As we expected, if we assume that in $[\delta b_n, 1-\delta b_n]$, the density $f_n(t, Q_{\alpha,n}(t))$ is bounded from below by a strictly positive number, then after canceling this quantity in both sides of equation (13), we have an approximation of $\hat{Q}_{\alpha,n}(t)$ independent of nuisance function $f_n(t, Q_{\alpha,n}(t))$. This is different from the short-memory case where it is shown that the SCB depends on $f_n(t, Q_{\alpha,n}(t))$. For stationary long memory data, similar results was obtained by Csörgö and Kulik (2008), among others. Once we establish Theorem 3, we find that the bias of $\hat{Q}_{\alpha,n}(t)$ is of the order $O(b_n^2)$ while the standard deviation of $\frac{1}{nb_n}\sum_{i=1}^{n} V_{in}K_{b_n}(i/n - t)\mathbf{z}_{i,n}(t)$ is of the order $O_p((nb_n)^{1/2-\gamma})$. Elementary calculations show that the optimal b_n to minimize the MSE of the estimates should be $b_n \sim n^{\frac{1/2-\gamma}{3/2+\gamma}}$, which lead the convergence rate

$$\varsigma_n = b_n^{3-1/p} + b_n^{\frac{1}{\gamma-1/2}-1/p}\log n + b_n^{\frac{3/2+\gamma}{1/2+1/p}-1} + b_n^{3/2+\gamma-1/p}. \tag{14}$$

Let $p \to \infty$. We find that ς_n will approach b_n^2, the order of MSE, if γ gets close to 0.5 or 1. In practice, Theorem 3 can be used to construct the SCB by simulating a large sample of i.i.d. copies of $\{\frac{1}{nb_n}\sum_{i=1}^n V_{in}K_{b_n}(i/n-t)\mathbf{z}_{i,n}(t)\}$ and calculate the empirical maximum deviations of the simulated sample. Theoretically, we will explore the limiting distributions of the SCB in the next section.

3.3 Maximum deviation

Many researchers have done excellent investigation on the maximum deviations of Gaussian processes. For instance, the extreme Gumble distribution for stationary Gaussian processes was obtained by Berman (1971) in i.i.d settings; Bickel and Rosenblatt (1973) is a good reference for this context and it also concludes a limiting distribution for maximum deviation of a type of non-stationary Gaussian process. Sun and Loader (1994) acquires a first order approximation for a general type of Gaussian process. In the next theorem, we find a limiting confidence band by referring to Sun and Loader's results and techniques:

Theorem 4. *Suppose $K(x)$ is non-decreasing when $x \le 0$, non-increasing when $x > 0$. $K(x)$ has bounded non-increasing first order derivative on $[0,1]$. (a): Suppose $a_j(t)$satisfies $\exists\ 0 < q \le Q < \infty$, such that $\frac{q}{(j+1)^\gamma} \le a_j(t) \le \frac{Q}{(j+1)^\gamma}$ $\forall j \ge 0$.*
Let $T_n = (\delta b_n, 1-\delta b_n)$ for some $\delta > 1$, $w = (\delta-1)/2$, $M = \sup_{x\in[-1,1]}|K'(x)|$, $\{\vartheta_i\}_{i=1}^\infty$ be a series of i.i.d Normal(0,1), and $V_{in} = \sum_{j=0}^\infty a_j(i/n)\vartheta_{i-j}$. Suppose (b) conditions of Theorem 3 hold, then we have,

$$\lim_{n\to\infty}\mathbb{P}\{\sup_{t\in T_n}\frac{nb_n|\hat{Q}_{\alpha,n}(t) - Q_{\alpha,n}(t) - b_n^2 Q''_{\alpha,n}(t)\mu_2/2|}{\sigma(t)\|\sum_{i=1}^n V_{in}K_{b_n}(i/n-t)\|} > \sqrt{2\log\frac{\kappa_n}{\pi\tau}}\} = \tau, \tag{15}$$

where κ_n satisfies: $\exists\ 0 < C_1 < C_2 < \infty$, such that

$$C_1/b_n \le \kappa_n \le C_2/b_n \tag{16}$$

and $1-\tau$ is the nominal coverage probability.

When n is large enough, we can find an explicit bound for C_1 and C_2 in (16): Let $a^* = \text{argmax}_{a\in[0,1]}(K'(a))^2(1-a)^{3-2\gamma}$, and define,

$$G(K(.)) = \frac{\mathbb{L}_{1-\gamma}}{(1-\gamma)^2}\int_{-1}^{1}\int_{-1}^{1} K(x)K(y)|x-y|^{1-2\gamma}dxdy,$$
$$\text{where } \mathbb{L}_{1-\gamma} = \frac{1}{3-2\gamma} + \int_0^\infty ((x+1)^{1-\gamma} - x^{1-\gamma})^2 dx.$$

Then we have,

$$C_1 = \frac{K'(a^*)^2 q^2 (1-a^*)^{3-2\gamma}}{(1-\gamma)^2(3-2\gamma)Q^2 M a_1 G(K(\cdot))^{1/2}},$$
$$C_2 = \frac{a_1 M Q}{qG(K(\cdot))^{1/2}},$$
$$\text{where } \ a_1 = (\frac{4w^{1-2\gamma}}{2\gamma-1} + \frac{w(2+w)^{2-2\gamma}}{(1-\gamma)^2} + \frac{2^{3-2\gamma}}{(3-2\gamma)(1-\gamma)^2})^{1/2}.$$

In Theorem 4, we make the assumption $\exists\ 0 < q \leq Q < \infty$, such that $\frac{q}{(j+1)^\gamma} \leq a_{j+1}(t) \leq \frac{Q}{(j+1)^\gamma}$ $\forall t \in T_n$ to make sure that the norm of the partial sum of V_{in} goes to infinity in a fairly stable rate as sample size increases. In general, due to the non-stationarity, the exact value of κ_n is hard to evaluate. However, the theorem obtains a bound for $\kappa_n b_n$, consequently assures the order of the width of the simultaneous confidence bands. The term $\sigma(t)$ can be estimated, say, by the local linear estimators. The term $\|\sum_{i=1}^n V_{in}K_{b_n}(i/n-t)\|$ determines the width of the SCB. Under our settings, we have the next corollary on the order of $\|\sum_{i=1}^n V_{in}K_{b_n}(i/n-t)\|$:

Corollary 1. *Under the conditions of Theorem 4,* $\|\sum_{i=1}^n V_{in}K_{b_n}(i/n-t)\| \sim (nb_n)^{3/2-\gamma}$.

A nicer form of κ_n and $\|\sum_{i=1}^n V_{in}K_{b_n}(i/n-t)\|$ can be obtained under slightly stronger assumptions:

Lemma 1. *Let* $\Gamma(i,j) = cov(V_{in}, V_{jn})$, *Let* b_n *satisfy* $b_n \to 0$, *and* $nb_n \to \infty$, *Then if (a) : the conditions of Theorem 3 hold; b)* $\Gamma(i,j) \asymp \breve{a}(i/n, j/n)|i-j|^{1-2\gamma}$, *where* $\breve{a}(x,y)$ *is Lipschitz continuous in both* x *and* y *for* $x, y \in (0,1)^2$, *and* $g(t,t)$ *has strictly positive lower bound and finite upper bound. Then we have*

$$\|\sum_{i=1}^n V_{in}K_{b_n}(i/n-t)\| \asymp \breve{a}^{1/2}(t,t)(nb_n)^{3/2-\gamma}(\int_{-1}^{1}\int_{-1}^{1}|x-y|^{1-2\gamma}K(x)K(y)dxdy)^{1/2},$$

and

$$\kappa_n \asymp \frac{1}{b_n}\left(\frac{\int_{-1}^{1}\int_{-1}^{1}|x-y|^{1-2\gamma}K'(x)K'(y)dxdy}{\int_{-1}^{1}\int_{-1}^{1}|x-y|^{1-2\gamma}K(x)K(y)dxdy}\right)^{1/2} := \frac{D}{b_n}$$

then we have

$$\lim_{n\to\infty}\mathbb{P}\{\sup_{t\in T_n}\frac{|\hat{Q}_{\alpha,n}(t)-Q_{\alpha,n}(t)-b_n^2Q''_{\alpha,n}(t)\mu_2/2|}{\breve{a}^{1/2}(t,t)\sigma(t)(\int_{-1}^{1}\int_{-1}^{1}|x-y|^{1-2\gamma}K(x)K(y)dxdy)^{1/2}(nb_n)^{1/2-\gamma}} > \sqrt{2\log\frac{D}{b_n\pi\tau}}\} = \tau. \tag{17}$$

Remark 3. *Note that condition (a) of Theorem 4 assures that $\kappa_n \sim 1/b_n$. In Lemma 1, we make assumptions on covariance structure which obtains the limit of κ_n. Thus we don't need (a), but only (b) of Theorem 4 to support Lemma 1.*

The following Corollary shows that if the functions $a_j(t)$ can be factorized as $a_j(t) = a(t)g(j)$, with innovation $\{\varepsilon_j\}_{-\infty}^{\infty}$ *i.i.d* with finite p moments, then a Gumble limiting distribution can be achieved.

