SIMULTANEOUS QUANTILE INFERENCE FOR NON-STATIONARY LONG-MEMORY TIME SERIES

WEICHI WU [1] AND ZHOU ZHOU

University College London and University of Toronto

April 20, 2017

Abstract

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

1 Introduction

There is an increasing need for non-stationary long-memory time series analyses in statistics and various applied fields, such as hydrology, geophysics, climate change, econometrics and quantitative finance. On the one hand, in the econometrics and quantitative finance literature, long memory has been empirically identified as one of the stylized facts for many financial time series data. We refer to Baillie (1996) and Henry and Zaffaroni (2003) for comprehensive reviews of long-memory processes in the finance and econometrics literature. In hydrology, Hurst (1951) found the well-known Hurst effect phenomenon in the geophysics record of water storage. In the geophysics literature, Haslett and Raftery (1989) assessed Ireland's wind power using a long-memory space-time model. In the climate change literature, numerous studies, such as Smith (1993), Eichner et al. (2003),

1. Corresponding author. Department of Statistics, University College, Gower Street, London WC1E 6BT UK.
E-mail : *w.wu@ucl.ac.uk*

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

Mills (2007), and Mann (2010), have investigated the long memory in surface temperature records.

On the other hand, it has long been recognized that the data-generating mechanisms do not remain unchanged for many financial, geophysical and engineering time series that span for at least moderately long periods of time. See, for instance, Cooley and Prescott (1976), Harvey (1989), Bekaert and Harvey (1995), Stock and Watson (1996), Orbe et al. (2005) and Ravn et al. (2008) for some representative papers in the finance and economics literature and Clarke (2007), Rea et al. (2011), and Kärner (2002) for some representative papers in the hydrology, geophysics and climate change literature. In the statistics literature, Mercurio and Spokoiny (2004), among others, proposed an approach for estimating and forecasting time-varying volatility. Dahlhaus and Rao (2006) and Fryzlewicz et al. (2008) analysed 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 this paper is to perform functional inference of the time-varying quantile curves for a class of non-stationary long-memory processes of the form

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

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

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

with i.i.d. η_i, and $\mu(t_i) = \mathbb{E}X_{i,n}$ is the deterministic trend function. In (1), long memory is introduced by allowing the coefficient functions $a_j(t)$ to decay slowly with j. The series $\{X_{i,n}\}$ is non-stationary since the functions $a_j(t)$ and $G(t, \cdot)$ vary with time t. In the following, we shall omit the subscript n in $X_{i,n}$ if no confusion arises. Indeed, $X_{i,n} = X_{i,n}(t_i)$ for some continuous time process $X_{i,n}(t)$. See (5) in Section 2 for the detailed definition of $X_{i,n}(t)$. Let $Q_{\alpha,n}(t)$ be the α_{th} quantile of $\{X_{0,n}(t)\}$ at time t, $0 \le t \le 1$. For a fixed $\beta \in (0, 1)$, we shall construct a $100(1-\beta)\%$ asymptotic simultaneous confidence band (SCB) for $Q_{\alpha,n}(t)$; i.e., we shall find random quantities $L_{\alpha,n}(t)$ and $U_{\alpha,n}(t)$ such that

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

Monitoring and inferring the quantile curves are very important tasks 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. VaR has become a widely used measure of market risk in risk management. We refer to Chapter 7 of Tsay (2010) 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, the simultaneous inference of $Q_{\alpha,n}(t)$ is a very important task because it allows the time-varying pattern of the market risk to be monitored with statistical guidance and confidence.

However, constructing quantile SCBs for non-stationary long-memory time series is a difficult problem. To our knowledge, there are currently no corresponding results in the literature. In general, the aforementioned problem can be solved if the following three tasks can be achieved. (i): Construct a uniform Bahadur representation for the quantile curves that approximates the deviation between the estimated $\hat{Q}_{\alpha,n}(t)$ and the true quantile $Q_{\alpha,n}(t)$ by linear forms of $\{X_{i,n}\}$ uniformly on $(0, 1)$. (ii): Approximate the partial sum process of the non-stationary long-memory process $\{X_{i,n}\}$ by a corresponding non-stationary 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 process of $\{X_{i,n}\}$. Note that due to long memory, the empirical process theories established for short-memory or independent data (see, for instance, Zhou (2010) and Pollard (1990)) 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 $Q_{\alpha,n}(t)$. The empirical process theory established here can further facilitate the asymptotic theory for a broad 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 widespread applications in statistics and probability and have received considerable attention in the literature. See, for instance, Komlós et al. (1975, 1976), Einmahl (1987a,b, 1989) and Zaitsev (2001, 2002a,b) for some thorough results for independent data; Dehling and Taqqu (1989) for a result on a class of stationary, long-range dependent empirical processes; 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 the 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 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. Thus, these results 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 non-stationary long-memory Gaussian processes. With the theoretical progress on issues (i)-(iii), in this paper, we construct SCBs for $Q_{\alpha,n}(t)$ with asymptotically correct coverage probabilities. The SCBs enable one to monitor and test the pattern and magnitude of the time-varying quantile curves, which, for instance, provides useful tools for the risk management of non-stationary long-memory financial time series.

The remainder of this 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. Some examples illustrating the theory are presented in Section 4. A discussion is provided in Section 5. Finally, proofs of the theoretical results are outlined in Section 6. The detailed proofs are relegated to the supplementary material.

2 Preliminaries

2.1 Notation

For a d-dimensional (random) vector $\mathbf{V} = (v_1, ..., v_d)^T$, write $|\mathbf{V}| = \sqrt{\sum_{i=1}^d v_i^2}$. A random vector $\mathbf{X}$ is said to be in $\mathcal{L}_p$, $p > 0$, if $\mathbb{E}(|\mathbf{X}|^p) < \infty$. In this case, let $\|\mathbf{X}\|_p =$

$(\mathbb{E}(|\mathbf{X}|^p))^{1/p}$ be its $\mathcal{L}_p$ norm, and write $\|\mathbf{X}\| := \|\mathbf{X}\|_2$ for short. Furthermore, for two series of real numbers x_n, y_n, denote $\lim_{n\to\infty} \frac{x_n}{y_n} = 1$ by $x_n \sim y_n$ for short. Write $f(x) \cong g(x)$ as $x \to \infty$ for $\lim_{x\to\infty} \left(\frac{f(x)}{g(x)}\right) = c$, where c is a finite non-zero constant. We say that $X_n = O_p(Y_n)$ if X_n is bounded by Y_n in probability and that $X_n = o_p(Y_n)$ if $X_n/Y_n \to_p 0$. To simplify the notation, we define that, for (random) vectors, $u = (u_1, .., u_n)^T, v = (v_1, ..., v_n)^T, \langle u, v\rangle = \sum_{i=1}^n u_i v_i$, and $|u|^2 := \langle u, u\rangle$. Let $\lfloor x \rfloor$ be the largest integer that is less than or equal to x. For any function of time $a(t)$, define $\dot{a}(t)$ as its partial derivative with respect to time t. Let $\mathcal{B}$ be the lag operator. For two functions $f(t)$ and $g(t)$, $t \in [0, 1]$, we write $f(t) \propto g(t)$ if some constant $C \neq 0$ exists such that $f(t) = Cg(t)$ for $t \in [0, 1]$. Write $t_i = i/n$. For an interval $\mathcal{I} \in \mathbb{R}$, denote by $\mathcal{C}^i\mathcal{I}, i \in \mathbb{N}$, the collection of functions that have i_{th}-order continuous derivatives on $\mathcal{I}$. Let $\mathbf{1}(\cdot)$ be the usual indicator function.

2.2 Assumptions

Suppose that we observe

$$X_{i,n} := \sum_{j=0}^{\infty} a_j(t_i)\varepsilon_{i-j,n} + \mu(t_i), 1 \leq i \leq n, \tag{4}$$

where the innovations $\varepsilon_{i,n} = G(t_i, \eta_i)$, $G(\cdot, \cdot)$ is a measurable function, $\{\eta_i\}_{i=-\infty}^{\infty}$ are i.i.d. random variables, and $\mathbb{E}(G(t, \eta_0)) = 0$ for $t \in [-\infty, 1]$. Observe that $X_{i,n} = X_{i,n}(t_i)$ with

$$X_{i,n}(t) = \sum_{j=0}^{\infty} a_j(t)G(t - t_j, \eta_{i-j}) + \mu(t), 0 \leq t \leq 1, \tag{5}$$

where $\mu(t) = \mathbb{E}(X_{i,n}(t))$ is a deterministic trend function that does not depend on n.

