

## NONPARAMETRIC INFERENCE OF QUANTILE CURVES FOR NONSTATIONARY TIME SERIES<sup>1</sup>

BY ZHOU ZHOU

*University of Toronto*

The paper considers nonparametric specification tests of quantile curves for a general class of nonstationary processes. Using Bahadur representation and Gaussian approximation results for nonstationary time series, simultaneous confidence bands and integrated squared difference tests are proposed to test various parametric forms of the quantile curves with asymptotically correct type I error rates. A wild bootstrap procedure is implemented to alleviate the problem of slow convergence of the asymptotic results. In particular, our results can be used to test the trends of extremes of climate variables, an important problem in understanding climate change. Our methodology is applied to the analysis of the maximum speed of tropical cyclone winds. It was found that an inhomogeneous upward trend for cyclone wind speeds is pronounced at high quantile values. However, there is no trend in the mean lifetime-maximum wind speed. This example shows the effectiveness of the quantile regression technique.

**1. Introduction.** After fitting a nonparametric model, one often asks whether it can be simplified into certain parametric, semiparametric or more parsimonious nonparametric forms. Recently, there has been an enormous interest in developing nonparametric specification tests; see, for example, Hall and Titterton (1988), Eubank and Speckman (1993), Härdle and Mammen (1993), Ingster (1993), Zheng (1996), Hart (1997), Stute (1997), Xia (1998), Horowitz and Spokoiny (2001), Fan, Zhang and Zhang (2001) and Fan and Jiang (2007) among others. Many of the previous results concern nonparametric inference of the (conditional) mean or density functions for independent data.

The primary goal of this paper is to perform nonparametric specification tests of quantile curves for a class of nonstationary processes that can be called locally stationary processes [Draghicescu, Guillas and Wu (2009) and Zhou and Wu (2009)]. Conceptually, the local stationarity is characterized by the smoothly time-varying data generating mechanisms of the processes. More precisely, let  $\{X_{i,n}\}_{i=1}^n$  be the observed sequence. We shall adopt the following formulation:

$$(1) \quad X_{i,n} = G(i/n, \mathcal{F}_i), \quad i = 1, 2, \dots, n,$$

---

Received August 2008; revised October 2009.

<sup>1</sup>Supported in part by the NSF Grant DMS-04-78704.

*AMS 2000 subject classifications.* Primary 62G10; secondary 60F17.

*Key words and phrases.* Simultaneous confidence band, integrated squared difference test, quantile estimation, nonstationary nonlinear time series, local stationarity, Gaussian approximation, climate change.

where  $\mathcal{F}_i = (\dots, \varepsilon_{i-1}, \varepsilon_i)$ ,  $\varepsilon_i$ ,  $i \in \mathbb{Z}$ , are independent and identically distributed (i.i.d.) random variables, and  $G: [0, 1] \times \mathbb{R}^\infty \mapsto \mathbb{R}$  is a measurable function such that  $\zeta_i(t) := G(t, \mathcal{F}_i)$  is a properly defined random variable for all  $t \in [0, 1]$ . Here, local stationarity means that the random function  $\zeta_i(t)$  is smooth in  $t \in [0, 1]$  in an appropriate sense. Process (1) covers a wide range of nonstationary linear and nonstationary nonlinear processes and it naturally extends many existing stationary time series models into the nonstationary setting. See Zhou and Wu (2009) for more discussion and examples. In the sequel, for notational convenience, we shall write  $X_{i,n}$  as  $X_i$ .

The paper is motivated by the important problem of understanding the trends of extremes and variability of climate variables. As stated in Katz and Brown (1992), “*understanding climate change demands attention to changes in climate variability and extremes.*” By interpreting climate extremes as upper and lower quantiles and climate variability as interpercentile ranges, the nonparametric quantile estimation [Koenker (2005)] provides a simple and effective means to address the latter problem. On the other hand, climatologists often want to know whether a simple linear or quadratic function is appropriate to describe the trends. The parametric description of the trends is more interpretable and efficient than its nonparametric counterparts if the parametric model is correctly specified. To address the latter issue, it is necessary to develop nonparametric specification tests for the quantile curves.

For independent data, there have been a few results on nonparametric specification tests of quantile curves. He and Zhu (2003) and Kim (2007) proposed tests based on the cusum process of the residuals and the Rao-score statistics, respectively. The tests use fitted values and residuals under the null hypothesis (namely the parametric model) and therefore enjoy the advantage of avoiding the need for nonparametric function estimation. On the other hand, however, the above tests are sensitive only to a restrictively small class of alternatives; namely alternatives in the neighborhood of the parametric null. An alternative specification test which achieves minimax rate over a large class of smooth functions is proposed in Horowitz and Spokoiny (2002). It seems that the test is tailored to the specific problem of testing linearity. For other contributions, see Zheng (1998), Rosenkrantz (2000) and Wang (2007, 2008).

To our knowledge, testing parametric functional forms of quantile curves under the time series setting has not yet been considered in the literature. In this paper, we shall conduct the specification tests by directly comparing the nonparametric quantile estimates and the parametric null. In principle, if some global measure such as the  $\mathcal{L}^\infty$  norm or the  $\mathcal{L}^2$  norm is large for the difference between the nonparametric fits and the parametric null, then there is evidence against the parametric null hypothesis. This approach is intuitively plausible and it is easy to understand. For nonparametric inference of the mean functions, this very idea has been widely applied; see, for example, Eubank and Speckman (1993) and Härdle and Mammen (1993) among others.

Nevertheless, obtaining the asymptotic distributions of appropriately normalized global measures of the deviations of the quantile curves has been a very difficult problem when dependence is present. Generally speaking, the latter problem can be solved if we have (i) uniform Bahadur representations of the estimated quantile curves and (ii) a sharp Gaussian approximation result for the weighted empirical processes of the nonstationary time series  $(X_i)_{i=1}^n$ . Recently, Zhou and Wu (2009) obtained a sharp uniform Bahadur representation for the local linear quantile estimates of locally stationary time series. On the other hand, Wu and Zhou (2010) established Gaussian approximation results for partial sums of nonstationary time series with nearly optimal rates. An extension of the latter result addresses the above issue (ii); see Theorem 4 below.

With the recent progress in Bahadur representations and Gaussian approximations for locally stationary time series, we are able to construct in this paper simultaneous confidence bands (SCB) (based on the  $\mathcal{L}^\infty$  norm) and integrated squared difference tests (ISDT) (based on the  $\mathcal{L}^2$  norm) for the quantile curves as tools for the nonparametric inference. The SCB is shown to asymptotically achieve the correct coverage probability. Moreover, we prove that the ISDT asymptotically attain the correct type I error rate and are asymptotically optimal in terms of rates of convergence for nonparametric hypothesis testing in the sense of Ingster (1993). Our results shed new light on nonparametric specification tests of  $M$ -type estimates for nonstationary time series.

The rest of the paper is structured as follows. Section 2 introduces the local linear quantile estimates and the dependence measures. Section 3 presents the asymptotic results on the nonparametric specification tests. In particular, asymptotic results on the SCB and ISDT are presented in Sections 3.2 and 3.3, respectively. A wild bootstrap procedure is introduced in Section 4 to address the issue of slow convergence of the SCB and ISDT tests. Section 4 also contains discussions on bandwidth selection and nuisance parameter estimation. Section 5 presents two simulation studies to compare accuracy and sensitivity of various nonparametric tests. The global tropical cyclone data is studied in Section 6. Proofs are given in Section 7.

**2. Preliminaries.** We now introduce some notation. For a vector  $\mathbf{v} = (v_1, v_2, \dots, v_p) \in \mathbb{R}^p$ , let  $|\mathbf{v}| = (\sum_{i=1}^p v_i^2)^{1/2}$ . For a  $p \times p$  matrix  $A$ , define  $|A| = \sup\{|A\mathbf{v}| : |\mathbf{v}| = 1\}$ . For a random vector  $\mathbf{V}$ , write  $\mathbf{V} \in \mathcal{L}^q$  ( $q > 0$ ) if  $\|\mathbf{V}\|_q := [\mathbb{E}(|\mathbf{V}|^q)]^{1/q} < \infty$  and  $\|\mathbf{V}\| = \|\mathbf{V}\|_2$ . Denote by  $\Rightarrow$  the weak convergence. For an interval  $\mathcal{I} \subset \mathbb{R}$ , denote by  $\mathcal{C}^i \mathcal{I}$ ,  $i \in \mathbb{N}$ , the collection of functions that have  $i$ th order continuous derivatives on  $\mathcal{I}$ , and, for  $\mathcal{D} \subset \mathbb{R}^d$ , let  $\mathcal{CD}$  be the collection of real-valued functions that are continuous on  $\mathcal{D}$ . A function  $f : \mathbb{R}^d \rightarrow \mathbb{R}$  is said Lipschitz continuous on  $\mathcal{D} \subset \mathbb{R}^d$  if there exists a finite constant  $C$ , such that  $|f(x_1) - f(x_2)| \leq C|x_1 - x_2|$  for all  $x_1, x_2 \in \mathcal{D}$ . For  $x \in \mathbb{R}$ , define  $x^+ = \max(x, 0)$ . The symbol  $C$  denotes a finite generic constant which may vary from line to line.

2.1. *Local linear quantile estimator.* Recall  $\zeta_i(t) = G(t, \mathcal{F}_i)$ . Let  $F(t, x) = \mathbb{P}(\zeta_i(t) \leq x)$ ,  $x \in \mathbb{R}$ , be the cumulative distribution function (cdf) of  $\zeta_i(t)$ ,  $t \in [0, 1]$ ; let  $Q_\alpha(t)$  be the  $\alpha$ th quantile function of  $\zeta_i(t)$ ,  $\alpha \in (0, 1)$ , namely  $Q_\alpha(t) = \inf_x \{F(t, x) \geq \alpha\}$ . Suppose  $Q_\alpha(t)$  is smooth on  $[0, 1]$ . As  $Q_\alpha(t_1) \approx Q_\alpha(t) + (t_1 - t)Q'_\alpha(t)$  for  $t_1$  close to  $t$ ,  $Q_\alpha(t)$  and  $Q'_\alpha(t)$  can be estimated by the local linear approach [Koenker (2005)]

$$(\hat{Q}_{\alpha, b_n}(t), \hat{Q}'_{\alpha, b_n}(t)) = \arg \min_{(\beta_0, \beta_1)} \sum_{i=1}^n \rho_\alpha(X_i - \beta_0 - \beta_1(t_i - t)) K_{b_n}(t_i - t),$$

where  $t_i = i/n$ ,  $\rho_\alpha(x) = \alpha x^+ + (1 - \alpha)(-x)^+$  is the check function. Here  $K$  is a kernel function,  $K_{b_n}(\cdot) = K(\cdot/b_n)$  and  $b_n = b_n(\alpha) > 0$  is the bandwidth depending on  $\alpha$ . We shall omit the subscript  $b_n$  in  $\hat{Q}$  and  $\hat{Q}'$  hereafter if no confusion will be caused.

2.2. *The dependence measures.* It is shown [Condition (B1) of Zhou and Wu (2009)] that the asymptotic behavior of the local linear quantile estimator is determined by the dependence structure of

$$F_k(t, x, \mathcal{F}_i) := \frac{\partial^k \mathbb{P}\{G(t, \mathcal{F}_{i+1}) \leq x | \mathcal{F}_i\}}{\partial x^k}, \quad k = 0, 1, 2, 3.$$

To quantify the dependence structure of the above processes, let us consider a generic nonlinear system  $\{H(t, x, \mathcal{F}_i)\}_{i \in \mathbb{Z}}$ , where  $(t, x) \in [0, 1] \times \mathbb{R}$  and  $H: [0, 1] \times \mathbb{R} \times \mathbb{R}^\infty \mapsto \mathbb{R}$  is a measurable function such that  $H(t, x, \mathcal{F}_i)$  is well defined for all  $(t, x) \in [0, 1] \times \mathbb{R}$ . Let  $\varepsilon'_i$ ,  $i \in \mathbb{Z}$ , be an i.i.d. copy of  $\varepsilon_j$ ,  $j \in \mathbb{Z}$ . For  $k \geq 0$ , let  $\mathcal{F}_k^* = (\mathcal{F}_{-1}, \varepsilon'_0, \varepsilon_1, \dots, \varepsilon_k)$  and define the physical dependence measure

$$(2) \quad \delta_H(k, p) = \sup_{(t, x) \in [0, 1] \times \mathbb{R}} \|H(t, x, \mathcal{F}_k) - H(t, x, \mathcal{F}_k^*)\|_p.$$

Here we recall  $\|\cdot\|_p = [\mathbb{E}(|\cdot|^p)]^{1/p}$ . Note that  $\delta_H(k, p)$  measures the overall dependence of  $H(t, x, \mathcal{F}_k)$  on the input  $\varepsilon_0$ . The physical dependence measures are by their definition closely related to the data generating mechanism and hence are easy to work with; see Section 4 of Zhou and Wu (2009) for the related calculations for locally linear and nonlinear time series models.