Corollary 2. *Suppose $a_j(t)$ satisfy $a_j(t) = a(t)g(j)$, $g_j = \frac{(1-\gamma)}{(j+1)^{\gamma}}(1+O(1/j))$. Suppose the innovation $\varepsilon_i's$ are i.i.d with variance 1, i.e, $X_{in}(t) = \sum_{j=0}^{\infty}a_j(t)\varepsilon_{i-j}$, and the kernel function $K(\cdot)$ have bounded first order derivative on $[0,1]$. In addition, define:*

$$\kappa_K^2 = \int_R\int_R K(x)K(y)|x-y|^{1-2\gamma}dxdy. \tag{18}$$

$$D_K = \int_R\int_R K'(x)K'(y)|x-y|^{1-2\gamma}dxdy. \tag{19}$$

For $m \geq 3$, define $B_K(m) = \sqrt{2\log m} + \frac{1}{2\sqrt{2\log m}}(\log C_k - 2\log 2 - 2\log\pi)$, $C_K = D_K/\kappa_k^2$. Suppose the conditions of Theorem 1 hold. In addition, suppose

$$n^{\gamma-1}\log n/b_n^{3/2-\gamma} = o((\log n)^{-1/2}).$$

Then we have:

$$\lim_{n\to\infty}\left(\mathbb{P}\Big[\max_{t\in[b_n,1-b_n]}\left|(1-\gamma)\left(\frac{\hat{Q}_{\alpha,n}(t)-Q_{\alpha,n}(t)-b_n^2 Q''_{\alpha,n}(t)\mu_2/2}{\mathbb{L}_{1-\gamma}^{1/2}\kappa_k(nb_n)^{1/2-\gamma}a(t)}\right)\right|-B_K(1/b_n)$$

$$\leq \{2\log(1/b_n)\}^{-1/2}u\Big]\right)=exp\{-2exp(-u)\}, \qquad (20)$$

where $\mathbb{L}_{1-\gamma}=\frac{1}{3-2\gamma}+\int_0^\infty((x+1)^{1-\gamma}-x^{1-\gamma})^2dx.$

3.4 Bandwidth Selection

In theorem 4, the bias term includes $Q''_{\alpha,n}(t)$, and the variance term includes $\|\sum_{i=1}^n X_iK_{b_n}(i/n-t)\|$. Both of them are very hard to estimate due to the non-stationarity of our model. In our simulation, we use a slightly modified data-driven bandwidth selection procedure, which is originally suggested by Cai and Xu (2009). Specifically, we shall find the bandwidth b_n by minimizing the so called 'non-parametric version of bias-correct AIC':

$$AIC(b_n)=\log\{\hat{\sigma}^2_{\alpha,b_n}\}+2(p_{b_n}+1)/(n-p_{b_n}+2). \qquad (21)$$

Where $\hat{\sigma}^2_{\alpha,b_n}=\sum_{i=\delta b_n}^{1-\delta b_n}\rho_\alpha(X_i-\hat{Q}_{\alpha,n}(i/n))/(1-2\delta b_n)$, δ is a constant which is larger than 1 (we take $\delta=1.1$ in our simulations). The 'degrees of freedom' p_{b_n} can be calculated as follows: for any fixed bandwidth b_n,

(i) Use local linear estimator to get estimates of quantiles $\hat{Q}_{\alpha,n}(i/n)$ for each $i\in[1,n]$.

(ii) For each i, calculate $\eta_i(s)=\mathbf{1}(X_i\leq\hat{Q}_{\alpha,n}(s/n)+1/\sqrt{nb_n})-\mathbf{1}(X_i\leq\hat{Q}_{\alpha,n}(s/n))$ for $s=1,2,..,n$.

(iii) For each $s=1,2,..,n$, calculate $\Delta(s):=\sum_{i=1}^n\eta_i(s)K_{b_n}(i/n-s/n)/\sqrt{nb_n}$. Following Cai and Xu (2009), we calculate $\Delta^{-1}(s)$.

(iv) We set $p_{b_n}=\sum_{i=\delta b_n}^{1-\delta b_n}K(0)/(\Delta(i)nb_n)$.

4 Examples

In the following, suppose the assumptions of innovation, including (A1),(A2), and i), ii) of proposition 1, and the assumption of kernel $K(\cdot)$, (A5) hold. We shall check condition (A3) so that our theories apply to the general examples.

Example 1. *Consider the fractionally integrated model:* $(1-\mathcal{B})^d X_{i,n}(t) = a(t)G(i/n, \eta_i)$, *where* $G(i/n, \eta_i)$ *satisfies condition (A1), (A2), i), ii) of proposition 1.* $a(t)$ *is a smooth function of* t*, which has bounded first derivative. Let* $\gamma = 1-d$*,* p *and* γ *satisfy condition (A3). Then the theory established in this paper can be used to obtain the simultaneous confidence bands of the quantile curves of* $X_{i,n}$*. Note that if* $G(t,\eta_i) \equiv \eta_i$*, where* $\{\eta_i\}_{i=1}^{\infty}$ *are i.i.d with mean* 0 *and variance* 1*, then we get a time varying Fractional ARIMA(0,d,0) model.*

Example 2. *Consider the locally stationary Fractional ARIMA(p,d,q) Model:* $\Phi^p(\mathcal{B},t)(1-\mathcal{B})^d X_n(t) = \Theta^q(\mathcal{B},t)\varepsilon_n$*, where* $\varepsilon_i's$ *are i.i.d random variables with mean* 0 *and variance* 1*,* $\Phi^p(z,t) = 1+\phi_1(t)z+...+\phi_p(t)z^p$ *and* $\Theta^q(z,t) = 1+\theta_1(t)z+...+\theta_q(t)z^q$ *are polynomials with degrees of freedom* p *and* q*, respectively, and* $0 < d < 1/2$*. Suppose that the coefficients of the two polynomials are twice differentiable in* t*. Then define a polynomial with* $p+q$ *degrees of freedom:*

$$\Xi^{p+q}(z,t) = (\dot{\Theta^q})(\mathcal{B},t)\Phi(\mathcal{B},t) - \Theta^q(\mathcal{B},t)\dot{\Phi}(\mathcal{B},t).$$

where the notation $'.'$ *means taking partial derivative with respect to time* t*. Suppose that for all* $t \in [0,1]$*,* $\Phi^p(z,t)$ *and* $\Theta^q(z,t)$*,* $\Xi^{p+q}(z,t)$ *do not have same roots, and* $\Phi^p(z,t)$ *does not have roots in the unit disk* $\{|z| \leq 1\}$*. Let* $G(z,t) = \frac{\Theta^q(z,t)}{\Phi^p(z,t)}(1-z)^{-d} := \sum_{j=0}^{\infty} c_j(t)z^j$*, then* $G(z,t)$ *is analytic in the circle* $\{|z| \leq R(t)\}$ *for some* $R(t) > 1$*. Now suppose there exists a number* Q *such that* $1 < Q < R(t)$ *for all* $t \in [0,1]$*. Consequently,* $\dot{G}(z,t) := \frac{\partial}{\partial t}G(z,t)$ *is also analytic with convergence radius* $r(t)$ *for some* $r(t) > 1$*. We also assume that there is a number* q *such that* $1 < q < r(t)$ *for all* $t \in [0,1]$*. Then condition A3) is satisfied with* $\gamma = 1-d$*. In addition, if the innovation has moment* p *such that* $\gamma < \frac{2+1/(2p)}{2+1/p}$*, then our theory for the quantile applies in this case.*

To see that condition A3) holds for example 2, we first carefully check Kokoszka and Taqqu (1994) and conclude that $|c_j(t)| \leq KQ^{-j}$ for all t and some large enough constant K.

Then apply lemma 3.2 in Kokoszka and Taqqu (1995), we conclude that $|c_j(t)| \leq Cj^{d-1}$ for all $t \in [0,1]$ and large enough constant C. Similarly consider the FARIMA$(2p, d, p+q)$ model

$$[\Phi^p(\mathcal{B}, t)]^2(1-\mathcal{B})^d X_n(t) = \{(\dot{\Theta^q})(\mathcal{B}, t)\Phi^p(\mathcal{B}, t) - \Theta^q(\mathcal{B}, t)\dot{\Phi}^p(\mathcal{B}, t)\}\varepsilon_n.$$

Then similarly we can get $|\dot{c}_j(t)| \leq Cj^{d-1}$. $\square$

Remark 4. *Consider the time varying FARIMA(p, d, q) model:*

$$(\sum_{j=0}^{p} \alpha_j(i/n)\mathcal{B}^j)(1-\mathcal{B})^d X_{i,n} = (\sum_{k=0}^{q} \beta_k(i/n)\mathcal{B}^k)\sigma(i/n)\bar{\eta}_i, \ 0 < d < 1/2, \tag{22}$$

where $\mathcal{B}$ is the back shift operator, $\{\bar{\eta}_i\}_{i=-\infty}^{\infty}$ is some i.i.d random variables with 0 mean and variance 1. It can be shown that representation (22) has a MA representation:

$$X_{i,n} := \sum_{j=0}^{\infty} a_{i,n}(j)\bar{\eta}_{i-j}. \tag{23}$$

It can also be shown, similarly to Dahlhaus and Polonik (2009), that we cannot find functions $a(j,t)$ such that $a(j, i/n) = a_{i,n}(j)$. However, consider the following locally stationary FARIMA model:

$$(\sum_{j=0}^{p} \alpha_j(t)\mathcal{B}^j)(1-\mathcal{B})^d X_i(t) = (\sum_{k=0}^{q} \beta_k(t)\mathcal{B}^k)\sigma(t)\bar{\eta}_i, \ 0 < d < 1/2. \tag{24}$$

Note that $\mathcal{B}$ only affects i, but not t. Under some regularity conditions, it has MA representation $J_{i,n} = \sum_{j=0}^{\infty} \tilde{a}(j,t)\bar{\eta}_{i-j}$ for some MA coefficients $\tilde{a}(j,t)$ satisfying (A3). We have discussed such conditions in example 2. It has been show that, under short range dependence, the time varying AR model can be well approximated by locally stationary AR model. See Zhang and Wu (2012), Zhou (2013). We shall show that, with long range dependence, time varying FARIMA model can still be well approximated by locally stationary FARIMA model.

Proposition 2. *Consider Model (22) and Model (24). Suppose that:*
a) The start point $(X_{p,n}, ..., X_{1,n})' \in \mathcal{L}_2$.
b) The coefficients $\alpha_j(\cdot)$, $\beta_k(\cdot)$, $j = 1, ..., p$, $k = 1, ..., q$, are Lipschitz continuous on $(0,1)$,

$\sigma(\cdot)$ is bounded, Lipschitz continuous in $\mathbb{R}$.
c) $\sum_{j=1}^p a_j(t)z^j \neq 1$ for all $|z| \leq 1+c$ with $c > 0$ uniformly in $t \in (0,1)$. Then we have, for some constant $C > 0$,

$$\max_{1\leq i\leq n} \|X_{in} - X_i(i/n)\| \leq Cn^{d-1/2}. \tag{25}$$

If the $\sigma(\cdot)$ is constant, then $\max_{1\leq i\leq n} \|X_{in} - X_i(i/n)\| \leq Cn^{-1}$.