Remark 1. *Note that in (4), the innovations $\varepsilon_{i,n} = G(t_i, \eta_i)$ are independent but non-identically distributed. Allowing the innovations of the process to be non-stationary is very important for the quantile analysis of non-stationary time series since under this setting, the marginal distributions of $X_{i,n}$ are able to arbitrarily change over time. To observe this* *process, simply compare the following two simple models:*

$$X_{i,n}(t) = \sum_{j=0}^{\infty} \frac{a(t)}{(j+1)^{\beta}} G(t - t_j, \eta_{i-j}), \tag{6}$$

$$Y_i(t) = \sum_{j=0}^{\infty} \frac{a(t)}{(j+1)^{\beta}} \zeta_{i-j}, \tag{7}$$

where ζ_i are i.i.d. random variables with finite variance and $\beta > 1/2$. Let $Q_{Y,\alpha}(t)$ represent the α_{th} quantile curve of $Y_0(t)$. Define $Z = \sum_{j=0}^{\infty} \frac{1}{(j+1)^{\beta}} \zeta_{-j}$, and let Z_α be Z's α_{th} quantile. It is clear 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 $Q_{Y,b}(t) - Q_{Y,c}(t) \propto a(t)$. Consequently, we have that

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

The above restriction on the shapes of the quantile curves makes model (7) less useful for quantile analysis in many cases. In particular, under model (7), if the a_{th} and b_{th} quantile curves remain unchanged across time for some $a < b$, then (8) implies that the d_{th} quantile curve should also be a constant function over time for any $d \in (0, 1)$. However, in many practical situations, it is possible that some quantile curves remain constant while others exhibit interesting patterns of changes over time. Meanwhile, note that the set up for $X_{i,n}$ in (4) does not impose any restrictions on the shapes of the quantile curves.

Remark 2. *Traditionally, for the second-order stationary process X_i, it possesses long memory if $\sum_{j=-\infty}^{\infty} |\Gamma(j)| = \infty$, where $\Gamma(j) := \mathrm{Cov}(X_1, X_{1+j})$ is the autocovariance function. For non-stationary time series, one can extend the aforementioned classic definition of long memory and define the following uniform long-memory property of non-stationary time series:*

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

$$\lim_{n \to \infty} \sum_{j=-\infty}^{\infty} |\mathrm{Cov}(X_{i,n}, X_{(i+j),n})| = \infty, \tag{9}$$

where we set $X_{i,n} = 0$ if $i \leq 0$ or $i > n$ for convenience.

A simple sufficient condition for the process $\{X_{i,n}\}$ defined in (1) to be a uniform long-memory process is that, uniformly in t and j, $c \le a_j(t)/j^{d(t)} \le C$ or $-C \le a_j(t)/j^{d(t)} \le -c$ for some positive and finite constants c and C while $1/2 < d(t) < 1$. Here, $d(t)$ is called a (time-varying) long-memory parameter. Note that the above condition is not necessary. For example, the quantities $a_j(t)$ need not share the same sign for a fixed t; see Example 4.

Our objective is to estimate the α_{th} quantile $Q_{\alpha,n}(t)$ of $X_{0,n}(t)$. We have several assumptions, as follows:

(A0) For fixed α, $Q_{\alpha,n}(t)$, $\dot{Q}_{\alpha,n}(t) := \frac{\partial Q_{\alpha,n}(t)}{\partial t}$ and $\ddot{Q}_{\alpha,n}(t) := \frac{\partial^2 Q_{\alpha,n}(t)}{\partial t^2}$ are bounded on [0,1]. We also assume that $\dot{\mu}(t) := \frac{\partial}{\partial t}\mu(t)$ is bounded on [0,1].

(A1) A positive constant C exists such that $\|G(t,\eta_0)\|_p \le C$, and for $t, s \in (-\infty, 1]$, $\|G(t,\eta_0) - G(s,\eta_0)\|_p \le C|t-s|$ for some $p \ge 2$, $\mathbb{E}[G(t,\eta_0)] = 0$, $Var[G(t,\eta_0)] = \sigma^2(t)$, with $|\dot{\sigma}^2(t)|$ bounded for $t \in (-\infty, 1]$.

(A2) Let $g(t,x)$ be the density of $G(t,\eta_i)$. We require that $|\frac{\partial^r}{\partial x^r} g(t,x)|$ and $|\frac{\partial}{\partial t} g(t,x)|$ are bounded and integrable for $r = 0, 1, ..., l$, $l \ge 3$, $t \in (-\infty, 1]$ and $x \in \mathbb{R}$.

(A3) Coefficients $a_j(t)$ satisfy $|a_j(t)| = O(\frac{1}{(j+1)^\gamma})$ $\forall t \in [0,1], j \in \mathbb{N}$, $1/2 < \gamma < 1$. In addition, $a_j(t)$ has derivative $\dot{a}_j(t) := \frac{\partial a_j(t)}{\partial t}$ such that $\dot{a}_j(t) = O(\frac{1}{(j+1)^\gamma})$ for all $t \in [0,1]$. Without loss of generality, let $a_0(t) \equiv 1$ for all $t \in [0,1]$. For any $t \in [0,1]$, the series $\{X_{i,n}(t) - \mu(t)\}_{i=1}^n$ has time-invariant long-memory parameter $d(t) = 1 - \gamma$.

(A4) $K(\cdot) \in \mathbf{K}$, where $\mathbf{K}$ is the collection of kernel functions that are symmetric with support [-1,1] and are in $\mathcal{C}^1[-1,1]$. We write $K_{b_n}(\cdot) = K(\cdot/b_n)$ for short, where b_n is a bandwidth.

Note that (A4) implies that uniformly on any closed interval of $(0,1)$,

$$\begin{aligned} \Sigma_n(t) :&= \sum_{i=1}^n (1, (t_i - t)/b_n)^T (1, (t_i - t)/b_n) K_{b_n}(t_i - t) \\ &= n b_n \mu_K + O(1), \end{aligned} \tag{10}$$

where $\mu_K = \mathrm{diag}(1, \mu_2)$ and $\mu_2 = 2\int_0^1 u^2 K(u) du$.

Condition (A0) places some requirements on the smoothness of $Q_{\alpha,n}(t)$ to perform the local linear quantile regression. (A1) and (A2) make some assumptions on the tail behaviour of the innovations $\{\varepsilon_{i,n}\}_{i=-\infty}^{\infty}$ for technical convenience. 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 a sub-series of $\{X_{i,n}\}$ observed near some $t_0 \in [0,1]$, e.g. $\{X_{i,n}, |i/n - t_0| \le b_n\}$ for some $b_n \to 0$, then the sub-series is approximately stationary. (A3) also assumes that the long-memory parameter $d(t)$ is time invariant. Conditions (A0)-(A3) together imply that the density function of the process $X_{i,n}(t)$ is smooth in time; see Lemma 2. Such smoothness assumptions are also made when investigating the quantile curves of non-stationary and short-range dependent time series; see Zhou and Wu (2009). Condition (A4) makes mild assumptions on the kernel function $K(\cdot)$, which consequently results in the convergence of $\Sigma_n(t)$ in equation (10).

3 Main Results

Let $F_n(t,x) = \mathbb{P}(X_{i,n}(t) \le x)$, and let $f_n(t,x) = \frac{\partial}{\partial x} F_n(t,x)$. The existence of $f_n(t,x)$ is supported by Lemma 2. Define the quantiles $Q_{\alpha,n}(t) = \inf_x\{F_n(t,x) \ge \alpha\}$. We estimate $Q_{\alpha,n}(t)$ and $\dot{Q}_{\alpha,n}(t)$ by

$$\Big(\hat{Q}_{\alpha,n,b_n}(t), \hat{\dot{Q}}_{\alpha,n,b_n}(t)\Big) = \arg\min_{\beta_0,\beta_1} \sum_{i=1}^{n} \rho_\alpha\Big(X_{i,n} - \beta_0 - \beta_1(t_i - t)\Big) K_{b_n}(t_i - t), \qquad (11)$$

where $\rho_\alpha(x) = \alpha x^+ + (1-\alpha)(-x)^+$ is the check function in Koenker (2005). Equation (11) defines the local linear quantile estimators. In addition, the following notation is required for the main results. Let $\Psi_\alpha(x) := \alpha - \mathbf{1}(x \le 0)$ be the left derivative of $\rho_\alpha(x)$. Define $\hat{\theta}_{\alpha,n}(t) = \Big(\hat{\theta}_{\alpha,n,1}(t), \hat{\theta}_{\alpha,n,2}(t)\Big)^T := \Big(\hat{Q}_{\alpha,n}(t) - Q_{\alpha,n}(t), b_n(\hat{\dot{Q}}_{\alpha,n}(t) - \dot{Q}_{\alpha,n}(t))\Big)^T$. Let $\mathbf{z}_{i,n}(t) = \big(1, (t_i - t)/b_n\big)^T$. Let $S_{\alpha,n}(t) = S_{\alpha,n}\left(t, (0,0)^T\right)$, where for $\theta = (\theta_1, \theta_2)^T$, we define

$$S_{\alpha,n}(t,\theta) = \sum_{i=1}^{n} \Psi_\alpha\left(X_{i,n} - Q_{\alpha,n}(t) - (t_i - t)\dot{Q}_{\alpha,n}(t) - \theta^T \mathbf{z}_{i,n}(t)\right) K_{b_n}(t_i - t)\mathbf{z}_{i,n}(t). \quad (12)$$