We shall call the system  $\{H(\cdot, \cdot, \mathcal{F}_i)\}_{i \in \mathbb{Z}}$  uniformly geometric moment contracting of order  $p$  [UGMC( $p$ )] if  $\delta_H(k, p)$  decays exponentially with respect to  $k$ ; namely

$$(3) \quad \delta_H(k, p) = O(\chi^k), \quad 0 < \chi < 1.$$

Following Section 4 of Zhou and Wu (2009), condition (3) is readily verifiable for a large class of nonstationary nonlinear processes and nonstationary linear models. All our results will be presented in terms of the physical dependence measures and the UGMC conditions.

Note that UGMC(2) is stronger than the stability condition of Zhou and Wu (2009). Analogous results of this paper can be proved with  $\delta_H(k, p)$  decaying algebraically at a sufficiently fast rate. However, the technical details are much lengthier and we chose to use the UGMC condition for clarity of presentation.

**3. Main results.**

3.1. *Assumptions.* We shall make the following assumptions on the process  $(X_i)$  and the kernel  $K$ .

- (A1)  $f(t, x)$  is Lipschitz continuous on  $[0, 1] \times \mathbb{R}$ , where  $f(t, \cdot)$  is the density function of  $\zeta_i(t)$ . Assume  $\inf_{t \in [0, 1]} f(t, Q_\alpha(t)) > 0$ .
- (A2) Let  $J(t, x, \mathcal{F}_i) = I\{G(t, \mathcal{F}_i) \leq x\}$  and

$$(4) \quad \sigma^2(t) = \sum_{i=-\infty}^{\infty} \text{cov}[J(t, Q_\alpha(t), \mathcal{F}_0), J(t, Q_\alpha(t), \mathcal{F}_i)].$$

Assume  $\sigma(t)$  is Lipschitz continuous on  $[0, 1]$  and  $\inf_{t \in [0, 1]} \sigma(t) > 0$ .

- (A3) (Stochastic Lipschitz continuity condition.) There exists  $q \geq 1$ , such that  $\|\zeta_i(t_1) - \zeta_i(t_2)\|_q \leq C|t_1 - t_2|$  holds for all  $t_1, t_2 \in [0, 1]$ .
- (A4) Assume that, for  $k = 1, 2, 3$ ,  $F_k(t, x, \mathcal{F}_i) := \partial^k F(t, x, \mathcal{F}_i) / \partial x^k$  exists and  $\sum_{k=0}^{\infty} \delta_{F_i}(k, 4) < \infty$ , for  $i = 1, 2, 3$ .
- (A5)  $F(t, x, \mathcal{F}_i)$  is UGMC(4).
- (A6) There exists  $C_0 < \infty$  such that  $\sup_{(t,x) \in [0, 1] \times \mathbb{R}} F_1(t, x, \mathcal{F}_i) < C_0$  almost surely.
- (K1)  $K(\cdot) \in \mathcal{K}$ , where  $\mathcal{K}$  is the collection of density functions  $K$  such that  $K$  is symmetric with support  $[-1, 1]$  and  $K \in C^1[-1, 1]$ . For  $K(\cdot) \in \mathcal{K}$  and  $j \geq 1$ , define

$$\phi_K = \int_{-1}^1 K^2(x) dx, \quad \mathfrak{C}_K = \frac{\int_{-1}^1 |K'(x)|^2 dx}{\phi_K},$$

$$\mu_{j,K} = \int_{-1}^1 x^j K(x) dx.$$

We shall omit the subscript  $K$  in the sequel if no confusion will be caused.

A few remarks on the above regularity conditions are in order. The function  $\sigma^2(t)$  in condition (A2) is called the long-run variance function of  $\{J(t, Q_\alpha(t), \mathcal{F}_i)\}_{i=-\infty}^{\infty}$  to account for the dependence of the series. Note that  $\sigma^2(t)$  is well defined even for heavy-tailed processes  $(X_i)$  since  $0 \leq I\{G(t, \mathcal{F}_i) \leq x\} \leq 1$ . Condition (A3) means local stationarity and it asserts smoothness of  $G(t, \mathcal{F}_i)$  with respect to time  $t$ . Conditions (A4) and (A5) assert that processes  $\{F_k(t, x, \mathcal{F}_i)\}$ ,  $k = 0, 1, 2, 3$ , are short range dependent (SRD). Conditions (A3)–(A5) can be verified for a large class of nonstationary linear and nonlinear processes by the arguments in Section 4 of Zhou and Wu (2009). Condition (A6) is mild and it means that the conditional density function of  $J(t, x, \mathcal{F}_i)$  is bounded. A popular choice of the kernel function is the Epanechnikov kernel  $K(x) = 3 \max(0, 1 - x^2)/4$ .

3.2. *Simultaneous confidence bands.* The simultaneous confidence band (SCB) is a classic tool for nonparametric inference. To construct a  $100(1 - \beta)\%$  SCB for  $Q_\alpha(\cdot)$ , one finds two functions  $l$  and  $u$  depending on  $(X_i)_{i=1}^n$ , such that

$$\lim_{n \rightarrow \infty} \mathbb{P}(l(t) \leq Q_\alpha(t) \leq u(t) \text{ for all } t \in (0, 1)) = 1 - \beta.$$

A candidate function for  $Q_\alpha(\cdot)$  is rejected at level  $\beta$  if it is not fully contained in the SCB. The SCB provides appreciable direct visual information on the overall variability of the fitted curves. See, for example, Bickel and Rosenblatt (1973) for the inference of density functions; Eubank and Speckman (1993), Sun and Loader (1994), Neumann and Kreiss (1998), Wu and Zhao (2007) and Zhao and Wu (2008) for the inference of (conditional) mean functions and Fan and Zhang (2000) for the inference of coefficient functions of varying coefficient models. In this section, we shall establish the asymptotic theory for the maximal absolute deviation of  $\hat{Q}_\alpha(t)$  from  $Q_\alpha(t)$  on  $(0, 1)$ . The theoretical results facilitate construction of a SCB of  $Q_\alpha(t)$  which asymptotically achieves the nominal coverage probability.

**THEOREM 1.** *Assume  $Q_\alpha(\cdot) \in \mathcal{C}^3[0, 1]$  and conditions (A1)–(A6) and (K1) hold. Further assume  $\sqrt{nb_n}/\log^5 n \rightarrow \infty$  and  $nb_n^7 \log n \rightarrow 0$ , then we have*

$$(5) \quad \lim_{n \rightarrow \infty} \mathbb{P} \left[ \sup_{t \in \mathcal{I}_n} \left\{ \frac{\sqrt{nb_n} f(t, Q_\alpha(t))}{\sqrt{\phi} \sigma(t)} \times |\hat{Q}_\alpha(t) - Q_\alpha(t) - \mu_2 b_n^2 Q_\alpha''(t)/2| \right\} - B(m^*) \leq \frac{x}{\sqrt{2 \log m^*}} \right] = e^{-2e^{-x}},$$

where  $\mathcal{I}_n = [b_n, 1 - b_n]$ ,  $m^* = 1/b_n$  and

$$B(m^*) = (2 \log m^*)^{1/2} + (2 \log m^*)^{-1/2} [\log \mathfrak{C} - 2 \log \pi - 2 \log 2]/2.$$

For a fixed level  $\beta$ , Theorem 1 implies that one can construct a  $100(1 - \beta)\%$  simultaneous confidence band for  $Q_\alpha(t)$

$$(6) \quad \hat{Q}_\alpha(t) - \mu_2 b_n^2 Q_\alpha''(t)/2 \pm \frac{\sqrt{\phi} \sigma(t)}{\sqrt{nb_n} f(t, Q_\alpha(t))} \left( B_K(m^*) + \frac{u_\beta}{\sqrt{2 \log m^*}} \right),$$

where  $u_\beta = -\log \log[(1 - \beta)^{-1/2}]$ . By Theorem 1, SCB (6) asymptotically achieves the right coverage probability  $1 - \beta$ . Note that  $V(t) := \frac{\sqrt{\phi} \sigma(t)}{\sqrt{nb_n} f(t, Q_\alpha(t))}$  is the asymptotic standard deviation of  $\hat{Q}_\alpha(t)$  [Theorem 1 of Zhou and Wu (2009)].

The following theorem concerns the local power of the SCB (6).

**THEOREM 2.** *Suppose  $Q_\alpha(t) = Q_\alpha^o(t) + \gamma_n \eta(t) + o(\gamma_n)$ , where  $Q_\alpha^o(t), \eta(t) \in \mathcal{C}[0, 1]$ ,  $\gamma_n = 1/\sqrt{-2nb_n \log b_n}$  and  $o(\gamma_n)$  is uniform in  $t$  on  $[0, 1]$ . Then under conditions of Theorem 1, we have*

$$(7) \quad \lim_{n \rightarrow \infty} \mathbb{P} \left[ \sup_{t \in \mathcal{I}_n} \left\{ \frac{\sqrt{nb_n} f(t, Q_\alpha(t))}{\sqrt{\phi} \sigma(t)} \right. \right. \\ \left. \left. \times |\hat{Q}_\alpha(t) - Q_\alpha^o(t) - \mu_2 b_n^2 Q_\alpha''(t)/2| \right\} - B(m^*) \leq \frac{x}{\sqrt{2 \log m^*}} \right] = e^{-s(\eta) e^{-x}},$$

where

$$s(\eta) = \int_0^1 \exp \left\{ \frac{\eta(t) f(t, Q_\alpha(t))}{\sqrt{\phi} \sigma(t)} \right\} dt + \int_0^1 \exp \left\{ \frac{-\eta(t) f(t, Q_\alpha(t))}{\sqrt{\phi} \sigma(t)} \right\} dt.$$

Theorem 2 follows from similar arguments to those in the proof of Theorem 1 and Theorem A1 of Bickel and Rosenblatt (1973). Details are omitted.

Theorem 2 implies that SCB (6) can detect alternatives with the rate  $\gamma_n$ . For the  $100(1 - \beta)\%$  SCB (6), the asymptotic power of the test  $H_0: Q_\alpha(t) = Q_\alpha^o(t)$  versus  $H_a: Q_\alpha(t) \neq Q_\alpha^o(t)$  is

$$(8) \quad 1 - (1 - \beta)^{s(\eta)/2}$$

under the local alternatives specified in Theorem 2. Since  $s(\eta) \geq 2$  and  $s(\eta) = 2$  if and only if  $\eta(t) \equiv 0$ , our test based on the SCB is always asymptotically unbiased under such alternatives. In the case of density function inference for i.i.d. data, the same result was obtained by Bickel and Rosenblatt (1973).

**3.2.1. Optimality of the SCB.** If the local linear smoothing technique is adopted and the bandwidth series ( $b_n$ ) is fixed, then SCB (6) is optimal in the sense that asymptotically it covers the minimum area. To see this, a Lagrange multiplier argument can be implemented. A similar argument can be found in Zhou and Wu (2010) for nonparametric inference of time-varying coefficients in functional linear models. Note by equations (46) and (47) in Section 7, we have under conditions of Theorem 1,

$$(9) \quad \sup_{t \in \mathcal{I}_n} |[\hat{Q}_\alpha(t) - Q_\alpha(t) - \mu_2 b_n^2 Q_\alpha''(t)/2]/V(t) - \Theta_n(t)| = o_{\mathbb{P}}(\log^{-1/2} n),$$

where  $\Theta_n(t) = \sum_{i=1}^n V_i K_{b_n}(t_i - t)/\sqrt{\phi n b_n}$  with  $(V_i)_{i=1}^n$  i.i.d. standard normal. In other words, the simultaneous fluctuations of  $[\hat{Q}_\alpha(t) - Q_\alpha(t) - \mu_2 b_n^2 Q_\alpha''(t)/2]/V(t)$  are asymptotically equivalent to those of  $\Theta_n(t)$ . Let  $s_i = (2i - 1)b_n$ ,  $i = 1, \dots, g_n$ , where  $g_n = \lfloor 1/(2b_n) \rfloor$ . From (9),  $\hat{Q}_\alpha(s_i) - Q_\alpha(s_i) - \mu_2 b_n^2 Q_\alpha''(s_i)/2$  are asymptotically independent  $N(0, V^2(s_i))$ . Suppose a band

$$l(s_i) \leq \hat{Q}_\alpha(s_i) - Q_\alpha(s_i) - \mu_2 b_n^2 Q_\alpha''(s_i)/2 \leq u(s_i), \quad i = 1, 2, \dots, g_n,$$

achieves a preassigned coverage probability  $1 - \beta$ . From the above discussion, we see that the coverage probability restriction can be asymptotically written as

$$c(n, b_n) := \prod_{i=1}^{g_n} [\Phi(u(s_i)/V(s_i)) - \Phi(l(s_i)/V(s_i))] = 1 - \beta,$$

where  $\Phi(\cdot)$  is the normal cumulative distribution function (cdf). In order to achieve the minimum average length, one minimizes the following Lagrange multiplier:

$$(10) \quad \sum_{i=1}^{g_n} [u(s_i) - l(s_i)] - \lambda \{\log[c(n, b_n)] - \log(1 - \beta)\}.$$

Simple calculations show that the minimum is achieved at  $u(s_i) = -l(s_i) = g(n, b_n, \beta)V(s_i)$ , where  $g(n, b_n, \beta)$  is a deterministic function. The important message here is that the asymptotically optimal SCB at each time point  $t$  should have length proportional to the asymptotic standard deviation of  $\hat{Q}_\alpha(t)$ , which is the case in our construction.