Example 3. *Consider the locally stationary FARIMA $(0,d,1)$ model of (24), then as shown in Palma (2010),*

$$Cov(X_s, X_m) \asymp g(s/n, m/n)|s-m|^{2d-1}, \tag{26}$$

for some $\mathcal{C}^1$ function $g(\cdot,\cdot)$ on $(0,1)\times(0,1)$. In addition, $g(t,t) > 0$ for $t \in (0,1)$, and $0 < d < 1/2$. As we discussed in example 2, our Theorems 1–4 hold for this model. In addition, the conditions of Lemma 1 are also satisfied due to the covariance structure (26), thus we can compute the asymptotic simultaneous confidence band via Lemma 1 if consistent estimates of $g(t,t)$ is provided. Note that now $g^{1/2}(t,t)$ plays the same role as $\sigma(t)\breve{a}^{1/2}(t,t)$ in Lemma 1.

Example 4. *Consider the locally stationary Gegenbauer ARMA process:*

$$(\sum_{j=0}^p \alpha_j(t)\mathcal{B}^j)(1-2\xi\mathcal{B}+\mathcal{B}^2)^\lambda X_i(t) = (\sum_{k=0}^q \beta_k(t)\mathcal{B}^k)\sigma(t)\bar{\eta}_i. \tag{27}$$

where $0 < \lambda < 0.25$, $|\xi| < 1$, $\bar{\eta}_i$ are i.i.d mean 0 and variance 1. Write $\Phi^p(z,t) = \sum_{j=0}^p \alpha_j(t)z^j$, $\Theta^q(z,t) = \sum_{k=0}^q \beta_k(t)z^k$. Suppose that $\Phi^p(z,t), \Theta^q(z,t)$ satisfies the same conditions of which is listed in example 2. Gegenbauer ARMA process is considered by Gray, Zhang, and Woodward (1989).

By our settings, model (4) can be rewritten as $(1-2\xi\mathcal{B}+\mathcal{B}^2)^\lambda X_i(t) = \sum_{j=0}^\infty c_j(t)\bar{\eta}_{i-j}$, where $c_j(t)$ is a $\mathcal{C}^1$ function such that $\sum_{j=0}^\infty(|c_j(t)| + |\dot{c}_j(t)|) < \infty$. Let $z_1 = cos\theta + i\sin\theta$,

$z_2 = \bar{z}_1$, z_1, z_2 be the solution of $1 - 2\xi z + z^2 = 0$. Hence

$$
\begin{aligned}
X_i(t) &= \sum_{j=0}^{\infty} \psi(j) z_1^j \mathcal{B}^j \sum_{k=0}^{\infty} \psi(k) z_2^k \mathcal{B}^k \sum_{l=0}^{\infty} c_l(t) \bar{\eta}_{i-l} \\
&= \sum_{j=0}^{\infty} \sum_{k=0}^{j} c_k(t) \sum_{s=0}^{j-k} \psi(s) z_1^s \psi(j-k-s) z_2^{j-k-s} \bar{\eta}_{i-j} \qquad (28)
\end{aligned}
$$

where $\psi(j) = \frac{\Gamma(d+j)}{\Gamma(d)\Gamma(j+1)} \asymp j^{\lambda-1}$. Then

$$
a_j(t) = \sum_{k=0}^{j} c_k(t) \sum_{s=0}^{j-k} \psi(s) z_1^s \psi(j-k-s) z_2^{j-k-s} := \sum_{k=0}^{j} c_k(t) \nu_{j-k}, \qquad (29)
$$

Since for $k \leq j$,

$$
\sum_{s=0}^{j-k} |\psi(s) z_1^s \psi(j-k-s) z_2^{j-k-s}| \leq C|(j-k)^{\lambda-1} + \sum_{s=1}^{j-k-1} s^{\lambda-1}(j-k-s)^{\lambda-1}|, \qquad (30)
$$

we have that $\nu_{k-j} = O((j-k)^{2\lambda-1})$. By $\sum_{j=0}^{\infty}(|c_j(t)| + |\dot{c}_j(t)|) < \infty$ and summation by parts, we have that $|a_j(t)| = O(j^{2\lambda-1})$. Similarly $|\dot{a}_j(t)| = O(j^{2\lambda-1})$. Then condition A3) is satisfied with $\gamma = 1 - 2\lambda$. In addition, if the innovation has moment p such that $\gamma < \frac{2+1/(2p)}{2+1/p}$, then our theory for the quantile applies in this case. In this example, for fixed t, $a_j(t)'s$ don't necessarily have the same sign. To see this, just note that if $j-k$ is odd, then $\nu_j = \sum_{s=0}^{j-k} cos((j-k-2s)\theta)\psi(s)\psi(j-k-s)$, and if $j-k$ is even, $\nu_j = 1 + 2\sum_{s=0}^{(j-k)/2-1} cos((j-k-2s)\theta)\psi(s)\psi(j-k-s)$, and $cos(\cdot)$ is a periodic function.

Example 5. *Consider the locally stationary seasonal Fractional ARIMA(p,d,s,q) Model:* $\Phi^p(\mathcal{B}, t)(1-\mathcal{B}^s)^d X_n(t) = \Theta^q(\mathcal{B}, t)\varepsilon_n$, *where* $\varepsilon_i's$ *are i.i.d random variables with mean* 0 *and variance* 1, $\Phi^p(z,t)$ *and* $\Theta^q(z,t)$ *satisfy the similar conditions as that in example 2. By Baillie (1996), for seasonal Fractional ARIMA(0,d,s,0), it has MA representation* $X_i = \sum_{j=0}^{\infty} \psi_j \varepsilon_{i-j}$, *such that* $\psi_j \asymp j^{d-1}$. *Then by similar argument as in example 4, the locally stationary seasonal Fractional ARIMA(p,d,s,q) has locally stationary MA representation* $X_i(t) = \sum_{j=0}^{\infty} a_j(t)\varepsilon_{i-j}$, *such that* $|a_j(t)| = O(j^{d-1})$, $|\dot{a}_j(t)| = O(j^{d-1})$. *Seasonal fractional ARMA is considered by Porter-Hudak, S.(1990) to model the monetary aggregates.*

5 Discussions

There are a small number of recent papers discussing non-stationary time series with long memory; see for instance Palma and Olea (2010), Palma (2010), Leipus and Surgailis (2013). Most of the aforementioned papers consider mean or spectral analysis of the series. We noticed that time varying long memory parameters $d(t)$ are allowed in some of the papers. Among them, Beran (2009) proposes a nonparametric method of estimating the time varying long memory parameters. Roueff and von Sachs (2011) advances a semi-parametric method of estimating time varying long memory parameters $d(t)$ and investigate its asymptotic behavior. Leipus and Surgailis (2013) studies the limiting behavior of the partial sums of linear process with time changing $d(t)$. Palma (2010) suggests a method of estimating sample mean for locally stationary processes with time-changing $d(t)$. In this paper we only considered the case where the memory parameter is a constant over time. However, note that in the context of simultaneous inference of quantile curves of non-stationary long memory processes, the stochastic variabilities of the estimated quantiles on $\mathcal{T}_n$ asymptotically dominates those on $(0,1)/\mathcal{T}_n$, where $\mathcal{T}_n = \{t : d(t) = d, \text{ and } d = \max_{s\in[0,1]} d(s)\}$. In many cases in practice, $\mathcal{T}_n$ can be assumed to be a collection of finite many non-overlapping subintervals of $(0,1)$. Hence for many scenarios where time-varying memory parameters are allowed, the construction of SCB for the quantile curves is essentially the same as that considered in this paper since one only needs to focus on $\mathcal{T}_n$, the time intervals where $d(t)$ attains its maximum. Note that the memory parameter does not change on $\mathcal{T}_n$. The major difficulties in the time-varying $d(t)$ case are to estimate $max_{s\in[0,1]}d(s)$ and determine $\mathcal{T}_n$, which we shall leave as a rewarding future work.

6 Proofs

In the following proofs, we shall only prove the case where $\alpha = 1/2$ as the proofs of the other quantiles follow by the same arguments. We shall also omit the subscript α if no confusions arise. In the proofs the constant C means a generic finite constant which may vary from place to place.

Lemma 2. *Let* $\bar{e}_{i,j} = \sum_{s=0}^{i-j} a_s(i/n)G(\frac{i-s}{n}, \eta_{i-s})$*, and* $f_{\bar{e}_{i,j}}(x)$ *be its density. Then under condition (A1), (A2), (A3), we have* $f^{(k)}_{\bar{e}_{i,j}}(x) \le M < \infty$ *for* $k = 0, 1, \cdots, r-1$*,* $1 \le i \le n$*,* $-\infty \le j < i$*.*

proof. See supplementary meterials.□

Lemma 3. *Let* $\Upsilon_n(t)$ *be a sequence of random variables and be once differentiable in* t*,* $t \in [0,1]$*. Let* p *be a positive constant such that* $p \ge 1$*, then if for any* $t \in [0,1]$*,* $\|\Upsilon_n(t)\|_p = O(m_n)$*,* $\|\Upsilon'_n(t)\|_p = O(l_n)$*, and* m_n, l_n *are sequences of real numbers, then* $\sup_{t\in[0,1]} |\Upsilon_n(t)| = O_p(m_n(\frac{m_n}{l_n})^{-\frac{1}{p}})$*. In particular, if* $p = 2$*, we have* $\sup_{t\in[0,1]} |\Upsilon_n(t)| = O_p(\sqrt{m_n l_n})$*.*

Proof. See supplementary meterials. □

Lemma 4. *Let* $\mathcal{F}_j$ *be* σ*-field generated by* $(..., \varepsilon_{j-1}, \varepsilon_j)$*. For* $j \in \mathbb{Z}$*, define* $\mathcal{P}_j \cdot = \mathbb{E}(\cdot|\mathcal{F}_j) - \mathbb{E}(\cdot|\mathcal{F}_{j-1})$ *be the projection operator on* $\mathcal{L}_1$*. Let* $\eta_{ni}(t) = \Psi(Y_i(t)) - \Psi(Y_i(t) - \alpha_{ni})$ *where* $Y_i(t) = X_i - Q(t) - (i/n - t)Q'(t)$*, and* $|\alpha_{ni}| \le C < \infty$ *be a number array. Then assume (A0)(A1)(A2)(A3), we have, for some* $p \ge 2$*,* $\|\mathcal{P}_j\eta_{ni}(t)\|_p = O(|\alpha_{ni}|\mathbf{1}(i-j \ge 0)/|i-j+1|^{\gamma})$*, where* $\mathbf{1}(\cdot)$ *is indicator function, and* $\|\mathcal{P}_j\eta_{ni}(t)\|_p = 0$ *if* $i < j$*.*