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), and Wu (2007), among others. In the local polynomial quantile regression literature, Chaudhuri (1991) provided a Bahadur representation for i.i.d. d-dimensional observations, and Zhou and Wu (2009) provided a Bahadur representation for non-stationary series with short-range dependence. For linear models with stationary long memory and heavy-tailed errors, Zhou and Wu (2011) provided a Bahadur representation for regression parameters estimated using a general convex check function ρ (which includes OLS and quantile regression). For functionals of Gaussian dependent sequences, Coeurjolly (2008) obtained a 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 that $b_n \to 0$, $nb_n / \log^2 n \to \infty$, $(nb_n)^{1/2-\gamma}(b_n)^{-1/p} \to 0$, and $\inf_n \inf_{t\in[0,1]} f_n(t, Q_{\alpha,n}(t)) \geq \eta > 0$ for some positive constant η. Assume (A0)-(A4). Then, we have the following uniform Bahadur representation:*

$$\begin{aligned}
&\sup_{t\in T_n} |f_n(t, Q_{\alpha,n}(t))\mu_K \hat{\theta}_{\alpha,n}(t) - S_{\alpha,n}(t)/(nb_n)| \\
&= O_p\left((\pi_n)^{1/2} \log n/\sqrt{nb_n} + (nb_n)^{1/2-\gamma}\pi_n b_n^{-1/p} + b_n\pi_n + (\pi_n)^2\right), \qquad (13)
\end{aligned}$$

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 approximated by $S_{\alpha,n}(t)$, which has a considerably simpler mathematical form. Consequently, Theorem 1 enables us to construct the SCBs of $Q_\alpha(t)$ over $t \in T_n$ via a Gaussian process approximating $\{S_{\alpha,n}(t), t \in T_n\}$. The Gaussian approximation can be obtained using Theorem 2 and Theorem 3, as follows.

3.2 Gaussian Approximation

Theorem 2. *Under conditions (A0)-(A4), on a possibly richer probability space, there exists* $Y_{k,n} = \sum_{j=0}^{\infty} a_j(t_k)\sigma(t_{k-j})\upsilon_{k-j} + \mu(t_k)$*, where the random variables* υ_i *are i.i.d.* $N(0,1)$*, such that*

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

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 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 the strong approximation of the partial sum of independent random variables. Zaitsev (2001, 2002a,b) extended the previous univariate results to the multi-dimensional case. In the context of non-stationary short-range dependent processes, Wu and Zhou (2011) acquired a Gaussian approximation result of the partial sums with O_p bounds. For stationary long-memory processes, Wang et al. (2003) proposed a strong approximation result. For more details about the strong approximation, see Csörgő 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 long-memory Gaussian processes:

Theorem 3. *Suppose that the conditions of Theorem 1 hold. Suppose that* $0 < \iota_1 < \iota_2 < 1$ *exist such that* $n^{-\iota_2} = o(b_n)$ *and* $b_n = o(n^{-\iota_1})$*. Assume that* $\gamma > \frac{1}{2} + \frac{1}{p}$ *and* $(nb_n)^{\frac{1}{2}-\gamma} b_n^{-\frac{2}{p}} = o(1)$*,* $b_n^3 (nb_n)^{\gamma-\frac{1}{2}} = o(1)$*,* $b_n^{-\frac{1}{p}}(nb_n)^{\gamma-1} = o(1)$*, and* $n^{\frac{(\frac{1}{2}-\frac{1}{p})(\frac{1}{2}-\gamma)}{\frac{1}{2}+\frac{1}{p}}} b_n^{\gamma-3/2} = o(1)$*. Then, on a possibly richer probability space, a sequence of i.i.d. standard normal random variables* $\{\vartheta_i\}_{i=-\infty}^{\infty}$ *exists such that for* $V_{i,n} = \sum_{j=0}^{\infty} a_j(t_i)\vartheta_{i-j}$*, we have*

$$\sup_{t\in T_n}\left| f_n(t, Q_{\alpha,n}(t))\left(\mu_K \hat{\theta}_{\alpha,n}(t) - \frac{\sigma(t)}{nb_n}\sum_{i=1}^{n} V_{i,n} K_{b_n}(t_i - t)\mathbf{z}_{i,n}(t) - \frac{b_n^2 \ddot{Q}_{\alpha,n}(t)(\mu_2, 0)^T}{2}\right)\right| = O_p(\varsigma_n), \tag{14}$$

where

$$\varsigma_n = \zeta_n + K_n^p/nb_n, K_n^p = nb_n^4 + \log n\sqrt{nb_n} + b_n^{-1/p} g_n + (b_n)^{1-\frac{1}{p}}(nb_n)^{3/2-\gamma},$$
$$g_n = (nb_n)^{2-2\gamma}\left(\log(nb_n)\mathbf{1}(\gamma = 3/4) + \mathbf{1}(\gamma < 3/4)\right) + (nb_n)^{1/2}\mathbf{1}(\gamma > 3/4),$$
$$\zeta_n = (\pi_n)^{1/2}\log n/\sqrt{nb_n} + (nb_n)^{1/2-\gamma}\pi_n b_n^{-1/p} + b_n\pi_n + (\pi_n)^2 + n^{\frac{1/2-\gamma}{1/2+1/p}}/(b_n) + n^{1/2-\gamma}b_n^{-1/p},$$
$$\pi_n = (nb_n)^{-1/2}\left(\log n + (nb_n^5)^{1/2} + (nb_n)^{1-\gamma}b_n^{-1/p}\right).$$

Straightforward calculations show that the sequence ς_n satisfies the property that $\frac{\varsigma_n}{(nb_n)^{1/2-\gamma}} \to 0$ as $n \to \infty$.

This theorem follows from Theorem 2, Lemma A1, and Lemma C2 in the supplementary material. Since the density $f_n(t, Q_{\alpha,n}(t))$ is uniformly bounded from below by a strictly positive number (see Lemma 3 for a detailed discussion on the uniform lower bound of $f_n(t, Q_{\alpha,n}(t))$), then after cancelling this quantity on both sides of equation (14), we have an approximation of $\hat{\theta}_{\alpha,n}(t)$ that is independent of the nuisance function $f_n(t, Q_{\alpha,n}(t))$. This differs 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 were obtained by Csörgő and Kulik (2008), among others. Once we establish Theorem 3, we find that the bias of $\hat{Q}_{\alpha,n}(t)$ is on the order of b_n^2, while the standard deviation of $\frac{1}{nb_n}\sum_{i=1}^n V_{i,n}K_{b_n}(t_i - t)\mathbf{z}_{i,n}(t)$ is on the order of $(nb_n)^{1/2-\gamma}$. Straightforward calculations show that the optimal b_n to minimize the MSE of the estimates should be on the order of $n^{\frac{1/2-\gamma}{3/2+\gamma}}$, which is feasible when additionally assuming $\frac{3}{p} < \gamma < \frac{4+\frac{1}{p}}{4+\frac{2}{p}}$ and further leads to the convergence rate

$$\varsigma_n = b_n^{3-1/p} + b_n^{\frac{1}{\gamma-1/2}-1/p} + b_n^{\frac{3/2+\gamma}{1/2+1/p}-1} + b_n^{3/2+\gamma-1/p}. \tag{15}$$

Let $p \to \infty$ at the rate of $\log n$. We find that if γ becomes close to either 0.5 or 1, ς_n will approach b_n^2, which is on the order of the square root of the MSE except for a factor of multiplicative logarithms (due to the extra factor of $p^{1/2}$ in the approximating order of Lemmas B4, B5, C1 and C2 in the supplementary material; proof of Theorem 3). In practice, if $\{a_j(t), t \in [0,1]\}_{j=1}^{\infty}$ can be estimated consistently, then Theorem 3 can be used to construct the SCB of $Q_{\alpha,n}(t)$ by generating a large sample of i.i.d. copies of $\{\frac{1}{nb_n}\sum_{i=1}^n V_{i,n}K_{b_n}(t_i - t)\}$ and calculating the empirical maximum deviations of the simulated samples. Theoretically, Theorem 3 can be used to explore the limiting distribution of the SCB, which we will discuss in the next section.