3.3. *The integrated squared difference test (ISDT).* Another popular basis for nonparametric inference is  $\mathcal{L}^2$ -distance based tests. In general, one calculates a  $\mathcal{L}^2$  norm related distance between the fitted nonparametric curve and the parametric null, and a large distance indicates violation of the null hypothesis. Most of the existing results on the  $\mathcal{L}^2$  type tests are for independent data. See, for instance, Bickel and Rosenblatt (1973), Härdle and Mammen (1993), Zheng (1996), Fan, Zhang and Zhang (2001), Zhang and Dette (2004) and Fan and Jiang (2005) among others.

A simple way to construct a  $\mathcal{L}^2$  type test is to use the statistic  $T_n = \int_0^1 [\hat{Q}_\alpha(t) - Q_\alpha(t)]^2 \pi(t) dt$ . However, the bias of the local linear estimate  $\hat{Q}_\alpha(t)$  is of order  $O(b_n^2 + \frac{1}{nb_n})$  and is not negligible for the asymptotic analysis. The extra bias term complicates the asymptotic distribution and reduces the precision of the test  $T_n$ . See also Härdle and Mammen (1993) for a related discussion in the case of conditional mean inference. To overcome the disadvantage, we shall use the following jackknife bias reduction technique [Wu and Zhao (2007)]:

$$(11) \quad \tilde{Q}_{\alpha, b_n}(t) = 2\hat{Q}_{\alpha, b_n}(t) - \hat{Q}_{\alpha, \sqrt{2}b_n}(t).$$

It can be shown that bias of  $\tilde{Q}_\alpha(t)$  is of order  $o(b_n^2 + \frac{1}{nb_n})$  uniformly on  $[0, 1]$  if  $Q_\alpha(t) \in \mathcal{C}^2[0, 1]$ . Note that using (11) is equivalent to using the second-order kernel

$$K^*(x) := 2K(x) - K(x/\sqrt{2})/\sqrt{2}.$$

To test  $H_0: Q_\alpha(t) = Q_\alpha^o(t)$  versus  $H_a: Q_\alpha(t) \neq Q_\alpha^o(t)$ , we propose the following test statistic

$$(12) \quad T_n^* = \int_{\mathcal{T}_n^*} [\tilde{Q}_\alpha(t) - Q_\alpha^o(t)]^2 \pi(t) dt,$$

where  $T_n^* = [\sqrt{2}b_n, 1 - \sqrt{2}b_n]$  and the weight  $\pi(t)$  are assumed to be nonnegative and Lipschitz continuous in  $t \in [0, 1]$ . The following theorem establishes the asymptotic normality of  $T_n^*$ .

**THEOREM 3.** *Assume that  $Q_\alpha(t) \in C^2[0, 1]$ ; that conditions (A1)–(A6) and (K1) hold and that  $Q_\alpha(t) = Q_\alpha^o(t) + \varrho_n \eta(t) + o(\varrho_n)$ , where  $Q_\alpha^o(t), \eta(t) \in C[0, 1]$ ,  $\varrho_n = n^{-1/2}b_n^{-1/4}$  and  $o(\varrho_n)$  is uniform in  $t$  on  $[0, 1]$ . Further assume  $nb_n^4/\log^{10}n \rightarrow \infty$  and  $nb_n^{9/2} = O(1)$ . We have*

$$(13) \quad n\sqrt{b_n}T_n^* - \frac{1}{\sqrt{b_n}}K^* \star K^*(0) \int_0^1 \pi^*(t) dt - \int_0^1 \eta^2(t)\pi(t) dt \\ \Rightarrow N\left(0, 2 \int_{\mathbb{R}} [K^* \star K^*(t)]^2 dt \int_0^1 \pi^*(t)^2 dt\right),$$

where  $\star$  denotes the convolution operator and  $\pi^*(t) = \pi(t)\sigma^2(t)/f^2(t, Q_\alpha(t))$ .

When  $\eta(t) \equiv 0$ , Theorem 3 unveils the asymptotic null distribution of  $T_n^*$ . The ISDT can detect alternatives with the rate  $\varrho_n$ . Under the local alternatives specified in Theorem 3, simple calculations based on (13) show that the asymptotic power of the ISDT with level  $\beta$  equals

$$(14) \quad \Phi\left(\frac{\int_0^1 \eta^2(t)\pi(t) dt}{[2 \int_{\mathbb{R}} [K^* \star K^*(t)]^2 dt \int_0^1 \pi^*(t)^2 dt]^{1/2}} - z_{1-\beta}\right),$$

where  $\Phi(\cdot)$  and  $z_{1-\beta}$  denote the cumulative distribution function and the  $1 - \beta$  quantile of the standard normal distribution. Therefore, a simple use of the Cauchy–Schwarz inequality shows that choosing weights  $\pi(t)$  proportional to  $\eta^2(t)f^4(t, Q_\alpha(t))/\sigma^4(t)$  maximizes the above asymptotic power. Of course, in real applications it is difficult to specify  $\eta(t)$ . Therefore, one can simply choose  $\pi(t) \equiv 1$ . On the other hand, we suggest choosing  $\pi(t) = f^2(t, \hat{Q}_\alpha(t))/\hat{\sigma}^2(t)$ , which leads to an easier implementation of the bootstrap. See the discussions in Section 4.1 for more details.

**COROLLARY 1.** *Under conditions of Theorem 3,  $T_n^*$  can detect alternatives with departure rate  $n^{-4/9}$  if the bandwidth  $b_n = O(n^{-2/9})$ .*

**REMARK 1.** If  $X_i$  is distributed as  $N(\mu(t_i), \sigma^2)$ , and the  $X_i$ 's are independent, then inference of quantile curves is equivalent to inference of the mean function  $\mu(\cdot)$ . In this case, Ingster (1993) and Lepski and Spokoiny (1999) proved that the optimal rate for testing  $H_0: Q_\alpha(t) = Q_\alpha^o(t)$  is  $n^{-4/9}$ . Hence, the integrated squared difference test  $T_n^*$  is optimal in the sense that it achieves the optimal rate of convergence. Furthermore, Corollary 1 implies that for nonparametric quantile function testing, weak dependence and local stationarity do not deteriorate the optimal rate. However, it should be noted that when there is long memory, the rate will be deteriorated.

REMARK 2. Since  $\gamma_n \gg \varrho_n$ , the ISDT test  $T_n^*$  dominates the test based on SCB (6) if bandwidths of the same order are used for the tests; namely  $T_n^*$  is asymptotically more powerful. Therefore a general rule of thumb is to use the SCB when one wants to explore the overall pattern of the quantile curves and to implement the ISDT test when one is interested in verifying a specific parametric null. On the other hand, as stated in Härdle and Mammen (1993), “*certainly from a more data analytic point of view distances would be more satisfactory which reflect similarities in the shape of the regression functions.*” For a moderate sample size, intuitively, the ISDT test would be more powerful compared to the SCB test if the true quantile curve differs from the null in a systematic and even way; while the SCB test is better when the latter difference is abrupt or “bumpy,” in which case the  $\mathcal{L}^2$  norm does not reflect the characteristics of the difference. Hence, if there is some prior knowledge on the shape of the discrepancy, one could select a test accordingly. In Section 5, we shall conduct a simulation study to compare the powers of the two tests under various alternatives.

#### 4. Implementation.

4.1. *A wild bootstrap procedure.* The asymptotic results in Section 3 are based on the uniform Gaussian approximations of  $\hat{Q}_\alpha(t) - Q_\alpha(t)$  on  $(0, 1)$ . It is known that the convergence rates of the  $\mathcal{L}^\infty$  and  $\mathcal{L}^2$  norms of the corresponding Gaussian processes are very slow (see also proofs in Section 7). For example, when  $b_n(\alpha) = O(n^{-1/5})$ , the convergence rates of the  $\mathcal{L}^\infty$  and  $\mathcal{L}^2$  norms are  $1/\log^{1/2} n$  and  $n^{-1/10}$ , respectively. Therefore, for moderate sample sizes, tests based on the asymptotic theory are not reliable.

For nonparametric inference, the bootstrap is a classic tool for achieving faster convergence. It is impossible to have a complete list of literature here and we shall only mention several representatives. Among others, Mammen (1993), Härdle and Mammen (1993), Chapter 8 of Shao and Tu (1995), Stute, Gonzalez Manteiga and Presedo Quindimil (1998), Neumann and Kreiss (1998) and Fan and Jiang (2007) considered wild bootstrap inference for conditional mean regression for independent data; Jhun (1988), Faraway and Jhun (1990) and Hall (1993) considered bootstrap inference of the density function for i.i.d. data. For nonparametric inference of dependent data, among others, Politis and Romano (1994) proposed a stationary bootstrap for simultaneous inference of the spectral density functions of weakly dependent stationary time series and Wu and Zhao (2007) used a wild bootstrap technique to test the mean function under stationary errors. For other contributions, see Barrio and Matrán (2000).

On the other hand, there is also large literature on bootstrap methods for parametric quantile regression. See, for example, Chapter 3.9 of Koenker (2005) and references therein for independent data and Fitzenberger (1998) for the moving block bootstrap for strong mixing samples.

Despite the huge literature on bootstrap strategies for nonparametric inference and parametric quantile regression, there have been few results on bootstrap methods for nonparametric quantile inference for nonstationary time series. The major difficulty lies in the fact that joint distributions of subseries within different time spans can be drastically different for a nonstationary process; therefore it is challenging to capture the variability of the process in every local structure in order to make valid nonparametric inferences.

To circumvent the above difficulties, here we shall adopt a different wild bootstrap technique. A similar technique was proposed in Wu and Zhao (2007) for nonparametric inference of the mean function under stationary errors. The key idea is still uniform Gaussian approximation of  $\hat{Q}_\alpha(t) - Q_\alpha(t)$  on  $(0, 1)$ . However, instead of resorting to asymptotic theory, we shall directly simulate the finite sample  $\mathcal{L}^2$  and  $\mathcal{L}^\infty$  norms of the Gaussian processes. More precisely, let us assume the bandwidth  $b_n(\alpha) = O(n^{-1/5})$  and  $Q_\alpha(t) \in \mathcal{C}^3[0, 1]$  for illustrative purposes. Then by the proofs of Theorems 1 and 3 in Section 7, we can obtain after elementary calculations that

$$(15) \quad \sup_{t \in \mathcal{T}_n^*} \left| \frac{f(t, Q(t))}{\sigma(t)} [\tilde{Q}(t) - Q(t)] - \mathfrak{X}_n(t) \right| = O_{\mathbb{P}}(n^{-11/20} \log^2 n)$$

and

$$(16) \quad \left| \int_{\mathcal{T}_n^*} [\tilde{Q}_\alpha(t) - Q_\alpha(t)]^2 \tilde{\pi}(t) dt - \int_{\mathcal{T}_n^*} \mathfrak{X}_n(t)^2 dt \right| = O_{\mathbb{P}}(n^\beta \log^{5/2} n),$$

where

$$\mathfrak{X}_n(t) = \sum_{i=1}^n V_i K_{b_n}^*(t_i - t) / (nb_n)$$

with  $(V_i)$  i.i.d. standard normal,  $\tilde{\pi}(t) = f^2(t, Q_\alpha(t)) / \sigma^2(t)$  and  $\beta = -0.95$ . Note that here the bias-corrected estimator  $\tilde{Q}(\cdot)$  is used in (15) in order to avoid estimating the unpleasant bias term  $\mu_2 b_n^2 Q_\alpha''(t) / 2$  of the SCB.