Proof. See supplementary meterials. □

For $t \in (0,1)$, let $s_n(t) = \max(\lfloor nt - nb_n \rfloor, 1), l_n(t) = \min(\lfloor nt + nb_n \rfloor, n)$ and

$$\mathcal{N}_n(t) = \{i \in \mathbb{N} : s_n(t) \le i \le l_n(t)\}. \tag{31}$$

Let

$$M_n(t,\theta) = \sum_{i=1}^{n} \{\Psi(Y_i(t) - \theta^T \mathbf{z}_{i,n}(t)) - \mathbb{E}[\Psi(Y_i(t) - \theta^T \mathbf{z}_{i,n}(t))|\mathcal{F}_{i-1}]\}\mathbf{v}_{i,n}(t), \tag{32}$$

$$N_n(t,\theta) = \sum_{i=1}^{n} \{\mathbb{E}[\Psi(Y_i(t) - \theta^T \mathbf{z}_{i,n}(t))|\mathcal{F}_{i-1}] - \mathbb{E}\Psi(Y_i(t) - \theta^T \mathbf{z}_{i,n}(t))\}\mathbf{v}_{i,n}(t). \tag{33}$$

Where $\mathcal{F}_i$ is the σ field generated by $(..., \varepsilon_{i-1}, \varepsilon_i)$, and $\mathbf{v}_{i,n}(t) = K_{b_n}(i/n - t)\mathbf{z}_{i,n}(t)$. So

$$S_n(t,\theta) - \mathbb{E}S_n(t,\theta) = M_n(t,\theta) + N_n(t,\theta). \tag{34}$$

Recall that $\mathbf{z}_{in}(t) = (1, \frac{i/n-t}{b_n})^T$. In the following proofs, write X_i for X_{in}, Y_i for Y_{in} for short if no confusions arise.

Theorem 5. *Assume (A0)-(A4). Let* $\{\gamma_n\}_{n\in N}$ *be positive with* $\gamma_n \to 0$, $nb_n\gamma_n \to \infty$, *we have*

$$\sup_{t\in[0,1],|\theta|\le\gamma_n} |M_n(t,\theta) - M_n(t,0)| = O_p((nb_n\gamma_n)^{1/2}\log n + n^{-3}),$$

where $M_n(t)$ *is defined in (32).*

Proof. Write $\delta_{ni}(t) = Q(t) + (i/n - t)Q'(t)$, then $Y_i(t) = X_i - \delta_{ni}(t)$. Let $\mathcal{N}_n(t-) = (0, nt]\bigcap\mathcal{N}_n(t)$, $\mathcal{N}_n(t+) = (nt, n]\bigcap\mathcal{N}_n(t)$, $\eta_i(t,\theta) = \Psi(Y_i(t) - \theta^T\mathbf{z}_{i,n}(t)) - \mathbb{E}[\Psi(Y_i(t) - \theta^T\mathbf{z}_{i,n}(t))|\mathcal{F}_{i-1}]$. Let $M_{n,1}(t,\theta) = \sum_{i=1}^n \eta_i(t,\theta)v_{i,1}(t)$, $M_{n,2}(t,\theta) = \sum_{i\in\mathcal{N}_n(t+)}^n \eta_i(t,\theta)v_{i,2}(t)$, $M_{n,3}(t,\theta) = \sum_{i\in\mathcal{N}_n(t-)}^n \eta_i(t,\theta)v_{i,2}(t)$. Then $M_n(t,\theta) = (M_{n,1}(t,\theta), M_{n,2}(t,\theta) + M_{n,3}(t,\theta))$. It is sufficient to show

$$\sup_{t\in[0,1],|\theta|\le\gamma_n} |M_{n,j}(t,\theta) - M_{n,j}(t,0)| = O_p((nb_n\gamma_n)^{1/2}\log n + n^{-3}) \tag{35}$$

for $j = 1, 2, 3$. We only proof when $j = 2$ here. The situations of $j = 1, 3$ are similar.

For a sequence of real number $(g_n)_{n\in N} \to \infty$, with $g_n \ge 3$ for all n, define $\mu_n = (nb_n\gamma_n)^{1/2}g_n/\log g_n$, $\phi_n = (nb_n\gamma_n)^{1/2}g_n\log n$, $a_i(t,\theta) = \Psi(Y_i(t) - \theta^T\mathbf{z}_{i,n}(t))v_{i,2}(t)$, and

$$\begin{aligned}
A_n(t) &= \max_{i\in\mathcal{N}_n(t+)} \sup_{|\theta|\le\gamma_n} |a_i(t,\theta) - a_i(t,0)|,\\
U_n(t) &= \sum_{i\in\mathcal{N}_n(t+)} \mathbb{E}\{[\Psi(Y_i(t) + |\mathbf{z}_{i,n}(t)|\gamma_n)) - \Psi(Y_i(t) - |\mathbf{z}_{i,n}(t)|\gamma_n)]^2|\mathcal{F}_{i-1}\}v_{i,2}^2\\
&:= \sum_{i\in\mathcal{N}_n(t+)} U_{ni}(t).
\end{aligned}$$

By monotonicity of $\Psi(.)$,

$$\sup_{|\theta|\le\gamma_n} \sum_{i\in\mathcal{N}_n(t+)} \mathbb{E}[(\eta_i(t,\theta) - \eta_i(t,0))^2 v_{i,2}^2|\mathcal{F}_{i-1}] \le U_n(t),\ \forall t\in(0,1). \tag{36}$$

By elementary calculations, we have $|U_{ni}(t)| \le \sup_{x\in R} f_{\varepsilon_i}(x) \times 2|\mathbf{z}_{i,n}(t)|\gamma_n \le 2C|\mathbf{z}_{i,n}(t)|\gamma_n$.

Hence

$$\mathbb{E}[\sup_{t\in[0,1]} U_n(t)] \le C\sum_{i=1}^{n} 2|\mathbf{z}_{i,n}(t)|\gamma_n v_{i,2}^2(t) \le Cnb_n\gamma_n. \tag{37}$$

Thus

$$\mathbb{P}[\sup_{t\in[0,1]} U_n(t) \ge \mu_n^2] \le \mu_n^{-2}\mathbb{E}[\sup_{t\in[0,1]} U_n(t)] = O(g_n^{-1}\log g_n) = o(1). \tag{38}$$

Similarly, by using $|A_n(t)| \le 2$, we can get

$$\mathbb{P}[\sup_{t\in[0,1]} A_n(t) \ge \mu_n] = o(1). \tag{39}$$

Then by similar argument as in Zhou and Wu (2009), Theorem 5 is established. More detailed version of proof is in supplemental materials. □

Theorem 6. *Assume (A0)-(A4).* $\sup_{t\in T_n,|\theta|\le\gamma_n} |N_n(t,\theta) - N_n(t,0)| = O_p((nb_n)^{3/2-\gamma} b_n^{-1/p}\gamma_n)$, *where* $N_n(t,\theta)$ *is defined in (33),* $T_n = (\delta b_n, 1-\delta b_n)$, $\delta > 1$.

Proof. Write $\eta_{ni}(t) = \Psi(Y_i(t) - \theta^T z_{i,n}(t)) - \Psi(Y_i(t))$, then we have

$$N_n(t,\theta) - N_n(t,0) = \sum_{i=1}^{n}\sum_{j=-\infty}^{i-1} \mathcal{P}_j \eta_{ni}(t)\mathbf{v}_{i,n}(t) = \sum_{j=-\infty}^{n-1}\sum_{i=j+1}^{n} \mathcal{P}_j \eta_{ni}(t) v_{i,n}(t). \tag{40}$$

By the Burkholder inequality and Triangular inequality, we have

$$\|N_n(t,\theta) - N_n(t,0)\|_p^2 \le \sum_{j=-\infty}^{n-1} (\sum_{i=j+1}^{n} \|\mathcal{P}_j \eta_{ni}(t) v_{i,n}(t)\|_p)^2. \tag{41}$$

By Lemma 4,$|\theta^T z_{i,n}(t)| \le |\theta^T||z_{i,n}(t)|$, $|\theta| \le \gamma_n$, the fact that $\exists M < \infty$, such that $|z_{i,n}(t)| \le M$, $|\mathbf{v}_{i,n}(t)| \le M$, and $|\mathbf{v}_i(t)| = 0$ if $i \bar{\in} \mathcal{N}(t)$, we get

$$\|N_n(t,\theta) - N_n(t,0)\|_p^2 \le C\sum_{j=-\infty}^{l_n(t)-1} (\sum_{i=j+1}^{l_n(t)} \mathbf{1}(i \in \mathcal{N}_n(t))|i-j+1|^{-\gamma})^2 \gamma_n^2. \tag{42}$$

By mean value theorem, we have

$$\|N_n(t,\theta) - N_n(t,0)\|_p = O((nb_n)^{3-2\gamma}\gamma_n). \tag{43}$$

Let $\zeta_{ni}(t,\theta) = f_{a_0\varepsilon_i}(-\sum_{s=1}^{\infty} a_s(i/n)\varepsilon_{i-s} + \delta_{ni}(t) + \theta^T z_{i,n}(t))$, where as before, $\delta_{ni}(t) = Q(t) + (i/n - t)Q'(t)$. Then similar to Lemma 4, one can show that

$$\|\mathcal{P}_j(\zeta_{ni}(t,\theta) - \zeta_{ni}(t,0))\|_p = O(|\theta|\mathbf{1}(i-j+1 \geq 0)/|i-j+1|^{\gamma}), \tag{44}$$

$$\|\mathcal{P}_j(\zeta_{ni}(t,\theta))\|_p = O(\mathbf{1}(i-j+1 \geq 0)/|i-j+1|^{\gamma}). \tag{45}$$

By Lebesgue's dominated convergence theorem (LDCT), similarly to how we deal with $\|N_n(t,\theta) - N_n(t,0)\|_p$, with triangular inequality and the fact that $\exists C < \infty$, such that $|\mathbf{v}'_{i,n}(t)| \leq C/b_n,\ |\mathbf{z}'_{i,n}(t)| \leq C/b_n$, we get

$$\|N'_n(t,\theta) - N'_n(t,0)\|_p = O((nb_n)^{3/2-\gamma}b_n^{-1}\gamma_n) \tag{46}$$

Then by equation (43)(46) and Lemma 3, we get the proof. Details are in supplemental materials. $\square$

Theorem 7. *Define* $\hat{\theta}(t) = [\hat{Q}(t) - Q(t),\ b_n(\hat{Q}'(t) - Q(t)')]^T,\ T_n = [\delta b_n, 1-\delta b_n],\ \delta > 1,\ \pi_n = (nb_n)^{-1/2}[\log n + (nb_n^5)^{1/2} + (nb_n)^{1-\gamma}b_n^{-1/p}],\ \inf_{t\in[0,1]} f(t, Q_\alpha(t)) > 0$. *Then under (A0)-(A5),* $\sup_{t\in T_n} |\hat{\theta}(t)| = O_p(\pi_n)$.