3.3 Maximum Deviation

Many researchers have conducted excellent investigations on the maximum deviations of Gaussian processes. For instance, the extreme Gumbel distribution for stationary Gaussian processes was obtained by Berman (1972) in i.i.d. settings. Bickel and Rosenblatt (1973) is a good reference for this context, and it also concludes with a limiting distribution for maximum deviations of a type of non-stationary Gaussian processes. Sun and Loader (1994) acquired a first-order approximation of the maximum deviations for a general type of Gaussian processes. In the next theorem, we find a limiting confidence band by referring to Sun and Loader's results and techniques:

Theorem 4. *Suppose that $K(x)$ is non-decreasing when $x \leq 0$, non-increasing when $x > 0$, and has a bounded non-increasing first-order derivative on $[0, 1]$. Let $T_n = [\delta b_n, 1 - \delta b_n]$ for some $\delta > 1$, $\{\vartheta_i\}_{i \in \mathbb{Z}}$ be a series of i.i.d. N(0,1), $V_{i,n} = \sum_{j=0}^{\infty} a_j(t_i)\vartheta_{i-j}$, and $\breve{S}_n(t) = \sum_{i=1}^{n} V_{i,n} K_{b_n}(t_i - t)$. Then, assuming that (a): the conditions of Theorem 3 hold, we have*

$$\lim_{n\to\infty} \mathbb{P}\left(\sup_{t \in T_n} \frac{nb_n|\hat{Q}_{\alpha,n}(t) - Q_{\alpha,n}(t) - b_n^2 \ddot{Q}_{\alpha,n}(t)\mu_2/2|}{\sigma(t)\|\sum_{i=1}^{n} V_{i,n} K_{b_n}(t_i - t)\|} > \sqrt{2 \log \frac{\kappa_n}{\pi\tau}}\right) = \tau, \tag{16}$$

where $\kappa_n = \int_{t \in T_n} \|\frac{\partial}{\partial t}(\frac{\breve{S}_n(t)}{\|\breve{S}_n(t)\|})\| dt$, and $1 - \tau$ is the nominal coverage probability. If we further assume that (b): $\exists\ 0 < \mathcal{L} \leq \mathcal{U} < \infty$ such that $\frac{\mathcal{L}}{(j+1)^\gamma} \leq a_j(t) \leq \frac{\mathcal{U}}{(j+1)^\gamma}\ \forall j \geq 0,\ t \in T_n$, then κ_n satisfies

$$C_1/b_n \leq \kappa_n \leq C_2/b_n \tag{17}$$

for some constants $0 < C_1 < C_2 < \infty$.

When n is sufficiently large, we can find explicit bounds for C_1 and C_2 in (17). Let

$a^* = \max_{a\in[0,1]} |K'(a)|(1-a)^{3/2-\gamma}$ and define

$$G(K(\cdot)) = \frac{\mathbb{L}_{1-\gamma}}{(1-\gamma)^2}\int_{-1}^{1}\int_{-1}^{1} K(x)K(y)|x-y|^{1-2\gamma}dxdy, \tag{18}$$

$$\text{where } \mathbb{L}_{1-\gamma} = \frac{1}{3-2\gamma} + \int_0^\infty \left((x+1)^{1-\gamma} - x^{1-\gamma}\right)^2 dx. \tag{19}$$

Then, we have

$$\begin{aligned} C_1 &= \frac{\mathcal{L}a^*}{\mathcal{U}G(K(\cdot))^{1/2}(1-\gamma)(3-2\gamma)^{1/2}}, \\ C_2 &= \frac{a_1 M\mathcal{U}}{\mathcal{L}G(K(\cdot))^{1/2}}, \\ \text{where } \ a_1 &= \left(\frac{4w^{1-2\gamma}}{2\gamma-1} + \frac{w(2+w)^{2-2\gamma}}{(1-\gamma)^2} + \frac{2^{3-2\gamma}}{(3-2\gamma)(1-\gamma)^2}\right)^{1/2}, \\ w &= (\delta-1)/2, \ M = \sup_{x\in[-1,1]} |K'(x)|. \end{aligned}$$

In Theorem 4, we impose assumption (b) to ensure that the norm of the partial sum of $V_{i,n}$ goes to infinity at a fairly stable rate as the sample size increases. In general, due to the non-stationarity, the exact value of κ_n is difficult to evaluate. However, the theorem obtains a bound for $\kappa_n b_n$, consequently ensuring the order of the width of the SCB. The term $\sigma(t)$ can be estimated, say, by the local linear estimators. The term $\|\sum_{i=1}^n V_{i,n}K_{b_n}(t_i - t)\|$ determines the width of the SCB. Under our setting, we have the next corollary on the order of $\|\sum_{i=1}^n V_{i,n}K_{b_n}(t_i - t)\|$.

Corollary 1. *Assume that the conditions of Theorem 4, including (a) and (b), hold. Then, $\|\sum_{i=1}^n V_{i,n}K_{b_n}(t_i - t)\|$ is on the order of $(nb_n)^{3/2-\gamma}$.*

The proof of Corollary 1 is relegated to Section 1 of the supplementary material. A nicer form of κ_n and $\|\sum_{i=1}^n V_{i,n}K_{b_n}(t_i - t)\|$ can be obtained under slightly stronger assumptions.

Lemma 1. *Let $\Gamma_n(i,j) = \mathrm{Cov}(V_{i,n}, V_{j,n})$. Assume the following:*

(a) The conditions of Theorem 3 hold.

(b) $\Gamma_n(i,j) \sim \breve{a}(t_i,t_j)(|i-j|+1)^{1-2\gamma}$, where $\breve{a}(x,y)$ is Lipschitz continuous in both x and y for $x, y \in [0,1]^2$.

(c) $\breve{a}(t,t)$ *has a strictly positive lower bound and a finite upper bound. Meanwhile,* $\frac{\partial^2}{\partial x^2}\breve{a}(x,y)$, $\frac{\partial^2}{\partial y^2}\breve{a}(x,y)$, *and* $\frac{\partial^2}{\partial x \partial y}\breve{a}(x,y)$ *are bounded.*

Then, we have

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

and

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

By combining the above with (16)*, we obtain*

$$\lim_{n\to\infty} \mathbb{P}\left(\sup_{t\in T_n} \frac{|\hat{Q}_{\alpha,n}(t) - Q_{\alpha,n}(t) - b_n^2 \ddot{Q}_{\alpha,n}(t)\mu_2/2|}{\breve{a}^{1/2}(t,t)\sigma(t) \left(\int_{-1}^{1} \int_{-1}^{1} |x-y|^{1-2\gamma} K(x)K(y) dxdy \right)^{1/2} (nb_n)^{1/2-\gamma}} \right.$$
$$\left. > \sqrt{2\log \frac{D}{b_n \pi \tau}} \right) = \tau. \tag{20}$$

The proof of Lemma 1 is relegated to Section 1 of the supplementary material.

Remark 3. *Note that condition (b) of Theorem 4 ensures that* κ_n *is on the order of* $1/b_n$*. In Lemma 1, we make an assumption on the covariance structure, which helps us evaluate the limit of* κ_n*. Thus, we only need (a) and not (b) of Theorem 4 to support Lemma 1. Note that there is no need to estimate* $\{a_j(t), t \in [0,1]\}_{j=0}^{\infty}$ *to apply Lemma 1. Rather, we need to estimate* $\breve{a}^{1/2}(t,t)\sigma(t)$ *to apply this lemma.*

The following corollary shows that if the functions $a_j(t)$ can be factorized as $a_j(t) = a(t)g(j)$ and the innovations $\{\varepsilon_j\}_{j=-\infty}^{\infty}$ are i.i.d. with a finite p_{th} moment, then a Gumbel limiting distribution can be achieved under certain conditions.