An important observation of (15) and (16) is that  $\mathfrak{X}_n(t)$  does not depend on the observations  $(X_i)$  and has a simple and explicit form. Therefore, one can generate a large sample of i.i.d. copies of  $\mathfrak{X}_n(t)$  and use the distributions of the  $\mathcal{L}^\infty$  and  $\mathcal{L}^2$  norms of the sample to approximate the distributions of the corresponding norms of  $\tilde{Q}(t) - Q(t)$  by virtue of (15) and (16). The following are the detailed procedures.

- (1) Choose bandwidth  $b_n$  according to the procedures in Section 4.3.
- (2) Obtain  $\tilde{Q}_\alpha(t)$  by (2) and (11).
- (3) Obtain estimate  $\hat{\sigma}(t)$ ,  $\hat{f}(t, Q_\alpha(t))$  from (19) and (20) below. Let  $\hat{\pi}(t) = \hat{f}^2(t, Q_\alpha(t)) / \hat{\sigma}^2(t)$ .
- (4S) Generate i.i.d. standard normal random variables  $V_i$ ,  $i = 1, 2, \dots, n$ . Calculate  $\varpi_{n,S} = \sup_{0 \leq t \leq 1} |\mathfrak{X}_n(t)|$ .

- (4I) Generate i.i.d. standard normal random variables  $V_i, i = 1, 2, \dots, n$ . Calculate  $\varpi_{n,I} = \int_0^1 \mathfrak{X}_n^2(t) dt$ .
- (5) Repeat (4)  $B$  times and obtain the estimated quantile  $\hat{q}_{1-\beta}$  of  $\varpi_n$ .
- (6S) The  $100(1 - \beta)\%$  SCB of  $Q_\alpha(t)$  can be constructed as  $\tilde{Q}_\alpha(t) \pm \hat{q}_{1-\beta} / \sqrt{\hat{\pi}(t)}$ . Accept the null hypothesis  $Q_\alpha(\cdot) = Q_\alpha^o(\cdot)$  at level  $1 - \beta$  if and only  $Q_\alpha^o(\cdot)$  is fully contained in the SCB.
- (6I) Accept the null hypothesis  $Q_\alpha(\cdot) = Q_\alpha^o(\cdot)$  at level  $1 - \beta$  if and only if  $\int_0^1 [\tilde{Q}_\alpha(t) - Q_\alpha^o(t)]^2 \hat{\pi}(t) dt \leq \hat{q}_{1-\beta}$ .

One should use steps (1)–(3), (4S), (5) and (6S) when performing hypothesis testing via the SCB. On the other hand, steps (1)–(3), (4I), (5) and (6I) should be adopted when testing via the ISDT. The number of replications  $B$  can be chosen as 2000. It is immediate to obtain  $p$ -values of the tests. For instance, the  $p$ -value of the squared difference test is  $\mathbb{P}(\varpi_{n,I} > \int_0^1 [\tilde{Q}_\alpha(t) - Q_\alpha^o(t)]^2 \hat{\pi}(t) dt)$ , which can be estimated by the bootstrap distribution of  $\varpi_{n,I}$ .

Let  $\{f(t, \theta)\}$  be a parametric family of functions that depends on  $t \in [0, 1]$  and  $\theta \in \Theta \subset \mathbb{R}^k$ . Often one wants to test  $Q_\alpha(t) = f(t, \theta)$  at level  $1 - \beta$  for some unknown  $\theta \in \Theta$ . Under the null hypothesis, we have a parametric model and one generally expects to obtain a root- $n$  consistent estimator  $\hat{\theta}$  of the true parameter value  $\theta_0$  by the parametric quantile regression method of Keonker (2005). Note that the convergence rates of our SCB and ISDT tests are always slower than  $\sqrt{n}$ . Therefore, if the null is true,  $f(t, \hat{\theta})$  can be treated as the true value of  $Q_\alpha(t)$  and one just needs to replace  $Q_\alpha^o(\cdot)$  in steps (1)–(6) by  $f(t, \hat{\theta})$  and the resulting testing procedures are still valid.

The following proposition validates the above procedure for the case  $\{f(t, \theta)\} = \{\theta^\top \mathbf{g}(t)\}$ , where  $\theta \in \mathbb{R}^k$  and  $\mathbf{g}(t) : [0, 1] \rightarrow \mathbb{R}^k$  is a known function. We shall first make the following constraint on  $\mathbf{g}$ :

- (B1) assume  $\mathbf{g}(\cdot) \in C[0, 1]$  and  $\mathcal{G} := \int_0^1 \mathbf{g}(t)\mathbf{g}^\top(t) dt$  is nonsingular.

PROPOSITION 1. Assume that  $Q_\alpha(t) = \theta_0^\top \mathbf{g}(t)$  for some  $\theta_0 \in \mathbb{R}^k$  and that conditions (A1), (A5) and (B1) hold. Then

$$(17) \quad |\hat{\theta}_\alpha - \theta_0| = O_{\mathbb{P}}(n^{-1/2}),$$

where  $\hat{\theta}_\alpha = \arg \min_{\theta} \sum_{i=1}^n \rho_\alpha(X_i - \theta^\top \mathbf{g}(i/n))$ .

4.2. Estimation of the density and long-run variance functions. We see from (15) and (16) that obtaining a good estimate of  $f(t, Q(t))/\sigma(t)$  is necessary and important for making our inferences. Here, we suggest using the estimation techniques in Zhou and Wu (2009), which are essentially local versions of the popular subsampling long-run variance estimator and kernel density estimator for stationary data. Since the time series is approximately stationary within comparatively small time spans, the methods are shown to be consistent. See Section 3.4 in Zhou

and Wu (2009) for more details. For the sake of completeness, we present the estimators here.

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

$$(18) \quad \mathcal{N}_n(t) = \{i \in \mathbb{N} : s_n(t) \leq i \leq l_n(t)\}.$$

Let  $Z_{i,\alpha} = \psi_\alpha(X_i - \hat{Q}_\alpha(i/n))$ . For a sequence  $m_n$  with  $m_n \rightarrow \infty$  and  $nb_n/m_n \rightarrow \infty$ , we shall estimate  $\sigma^2(t)$  by

$$(19) \quad \hat{\sigma}^2(t) = \frac{m_n}{|\mathcal{N}_n(t)| - m_n + 1} \sum_{j=s_n(t)}^{l_n(t)-m_n+1} \left( \frac{\sum_{i=j}^{j+m_n-1} Z_{i,\alpha}}{m_n} - \bar{Z}_n(t) \right)^2,$$

where  $\bar{Z}_n(t) = \sum_{i \in \mathcal{N}_n(t)} Z_{i,\alpha} / |\mathcal{N}_n(t)|$  and  $|\mathcal{N}_n(t)| = l_n(t) - s_n(t) + 1$  is the cardinality of  $\mathcal{N}_n(t)$ .

For  $f(t, Q_\alpha(t))$ , we shall use

$$(20) \quad \hat{f}(t, Q_\alpha(t)) = \frac{1}{|\mathcal{N}_n(t)|h_n} \sum_{i \in \mathcal{N}_n(t)} K_{h_n}^\#(\hat{Q}_\alpha(t) - X_i),$$

where  $K^\# \in \mathcal{K}$  is a kernel and  $h_n$  is the bandwidth satisfying  $h_n \rightarrow 0$  and  $nb_n h_n \rightarrow \infty$ .

We refer to Zhou and Wu (2009) for a discussion on the selection of the tuning parameters  $m_n$  and  $\tau_n$ .

4.3. *Bandwidth selection.* Choosing a good bandwidth  $b_n(\alpha)$  is important in practical applications. For quantile curve estimation, Zhou and Wu (2009) proposed the following way to choose the bandwidth based on modifications of existing bandwidth selectors for independent data. By Theorem 1 of the latter paper regarding asymptotic normality of  $\hat{Q}_\alpha(t)$ , we have

$$(21) \quad \frac{b_n^*(\alpha)}{b_n^{\text{ind}}(\alpha)} = \left[ \frac{\int_0^1 \sigma^2(t) dt}{\alpha(1-\alpha)} \right]^{1/5} := \rho^*(\alpha),$$

where  $b_n^*(\alpha)$  denotes the optimal weighted asymptotic mean integrated squared error (AMISE) bandwidth,  $b_n^{\text{ind}}(\alpha)$  is the optimal bandwidth obtained under independence and  $\rho^*(\alpha)$  is called the variance correction factor which accounts for the dependence. Note that  $\alpha(1-\alpha) = \text{Var}(J(t, Q_\alpha(t), \mathcal{F}_i))$ . For independent data, there have been many discussions on bandwidth selection for nonparametric quantile estimation; see, for instance, Yu and Jones (1998), Fan and Gijbels (1996) and Ghosh and Draghicescu (2002) among others. Hence, one could first select a bandwidth  $b_n^{\text{ind}}(\alpha)$  by treating the data as if they were independent. After that, the variance correction factor  $\rho^*(\alpha)$  can be estimated by the following:

$$\hat{\rho}^*(\alpha) = \frac{(\tilde{\sigma}^2)^{1/5}}{(\alpha(1-\alpha))^{1/5}},$$

where

$$\tilde{\sigma}^2 = \frac{\tilde{m}}{n - \tilde{m} + 1} \sum_{j=1}^{n-\tilde{m}+1} \left( \frac{1}{\tilde{m}} \sum_{i=j}^{j+\tilde{m}-1} \varsigma_{i,\alpha} - \bar{\varsigma}_{n,\alpha} \right)^2,$$

$\varsigma_{i,\alpha} = \psi_\alpha(X_i - \hat{Q}_{\alpha, b_n^{\text{ind}}(\alpha)}(t))$ ,  $\bar{\varsigma}_{n,\alpha} = \sum_{i=1}^n \varsigma_{i,\alpha} / n$  and  $\tilde{m} = \lfloor n^{1/3} \rfloor$ . It can be shown that  $\hat{\rho}^*(\alpha)$  is a consistent estimate of  $\rho^*(\alpha)$  and we shall point out that the selected bandwidth  $b_n^*(\alpha)$  typically varies with respect to  $\alpha$  [see also Yu and Jones (1998)]. Moreover, since the jackknife bias reduction technique reduces the bias of the local linear quantile estimates to second order, following Wu and Zhao (2007), we suggest using  $b_n^{\text{jack}}(\alpha) = 2b_n^*(\alpha)$  for the nonparametric estimation when the jackknife is implemented. We refer the interested reader to Sections 3.1.1 and 3.3 of Zhou and Wu (2009) for more details on the bandwidth selection.

The bandwidth selected for quantile curve estimation provides a reasonable starting point for nonparametric tests [Fan and Jiang (2007)]. For the SCB test, we suggest using the bandwidth  $b_n^{\text{jack}}(\alpha)$  following Eubank and Speckman (1993) and Wu and Zhao (2007). For the ISDT test  $T_n^*$ , Corollary 1 implies that the bandwidth which renders the optimal power is of order  $n^{-2/9} = n^{-1/5} n^{-1/45}$ . Following Fan and Jiang (2007), we suggest using the bandwidth  $b_n^{\text{jack}}(\alpha) \times n^{-1/45}$  for the ISDT test.

**5. Simulation studies.** In this section, we perform simulation studies to investigate the accuracy and power of the proposed tests for moderate sample sizes. Let us consider the following time-varying AR(1) model

$$(22) \quad G(t, \mathcal{F}_i) = a_0(t) + a_1(t)G(t, \mathcal{F}_{i-1}) + \delta(t)\varepsilon_i,$$

where  $a_0(t)$ ,  $a_1(t)$  and  $\delta(t)$  are continuous functions on  $[0, 1]$ ,  $\max_t |a_1(t)| < 1$ ,  $\min_t \delta(t) > 0$  and  $\varepsilon_i$  are i.i.d. with  $\|\varepsilon_i\|_p < \infty$  for some  $p > 0$ . We observe the time series  $X_i = G(i/n, \mathcal{F}_i)$ ,  $i = 1, 2, \dots, n$ .

It is clear that  $G(t, \mathcal{F}_i)$  has the representation

$$(23) \quad G(t, \mathcal{F}_i) = \frac{a_0(t)}{1 - a_1(t)} + \delta(t) \sum_{j=0}^{\infty} [a_1(t)]^j \varepsilon_{i-j}.$$

The UGMC and local stationarity conditions of (22) can be easily verified by the results in Section 4.1 of Zhou and Wu (2009).