Proof. After establishing (5) and (6), this theorem can be shown with similar method of Zhou and Wu (2009). Details are in Supplementary meterials. $\square$.

Proof of Theorem 1. Via carefully checking Zhou and Wu (2009), their Lemma 8 holds in our settings. By equation (34), Theorem 5, Theorem 6, Theorem 7, Lemma 8 in Zhou and Wu (2009), and similar argument as in the proof of Theorem 3 in Zhou and Wu (2009), Theorem 1 holds.$\square$

Theorem 8. *Let*

$$S_n(t) = \sum_{i=1}^{n} \Psi(X_{in} - \delta_{ni}(t))K_{b_n}(i/n - t)\mathbf{z}_{i,n}(t),$$

$b_n \to 0$ and $(nb_n)^{1/2-\gamma} b_n^{-1/p} \log(nb_n) = o(1)$. Then under conditions of Theorem 3, for some $p \geq 2$, $\mathbb{E}|\varepsilon_i|^p < \infty$, we have

$$\sup_{t \in T_n} |S_n(t) - \sum_{i=1}^{n} f_n(t, Q_{\alpha,n}(t)) \bar{X}_{in} K_{b_n}(i/n - t) \mathbf{z}_{in}(t) - nb_n^3 f_n(t, Q_{\alpha,n}(t)) Q''_{\alpha,n}(t)(\mu_2, 0)^T/2| = O_p(K_n^p), \tag{47}$$

where T_n, K_n^p, are defined in Theorem 3, $\bar{X}_{in} = X_{in} - \mathbb{E}X_{in}$.

Proof. This lemma can be shown directly by elementary calculations and similar method in the proof of Theorem 6 and proof of Lemma 3 in Zhou and Wu (2011). Details are in supplemental materials.

Proof of Theorem 2.

Note that $\sum_{k=1}^{n}(X_k - \mu(\frac{k}{n})) = \sum_{k=1}^{n}\sum_{j=0}^{\infty} a_j(\frac{k}{n})\varepsilon_{k-j} = \sum_{j=1}^{n}\varepsilon_j \sum_{k=j}^{n} a_{k-j}(\frac{k}{n}) + \sum_{j=0}^{\infty}\varepsilon_{-j}\sum_{k=1}^{n} a_{k+j}(\frac{k}{n})$.

Let $Z_k = \sum_{j=0}^{k} \varepsilon_{-j}$ with $Z_j = 0$ for $j \leq -1$, $W_j = \sum_{i=1}^{j} \varepsilon_i$ with $W_j = 0$ for $j \leq 0$, then by similar argument in Wang et al. (2003), we can get that

$$\sum_{k=1}^{n}(X_k - \mu(k/n)) = \sum_{j=1}^{n-1}[\sum_{k=j}^{n} a_{k-j}(k/n) - \sum_{k=j+1}^{n} a_{k-j-1}(k/n)]W_j + W_n a_0(1)$$
$$+ \sum_{j=0}^{N-1}\sum_{k=1}^{n}[a_{k+j}(k/n) - a_{k+j+1}(k/n)]Z_j + Z_N \sum_{k=1}^{n} a_{k+N}(k/n) + \sum_{j=N+1}^{\infty}\varepsilon_{-j}\sum_{k=1}^{n} a_{k+j}(k/n).$$

Redefine $\{\varepsilon_j, j = -\infty, .., \infty\}$ on a richer probability space. By Shao (1995), Theorem B, use condition (A3), $\forall \varepsilon > 0$, take $x = n^{1/p}\varepsilon^{1/p}$, we have that there exists independent centered normal random variables $\{\varrho_j, j = -\infty, .., \infty\}$, $j \in \mathbb{Z}$, with $\mathrm{var}(\varrho_j) = \sigma(j/n)$, such that for any n,

$$\zeta_n := \max_{1 \leq m \leq n} |\sum_{j=1}^{m}\varepsilon_j - \sum_{j=1}^{m}\varrho_j| = O_p(n^{1/p}),\ \zeta_n^* := \max_{1 \leq m \leq n} |\sum_{j=0}^{m}\varepsilon_{-j} - \sum_{j=0}^{m}\varrho_{-j}| = O_p(n^{1/p}).$$

Let $Y_i = \sum_{j=0}^{\infty} a_j(i/n)\varrho_{i-j} + \mu(i/n)$, $W_j^* = \sum_{j=1}^{n} \varrho_j$, $Z_k^* = \sum_{j=0}^{k} \varrho_{-j}$,

$$\sum_{k=1}^{n}(X_i - Y_i) = \sum_{j=1}^{n-1}[\sum_{k=j}^{n} a_{k-j}(k/n) - \sum_{k=j+1}^{n} a_{k-j-1}(k/n)][W_j - W_j^*] + [W_n - W_n^*]a_0(1)$$
$$+ \sum_{j=0}^{N-1}\sum_{k=1}^{n}[a_{k+j}(k/n) - a_{k+j+1}(k/n)][Z_j - Z_j^*] + [Z_N - Z_N^*]\sum_{k=1}^{n} a_{k+N}(k/n)$$
$$+ \sum_{j=N+1}^{\infty} \varepsilon_{-j} \sum_{k=1}^{n} a_{k+j}(k/n) - \sum_{j=N+1}^{\infty} \varrho_{-j} \sum_{k=1}^{n} a_{k+j}(k/n) \tag{48}$$
$$= A + B + C + D + E + F. \tag{49}$$

Via elementary calculations using condition (A3), we get

$$|A| \le C\zeta_n \sum_{j=1}^{n-1}[1/n \sum_{j=0}^{n-j-1} |j+1|^{-\gamma} + (n-j+1)^{-\gamma}] = O_p(n^{1/p+1-\gamma}). \tag{50}$$
$$|B| \le \zeta_n = O_p(n^{1/p}),\ |C| = O_p(n^{\alpha(1+1/p-\gamma)}). \tag{51}$$
$$|D| \le M\zeta_N^* \sum_{k=1}^{n}(k+N)^{-\gamma} = O_p(N^{1/p}(n+N)^{1-\gamma}) = O_p(n^{\alpha(1+1/p-\gamma)}). \tag{52}$$
$$|E| = O_p(n^{1+\alpha(1/2-\gamma)}),\ |F| = O_p(n^{1+\alpha(1/2-\gamma)}). \tag{53}$$

We finish the proof of theorem 2 by noting that $\alpha = \frac{1}{1/2+1/p}$ is chosen to make $1+\alpha(1/2-\gamma) = \alpha(1+1/p-\gamma)$, and $\varsigma_i = \varrho_i/\sigma(i/n)$. A detailed proof is in supplementary materials. □

Theorem 9. *Under conditions of Theorem 2, there exists a sequence of i.i.d. standard normal random variables* ϑ_i*, such that* $X_{in}^* := \sum_{j=0}^{\infty} a_j(i/n)\sigma(i/n)\vartheta_{i-j} + \mu(i/n)$ *and*

$$\sup_{t\in(0,1)} |\sum_{i=1}^{n}(X_{in}^* - X_{i,n})K_{b_n}(i/n - t)| = O_p(n^{1+\frac{1/2-\gamma}{1/2+1/p}} + n^{3/2-\gamma}b_n^{1-1/p}) \tag{54}$$

.

proof. This theorem is an immediate result of Theorem 2, the summation by parts formula, the Burkholder's inequality and Lemma 3. Details are in supplemental materials. □

Proof of Theorem 3

By Theorem 1, Theorem 2, Theorem 8 and Theorem 9, we can easily see that

$$\sup_{t\in T_n}|f_n(t,Q_{\alpha,n}(t))(\mu_{\mathbf{k}}\hat{\theta}_{\alpha,n}(t)-\frac{1}{nb_n}\sum_{i=1}^{n}\sigma(i/n)V_iK_{b_n}(i/n-t)\mathbf{z}_{i,n}(t)$$
$$-\frac{b_n^2Q''_{\alpha,n}(t)(\mu_2,0)^T}{2})|=O_p(\varsigma_n), \quad (55)$$

Elementary calculations by using Lemma 3, we have

$$\sup_{t\in(0,1)}|\sum_{i=1}^{n}(\sigma(t)-\sigma(i/n))V_iK_{b_n}(i/n-t)|=O_p((nb_n)^{3/2-\gamma}b_n^{1-1/p}). \quad (56)$$

The theorem holds by combining equation (55) and equation (56). □.

Proof of Proposition 1. It is a direct result by applying LDCT. □

Proof of Corollary 2. This corollary can be shown by the next two lemmas and a similar argument as the one in Lemma 1 of Wu and Zhao (2007).

Lemma 5. *If* $\max_{1\le s\le n}|\sum_{i=1}^{s}(X_i-Y_i)|=O_p(G_n)$, *then* $\sup_{t\in T_n}|\sum_{i=1}^{n}(X_i-Y_i)K_{b_n}(i/n-t)|=O_p(G_n)$, *where* $G_n\to\infty$ *is a sequence of real number.*

proof. The lemma uses the summation by parts formula and it can be shown in the same way as Zhou (2010).

Lemma 6. *Under conditions of Corollary 2, there exists a sequence of i.i.d standard normal random variables* ε_i, $i=1,2,..,\infty$, *and an associate standard fractional Brownian motion* $B_H(t)$, *such that: for* $X_i=\sum_{j=0}^{\infty}g_j\varepsilon_{i-j}$, $S_{a,b}=\sum_{i=a}^{b}X_i$, *where* a, b *are positive integers,* $g_j=(1-\gamma)/(j+1)^{\gamma}(1+O(1/j))$, *we have*

$$\max_{a\le k\le b}|S_{a,k}-\mathbb{L}_{1-\gamma}^{1/2}(B_H(k)-B_H(a))|=O_{a.s}(\sqrt{n}\log n) \quad (57)$$

This lemma can be shown by slightly modifying the proof of Lemma 2, 3, 4, 6 of Konstantopoulos and Sakhanenko (2004). Details are omitted.