Corollary 2. *Assume that the conditions of Theorem 1 hold and that* $n^{\gamma-1}\log n / b_n^{3/2-\gamma} = o((\log n)^{-1/2})$*. Suppose that* $a_j(t) = a(t)g(j)$, $g_j = \frac{(1-\gamma)}{(j+1)^\gamma}(1 + O(1/j))$*, and the innovations* ε_i *are i.i.d. with mean 0 and variance 1 s.t.* $X_{i,n}(t) = X_i(t) = \sum_{j=0}^{\infty} a_j(t)\varepsilon_{i-j} + \mu(t)$*. In* *addition, define*

$$\kappa_K^2 = \int_{\mathbb{R}} \int_{\mathbb{R}} K(x)K(y)|x-y|^{1-2\gamma} dxdy, \tag{21}$$

$$D_K = \int_{\mathbb{R}} \int_{\mathbb{R}} K'(x)K'(y)|x-y|^{1-2\gamma} dxdy. \tag{22}$$

For $m \geq 3$*, define* $B_K(m) = \sqrt{2\log m} + \frac{1}{2\sqrt{2\log m}}(\log C_K - 2\log 2 - 2\log \pi)$*,* $C_K = D_K/\kappa_K^2$*. Then, we have*

$$\lim_{n\to\infty} \mathbb{P}\left(\sup_{t\in[b_n, 1-b_n]} \left| (1-\gamma)\left(\frac{\hat{Q}_{\alpha,n}(t) - Q_{\alpha,n}(t) - b_n^2 \ddot{Q}_{\alpha,n}(t)\mu_2/2}{\mathbb{L}_{1-\gamma}^{1/2} \kappa_K (nb_n)^{1/2-\gamma} a(t)} \right) \right| - B_K(1/b_n) \right.$$
$$\left. \leq \{2\log(1/b_n)\}^{-1/2} u \right) = \exp\{-2\exp(-u)\}, \tag{23}$$

where $\mathbb{L}_{1-\gamma}$ *is defined in* (19)*.*

The proof of Corollary 2 is relegated to Section 1 of the supplementary material.

4 Examples

In this section, we assume that conditions (A0), (A1), (A2) and (A4) hold. To apply our theory to the general examples, we shall evaluate condition (A3).

Example 1. *Consider the following fractionally integrated model:* $X_{i,n}(t) = (1-\mathcal{B})^{-d} a(t) G(t_i, \eta_i)$*. Assume that* $a(t)$ *is a smooth function of* t*, which has a bounded first-order derivative. Let* $\gamma = 1-d$*, where* p *and* γ *satisfy the bandwidth conditions in Theorem 3. Then, the theory established in this paper can be used to obtain the SCBs of the quantile curves of* $X_{i,n}$*. Note that if* $G(t, \eta_i) \equiv \eta_i$*, where innovations* $\{\eta_i\}_{i=1}^{\infty}$ *are i.i.d. with mean* 0 *and variance* 1*, then the model is reduced to a locally stationary fractional ARIMA(0,d,0) model.*

Example 2. *Consider the locally stationary fractional ARIMA(p,d,q) model, as follows:* $\Phi^p(\mathcal{B}, t)X_i(t) = \Theta^q(\mathcal{B}, t)(1-\mathcal{B})^{-d}\varepsilon_i$*, where the random variables* ε_i *are i.i.d. 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* p *and* q*, respectively; and* $0 < d < 1/2$*. Suppose that* $\{\phi_i(t), 1 \leq$

$i \le p\}$ and $\{\theta_j(t), 1 \le j \le q\}$ are twice differentiable in t. Define a polynomial with $p+q$ degrees of freedom:

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

Suppose that for all $t \in [0,1]$, $\Phi^p(z,t)$ and $\Theta^q(z,t)$, $\Phi^p(z,t)$ and $\Xi^{p+q}(z,t)$ do not have the same roots, and $\Phi^p(z,t)$ does not have roots in the unit disk $\{|z| \le 1\}$. Let $G(z,t) = \frac{\Theta^q(z,t)}{\Phi^p(z,t)} := \sum_{j=0}^{\infty} c_j(t) z^j$; then, $G(z,t)$ is analytic in the circle $\{|z| \le R(t)\}$ for some $R(t) > 1$. Now suppose that a number $\mathcal{Q}_1$ exists such that $1 < \mathcal{Q}_1 < R(t)$ 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 a number $\mathcal{Q}_2$ exists such that $1 < \mathcal{Q}_2 < r(t)$ for all $t \in [0,1]$. Then, condition (A3) is satisfied with $\gamma = 1-d$. In addition, if the innovation has a finite p_{th} moment such that p and γ satisfy the bandwidth conditions in Theorem 3, then our theory for the quantile curves applies to this case.

To demonstrate that condition (A3) holds for Example 2, we first carefully check Kokoszka and Taqqu (1994) and conclude that $|c_j(t)| \le C_1 \mathcal{Q}_1^{-j}$ for all t and some sufficiently large constant C_1. Then, by applying Lemma 3.2 in Kokoszka and Taqqu (1995), we conclude that $|a_j(t)| \le C_2 j^{d-1}$ for all $t \in [0,1]$ and some sufficiently large constant C_2. By applying similar arguments to the following locally stationary fractional ARIMA$(2p, d, p+q)$ model,

$$(\Phi^p(\mathcal{B},t))^2 X_i(t) = \{\dot{\Theta}^q(\mathcal{B},t)\Phi^p(\mathcal{B},t) - \Theta^q(\mathcal{B},t)\dot{\Phi}^p(\mathcal{B},t)\}(1-\mathcal{B})^{-d}\varepsilon_i,$$

we have that $|\dot{a}_j(t)| \le C_3 j^{d-1}$ for all $t \in [0,1]$ and some sufficiently large constant C_3. □

Remark 4. *Consider the time-varying fractional ARIMA(p,d,q) model with $0 < d < 1/2$,*

$$(\sum_{j=0}^{p} \alpha_j(t_i)\mathcal{B}^j) Z_{i,n} = (\sum_{k=0}^{q} \beta_k(t_i)\mathcal{B}^k)(1-\mathcal{B})^{-d}\sigma(t_i)\bar{\eta}_i, \tag{24}$$

where innovations $\{\bar{\eta}_i\}_{i=-\infty}^{\infty}$ are i.i.d. random variables with mean 0 and variance 1 and

$\alpha_0(\cdot) = \beta_0(\cdot) \equiv 1$. *It can be shown that model (24) has an MA representation:*

$$Z_{i,n} = \sum_{j=0}^{\infty} a_{i,n}(j)\bar{\eta}_{i-j}. \tag{25}$$

It can also be shown that, similar to Dahlhaus and Polonik (2009), we cannot find functions $a_j(t)'s$ *that satisfy condition (A3) such that* $a_j(t_i) = a_{i,n}(j)$. *However, consider the following locally stationary fractional ARIMA model with* $0 < d < 1/2$*:*

$$(\sum_{j=0}^{p} \alpha_j(t)\mathcal{B}^j)X_i(t) = (\sum_{k=0}^{q} \beta_k(t)\mathcal{B}^k)(1-\mathcal{B})^{-d}\sigma(t)\bar{\eta}_i. \tag{26}$$

Note that $\mathcal{B}$ *only affects* i *and not* t. *Under some regularity conditions,* (26) *has an MA representation* $X_{i,n}(t) = \sum_{j=0}^{\infty} \tilde{a}_j(t)\bar{\eta}_{i-j}$ *for some MA coefficients* $\tilde{a}_j(t)'s$ *satisfying (A3). We have discussed such conditions in Example 2. It has been shown that under short-range dependence, time-varying AR models can be well approximated using a locally stationary AR model. See, for instance, Zhang and Wu (2012) and Zhou (2013). Proposition 1 shows that with long-range dependence, the time-varying fractional ARIMA model can still be well approximated by a locally stationary fractional ARIMA model.*

Proposition 1. *Consider model (24) and model (26). Suppose the following:*
(a) The start point $(Z_{p,n}, ..., Z_{1,n})^T \in \mathcal{L}_2$.
(b) The coefficients $\{\alpha_j(\cdot), \beta_k(\cdot), j = 1, ..., p, k = 1, ..., q\}$ *and* $\sigma(\cdot)$ *are Lipschitz continuous on* $[0, 1]$.
(c) $\sum_{j=1}^{p} \alpha_j(t)z^j \neq -1$ *for all* $|z| \leq 1 + c$ *with some* $c > 0$ *uniformly for* $t \in [0, 1]$. *Then, we have for some constant* $C > 0$

$$\max_{1\leq i\leq n} \|Z_{i,n} - X_i(t_i)\| \leq Cn^{d-1/2}. \tag{27}$$

In addition, if $\sigma(\cdot)$ *is constant, then we have* $\max_{1\leq i\leq n} \|Z_{i,n} - X_i(t_i)\| \leq Cn^{-1}$.

The proof of Proposition 1 is relegated to Section 1 of the supplementary material.