*5.1. Accuracy of the SCB and ISDT tests.* In this subsection, we describe a simulation study to compare the accuracy of the asymptotic and bootstrap tests for both light tailed and heavy tailed processes. To this end, we shall use the time-varying AR(1) process (22) with  $a_0(t) = 0$ ,  $a_1(t) = \sin(2\pi t)/2$  and  $\delta(t) = \exp((t - 1/4)^2)$ . Consider the following two scenarios:

$$(a) \ \varepsilon_i \sim N(0, 1); \quad (b) \ \varepsilon_i \sim S\alpha S(1.8).$$

Here,  $S\alpha S(1.8)$  denotes the standard symmetric  $\alpha$  stable distribution with index 1.8 which has the characteristic function  $\exp(-|t|^{1.8})$ ,  $t \in \mathbb{R}$ . It is easy to show that the  $S\alpha S(1.8)$  distribution has mean 0 and infinite variance. Therefore, scenarios (a) and (b) represent light tailed and heavy tailed processes, respectively. Elementary calculations show that under scenarios (a) and (b),

$$(24) \quad Q_\alpha(t) = \frac{\delta(t)}{[1 - |\alpha_1(t)|^\nu]^{1/\nu}} Q_\alpha^\varepsilon,$$

where  $Q_\alpha^\varepsilon$  is the  $\alpha$ th quantile of  $\varepsilon_i$  and  $\nu = 2$  and 1.8 in scenarios (a) and (b), respectively.

We shall compare type I error rates of the following six tests: the asymptotic SCB test (AS) based on Theorem 1; the asymptotic ISDT test (AI) based on Theorem 3; the bootstrap SCB test (BS); the bootstrap ISDT test (BI); the asymptotic point-wise confidence band (PC) and the Bonferroni confidence band (BF) based on Theorem 1 of Zhou and Wu (2009). The Bonferroni confidence band is simply the point-wise confidence band at level  $\beta/n$ , where  $\beta$  is the desired level. We generate time-varying AR(1) processes under scenarios (a) and (b) with length  $n = 300$  and perform the above six tests at the nominal level 5% for the following four quantile curves  $\alpha = 0.5, 0.75, 0.9$  and 0.95. Bandwidths are chosen according to Section 4.3 and the critical values  $\hat{q}_{0.95}$  of the bootstrap tests are based on 2000 bootstrap samples. The simulated type I error rates with 1000 replicates are shown in Table 1 below.

It is clear from the output of Table 1 that point-wise confidence bands are not appropriate for nonparametric inference. The Bonferroni confidence band test is too conservative, namely the band is too wide. On the other hand, the asymptotic SCB test (AS) is conservative and the asymptotic ISDT test (AI) tends to slightly inflate the type I error. As discussed in Section 4.1, the loss of accuracy of the asymptotic tests is due to their slow convergence rates.

TABLE 1  
Simulated type I error rates for the six tests with nominal level 5% under scenarios (a) and (b).  
Series length  $n = 300$  with 1000 replicates

Test	$\alpha = 0.5$		$\alpha = 0.75$		$\alpha = 0.9$		$\alpha = 0.95$	
	(a)	(b)	(a)	(b)	(a)	(b)	(a)	(b)
AS	2.2%	2.6%	2.7%	2.8%	2.1%	9.9%	2%	21.7%
AI	8%	6.5%	7.2%	6.9%	7.4%	10.4%	6.7%	21.4%
BS	4.8%	6%	4.4%	5.4%	4%	10.4%	4%	23.8%
BI	5.6%	5.3%	5.2%	5.2%	5.2%	9.3%	5.5%	20.2%
BF	1.5%	0.8%	1.7%	1.6%	1.3%	5.5%	0.2%	14.1%
PC	61.3%	61%	53.9%	53.1%	45.4%	40.8%	45.7%	44.2%

For the bootstrap tests, the nominal type I error is achieved except for extreme quantiles of heavy tailed processes. It is not difficult to see that data is relatively sparse at high quantiles and very large jumps occur more frequently for heavy tailed processes. These facts suggest that for extreme quantile inference of heavy tailed processes, a relatively large sample size is needed in order to achieve the desired accuracy. We performed the bootstrap tests at 5% level for scenario (b) for  $n = 500$  and with 1000 replicates. Simulated type I errors of the BS and BI tests were 0.061 and 0.056, respectively, for the 90% quantile curve. However, for the 95% quantile, the corresponding simulated type I errors were 0.139 and 0.131, respectively, which were still way larger than the nominal. When  $n$  was increased to 1000, accuracy of the bootstrap tests were achieved for the 95% quantile curve under scenario (b).

*5.2. Power comparison of the SCB and ISDT tests.* As discussed in Remark 2, for a moderate sample size, power of the SCB and ISDT tests are greatly influenced by the shape of  $Q_\alpha(t) - Q_\alpha^o(t)$ . Recall that  $Q_\alpha^o(t)$  is the hypothesized value of  $Q_\alpha(t)$ . In this subsection, we present a simulation to compare the power performance of the above two tests under various shapes of  $Q_\alpha(t) - Q_\alpha^o(t)$ . To this end, we consider model (22) with  $a_0(t) = \varphi(t)(1 - a_1(t))$ ,  $a_1(t) = \sin(2\pi t)/2$ ,  $\delta(t) = \exp((t - 1/4)^2)$  and  $\varepsilon_i$  i.i.d.  $N(0, 1)$ . Then

$$(25) \quad Q_\alpha(t) = \varphi(t) + Q_\alpha^o(t) \quad \text{where } Q_\alpha^o(t) = \frac{\delta(t)}{[1 - a_1^2(t)]^{1/2}} Q_\alpha^\varepsilon.$$

We test the hypothesis  $H_0: Q_\alpha(\cdot) = Q_\alpha^o(\cdot)$  versus  $H_a: Q_\alpha(\cdot) \neq Q_\alpha^o(\cdot)$ . Consider the following two situations:

$$(i) \varphi(t) = c_1; \quad (ii) \varphi(t) = c_2 \exp(-c_3(t - 1/2)^2),$$

where  $c_i$ ,  $i = 1, 2, 3$ , are positive constants. Cases (i) and (ii) correspond to flat and bumpy differences of the true and hypothesized quantile curves, respectively. Note that as  $c_3$  gets larger, we observe sharper peaks in  $\varphi(t)$ . In our simulations, we follow steps (1) to (6) in Section 4.1 and perform the SCB and ISDT tests at 5% level with  $n = 300$  and  $\alpha = 0.5, 0.75, 0.9$  and  $0.95$ . Various values of  $c_1$  and  $(c_2, c_3)$  are investigated and for each of the values we perform 1000 replicates and record the simulated probability of rejecting the null hypothesis. The simulated power curves of case (i) and (ii) are shown in Figures 1 and 2, respectively.

Generally, the displays in Figures 1 and 2 are consistent with our arguments in Remark 2 that for a moderate sample size the ISDT test is more powerful when  $\varphi(\cdot)$  is flat and the SCB test performs better when  $\varphi(\cdot)$  changes abruptly. On the other hand, when the peak of  $\varphi(\cdot)$  is not sharp enough; namely when  $c_3$  is relatively small, we see from Figure 2 that the ISDT test is still more powerful than the SCB test, which can be explained by the fact that the ISDT test asymptotically dominates the SCB test.

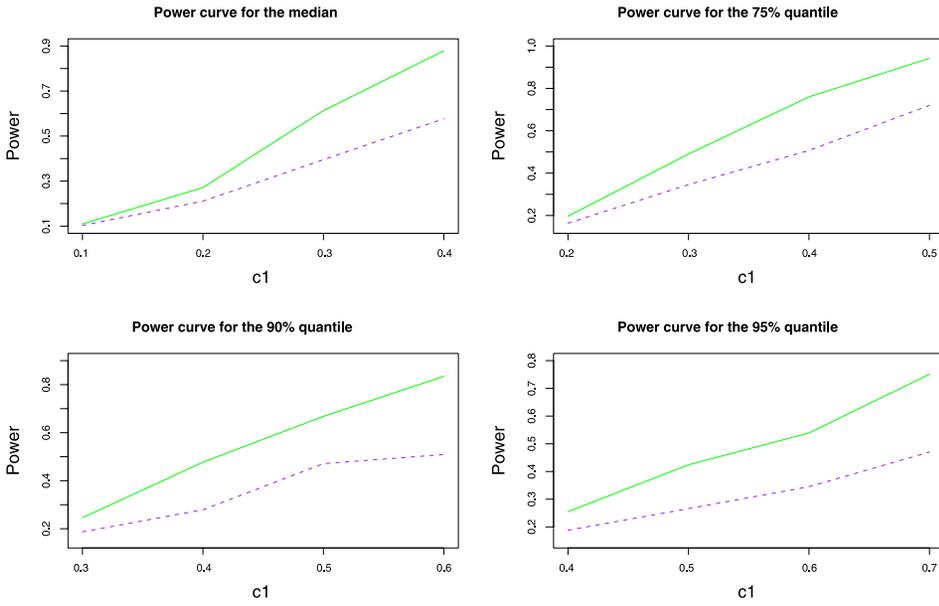


FIG. 1. Power curves for the 50%, 75%, 90% and 95% quantiles under case (i) of Section 5.2. The solid lines are the power curves for the ISDT test and the dashed lines are the power curves for the SCB test.

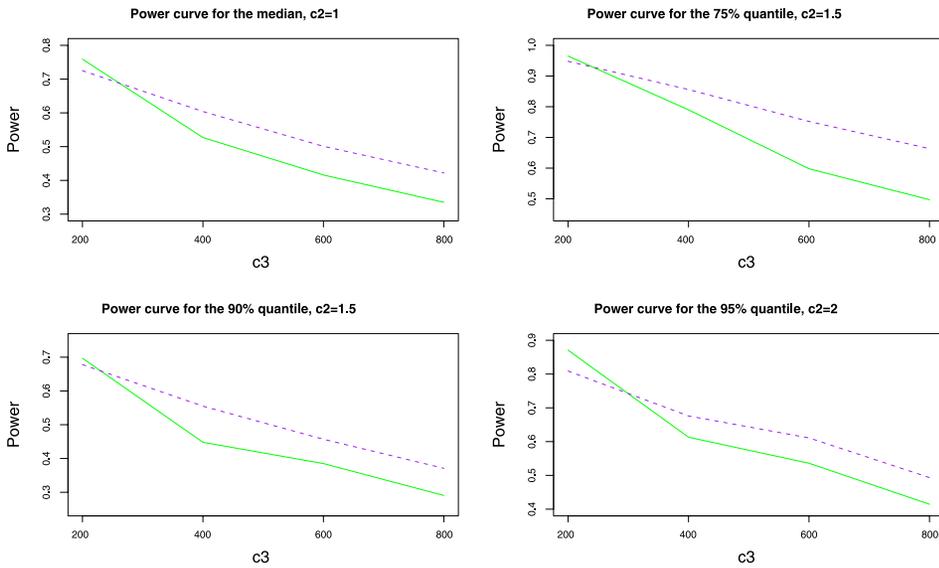


FIG. 2. Power curves for the 50%, 75%, 90% and 95% quantiles under case (ii) of Section 5.2. The solid lines are the power curves for the ISDT test and the dashed lines are the power curves for the SCB test.

**6. The global tropical cyclone data.** One of the most important consequences of global warming is the increase of ocean temperature. Theoretical arguments and modeling studies indicate that tropical cyclone winds should increase with increasing ocean temperature [Elsner, Kossin and Jagger (2008)]. Meanwhile, climatologists are very interested in finding empirical evidences on the change of intensity of tropical cyclone winds. In a recent paper, Elsner, Kossin and Jagger (2008) tackled the latter problem partially by fitting *linear* trends for quantiles of the global tropical cyclone data. The data set contains satellite-derived lifetime-maximum wind speeds of 2098 tropical cyclones over the globe during the period 1981–2006. It is available at James Elsner’s website at <http://myweb.fsu.edu/jelsner/extspace/globalTCmax4.txt>. We shall refer to Elsner, Kossin and Jagger (2008) for a detailed description on how the data are obtained and the related issues. Figure 3 shows the time series plot of the data. Significant linearly increasing trends were found in high quantiles of the global tropical cyclone data in Elsner, Kossin and Jagger (2008). In other words, the worst tropical cyclones are getting stronger over the globe.

In this section, we are mainly interested in the following issues. First, we shall compare our quantile-based tests with the mean-based approaches to see whether the increase of intensity of the worst tropical cyclones is due to a shift of mean. Second, we shall test whether the linear trend used in Elsner, Kossin and Jagger (2008) is sufficient to describe the change of the high quantiles of the global cyclone winds. This is an important issue because the parametric linearity assumption implies homogeneous change of the wind speeds. However, much complex yet important information on the dynamics of high wind speeds may be buried under this assumption.