Proof of Theorem 4.

Let $\{\vartheta_j\}_{-\infty}^{\infty}$ are $i.i.d$ standard normal. Write V_{in} as V_i for short, then $V_i=\sum_{j=0}^{\infty}a_j(i/n)\vartheta_{i-j}$.

For any $\beta > 1$, Define

$$\widetilde{V}_i = \sum_{j=0}^{n^\beta} a_j(i/n)\vartheta_{i-j} = \sum_{j=i-n^\beta}^{i} a_{i-j}(i/n)\vartheta_j,\ \widetilde{S}_n(t) = \sum_{i=1}^{n} K_{b_n}(i/n-t)\widetilde{V}_i, \tag{58}$$

$$S_n(t) = \sum_{i=1}^{n} K_{b_n}(i/n-t)V_i,\ T_n = [\delta b_n, 1-\delta b_n],\ \text{where } \delta > 1. \tag{59}$$

Then elementary calculations with Lemma 3 shows that

$$\sup_{t\in T_n} |\widetilde{S}_n(t) - S_n(t)| = O_p(n^{\frac{(1-2\gamma)\beta}{2}+1} b_n^{1/2}). \tag{60}$$

Take $\beta > 1$ such that $\frac{(1-2\gamma)\beta}{2} + 1 < 0$. Let $N = \lfloor n^\beta \rfloor$. Apply Sun and Loader (1994): Let

$$T_j(t) = \begin{cases} \sum_{i=1}^{N+j} a_{i-j}(i/n)K_{b_n}(i/n-t) \text{ if } 1-N \le j \le n-N \\ \sum_{i=1}^{n} a_{i-j}(i/n)K_{b_n}(i/n-t) \text{ if } 1+n-N \le j \le n \end{cases}, \tag{61}$$

$$T(t) = (T_{1-N}(t), .., T_{n-N}(t), T_{n-N+1}(t), .., T_n(t)). \tag{62}$$

Then we have

$$\mathbb{P}\{\sup_{t\in T_n} \frac{\tilde{S}_n(t)}{\|\tilde{S}_n(t)\|} > c\} = \tau, \qquad \kappa_0 = \int_{\delta b_n}^{1-\delta b_n} \|[\frac{T(t)}{\|T(t)\|}]'\| dt. \tag{63}$$

Where $\tau = \frac{\kappa_0}{\pi} exp(-c^2/2) + 2(1-\Phi(c)) + O(e^{-c^2/2})$, and $\Phi(\cdot)$ is CDF of $N(0,1)$. However, by $\langle T'(t), T(t)\rangle = [\|T(t)\|^2]'/2 = \|T(t)\|\|T(t)\|'$, we have

$$\|[\frac{T(t)}{\|T(t)\|}]'\| = \frac{\|T'(t)\|}{\|T(t)\|}\sqrt{1-[\langle\frac{T'(t)}{\|T'(t)\|}, \frac{T(t)}{\|T(t)\|}\rangle]^2}. \tag{64}$$

note that for $j \ge \lfloor nt \rfloor + 1, i \ge j, i/n - t \ge 0$, then $K'_{b_n}(i/n-t) \le 0$, so we have, by Cauchy inequality,

$$\langle T'(t), T(t)\rangle \le (\|T(t)\|^2 - \sum_{j=\lfloor nt\rfloor+2}^{n}[\sum_{i=1}^{n} a_{i-j}(i/n)K_{b_n}(i/n-t)]^2)^{1/2}$$
$$\times (\|T'(t)\|^2 - \sum_{j=\lfloor nt\rfloor+2}^{n}[\sum_{i=1}^{n} a_{i-j}(i/n)K'_{b_n}(i/n-t)]^2)^{1/2}. \quad (65)$$

As a consequence, by $\lfloor nt - nb_n\rfloor \le i \le \lfloor nt + nb_n\rfloor$ for $t \in T_n$ and $a_{i-j}(t) = 0$ for $i \le j-1$,

$$0 \le [\langle \frac{T'(t)}{\|T'(t)\|}, \frac{T(t)}{\|T(t)\|}\rangle]^2 \quad \le \{1 - \frac{\sum_{j=\lfloor nt\rfloor+2}^{\lfloor nt+nb_n\rfloor}[\sum_{i=j}^{\lfloor nt+nb_n\rfloor} a_{i-j}(i/n)K'_{b_n}(i/n-t)]^2}{\|T'(t)\|^2}\}. \quad (66)$$

Pick some $0 < a < 1$ with $K'(a) < 0$, then $|K'(x)| \ge |K'(a)|\ \forall x \in [a,1]$. Choose n large such that $\lfloor anb_n\rfloor \ge 2$, then

$$\sum_{j=\lfloor nt\rfloor+2}^{\lfloor nt+nb_n\rfloor}[\sum_{i=j}^{\lfloor nt+nb_n\rfloor} a_{i-j}(i/n)K'_{b_n}(i/n-t)]^2 \ge (K'(a))^2q^2 \sum_{j=\lfloor nt+a_nb_n\rfloor}^{\lfloor nt+nb_n\rfloor}[\sum_{i=j}^{\lfloor nt+nb_n\rfloor}\frac{1}{(i-j+1)^\gamma}]^2/b_n^2 \quad (67)$$

$$\ge \frac{(K'(a))^2q^2}{(1-\gamma)^2(3-2\gamma)}(((1-a)nb_n+1)^{3-2\gamma}-1)/b_n^2 + O((nb_n)^{2-2\gamma}/b_n^2). \quad (68)$$

To continue , we need the following:

$$q^2(nb_n)^{3-2\gamma}G(K(.)) + O(n^{(1-2\gamma)\beta+2}b_n^2) + O((\log n)^{-1}(nb_n)^{3-2\gamma}) \le \|T(t)\|^2$$
$$\le Q^2(nb_n)^{3-2\gamma}G(K(.)) + O(n^{(1-2\gamma)\beta+2}b_n^2) + O((\log n)^{-1}(nb_n)^{3-2\gamma}). \quad (69)$$

By definition, $\|T(t)\|^2 = \|\widetilde{S}_n(t)\|^2$, and from before, $\|S_n(t) - \widetilde{S}_n(t)\| = O(n^{\frac{(1-2\gamma)\beta}{2}+1}b_n)$. On the other hand, let $\underline{S}_n(t) = \sum_{i=1}^n K_{b_n}(i/n-t)\underline{X}_i$, $\underline{X}_i = \sum_{j=0}^\infty \frac{1}{(j+1)^\gamma}\vartheta_{i-j}$, $H = 3/2-\gamma$, use the result of Lemma 5, Lemma 6 and similar arguments in the proof of Wu and Zhao (2007), we can get that

$$|\frac{(1-\gamma)\underline{S}_n(t)}{(nb_n)^{3/2-\gamma}\mathbb{L}_{1-\gamma}^{1/2}} - \frac{\int_0^n K(\frac{\lfloor 1+u\rfloor/n}{b_n}-t)dB_H(u)}{(nb_n)^{3/2-\gamma}}| = O_{a.s}(\log n)^{-1/2}.$$

and

$$\sup_{t\in T_n}|\frac{\int_0^n K(\frac{\lfloor 1+u\rfloor/n}{b_n}-t)dB_H(u)}{(nb_n)^{3/2-\gamma}}-\int_R K(u-t)dB_H(u)|=O_{a.s}((\log n)^{-1/2}).$$

while $\|\int_R K(u-t)dB_H(u)\|^2=\int_{-1}^1\int_{-1}^1 K(x)K(y)|x-y|^{1-2\gamma}dxdy$. For $\|T'(t)\|$, it is easy to see that $\|T'(t)-S_n'(t)\|=O(n^{\frac{(1-2\gamma)\beta}{2}+1})$. Let $M=\sup_{x\in[-1,1]}|K'(x)|$, $w=(\delta-1)/2$, then elementary calculations show that

$$\|S_n'(t)\|^2\le b_n^{-2}M^2Q^2((\frac{4w^{1-2\gamma}}{2\gamma-1}+\frac{w(2+w)^{2-2\gamma}}{(1-\gamma)^2})(nb_n)^{3-2\gamma}(1+O(\frac{1}{nb_n}))+\frac{(2nb_n+2)^{3-2\gamma}}{(1-\gamma)^2(3-2\gamma)}).$$

On the other hand, let a be the number that we choose in equation (67) and similarly to equation (68),

$$\|S_n'(t)\|^2\ge\frac{(K'(a))^2q^2}{(1-\gamma)^2(3-2\gamma)}(((1-a)nb_n+1)^{3-2\gamma}-1)/b_n^2+O((nb_n)^{2-\gamma}/b_n^2). \qquad (70)$$

Then we have

$$\begin{aligned}&O(n^{(1-2\gamma)\beta+2})+\frac{(K'(a))^2q^2}{(1-\gamma)^2(3-2\gamma)}(((1-a)nb_n+1)^{3-2\gamma}-1)/b_n^2+O((nb_n)^{2-\gamma}/b_n^2)\\&\le\|T_n'(t)\|^2\le O(n^{(1-2\gamma)\beta+2})+b_n^{-2}M^2Q^2((\frac{4}{2\gamma-1}w^{1-2\gamma}+\frac{w(2+w)^{2-2\gamma}}{(1-\gamma)^2})(nb_n)^{3-2\gamma}\\&(1+O((nb_n)^{-1}))+\frac{(2nb_n+2)^{3-2\gamma}}{(3-2\gamma)(1-\gamma)^2}). \qquad (71)\end{aligned}$$

The theorem now can be shown by Sun and Loader (1994), combined with equation (63), (64), (66), (68), (69), (71), and the fact that $\frac{\tilde{S}_n(t)}{\|\tilde{S}_n(t)\|}=\frac{S_n(t)+O_p(n^{\frac{(1-2\gamma)\beta}{2}+1}b_n^{1/2})}{\|S_n(t)\|+O(n^{\frac{(1-2\gamma)\beta}{2}+1}b_n)}$. $\square$

Proof of Corolary 1.

This corollary can be shown by equation (60), (69). $\square$

Proof of lemma 1.