Example 3. *Consider the locally stationary fractional ARIMA* $(0, d, 1)$ *model (26). Palma*

(2010) showed that

$$\text{Cov}(X_s(t_s), X_m(t_m)) \sim g(t_s, t_m)|s-m|^{2d-1} \tag{28}$$

for some $\mathcal{C}^1$ *function* $g(\cdot,\cdot)$ *on* $[0,1]\times[0,1]$ *and* $d \in (0, 1/2)$. *In addition,* $g(t,t) > 0$ *for* $t \in [0,1]$. *Assume that* $g(x,y)$ *is smooth such that* $\frac{\partial^2}{\partial x^2} g(x,y)$, $\frac{\partial^2}{\partial y^2} g(x,y)$, *and* $\frac{\partial^2}{\partial x \partial y} g(x,y)$ *exist and are all bounded. Then, as we discussed in Example 2, our Theorems 1–4 hold for this model provided that the bandwidth conditions in Theorem 3 hold. In addition, the conditions of Lemma 1 are also satisfied due to the covariance structure (28). Thus, we can compute the asymptotic SCB via Lemma 1 if consistent estimates of* $g(t,t)$ *are provided. Note that* $g^{1/2}(t,t)$ *now 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:*

$$(\sum_{j=0}^{p} \alpha_j(t)\mathcal{B}^j)(1 - 2\xi\mathcal{B} + \mathcal{B}^2)^\lambda X_i(t) = (\sum_{k=0}^{q} \beta_k(t)\mathcal{B}^k)\sigma(t)\bar{\eta}_i, \tag{29}$$

where $\alpha_0(\cdot) = \beta_0(\cdot) \equiv 1$, $0 < \lambda < 0.25$, $|\xi| < 1$ *and innovations* $\bar{\eta}_i$ *are i.i.d. with mean* 0 *and variance* 1. *Write* $\Phi^p(z,t) = \sum_{j=0}^{p} \alpha_j(t) z^j$ *and* $\Theta^q(z,t) = \sum_{k=0}^{q} \beta_k(t) z^k$. *Suppose that* $\Phi^p(z,t)$ *and* $\Theta^q(z,t)$ *satisfy the same conditions as those listed in Example 2. The Gegenbauer ARMA process was considered by Gray et al. (1989) and Gray et al. (1994).*

Under our settings, model (29) 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 $z_1 = \cos\theta + i\sin\theta$, $z_2 = \bar{z}_1$ such that z_1 and z_2 are the solutions of $1 - 2\xi z + z^2 = 0$. Hence, we obtain

$$\begin{aligned} X_i(t) &= \sum_{j=0}^{\infty} \psi(j) z_1^j \mathcal{B}^j \sum_{k=0}^{\infty} \psi(k) z_2^k \mathcal{B}^k \sum_{l=0}^{\infty} c_l(t)\bar{\eta}_{i-l} \\ &= \sum_{j=0}^{\infty} \sum_{k=0}^{j} c_k(t) \sum_{s=0}^{j-k} \psi(s) z_1^s \psi(j-k-s) z_2^{j-k-s} \bar{\eta}_{i-j}, \end{aligned} \tag{30}$$

where $\psi(j) = \frac{\Gamma(\lambda+j)}{\Gamma(\lambda)\Gamma(j+1)} \cong j^{\lambda-1}$. (30) implies that

$$a_j(t) := \sum_{k=0}^{j} c_k(t) \sum_{s=0}^{j-k} \psi(s) z_1^s \psi(j-k-s) z_2^{j-k-s} := \sum_{k=0}^{j} c_k(t)\nu_{j-k}. \tag{31}$$

Then, we have that for $k \leq j$,

$$\sum_{s=0}^{j-k} |\psi(s) z_1^s \psi(j-k-s) z_2^{j-k-s}| \leq C|(j-k)^{\lambda-1} + \sum_{s=1}^{j-k-1} s^{\lambda-1}(j-k-s)^{\lambda-1}|, \tag{32}$$

which leads to $\nu_{j-k} = O((j-k)^{2\lambda-1})$. By $\sum_{j=0}^{\infty}(|c_j(t)| + |\dot{c}_j(t)|) < \infty$ and the summation by parts formula, we have that $|a_j(t)| = O(j^{2\lambda-1})$. Similar arguments yield that $|\dot{a}_j(t)| = O(j^{2\lambda-1})$. Then, condition (A3) is satisfied with $\gamma = 1 - 2\lambda$. In addition, if the bandwidth conditions for Theorem 3 are satisfied, then our theory for the quantile curves applies to model (29). Note that for fixed t, the quantities $a_j(t)$ do not necessarily have the same sign. To illustrate this fact, simply note that if $j-k$ is odd, then $\nu_{j-k} = \sum_{s=0}^{j-k} \cos((j-k-2s)\theta)\psi(s)\psi(j-k-s)$; if $j-k$ is even, then $\nu_{j-k} = 1 + 2\sum_{s=0}^{(j-k)/2-1} \cos((j-k-2s)\theta)\psi(s)\psi(j-k-s)$, and $\cos(\cdot)$ is a periodic function.

Example 5. *Consider the locally stationary seasonal fractional ARIMA(p,d,s,q) model:* $\Phi^p(\mathcal{B}, t)X_i(t) = \Theta^q(\mathcal{B}, t)(1-\mathcal{B}^s)^{-d}\varepsilon_i$, *where the random variables* ε_i *are i.i.d. with mean* 0 *and variance* 1, *and* $\Phi^p(z,t)$ *and* $\Theta^q(z,t)$ *satisfy the same conditions as those listed in Example 2. According to Baillie (1996), the seasonal fractional ARIMA(0,d,s,0) model has an MA representation* $X_i = \sum_{j=0}^{\infty} \psi_j \varepsilon_{i-j}$ *such that* $\psi_j = O(j^{d-1})$. *Then, through similar arguments to those in Example 4, the locally stationary seasonal fractional ARIMA(p,d,s,q) model has a locally stationary MA representation* $X_i(t) = \sum_{j=0}^{\infty} a_j(t)\varepsilon_{i-j}$ *such that* $|a_j(t)| = O(j^{d-1})$ *and* $|\dot{a}_j(t)| = O(j^{d-1})$. *The seasonal fractional ARIMA model is considered by Porter-Hudak (1990) to model monetary aggregates.*

5 Discussion

A small number of recent papers discuss non-stationary time series with long memory; see, for instance, Palma and Olea (2010), Palma (2010), and Leipus and Surgailis

(2013). The majority of the aforementioned papers consider mean or spectral analysis of the series. We observed that the time-varying long-memory parameter $d(t)$ is allowed in some of the papers. Among them, Beran (2009) proposed a nonparametric method for estimating the time-varying long-memory parameters. Roueff and Von Sachs (2011) advanced a semi-parametric method for estimating time-varying long-memory parameters $d(t)$ and investigated its asymptotic behaviour. Leipus and Surgailis (2013) studied the limiting behaviour of the partial sums of a linear process with time-varying $d(t)$. Palma (2010) proposed a method for estimating the sample mean for locally stationary processes with time-varying $d(t)$. In this paper, we only considered the case in which the memory parameter is a constant over time. However, our results can readily be extended to a broad class of non-stationary long-memory processes with time-varying memory parameters. Note that in the context of simultaneous inference of quantile curves of non-stationary long-memory processes, the stochastic variability 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 finitely many non-overlapping subintervals of $(0,1)$. Hence, for many scenarios in which time-varying memory parameters are allowed, the construction of SCB for the quantile curves is essentially the same as those 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 estimating $max_{s\in[0,1]}d(s)$ and determining $\mathcal{T}_n$, which we shall leave as a rewarding future work.

6 Proofs

In the following proofs, we shall only show the case where $\alpha = 1/2$ because the proofs of the other quantiles follow by the same arguments. We shall also omit the subscript α if no confusion arises. We also omit the subscript n from $Q_{\alpha,n}(\cdot)$, $\dot{Q}_{\alpha,n}(\cdot)$ and $\ddot{Q}_{\alpha,n}(\cdot)$. In the proofs, the constant C represents a generic finite constant that may vary from place to place. We also write X_i for $X_{i,n}(t_i)$ if no confusion arises. Define $Y_i(t) = X_i - \delta_{ni}(t)$, where $\delta_{ni}(t) = Q(t) + (t_i - t)\dot{Q}(t)$. Let $\mathcal{F}_j$ be the σ-field generated by $(..., \eta_{j-1}, \eta_j)$.

Lemma 2. *Suppose that conditions (A0), (A1), (A2) and (A3) hold. Then, we have that*

$f_n(t,x) \in \mathcal{C}([0,1]\times\mathbb{R})$, where $f_n(t,\cdot)$ (defined in Section 3) is the density of $X_{i,n}(t)$. Furthermore, $\frac{\partial}{\partial t}f_n(t,x)$ and $\frac{\partial}{\partial x}f_n(t,x)$ are bounded.

As in our comment on the assumptions, conditions (A0)-(A3) ensure the smoothness of the density function of the time series $X_{i,n}(t)$. Lemma 2 formally states this result and is important for the proof of Theorem 3. The proof of Lemma 2 is relegated to Section 1 of the supplementary material.