To address the first issue, we follow the procedures in Zhou and Wu (2010) and provide a SCB for the mean curve of the global tropical cyclone series. The bandwidth is chosen as 0.16. The left panel of Figure 4 shows the 95% SCB. The height of the horizontal line is the average of all cyclone wind speeds. It is clear that

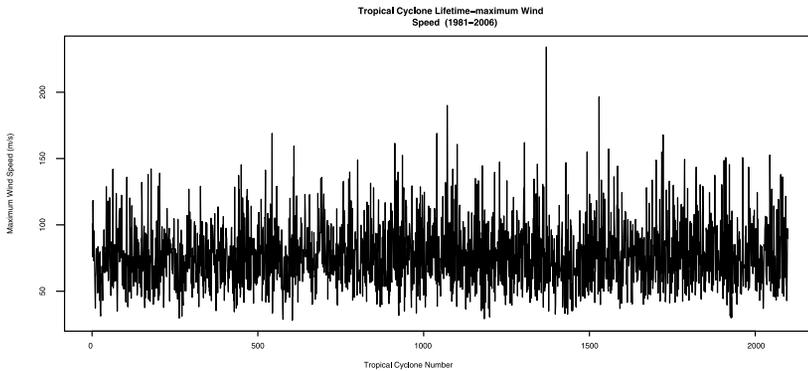


FIG. 3. *Time series plot of tropical cyclone lifetime-maximum wind speed (1981–2006).*

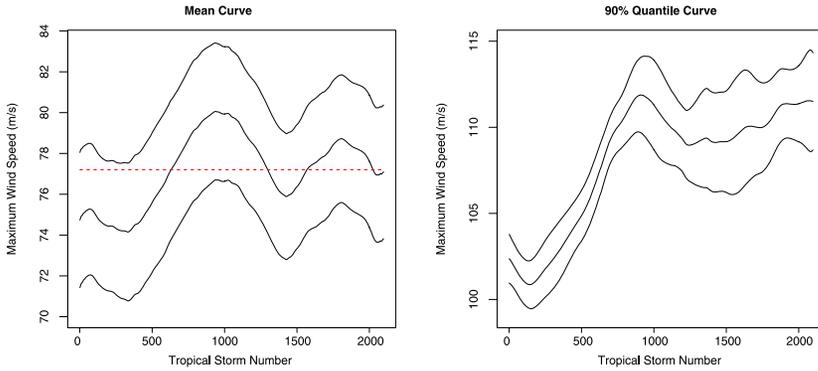


FIG. 4. Trends in mean and 90% quantile of global tropical cyclone lifetime-maximum wind speed, with 95% simultaneous confidence bands.

the horizontal line is fully contained in the SCB. The  $p$ -value for testing constancy of the mean trend is 0.11. Therefore, statistically there is no trend in the average tropical cyclone speeds. In fact, one of the main reasons why scientists did not find interesting signals in the intensity of global tropical cyclone winds before Elsner, Kossin and Jagger (2008) is due to the focus on mean trends.

We performed the SCB and ISDT tests on the linearity of the 90% quantile curve to address the second issue. The bandwidths of the SCB and ISDT tests are chosen as 0.21 and 0.18, respectively. Both tests give  $p$ -values less than 0.001. The right panel of Figure 4 shows the 95% SCB. It is clear from the SCB that there is an inhomogeneous increasing trend of the 90% quantile curve. More specifically, the quantile curve underwent a sharp increase during the period November 1982 to September 1992, and after that the trend became flat. This information cannot be retrieved if the linear trend analysis in Elsner, Kossin and Jagger (2008) is adopted. Furthermore, the inhomogeneous trend in high quantiles provides climatologists with useful information on the underlying complex mechanisms producing the hurricane climate.

**7. Proofs.** Unless otherwise specified, we will only prove results for  $\alpha = 1/2$ , since results for other quantiles follow similarly. We shall also omit the subscript  $\alpha$  in the notation if no confusion will be caused.

Consider a system  $H(t, x, \mathcal{F}_i)$  defined in Section 2.2. Let  $g : [0, 1] \mapsto \mathbb{R}$  be a measurable function. Define  $\mathfrak{H}(t, \mathcal{F}_i) = H(t, g(t), \mathcal{F}_i)$ . Then  $Z_i := \mathfrak{H}(t_i, \mathcal{F}_i)$   $i = 1, 2, \dots, n$ , defines a nonstationary time series. Recall  $t_i = i/n$ . Let  $S_Z(i) = \sum_{j=1}^i Z_j$ ,  $i = 1, 2, \dots, n$ . The following invariance principle for  $(Z_i)_{i=1}^n$  plays an important role in both establishing the asymptotic theory of the testing methods and justifying the wild bootstrap procedures.

**THEOREM 4.** Assume that (i)  $\mathbb{E}\mathfrak{H}(t, \mathcal{F}_i) = 0$  for all  $t \in [0, 1]$ ; that (ii)  $\sup_{(t,x) \in [0,1] \times \mathbb{R}} \|H(t, x, \mathcal{F}_i)\|_4 < c_0$  for some finite constant  $c_0$ ; that (iii)  $H$  sat-

isfies UGMC(4) and that (iv) there exists  $q \geq 1/4$ , such that for all  $i \geq 0$

$$(26) \quad \|\mathcal{P}_0\{\mathfrak{H}(s_1, \mathcal{F}_i) - \mathfrak{H}(s_2, \mathcal{F}_i)\}\| \leq C|s_1 - s_2|^q$$

holds for all  $s_1, s_2 \in [0, 1]$ , where  $\mathcal{P}_k(\cdot) = \mathbb{E}(\cdot|\mathcal{F}_k) - \mathbb{E}(\cdot|\mathcal{F}_{k-1})$  for  $k \geq 0$ . Then on a richer probability space, there exists i.i.d. standard normal random variables

$(V_i)_1^n$  and a process  $S_Z^\circ(i)$  with  $\{S_Z^\circ(i)\}_{i=1}^n \stackrel{D}{=} \{S_Z(i)\}_{i=1}^n$ , such that

$$(27) \quad \max_{i \leq n} \left| S_Z^\circ(i) - \sum_{j=1}^i v(j/n) V_j \right| = o_{\mathbb{P}}(n^{1/4} \log^2 n),$$

where  $v(s) = [\sum_{k \in \mathbb{Z}} \text{cov}(H(s, g(s), \mathcal{F}_0), H(s, g(s), \mathcal{F}_k))]^{1/2}$ ,  $0 \leq s \leq 1$ .

PROOF. Note  $v(s) = \|\mathcal{P}_0 \sum_{k=0}^\infty \mathfrak{H}(s, \mathcal{F}_k)\|$ . Let  $\mathfrak{H}(t, \mathcal{F}_i) = 0$  if  $t > 1$ . By Corollary 1 of Wu and Zhou (2010), it follows that under conditions (i) and (ii), there exist i.i.d. standard normal random variables  $(V_i)_1^n$  and a process  $S_Z^\circ(i)$  with  $\{S_Z^\circ(i)\}_{i=1}^n \stackrel{D}{=} \{S_Z(i)\}_{i=1}^n$ , such that

$$(28) \quad \max_{i \leq n} \left| S_Z^\circ(i) - \sum_{j=1}^i \tilde{v}_j V_j \right| = o_{\mathbb{P}}(n^{1/4} \log^{3/2} n),$$

where  $\tilde{v}_j = \|\mathcal{P}_0 \sum_{k=0}^\infty \mathfrak{H}(t_j + t_k, \mathcal{F}_k)\|$ . Based on (26),

$$(29) \quad \|\mathcal{P}_0\{\mathfrak{H}(t_i, \mathcal{F}_k) - \mathfrak{H}(t_i + t_k, \mathcal{F}_k)\}\| \leq t_k^q$$

for  $0 \leq k \leq n - i$ . On the other hand, by Theorem 1 of Wu (2005), we obtain

$$(30) \quad \|\mathcal{P}_0\{\mathfrak{H}(t_i, \mathcal{F}_k) - \mathfrak{H}(t_i + t_k, \mathcal{F}_k)\}\| \leq 2\delta_H(k, 2)$$

for all  $k \geq 0$ . By (29), (30) and the fact that  $\delta_H(k, 2) \leq \delta_H(k, 4) = O(\chi^k)$  for some  $0 < \chi < 1$ , we have for all  $1 \leq j \leq n - \lceil n^{1/4} \rceil$ ,

$$\begin{aligned} (v(j/n) - \tilde{v}_j)^2 &\leq \left\| \mathcal{P}_0 \sum_{k=0}^\infty [\mathfrak{H}(t_j, \mathcal{F}_k) - \mathfrak{H}(t_j + t_k, \mathcal{F}_k)] \right\|^2 \\ &\leq \left( \sum_{k=0}^\infty \|\mathcal{P}_0[\mathfrak{H}(t_j, \mathcal{F}_k) - \mathfrak{H}(t_j + t_k, \mathcal{F}_k)]\| \right)^2 \\ &\leq \left( \sum_{k=0}^{k^*} t_k^q + \sum_{k=k^*}^\infty 2\chi^k \right)^2 = O(n^{-1/2} \log n^{5/2}), \end{aligned}$$

where  $k^* = \lfloor -\log n / \log \chi \rfloor$ . Note for  $n - \lceil n^{1/4} \rceil < j \leq n$ ,  $(v(j/n) - \tilde{v}_j)^2 = O(1)$ . Therefore,  $\sum_{j=1}^n (v(j/n) - \tilde{v}_j)^2 = O(n^{1/2} \log^{5/2} n)$ . Hence,

$$(31) \quad \begin{aligned} \max_{i \leq n} \left| \sum_{j=1}^i v(j/n) V_j - \sum_{j=1}^i \tilde{v}_j V_j \right| &= O_{\mathbb{P}}((n^{1/2} \log^{5/2} n)^{1/2} \log^{1/2} n) \\ &= o_{\mathbb{P}}(n^{1/4} \log^2 n). \end{aligned}$$

By (28) and (31), Theorem 4 follows.  $\square$

REMARK 3. The Gaussian approximation in Theorem 4 shows that partial sums of a short range dependent (SRD) locally stationary process can be approximated by weighted sums of i.i.d. standard normal random variables. We shall call  $v^2(t) = \sum_{k \in \mathbb{Z}} \text{cov}(\mathfrak{H}(t, \mathcal{F}_0), \mathfrak{H}(t, \mathcal{F}_k))$  the long-run variance of the system  $\{\mathfrak{H}(t, \mathcal{F}_i)\}$  at time  $t$ . The weight  $v(t)$  captures the local dependence of the series  $(Z_i)$  at  $t$ ; while the fluctuation of  $v(t)$  on  $[0, 1]$  is due to the nonstationarity of the series. Following the arguments in Wu and Zhou (2010), the bound  $o_{\mathbb{P}}(n^{1/4} \log^2 n)$  in (27) is optimal within a multiplicative logarithmic factor.

We shall state and prove the following Lemma 1 and Lemma 2 before we proceed to the proof of Theorem 1.

LEMMA 1. Assume  $Q(t) \in \mathcal{C}^2[0, 1]$ ; that conditions (A1)–(A6) and (K1) hold, and that  $b_n \rightarrow 0$  with  $nb_n^{3/2} \rightarrow \infty$ , then we have

$$\begin{aligned} & \sup_{t \in \mathcal{T}_n} |f(t, Q(t))[\hat{Q}(t) - Q(t)] - S_n(t) - \mathbb{E}S_n^*(t)| \\ &= O_{\mathbb{P}}\left(\frac{\pi_n^{1/2} \log n + b_n^{3/4}}{\sqrt{nb_n}} + b_n \pi_n + \pi_n^2\right), \end{aligned}$$

where  $\pi_n = (nb_n)^{-1/2}(\log n + (b_n)^{-1/4} + (nb_n^5)^{1/2})$ ,

$$(32) \quad S_n(t) = \sum_{i=1}^n \psi(X_i - Q(t_i))K_{b_n}(t - t_i)/(nb_n)$$

with  $\psi_{\alpha}(x) = \alpha - I\{x \leq 0\}$ , and

$$S_n^*(t) = \sum_{i=1}^n \psi(X_i - Q(t) - (t_i - t)Q'(t))K_{b_n}(t - t_i)/(nb_n).$$

PROOF. A careful check of the proof of Theorem 3 of Zhou and Wu (2009) shows that under conditions of Lemma 1

$$(33) \quad \begin{aligned} & \sup_{t \in \mathcal{T}_n} |f(t, Q(t))[\hat{Q}(t) - Q(t)] - S_n^*(t)| \\ &= O_{\mathbb{P}}\left(\frac{\pi_n^{1/2} \log n}{\sqrt{nb_n}} + b_n \pi_n + \pi_n^2\right). \end{aligned}$$