Write $\breve{S}_n(t)=\sum_{i=1}^n V_{in}K_{b_n}(i/n-t)$, note that by the lipschitz continuity and the boundedness of $\breve{a}(x,y)$,

$$\begin{aligned}
\|\breve{S}_n(t)\|^2 &= \mathbb{E}\left(\sum_{i=1}^{n}\sum_{j=1}^{n} V_{in}V_{jn}K_{b_n}(i/n-t)\right) \\
&= \sum_{i=1}^{n}\sum_{j=1}^{n} \Gamma(i,j)K_{b_n}(i/n-t)K_{b_n}(j/n-t) \\
&\asymp \sum_{i=1}^{n}\sum_{j=1}^{n} \breve{a}(i/n,j/n)|i-j|^{1-2\gamma}K_{b_n}(i/n-t)K_{b_n}(j/n-t) \\
&\asymp (nb_n)^{3-2\gamma}\int\int \breve{a}(b_nx+t,b_ny+t)|x-y|^{1-2\gamma}K(x)K(y)dxdy \\
&\asymp (nb_n)^{3-2\gamma}\int\int \breve{a}(t,t)|x-y|^{1-2\gamma}K(x)K(y)dxdy. \qquad (72)
\end{aligned}$$

Similarly, $\|\frac{\partial \breve{S}_n(t)}{\partial t}\|^2 \asymp (nb_n)^{3-2\gamma}b_n^{-2}\int\int \breve{a}(t,t)|x-y|^{1-2\gamma}K'(x)K'(y)dxdy$. By elementary calculus, we have

$$\begin{aligned}
\frac{\partial\|\breve{S}_n(t)\|^2}{\partial t} &= 2\sum_{i=1}^{n}\sum_{j=1}^{n}\Gamma(i,j)K'(\frac{i/n-t}{b_n})K(\frac{j/n-t}{b_n})/b_n \\
&= (nb_n)^{3/2-\gamma}/b_n\int\int \breve{a}(b_nx+t,b_ny+t)|x-y|^{1-2\gamma}K(x)K'(y)dxdy + o((nb_n)^{3-2\gamma}b_n^{-1}). \qquad (73)
\end{aligned}$$

Let $R(x,y) = |x-y|^{1-2\gamma}K(x)K'(y)$, since $R(x,y)+R(-x,-y)=0$, we have that $\int\int |x-y|^{1-2\gamma}K(x)K'(y)dxdy = 0$, then

$$\begin{aligned}
&\int\int \breve{a}(b_nx+t,b_ny+t)|x-y|^{1-2\gamma}K(x)K'(y)dxdy \\
&\asymp b_n\int\int \breve{a}(t,t)(x+y)|x-y|^{1-2\gamma}K(x)K'(y)dxdy. \qquad (74)
\end{aligned}$$

Combine with (73), we have $\frac{\partial\|\breve{S}_n(t)\|^2}{\partial t} = o((nb_n)^{3-2\gamma}b_n^{-1})$. Since $\frac{\partial\|\breve{S}_n(t)\|}{\partial t} = \frac{\partial\|\breve{S}_n(t)\|^2}{\partial t}/(2\|\breve{S}_n(t)\|)$, by (72),

$$\|\breve{S}_n(t)\| = (nb_n)^{3/2-\gamma}\breve{a}^{1/2}(t)(\int\int |x-y|^{1-2\gamma}K(x)K(y)dxdy)^{1/2}, \qquad (75)$$

we have

$$\frac{\partial \|\breve{S}_n(t)\|}{\partial t} = o((nb_n)^{3/2-\gamma}/b_n). \tag{76}$$

By the proof of Theorem 4,

$$\begin{aligned}\kappa_n &= \int_{t\in T_n} \|\frac{\partial}{\partial t}(\frac{\breve{S}_n(t)}{\|\breve{S}_n(t)\|})\|dt = \int_{t\in T_n} \|\frac{(\frac{\partial}{\partial t}\breve{S}_n(t))\|\breve{S}_n(t)\| - (\frac{\partial}{\partial t}\|\breve{S}_n(t)\|)\breve{S}_n(t)}{\|\breve{S}_n(t)\|^2}\|dt \\ &= \int_{t\in T_n}\left(\|\frac{\partial}{\partial t}\breve{S}_n(t)\|/\|\breve{S}_n(t)\|(1+o(1))\right)dt = \int_{t\in T_n}\left(\frac{D}{b_n}(1+o(1))\right)dt. \end{aligned} \tag{77}$$

The lemma follows by Theorem 4, (75) and (77) $\Box$.

Proof of Proposition 2.

The proof is motivated by Zhang and Wu (2012). Let $\mathbf{x}_i = (X_{in}, ..., X_{i-p+1,n})'$, $\mathbf{z}_i(t) = (z_i(t), ..., z_{i-p+1}(t))'$. Define a $p\times p$ matrix $\mathbf{A}(t)$, and $A_{ij}(t)$ be its entry in i_th row, j_{th} column. Then $A_{1,j}(t) = \alpha_j(t)$, $1 \le j \le p$, and $A_{j,j-1}(t) = 1$ for $j = 2, ..., p$. Let $\mathcal{F}_i = (\bar{\eta}_{-\infty}, ..., \bar{\eta}_{i-1}, \bar{\eta}_i)$. Define $H(t,\mathcal{F}_i) = (\sum_{k=0}^q \beta_k(t)\mathcal{B}^k)\sigma(t)(1-\mathcal{B})^{-d}\bar{\eta}_i$, $\mathbf{H}(t,\mathcal{F}_i) = (H(t,\mathcal{F}_i), 0, ..., 0)'$ be a $p\times 1$ vector. Similarly, $L_n(\mathcal{F}_i) = (\sum_{k=0}^q \beta_k(i/n)\mathcal{B}^k)(1-\mathcal{B})^{-d}\sigma(i/n)\bar{\eta}_i$, $\mathbf{L}_n(\mathcal{F}_i) = (L_n(\mathcal{F}_i), 0, ..., 0)'$ be a $p\times 1$ vector. Then model (22) can be written as

$$\mathbf{x}_i = \mathbf{A}(i/n)\mathbf{x}_i + \mathbf{L}_n(\mathcal{F}_i), \tag{78}$$

and model (24) can be written as

$$\mathbf{z}_i(t) = \mathbf{A}(t)\mathbf{z}_{i-1}(t) + \mathbf{H}(t,\mathcal{F}_i). \tag{79}$$

First,

$$\|\mathbf{H}(i/n,\mathcal{F}_i) - \mathbf{L}_n(\mathcal{F}_i)\| \le \sum_{k=0}^q |\beta_k(i/n)|\|[\sum_{j=0}^\infty \psi(j)(\sigma(\frac{i-j-k}{n}) - \sigma(\frac{i}{n}))\bar{\eta}_{i-k-j}]\|, \tag{80}$$

where $\psi(j) = \frac{\Gamma(d+j)}{\Gamma(d)\Gamma(j+1)} \asymp j^{d-1}$, $\Gamma(\cdot)$ is the usual gamma function. Then by the Lipschitz continuity, let $\alpha \in (0,1)$, for $0 \le k \le q$ and some large constant C,

$$\|[\sum_{j=0}^{\lfloor n^\alpha \rfloor} \psi(j)(\sigma(\frac{i-j-k}{n}) - \sigma(\frac{i}{n}))\bar{\eta}_{i-k-j}]\|^2 \le C \sum_{j=0}^{\lfloor n^\alpha \rfloor} \psi(j)^2 (\frac{j}{n})^2 \le C n^{\alpha(1+2d)-2}. \qquad (81)$$

On the other hand, similar argument yields that

$$\|[\sum_{j=\lfloor n^\alpha \rfloor}^{\infty} \psi(j)(\sigma(\frac{i-j-k}{n}) - \sigma(\frac{i}{n}))\bar{\eta}_{i-k-j}]\|^2 \le C n^{\alpha(2d-1)}. \qquad (82)$$

Take $\alpha = 1$, by triangular inequality, we have that

$$\max_{1 \le i \le n} \|\mathbf{H}(i/n, \mathcal{F}_i) - \mathbf{L}_n(\mathcal{F}_i)\| = O(n^{d-1/2}). \qquad (83)$$

For matrix $\mathbf{A}$, define $\rho(\mathbf{A}) = \sup\{|\mathbf{Av}| : |\mathbf{v}| = 1\}$. Let $\rho = \sup_{t \in (0,1)} \rho(\mathbf{A})$. By (c), $\rho < 1$. By (b), triangular inequality and induction, we have, for $k \ge 2$,

$$\max_{1 \le i \le n} \|\mathbf{x}_i - \mathbf{z}_i(i/n)\| \le \rho^k \max_{1 \le i \le n} \|\mathbf{x}_{i-k} - \mathbf{z}_{i-k}(i/n)\| + C \sum_{j=1}^{k-1} j \rho^j n^{d-1/2}. \qquad (84)$$

Let $k \to \infty$, (25) follows. If $\sigma(\cdot)$ is constant, $\mathbf{H}(i/n, \mathcal{F}_i) = \mathbf{L}_n(\mathcal{F}_i)$, then similarly $\max_{1 \le i \le n} \|X_{in} - X_i(i/n)\| \le C n^{-1}$ $\square$

REFERENCES

Baillie, R. T. (1994). Long memory processes and fractional integration in econometrics. *Journal of Econometrics* **73** 5–59.

Bekaert, G. and Harvey, C. R. (1995). Time-varying world market integration. *The Journal of Finance* **50** 403–444.

Beran, J. (2009). On parameter estimation for locally stationary long-memory processes. *J. of Statistical Planning and Inference* **139** 900-915.

Berman, S. M. (1971). Maximum and high level excursion of stationary Gaussian processes. *Transaction of American Mathematical Society* **160** 65-85.

Bickel, P. J. and Rosenblatt, M. (1973). On some global measures of the deviations of density function estimates. *Ann. Statist.* **1** 1071–1095.

Cai, Z. and Xu, X. (2009). Nonparametric Quantile Estimations for Dynamic Smooth Coefficient Models. *Journal of the American Statistical Association* **104** 371-383. *Ann. Statist.* **1** 1071–1095.

Csörgö, M. and Kulik, R. (2008). Reduction principles for quantile and Bahadur-Kiefer processes of long-range dependent linear sequences. *Probability Theory and Related Fields*, **142**, 339–366.

Csörgö, M. and Révész, P. (1981). Strong approximations in probability and statistics. *Academic Press.*

Chaudhuri, P. (1991). Nonparametric Estimates of Regression Quantiles and Their Local Bahadur Representation *Ann. Statist* **19** 760–777.