Lemma 3. *Assume (A0)-(A3). Assume that for a sufficiently large number M with $M \geq \sup_{t\in[0,1]}(|\mu(t)| + \max_n |Q_{\alpha,n}(t)|)$, $\eta > 0$ exists such that $\inf_{t\in(-\infty,1],|x|\leq M} g(t,x) \geq \eta$. Then, we have that a positive η_0 exists such that $\inf_n f_n(t, Q_{\alpha,n}(t)) \geq \eta_0 > 0$ for $t \in [0,1]$.*

When the density function of the time series evaluated at the considered quantile is bounded away from 0, the deviation between the estimated quantile and true quantile can be approximated by a certain Gaussian process independent of the density function. We show this effect in Theorem 3. In addition, Theorem 4, Lemma 1, and Corollaries 1 and 2 all assume that $\inf_n \inf_{t\in[0,1]} f_n(t, Q_{\alpha,n}(t)) \geq \eta_0 > 0$. Meanwhile, Lemma 3 provides a sufficient condition for $\inf_n \inf_{t\in[0,1]} f_n(t, Q_{\alpha,n}(t)) \geq \eta_0 > 0$. The proof of Lemma 3 is relegated to Section 1 of the supplementary material.

Lemma 4. *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 $p \geq 1$. Assume that for any $t \in [0,1]$, $\|\Upsilon_n(t)\|_p = O(m_n)$, $\|\dot{\Upsilon}_n(t)\|_p = O(l_n)$, and m_n, l_n are sequences of real numbers such that $m_n \leq M l_n$ for some large constant M. Then, we have $\sup_{t\in[0,1]} |\Upsilon_n(t)| = O_p(m_n(\frac{m_n}{l_n})^{-\frac{1}{p}})$. In particular, if $p = 2$, then we have $\sup_{t\in[0,1]} |\Upsilon_n(t)| = O_p(\sqrt{m_n l_n})$.*

Lemma 4 is of general interest. This lemma provides a convenient way to evaluate the probabilistic bound of the maximum of a series of random processes that are smooth in t, and it is important for the proofs of Theorems 1, 3 and 4. The proof of Lemma 4 is relegated to Section 1 of the supplementary material.

Proof of Theorem 1

The theorem follows from equation (S.26) in the supplementary material, Technical Lemma B3, Lemma B4, Lemma B5 in the supplementary material, Lemma 8 in Zhou and Wu (2009) and Taylor expansion. See Section 2 of the supplementary material for more details. □

Proof of Theorem 2

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

Define $Z_j = \sum_{i=0}^{j}\varepsilon_{-i,n}$ with $Z_j = 0$ for $j \le -1$ and $W_j = \sum_{i=1}^{j}\varepsilon_{i,n}$ with $W_j = 0$ for $j \le 0$. Although $\{W_j, j \in \mathbb{Z}\}$ and $\{Z_j, j \in \mathbb{Z}\}$ depend on n, we omit the subscript n to shorten the notation. Then, similar arguments as in Wang et al. (2003) lead to

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

for some integer N. Redefine $\{\varepsilon_{j,n}, j \in \mathbb{Z}\}$ on a richer probability space. By taking $x = n^{1/p}\varepsilon^{1/p}$, $\forall\varepsilon > 0$ in Theorem B of Shao (1995), and condition (A3), we have that independent centred normal random variables $\{\varrho_j, j \in \mathbb{Z}\}$ with $\mathrm{var}(\varrho_j) = \sigma^2(t_j)$ exist such that for any n,

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

Consequently, we obtain

$$\begin{aligned}\sum_{k=1}^{n}(X_k - Y_k) = &\sum_{j=1}^{n-1}\big(\sum_{k=j}^{n} a_{k-j}(t_k) - \sum_{k=j+1}^{n} a_{k-j-1}(t_k)\big)(W_j - W_j^*) + (W_n - W_n^*)a_0(1)\\ &+ \sum_{j=0}^{N-1}\sum_{k=1}^{n}\big(a_{k+j}(t_k) - a_{k+j+1}(t_k)\big)(Z_j - Z_j^*) + (Z_N - Z_N^*)\sum_{k=1}^{n} a_{k+N}(t_k)\\ &+ \sum_{j=N+1}^{\infty}\varepsilon_{-j}\sum_{k=1}^{n} a_{k+j}(t_k) - \sum_{j=N+1}^{\infty}\varrho_{-j}\sum_{k=1}^{n} a_{k+j}(t_k)\\ &:= A + B + C + D + E + F,\end{aligned} \tag{33}$$

where $Y_k = \sum_{j=0}^{\infty} a_j(t_k)\varrho_{k-j} + \mu(t_k)$, $W_j^* = \sum_{i=1}^{j} \varrho_i$, $Z_j^* = \sum_{j=0}^{j} \varrho_{-j}$. Let $N = \lfloor n^\alpha \rfloor + 1$ for some $\alpha > 1$. Direct calculations and condition (A3) lead to

$$|A| \le C\zeta_n \sum_{j=1}^{n-1} \left(\frac{1}{n} \sum_{j=0}^{n-j-1} |j+1|^{-\gamma} + (n-j+1)^{-\gamma} \right) = O_p(n^{1/p+1-\gamma}), \tag{34}$$

$$|B| \le \zeta_n = O_p(n^{1/p}),\ |C| = O_p(n^{\alpha(1+1/p-\gamma)}), \tag{35}$$

$$|D| \le M\zeta_N^* \sum_{k=1}^{n} (k+N)^{-\gamma} = O_p(N^{1/p}(n+N)^{1-\gamma}) = O_p(n^{\alpha(1+1/p-\gamma)}), \tag{36}$$

$$|E| = O_p(n^{1+\alpha(1/2-\gamma)}),\ |F| = O_p(n^{1+\alpha(1/2-\gamma)}). \tag{37}$$

Note that $\alpha = \frac{1}{1/2+1/p}$ is the solution of $1 + \alpha(1/2-\gamma) = \alpha(1+1/p-\gamma)$. Then, Theorem 2 follows from allowing $\upsilon_i = \varrho_i/\sigma(t_i)$. □

Proof of Theorem 3

Theorem 1, Theorem 2, and Lemmas C1 and C2 in the supplementary material lead to

$$\sup_{t \in T_n} \left| f_n(t, Q(t)) \left(\mu_K \hat{\theta}_{\alpha,n}(t) - \frac{1}{nb_n} \sum_{i=1}^{n} \sigma(t_i) V_i K_{b_n}(t_i - t) \mathbf{z}_{i,n}(t) - \frac{b_n^2 \ddot{Q}(t)(\mu_2, 0)^T}{2} \right) \right| = O_p(p^{1/2}\varsigma_n). \tag{38}$$

Lemma 4 and direct calculations lead to

$$\sup_{t \in (0,1)} \left| \sum_{i=1}^{n} (\sigma(t) - \sigma(t_i)) V_i K_{b_n}(t_i - t) \right| = O_p(p^{1/2}(nb_n)^{3/2-\gamma} b_n^{1-1/p}). \tag{39}$$

Then, Theorem 3 follows from equation (38) and equation (39). □

Key Idea of Proof of Theorem 4.

Let $S_n(t) = \sum_{i=1}^{n} K_{b_n}(t_i - t)V_i$. By changing the order of the summation, it can be shown that S_n can be approximated by $\sum_{j=1-N}^{n} T_j(t)\epsilon_j = \langle T(t), \epsilon\rangle$, where N is a sufficiently large constant, $T(t) = (T_{1-N}(t), ..., T_n(t))^T$ is a non-random coefficient vector, and $\{\epsilon_j, j \in \mathbb{Z}\}$ are i.i.d. standard normal random variables. Write $\epsilon = (\epsilon_{1-N}, ..., \epsilon_n)$. Proposition 1 in Sun and Loader (1994) shows that

$$\mathbb{P}(\sup_{t\in T_n} |\langle T(t)|T(t)|^{-1}, \epsilon\rangle| > c) = \tau, \tag{40}$$

where $\tau = \frac{\kappa_0}{\pi}\exp(-\frac{c^2}{2}) + 2(1-\Phi(c)) + o(\exp(-\frac{c^2}{2}))$ as $c \to \infty$, $\Phi(\cdot)$ is the CDF of $N(0,1)$, and $\kappa_0 = \int |\frac{\partial}{\partial x}(T(x)|T(x)|^{-1})|dx$. Then, the theorem follows from mathematically manipulating $T(t)$. The detailed proof of Theorem 4 is relegated to Section 4 of the supplementary material.

References

Baillie, R. T. (1996). 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. *Journal of Statistical Planning and Inference*, 139:900–915.

Berman, S. M. (1972). Maximum and High Level Excursion of a Gaussian Process with Stationary Increments. *The Annals of Mathematical Statistics*, 43(4):1247–1266.