Let  $\psi(i, t) = \psi(X_i - Q(t_i)) - \psi(X_i - Q(t) - (t_i - t)Q'(t))$ ,

$$M_n(t) = \sum_{i=1}^n \mathcal{P}_i \psi(i, t)K_{b_n}(t - t_i)/(nb_n),$$

$$N_n(t) = \sum_{i=1}^n [\mathbb{E}[\psi(i, t)|\mathcal{F}_{i-1}] - \mathbb{E}\psi(i, t)]K_{b_n}(t - t_i)/(nb_n).$$

Note the summands of  $M_n(t)$  form a triangular array of martingale differences and  $N_n(t)$  is differentiable with respect to  $t$ . Using similar chaining arguments as those in the proof of Lemmas 5 and 6 in Zhou and Wu (2009), we have

$$(34) \quad \sup_{t \in [0,1]} |M_n(t)| = O_{\mathbb{P}}\left(\frac{b_n \log n}{\sqrt{nb_n}}\right), \quad \sup_{t \in [0,1]} |N_n(t)| = O_{\mathbb{P}}\left(\frac{b_n^{3/4}}{\sqrt{nb_n}}\right).$$

Note  $S_n(t) - (S_n^*(t) - \mathbb{E}S_n^*(t)) = M_n(t) + N_n(t)$ . Therefore, by (33), (34) and the fact that  $b_n \log n = o(b_n^{3/4})$ , we show that Lemma 1 holds.  $\square$

LEMMA 2. *Let  $\{V_i\}$  be i.i.d. standard normal random variables. Assume condition (K1) and  $b_n \rightarrow 0$  with  $nb_n/(\log n)^2 \rightarrow \infty$ . Then for any  $x \in \mathbb{R}$*

$$\lim_{n \rightarrow \infty} \mathbb{P}\left[\frac{1}{\sqrt{\phi nb_n}} \sup_{t \in \mathcal{T}_n} \left| \sum_{i=1}^n V_i K_{b_n}(t_i - t) \right| - B(m^*) \leq \frac{x}{\sqrt{2 \log m^*}}\right] = e^{-2e^{-x}}.$$

Lemma 2 follows from classic results for extremes of Gaussian processes. See, for example, Bickel and Rosenblatt (1973) and Lemma 2 of Wu and Zhao (2007). Details are omitted.

PROOF OF THEOREM 1. Consider the system  $\{1/2 - J(t, x, \mathcal{F}_i)\}$ ,  $(t, x) \in [0, 1] \times \mathbb{R}$ . Recall  $J(t, x, \mathcal{F}_i) = I\{G(t, \mathcal{F}_i) \leq x\}$ . Let  $\mathfrak{J}(t, \mathcal{F}_i) = 1/2 - J(t, Q(t), \mathcal{F}_i)$ . We shall first show that  $\mathfrak{J}(t, \mathcal{F}_i)$  satisfies conditions of Theorem 4. Obviously  $\mathbb{E}\mathfrak{J}(t, \mathcal{F}_i) = 0$  and  $\|1/2 - J(t, x, \mathcal{F}_i)\|_4 \leq 1$  for all  $t \in [0, 1]$  and  $x \in \mathbb{R}$ . Based on (A5), condition (iii) of Theorem 4 holds. We now check (iv). Note for any  $i \geq 0$

$$(35) \quad \begin{aligned} \|\mathcal{P}_0\{\mathfrak{J}(s_1, \mathcal{F}_i) - \mathfrak{J}(s_2, \mathcal{F}_i)\}\| &\leq \|\mathfrak{J}(s_1, \mathcal{F}_i) - \mathfrak{J}(s_2, \mathcal{F}_i)\| \\ &= \|J(s_1, Q(s_1), \mathcal{F}_i) - J(s_2, Q(s_2), \mathcal{F}_i)\|. \end{aligned}$$

Based on (A6) and the smoothness condition on  $Q(t)$ ,

$$(36) \quad \|J(s_2, Q(s_1), \mathcal{F}_i) - J(s_2, Q(s_2), \mathcal{F}_i)\| \leq C|s_1 - s_2|^{1/2}.$$

On the other hand,

$$(37) \quad \|J(s_1, Q(s_1), \mathcal{F}_i) - J(s_2, Q(s_1), \mathcal{F}_i)\| \leq I + II,$$

where  $I = \|[J(s_1, Q(s_1), \mathcal{F}_i) - J(s_2, Q(s_1), \mathcal{F}_i)]I\{|\zeta_i(s_1) - \zeta_i(s_2)| \leq \delta\}\|$  and  $II = \|[J(s_1, Q(s_1), \mathcal{F}_i) - J(s_2, Q(s_1), \mathcal{F}_i)]I\{|\zeta_i(s_1) - \zeta_i(s_2)| > \delta\}\|$ . Recall  $\zeta_i(t) = G(t, \mathcal{F}_i)$ . Using condition (A6), we have for all  $\delta > 0$ ,

$$(38) \quad \begin{aligned} I &\leq \|I\{Q(s_1) - \delta \leq \zeta_i(s_1) \leq Q(s_1)\}\| + \|I\{Q(s_1) \leq \zeta_i(s_1) \leq Q(s_1) + \delta\}\| \\ &\leq C\delta^{1/2}. \end{aligned}$$

By condition (A3), there exists  $q \geq 1$ , such that

$$(39) \quad H \leq \frac{\|\zeta_i(s_1) - \zeta_i(s_2)\|_q^{q/2}}{\delta^{q/2}} \leq C \frac{|s_1 - s_2|^{q/2}}{\delta^{q/2}}.$$

Let  $\delta = |s_1 - s_2|^{q/(q+1)}$ , then (37), (38) and (39) imply that

$$(40) \quad \|J(s_1, Q(s_1), \mathcal{F}_i) - J(s_2, Q(s_1), \mathcal{F}_i)\| \leq C|s_1 - s_2|^s,$$

where  $s = \frac{q}{2(1+q)} \geq \frac{1}{4}$ . By (35), (36) and (40), we have condition (iv) of Theorem 4 holds. Therefore, Theorem 4 implies that there exist i.i.d. standard normal random variables  $(V_i)_1^n$ , such that

$$(41) \quad \max_{i \leq n} \left| S_{\mathfrak{J}}(i) - \sum_{j=1}^i \sigma(t_j) V_j \right| = o_{\mathbb{P}}(n^{1/4} \log^2 n),$$

where  $S_{\mathfrak{J}}(i) = \sum_{k=1}^i \mathfrak{J}(t_k, \mathcal{F}_k)$ . Recall  $\sigma(t)$  was defined in (4). Define

$$(42) \quad \Xi_n(t) = \sum_{i=1}^n \sigma(t_i) V_i K_{b_n}(t_i - t) / (nb_n),$$

$$(43) \quad \Omega_n(t) = \left\{ K_{b_n}(t_1 - t) + \sum_{i=2}^n |K_{b_n}(t_i - t) - K_{b_n}(t_{i-1} - t)| \right\} / (nb_n).$$

By the summation by parts formula and (41), we obtain

$$(44) \quad \begin{aligned} \sup_{t \in \mathcal{T}_n} |S_n(t) - \Xi_n(t)| &\leq \sup_{t \in \mathcal{T}_n} \Omega_n(t) \max_{i \leq n} \left| S_{\mathfrak{J}}(i) - \sum_{j=1}^i \sigma(t_j) V_j \right| \\ &= o_{\mathbb{P}} \left( \frac{n^{1/4} \log^2 n}{nb_n} \right). \end{aligned}$$

By the Lipschitz continuity of  $\sigma(t)$  in (A2), it is easy to see that

$$(45) \quad \sup_{t \in \mathcal{T}_n} |\Xi_n(t) - \Xi_n^*(t)| = O_{\mathbb{P}} \left( \frac{b_n \log n}{\sqrt{nb_n}} \right),$$

where  $\Xi_n^*(t) = \sigma(t) \sum_{i=1}^n V_i K_{b_n}(t_i - t) / (nb_n)$ . By Lemma 1, (44), (45) and the fact that  $b_n \log n = o(b_n^{3/4})$ , it follows that

$$(46) \quad \begin{aligned} \sup_{t \in \mathcal{T}_n} |f(t, Q(t))[\hat{Q}(t) - Q(t)] - \Xi_n^*(t) - \mathbb{E}S_n^*(t)| \\ = O_{\mathbb{P}} \left( \frac{\pi_n^{1/2} \log n + b_n^{3/4}}{\sqrt{nb_n}} + b_n \pi_n + \pi_n^2 + \frac{n^{1/4} \log^2 n}{nb_n} \right). \end{aligned}$$

It is easy to check that under bandwidth conditions of Theorem 1, the right-hand side of (46) is of order  $o_{\mathbb{P}}((nb_n \log n)^{-1/2})$ . On the other hand, by condition (A1) and a Taylor expansion, we have

$$(47) \quad \begin{aligned} \sup_{t \in \mathcal{T}_n} |\mathbb{E}S_n^*(t) - \mu_2 f(t, Q(t)) Q''(t) b_n^2 / 2| &= O\left(b_n^3 + \frac{1}{nb_n}\right) \\ &= o((nb_n \log n)^{-1/2}). \end{aligned}$$

By Lemma 2, (46) and (47), Theorem 1 follows.  $\square$

Let

$$(48) \quad \begin{aligned} S_n(t) &= \sum_{i=1}^n \psi(X_i - Q(t_i)) K_{b_n}^*(t - t_i) / (nb_n), \\ S_n^o(t) &= \sum_{i=1}^n \psi(X_i - Q(t) - (t_i - t) Q'(t)) K_{b_n}^*(t - t_i) / (nb_n). \end{aligned}$$

Define

$$(49) \quad \mathfrak{Y}_n(t) = \tilde{Q}(t) - Q(t) - \mathbb{E}S_n^o(t) / f(t, Q(t)).$$

We shall introduce several lemmas before proceeding to the proof of Theorem 3.

LEMMA 3. *Under the conditions of Theorem 3, we have*

$$(50) \quad \sqrt{n} \int_{\mathcal{T}_n^*} \mathfrak{Y}_n(t) \pi(t) dt \Rightarrow N\left(0, \int_0^1 \varpi^2(t) dt\right),$$

where  $\varpi(t) = \pi(t)\sigma(t)/f(t, Q(t))$ .

PROOF. By Lemma 1, we have

$$(51) \quad \begin{aligned} \sup_{t \in \mathcal{T}_n^*} |f(t, Q(t)) \mathfrak{Y}_n(t) - S_n(t)| \\ = O_{\mathbb{P}}\left(\frac{\pi_n^{1/2} \log n + b_n^{3/4}}{\sqrt{nb_n}} + b_n \pi_n + \pi_n^2\right). \end{aligned}$$

On the other hand, arguments similar to those in the proof of (44) and (45) lead to

$$(52) \quad \sup_{t \in \mathcal{T}_n^*} |S_n(t) - \sigma(t) \mathfrak{X}_n(t)| = o_{\mathbb{P}}\left(\frac{n^{1/4} \log^2 n}{nb_n} + \frac{b_n \log n}{\sqrt{nb_n}}\right),$$

where

$$(53) \quad \mathfrak{X}_n(t) = \sum_{i=1}^n V_i K_{b_n}^*(t_i - t) / (nb_n).$$

Recall  $(V_i)_1^n$  are i.i.d. standard normal random variables defined in (41). Hence, by (51), (52) and the bandwidth conditions of Theorem 3, it follows that

$$(54) \quad \sup_{t \in \mathcal{T}_n^*} \left| \mathfrak{Y}_n(t) - \frac{\sigma(t)}{f(t, Q(t))} \mathfrak{X}_n(t) \right| = o_{\mathbb{P}}(n^{-1/2}).$$

Furthermore, it is easy to obtain that

$$(55) \quad \sqrt{n} \int_{\mathcal{T}_n^*} \mathfrak{X}_n(t) \varpi(t) dt \Rightarrow N\left(0, \int_0^1 \varpi^2(t) dt\right).$$

By (54) and (55), this lemma follows.  $\square$

LEMMA 4. Recall (13) for the definition of  $\pi^*(t)$  and (53) for the definition of  $\mathfrak{X}_n(t)$ . We have

$$\begin{aligned} & n\sqrt{b_n} \int_{\mathcal{T}_n^*} \mathfrak{X}_n(t)^2 \pi^*(t) dt - \frac{1}{\sqrt{b_n}} K^* \star K^*(0) \int_0^1 \pi^*(t) dt \\ & \Rightarrow N\left(0, 2 \int_{\mathbb{R}} [K^* \star K^*(t)]^2 dt \int_0^1 \pi^*(t)^2 dt\right). \end{aligned}$$

PROOF. Define

$$\begin{aligned} I_n &= \sum_{i=1}^n R_i^2 \int_{\mathcal{T}_n^*} [K_{b_n}^*(t_i - t)]^2 \pi^*(t) dt / (nb_n^{3/2}), \\ II_n &= \sum_{1 \leq i \neq j \leq n} R_i R_j \int_{\mathcal{T}_n^*} K_{b_n}^*(t_i - t) K_{b_n}^*(t_j - t) \pi^*(t) dt / (nb_n^{3/2}). \end{aligned}$$