Cheung, Y. W. (1993). Long memory in foreign exchange rates. *Journal of Business and Economics statistics* **11** 93–101.

Clarke, R. T. (2007) Hydrological prediction in a non-stationary world. *Hydrol. Earth Syst. Sci.,* **11**, 408–417.

Coeurjolly, J. F. (2008). Bahadur representation of sample quantiles for functional of Gaussian dependent sequences under a minimal assumption. *Statistics and Probability letters* **78** 2485-2489.

Cooley, T. F. and Prescott, E. C. (1976). Estimation in the presence of stochastic parameter variation. *Econometrica* **44** 167–184.

Dahlhaus, R. (1997). Fitting time series models to non-stationary processes. *Ann. Statist.* **25** 1–37.

Dahlhaus, R. and Polonik, W. (2009). Empirical spectral processes for locally stationary time series. *Bernoulli* **15** 1–39.

Dahlhaus, R. and Subba Rao, S. (2006). Statistical Inference for time-varying ARCH processes. *Annals of Statistics* **34** 1074–1114.

Eichner, J. F., E. Koscielny-Bunde, A. Bunde, S. Havlin, and H. J. Schellnhuber (2003), Power-law persistence and trends in the atmosphere: A detailed study of long temperature records, *Phys. Rev. E.*, **68**, 046133

Einmahl, U. (1987a). A useful estimate in the multidimensional invariance principle. *Probab. Theory Related Fields* **76**, 81-101.

Einmahl, U. (1987b). Strong invariance principles for partial sums of independent random vectors. *Ann. Probab.* **15**, 1419-1440.

Einmahl, U. (1989). Extensions of results of Komlós, Major and Tusnády to the multivariate case. *J. Multivariate Anal.* **28** 20–68.

Fan, J. and Gijbels, I. (1996). Local Polynomial Modelling and Its Applications. Chapman & Hall, London.

Fleming, J. and Kirby, C. (2011). Long memory in volatility and trading volume. *Journal of Banking and Finance* **35** 1714–1726.

Fryzlewicz, P., Sapatinas, T. and Subba Rao, S. (2008) Normalized least-squares estimation in time-varying ARCH models. *Annals of statistics* **36** 742–786.

Gray, H.L., N.-F. Zhang, and W.A. Woodward, (1989). On generalized fractional processes. *Journal of Time Series Analysis* **10**, 233–257.

Härdle, W. (1989) Asymptotic maximal deviation of M-smoothers. *J. Multiv. Anal.* **29** 163–179.

Haslett J. , Raftery A. E. (1989) Space–time modeling with long– memory dependence: assessing Ireland's wind power resource. *Appl Stat.* **38** 1–50.

Harvey, C. R. (1989). Time-varying conditional covariances in tests of asset pricing models. *Journal of Financial Economics.* **24** 289–317.

Henry, M., Zaffaroni, P., (2003). The long range dependence paradigm for macroeconomics and finance. In: Doukhan, P., Oppenheim, G., Taqqu, M. S. (Eds.), *Theory and Applications of Long-Range Dependence*, Birkhauser, Boston, 417–438.

He, X. and Shao, Q. M. (1996). A general Bahadur representation of M-estimators and its application to linear regression with nonstochastic designs. *Ann. Statist.* **24** 2608–2630.

Ho, H. C. and Hsing, T. (1997). Limit theorems for functionals of moving averages. *The Annals of Probability* **25** 1636-1669.

Holton, G. (2003). *Value-at-Risk: Theory and Practice.* Academic Press.

Hurst, H. E. (1951), Long-term storage capacity of reservoirs, *Trans. Am. Soc. Civ. Eng.*, **116**, 770–799.

Jorion, P. (2006). *Value at Risk: The New Benchmark for Managing Financial Risk* (3rd ed.). McGraw-Hill.

Kärner, O. (2002) On nonstationarity and antipersistency in global temperature series. *Journal Of Geophysical Research*, **107** NO. D20.

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

Kokoszka, P. S, and Taqqu, M.S. (1994). Infinite variance stable ARMA processes. *J.of Time Series Analysis*, **15**, 203-220.

Kokoszka, P. S. and Taqqu, M.S. (1995). Fractional ARIMA with stable innovations. *Stochastic Process and their Applications*, **60**, 19-47.

Komlós, J., Major, P. and Tusnády, G. (1975). An approximation of partial sums of independent

RV's and the sample DF. I. *Z. Wahrsch. Verw. Gebiete* **32**, 111–131; II. (1976) **34**, 33–58.

Konstantopoulos, T. and Sakhanenko, T. (2004). Convergence and convergence rate to fractional Brownian motion for weighted random sums. *Sib. Èlektron. Mat. Izv.* **1** 47-63.

Leipus, R and Surgailis, D. (2013). Asymptotics of partial sums of linear processes with changing memory parameter. *Lithuanian Mathematical Journal.* **53** 196-219.

Mallat, S., Papanicolaou, G. and Z. Zhang (1998). Adaptive Covariance Estimation of Locally Stationary Processes. *Annals of Statistics* **26** 1–47.

Mann, Michael E. (2011) On long range dependence in global surface temperature series. *Climatic Change* **107** 267–276.

Mercurio, D. and Spokoiny, V. (2004). Statistical inference for time-inhomogeneous volatility models. *The Annals of Statistics* **32** 577–602.

Mills T (2007) Time series modelling of two millenia of northern hemisphere temperatures: long memory or shifting trends? *J R Stat Soc Ser A* **170** 83–94

Nason, G. P., Von Sachs, R. and Kroisandt, G. (2000). Wavelet processes and adaptive estimation of the evolutionary wavelet spectrum. *Journal of the Royal Statistical Society: Series B* **62** 271–292.

Neumann, M. Hand Von Sachs, R. (1997). Wavelet thresholding in anisotropic function classes and application to adaptive estimation of evolutionary spectra. *The Annals of Statistics* **25** 38–76.

Ombao, H., von Sachs, R. and W. Guo (2005). SLEX analysis of multivariate nonstationary time series. *J. Amer. Statist. Assoc.* **100** 519–531.

Orbe, S., Ferreira, E. and Rodrigues-Poo, J., (2005). Nonparametric estimation of time varying parameters under shape restrictions. *Journal of Econometrics* **126** 53–77.

Palma, W. (2010). On the sample mean of locally stationary long-memory processes. *J.of Statistical Planning and Inference.* **140** 3764-3774.

Palma, W. and Olea, R. (2010). An efficient estimator for Gaussian locally stationary processes. *The Annals of Statistics* **38** 2958–2997.

Porter-Hudak, S., (1990), An application of the seasonally fractionally differenced model to the monetary aggregates, *Journal of the American Statistical Association* **85**, 338–344.

Priestley, M. B. (1988). *Nonlinear and nonstationary time series analysis.* Academic Press, London.

Ravn, M., Schmitt-Grohe, S. and Uribe, M. (2008). The macroeconomics of subsistence points. *Macroeconomic Dynamics* **12** 136–147.

Rea, W., Reale, M., Brown, J. and Oxley, L. (2011) Long memory or shifting means in geophysical time series? *Mathematics and Computers in Simulation*, **81**, 1441–453.

Rossi, E. and de Magistris, P. S. (2013). Long memory and tail dependence in trading volume and volatility. *Journal of Empirical Finance* **22** 94–112.

Roueff, F. and von Sachs, R. (2011) Locally stationary long memory estimation. *Stochastic Processes and their Applications* **121** 813–844.

Samorodnitsky, G. and Taqqu, M.(1994). Stable non-Gaussian random processes: Stochastic models with infinite variance.Chapman & Hall (New York).

Shao, Q. M. (1995). *Strong approximation theorems for independent random variables and their applications*, J. Multivariate Anal.**52**, 107-130.

Smith, R. L., 1993: Long range dependence and global warming. *Statistics for the Environment*, V. Barnett and F. Turkman, Eds., John Wiley, 141–161.

Stock, J. H. and Watson, M. W. (1996). Evidence on structural instability in macroeconomic time series relations. *Journal of Business and Economic Statistics.* **14** 11–30.

Sun, J. (1993). Tail Probabilities of the Maxima of Gaussian Random Fields.*The Annals of Probability.* **21** 34-71.

Sun, J and Loader, C. R. (1994). Simultaneous Confidence Bands for Linear Regression and Smoothing. *The Annals of Statistics.* **22** 1328-1345.

Tsay, R. S. (2005). *Analysis of Financial Time Series, second edition.* Wiley, Hoboken, New Jersey.

Wang, Q. Y, Lin, Y. X, Gulati, C. M. (2003). Strong Approximation for Long Memory Processes with Applications. *Journal of Theoretical Probability.* **16**,377-389.

Wu, W. B. (2007). Strong invariance principles for dependent random variables. *The Annals of probability* **35** 2294–2320.

Wu, W. B. and Zhao, Z. B (2007). Inference of Trends in Time Series. *Journal of the Royal Statistical Society, Series B* **69** 391-410(20).

Wu, W. B. and Zhou, Z. (2011). Gaussian approximations for non-stationary multiple time series. *Statistica Sinica* **21** 1397–1413.

Xia, Y. (1998) Bias-corrected confidence bands in nonparametric regression. *J. R. Statist. Soc. B* **60** 797–811.

Zaitsev, A. (2001). Multidimensional version of a result of Sakhanenko in the invariance principle for vectors with finite exponential moments. I *Theory Probab. Appl.* **45**, 624-641; II (2002a) **46**, 490-514; III (2002b) **46**, 676-698.

Zhang, T. and Wu, W.B. (2012). Inference of time-varying regression models.*Annals of Statistics.*

40, 1376-1402.

Zhou, Z.(2010) Nonparametric Inference of Quantile Curves for Non-stationary Time Series. *The Annals of Statistics*. **38** 2187-2217.

Zhou, Z. and Wu, W. B. (2009). Local linear quantile estimation for non-stationary time series. *The Annals of Statistics*. **37** 2696–2729.

Zhou, Z. and Wu, W. B. (2010). Simultaneous Inference of Linear Models with Time-Varying Coefficients. *Journal of the Royal Statistical Society, Series B.* **72** 513–531.

Zhou, Z. and Wu, W. B. (2011). On linear models with long memory and heavy-tailed errors. *Journal of Multivariate Analysis*. **102** 349–362.

Zhou, Z. (2013). Inference for non-stationary time-series autoregression. *Journal of Time Series Analysis*. **34** 508–516.