Bickel, P. J. and Rosenblatt, M. (1973). On Some Global Measures of the Deviations of Density Function Estimates. *The Annals of Statistics*, 1(6):1071–1095.

Chaudhuri, P. (1991). Nonparametric Estimates of Regression Quantiles and Their Local Bahadur Representation. *The Annals of Statistics*, 19:760–777.

Clarke, R. T. (2007). Hydrological Prediction in a Non-stationary World. *Hydrology and Earth System Sciences*, 11:408–414.

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.

Csörgő, 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.

Csörgő, M. and Révész, P. (1981). *Strong Approximations in Probability and Statistics*. Academic Press.

Dahlhaus, R. and Polonik, W. (2009). Empirical Spectral Processes for Locally Stationary Time Series. *Bernoulli*, 15(1):1–39.

Dahlhaus, R. and Subba Rao S. (2006). Statistical Inference for Time-Varying ARCH Processes. *Annals of Statistics*, 34:1075–1114.

Dehling, H., and Taqqu, M. S. (1989). The Empirical Process of Some Long-Range Dependent Sequences with an Application to U-statistics. *Annals of Statistics*, 17:1767–1783.

Eichner, J. F., Koscielny-Bunde, E., Bunde, A., Havlin, S., and Schellnhuber, H.-J. (2003). Power-Law Persistence and Trends in the Atmosphere: a Detailed Study of Long Temperature Records. *Physical review. E, Statistical, nonlinear, and soft matter physics*, 68:046133.

Einmahl, U. (1987a). A Useful Estimate in the Multidimensional Invariance Principle. *Probability Theory and Related Fields*, 76(1):81–101.

Einmahl, U. (1987b). Strong Invariance Principles for Partial Sums of Independent Random Vectors. *The Annals of Probability*, 15(4):1419–1440.

Einmahl, U. (1989). Extensions of Results of Komlós, Major, and Tusnády to the Multivariate Case. *Journal of Multivariate Analysis*, 28(1):20–68.

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., Zhang, N.-F., and Woodward, W.A. (1989). On Generalized Fractional Processes. *Journal of Time Series Analysis*, 10:233–257.

Gray, H. L., Zhang, N.-F., and Woodward, W.A. (1994). On Generalized Fractional processes-a correction. *Journal of Time Series Analysis*, 15:561–562.

Härdle, W. (1989). Asymptotic Maximal Deviation of M-smoothers. *Journal of Multivariate Analysis*, 29(2):163–179.

Harvey, C. R. (1989). Time-Varying Conditional Covariances in Tests of Asset Pricing Models. *Journal of Financial Economics*, 24:289–317.

Haslett, J. and Raftery, A. (1989). Space-Time Modelling with Long-memory Dependence: Assessing Ireland's Wind Power Resource. *Applied Statistics*, 38:1–50.

He, X. and Shao, Q.-M. (1996). A General Bahadur Representation of M-estimators and Its Application to Linear Regression with Nonstochastic Designs. *The Annals of Statistics*, 24(6):2608–2630.

Henry, M. and Zaffaroni, P. (2003). The Long Range Dependence Paradigm for Macroeconomics and Finance. *Theory and Application of Long-Range Dependence (P. Doukhan, G.Oppenheim, M.S. Taqqu, eds). Birkhäuser, Boston,* 417–438.

Ho, H.-C. and Hsing, T. (1997). Limit Theorems for Functionals of Moving Averages. *The Annals of Probability*, 25(4):1636–1669.

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

Hurst, H. E. (1951). Long Term Storage Capacity of Reservoirs. *Transactions of the American Society of Civil Engineers*, 116:770–799.

Jorion, P. (2006). *Value at Risk, 3rd Ed.: The New Benchmark for Managing Financial Risk*, volume 9. McGraw Hill Professional.

Kärner, O. (2002). On Nonstationarity and Antipersistency in Global Temperature Series. *Journal of Geophysical Research D: Atmospheres*, 107.

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

Kokoszka, P. S. and Taqqu, M. S. (1994). Infinite Variance Stable ARMA Processes. *Journal of Time Series Analysis*, 15(2):203–220.

Kokoszka, P. S. and Taqqu, M. S. (1995). Fractional ARIMA with Stable Innovations. *Stochastic Processes 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. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, 32:111–131.

Komlós, J., Major, P., and Tusnády, G. (1976). An Approximation of Partial Sums of Independent RV'-s, and the Sample DF. II. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, 34:33–58.

Leipus, R. and Surgailis, D. (2013). Asymptotics of Partial Sums of Linear Processes with Changing Memory parameter. *Lithuanian Mathematical Journal*, 53(2):196–219.

Mann, M. E. (2010). On Long Range Dependence in Global Surface Temperature Series. *Climatic Change*, 107(3-4):267–276.

Mercurio, D. and Spokoiny, V. (2004). Statistical Inference for Time-Inhomogeneous Volatility Models. *Annals of Statistics*, 32:577–602.

Mills, T. C. (2007). Time Series Modelling of Two Millennia of Northern Hemisphere Temperatures: Long Memory or Shifting Trends? *Journal of the Royal Statistical Society. Series A: Statistics in Society*, 170:83–94.

Orbe, S., Ferreira, E., and Rodriguez-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. *Journal of Statistical Planning and Inference*, 140:3764–3774.

Palma, W. and Olea, R. (2010). An Efficient Estimator for Locally Stationary Gaussian Long-Memory Processes. *Annals of Statistics*, 38:2958–2997.

Pollard, D. (1990). Empirical Processes: Theory and Applications. *NSF-CBMS regional conference series in probability and statistics* **2**. Institute of Mathematical Statistics and American Statistical Association.

Porter-Hudak, S. (1990). An Application of the Seasonal Fractionally Differenced Model to the Monetary Aggregates. *Journal of the American Statistical Association*, 85(410):338–344.

Ravn, M. O., Schmitt-Grohé, S., and Uribe, M. (2008). Macroeconomics of Subsistence Points. *Macroeconomic Dynamics*, 12(S1):136–147.

Rea, W., Reale, M., Brown, J., and Oxley, L. (2011). Long Memory or Shifting Means in Geophysical Time Series? In *Mathematics and Computers in Simulation*, volume 81, pages 1441–1453.

Roueff, F. and Von Sachs, R. (2011). Locally Stationary Long Memory Estimation. *Stochastic Processes and their Applications*, 121:813–844.

Shao, Q. (1995). Strong Approximation Theorems for Independent Random Variables and Their Applications. *Journal of Multivariate Analysis*, 52(1):107–130.

Smith, R. L. (1993). Long Range Dependence and Global Warming. In Barnett, V. and Turkman, F., editors, *Statistics for the Environment.* John Wiley.

Stock, J. H. and Watson, M. W. (1996). Evidence on Structural Instability in Macroeconomic Time Series Relations. *Journal of Business & 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(3):1328–1345.

Tsay, R. S. (2010). *Analysis of Financial Time Series.* John Wiley & Sons.

Wang, Q., Lin, Y.-X., and Gulati, C. M. (2003). Strong Approximation for Long Memory Processes with Applications. *Journal of Theoretical Probability*, 16(2):377–389.

Wu, W. B. (2007). Strong Invariance Principles for Dependent Random Variables. *Annals of Probability*, 35:2294–2320.

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. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 60(4):797–811.

Zaitsev, A. Y. (2001). Multidimensional Version of a Result of Sakhanenko in the Invariance Principle for Vectors with Finite Exponential Moments. I. *Theory of Probability & Its Applications*, 45(4):624–641.

Zaitsev, A. Y. (2002a). Multidimensional Version of a Result of Sakhanenko in the Invariance Principle for Vectors with Finite Exponential Moments. II. *Theory of Probability & Its Applications*, 46(3):490–514.

Zaitsev, A. Y. (2002b). Multidimensional Version of a Result of Sakhanenko in the Invariance Principle for Vectors with Finite Exponential Moments. III. *Theory of Probability & Its Applications*, 46(4):676–698.

Zhang, T. and Wu, W. B. (2012). Inference of Time-varying Regression Models. *The Annals of Statistics*, 40(3):1376–1402.

Zhou, Z. (2010). Nonparametric Inference of Quantile Curves for Nonstationary Time Series. *The Annals of Statistics*, 38(4):2187–2217.

Zhou, Z. (2013). Heteroscedasticity and Autocorrelation Robust Structural Change Detection. *Journal of the American Statistical Association*, 108(502):726–740.

Zhou, Z. and Wu, W. B. (2009). Local Linear Quantile Estimation for Nonstationary Time Series. *Annals of Statistics*, 37:2696–2729.

Zhou, Z. and Wu, W. B. (2011). On Linear Models with Long Memory and Heavy-Tailed Errors. *Journal of Multivariate Analysis*, 102:349–362.