Then by the central limit theorem for  $I_n$ , it follows that

$$(56) \quad \begin{aligned} I_n &= \frac{1}{\sqrt{b_n}} K^* \star K^*(0) \int_0^1 \pi^*(t) dt + O_{\mathbb{P}}\left(\frac{1}{nb_n^{3/2}} + \frac{1}{\sqrt{nb_n}}\right) \\ &= \frac{1}{\sqrt{b_n}} K^* \star K^*(0) \int_0^1 \pi^*(t) dt + o_{\mathbb{P}}(1). \end{aligned}$$

On the other hand, by Theorem 2.1 of de Jong (1987), elementary but tedious calculations show that

$$(57) \quad II_n \Rightarrow N\left(0, 2 \int_{\mathbb{R}} [K^* \star K^*(t)]^2 dt \int_0^1 \pi^*(t)^2 dt\right).$$

Since  $n\sqrt{b_n} \int_{\mathcal{T}_n^*} \mathfrak{X}_n(t)^2 \pi^*(t) dt = I_n + II_n$ , by (56) and (57), the lemma follows.  $\square$

PROPOSITION 2. Let  $\tilde{T}_n = \int_{\mathcal{T}_n^*} \mathfrak{Y}_n^2(t) \pi(t) dt$ . Then under the conditions of Theorem 3, we have

$$(58) \quad \begin{aligned} n\sqrt{b_n}\tilde{T}_n - \frac{1}{\sqrt{b_n}}K^* \star K^*(0) \int_0^1 \pi^*(t) dt \\ \Rightarrow N\left(0, 2 \int_{\mathbb{R}} [K^* \star K^*(t)]^2 dt \int_0^1 \pi^*(t)^2 dt\right). \end{aligned}$$

PROOF. By (51) and (52), we have

$$(59) \quad \sup_{t \in \mathcal{T}_n^*} \left| \mathfrak{Y}_n(t) - \frac{\sigma(t)}{f(t, Q(t))} \mathfrak{X}_n(t) \right| = O_{\mathbb{P}}(\nu_n),$$

where  $\nu_n = \frac{\pi_n^{1/2} \log n + b_n^{3/4}}{\sqrt{nb_n}} + b_n \pi_n + \pi_n^2 + \frac{n^{1/4} \log^2 n}{nb_n}$ .

On the other hand, it is easy to show that

$$(60) \quad \sup_{t \in \mathcal{T}_n^*} |\mathfrak{X}_n(t)| = O_{\mathbb{P}}\left(\frac{\log^{1/2} n}{\sqrt{nb_n}}\right).$$

Therefore, (59) and (60) imply that

$$(61) \quad \sup_{t \in \mathcal{T}_n^*} |\mathfrak{Y}_n(t)| = O_{\mathbb{P}}\left(\frac{\log^{1/2} n}{\sqrt{nb_n}}\right).$$

Hence, by (59), (60), (61) and bandwidth conditions of Theorem 3,

$$(62) \quad \begin{aligned} & \left| \tilde{T}_n - \int_{\mathcal{T}_n^*} [\mathfrak{X}_n(t)]^2 \pi^*(t) dt \right| \\ &= \left| \int_{\mathcal{T}_n^*} \Delta_n(t) \left( \mathfrak{Y}_n(t) + \frac{\sigma(t)}{f(t, Q(t))} \mathfrak{X}_n(t) \right) \pi(t) dt \right| \\ &\leq C \sup_{t \in \mathcal{T}_n^*} |\Delta_n(t)| \sup_{t \in \mathcal{T}_n^*} \left[ |\mathfrak{Y}_n(t)| + \left| \frac{\sigma(t)}{f(t, Q(t))} \mathfrak{X}_n(t) \right| \right] \\ &= O_{\mathbb{P}}\left(\nu_n \left(\frac{\log^{1/2} n}{\sqrt{nb_n}}\right)\right) = o_{\mathbb{P}}\left(\frac{1}{n\sqrt{b_n}}\right), \end{aligned}$$

where  $\Delta_n(t) = \mathfrak{Y}_n(t) - \frac{\sigma(t)}{f(t, Q(t))} \mathfrak{X}_n(t)$ . Therefore, by Lemma 4, we illustrate that Proposition 2 holds.  $\square$

PROOF OF THEOREM 3. First, note

$$\tilde{Q}(t) - Q^o(t) = \mathfrak{Y}_n(t) + \mathbb{E}\mathcal{S}_n^o(t)/f(t, Q(t)) + \varrho_n(\eta(t) + o(1)).$$

Therefore,

$$\begin{aligned}
 T_n^* &= \tilde{T}_n + \int_{\mathcal{T}_n^*} [\mathbb{E}S_n^o(t)/f(t, Q(t)) + \varrho_n(\eta(t) + o(1))]^2 \pi(t) dt \\
 (63) \quad &+ 2 \int_{\mathcal{T}_n^*} \mathfrak{Y}_n(t) [\mathbb{E}S_n^o(t)/f(t, Q(t)) + \varrho_n(\eta(t) + o(1))] \pi(t) dt \\
 &:= \tilde{T}_n + I_n^* + II_n^*.
 \end{aligned}$$

Since  $Q(t) \in \mathcal{C}^2[0, 1]$ , simple calculations show that

$$(64) \quad \sup_{t \in \mathcal{T}_n^*} |\mathbb{E}S_n^o(t)| = o(b_n^2) + O\left(\frac{1}{nb_n}\right).$$

By the bandwidth conditions of Theorem 3, it follows that

$$(65) \quad I_n^* - \varrho_n^2 \int_{\mathcal{T}_n^*} \eta^2(t) \pi(t) dt = o(b_n^4 + b_n^2 \varrho_n) = o(\varrho_n^2).$$

By Lemma 3,

$$\begin{aligned}
 (66) \quad \varrho_n \int_{\mathcal{T}_n^*} \mathfrak{Y}_n(t) [\eta(t) + o(1)] \pi(t) dt &= O_{\mathbb{P}}(\varrho_n n^{-1/2}) \\
 &= o_{\mathbb{P}}(n^{-1} (b_n)^{-1/2}).
 \end{aligned}$$

Similarly,

$$(67) \quad \int_{\mathcal{T}_n^*} \mathfrak{Y}_n(t) [\mathbb{E}S_n^o(t)/f(t, Q(t))] \pi(t) dt = o_{\mathbb{P}}(n^{-1} (b_n)^{-1/2}).$$

Hence, Proposition 2, (63), (65), (66) and (67) imply that Theorem 3 holds.  $\square$

REMARK 4. Under the null hypothesis  $Q(t) = Q^o(t)$ , we see from the above proof that the bias  $\mathcal{B}_n(t)$  influences the asymptotic behavior of  $T_n^*$  through two terms  $\int_{\mathcal{T}_n^*} \mathcal{B}_n^2(t) \pi(t) dt$  and  $\int_{\mathcal{T}_n^*} \mathfrak{Y}_n(t) \mathcal{B}_n(t) \pi(t) dt$ , where  $\mathcal{B}_n(t) = \mathbb{E}S^o(t)/f(t, Q(t))$ . The jackknife bias reduction technique reduces those two effects to second order. However, it can be shown that if the original estimate  $\hat{Q}(t)$  is used in the  $\mathcal{L}^2$  test, then the first term is not negligible under the optimal bandwidth  $b_n = O(n^{-2/9})$ , which complicates the asymptotic analysis and reduces the precision of the test.

PROOF OF PROPOSITION 1. Proposition 1 follows from Lemmas 5 and 6 below.  $\square$

LEMMA 5. Let  $L_n = \sum_{i=1}^n \psi(e_i) \mathbf{g}(t_i)$ . Then under the conditions of Proposition 1, we have  $L_n = O_{\mathbb{P}}(\sqrt{n})$ .

PROOF. Let  $L_{n,k} = \sum_{i=1}^n \mathcal{P}_{i-k} \psi(e_i) \mathbf{g}(t_i)$ . Recall the operator  $\mathcal{P}_k$  is defined in Theorem 4. Then the summands of  $L_{n,k}$  are martingale differences. By the orthogonality, we have for  $k \geq 1$

$$\begin{aligned} \|L_{n,k}\|^2 &= \sum_{i=1}^n \|\mathcal{P}_{i-k} \psi(e_i) \mathbf{g}(t_i)\|^2 = \sum_{i=1}^n |\mathbf{g}(t_i)|^2 \|\mathcal{P}_{i-k} \psi(e_i)\|^2 \\ &\leq Cn\delta_F^2(k-1, 2). \end{aligned}$$

Similar arguments also imply  $\|L_{n,k}\|^2 \leq Cn$ . Therefore,

$$\|L_n\| = \sum_{k=0}^{\infty} \|L_{n,k}\| \leq C\sqrt{n} \left( 1 + \sum_{k=1}^{\infty} \delta_F(k-1, 2) \right) \leq C\sqrt{n}.$$

Hence, Lemma 5 follows.  $\square$

LEMMA 6. Under the conditions of Proposition 1, we have

$$(68) \quad \mathcal{H}_\alpha \tilde{\theta}_\alpha - \sum_{i=1}^n \psi_\alpha(e_i) \mathbf{g}(t_i) / \sqrt{n} = o_p(1),$$

where  $\tilde{\theta}_\alpha = \sqrt{n}(\hat{\theta}_\alpha - \theta_0)$ ,  $e_i = X_i - Q_\alpha(t_i)$  and

$$\mathcal{H}_\alpha = \int_0^1 \mathbf{g}(t) \mathbf{g}^\top(t) f(t, Q_\alpha(t)) dt.$$

PROOF. We shall omit the subscript  $\alpha$  in the proof. Let  $\mathbf{g}_n(t) = \mathbf{g}(t) / \sqrt{n}$ . By arguments similar to those of Lemma 3 of Wu (2007), we have for any fixed constant  $c$  and fixed  $\theta \leq c$ ,

$$(69) \quad \text{var} \left( \sum_{i=1}^n \eta_i(\theta) \right) = o(1),$$

where  $\eta_i(\theta) = \rho(e_i - \mathbf{g}_n^\top(t_i)\theta) - \rho(e_i) + \mathbf{g}_n^\top(t_i)\theta\psi(e_i)$ . On the other hand, elementary calculations based on the Taylor expansion show that

$$(70) \quad \begin{aligned} \sum_{i=1}^n \mathbb{E}[\eta_i(\theta)] &= \sum_{i=1}^n \left[ \frac{f(t_i, Q(t_i))}{2} |\mathbf{g}_n^\top(t_i)\theta|^2 + o(|\mathbf{g}_n^\top(t_i)\theta|^2) \right] \\ &= \frac{\theta^\top \mathcal{H} \theta}{2} + o(1). \end{aligned}$$

From equations (69), (70) and the convexity lemma [Pollard (1991), page 187], we obtain

$$(71) \quad \sup_{\theta \leq c} \left| \sum_{i=1}^n [\rho(e_i - \mathbf{g}_n^\top(t_i)\theta) - \rho(e_i) + \mathbf{g}_n^\top(t_i)\theta\psi(e_i)] - \frac{\theta^\top \mathcal{H} \theta}{2} \right| = o_p(1).$$

Now a standard argument using properties of convex functions will lead to (68). See, for example, the proofs of Theorems 2.2 and 2.4 in Bai, Rao and Wu (1992). Details are omitted.  $\square$

**Acknowledgments.** We are grateful to two referees, the Associate Editor and the editor for their many helpful comments.

## REFERENCES

- BAI, Z. D., RAO, C. R. and WU, Y. (1992).  $M$ -estimation of multivariate linear regression parameters under a convex discrepancy function. *Statist. Sinica* **2** 237–254. [MR1152307](#)
- BICKEL, P. J. and ROSENBLATT, M. (1973). On some global measures of the deviations of density function estimates. *Ann. Statist.* **1** 1071–1095. [MR0348906](#)
- DE JONE, P. (1987). A central limit theorem for generalized quadratic forms. *Probab. Theory Related Fields* **75** 261–277. [MR0885466](#)
- DEL BARRIO, E. and MATRÁN, C. (2000). The weighted bootstrap mean for heavy-tailed distributions. *J. Theoret. Probab.* **13** 547–569. [MR1778584](#)
- DRAGHICESCU, D., GUILLAS, S. and WU, W. B. (2009). Quantile curve estimation and visualization for non-stationary time series. *J. Comput. Graph. Statist.* **18** 1–20. [MR2511058](#)
- ELSNER, J. B., KOSSIN, J. P. and JAGGER, T. H. (2008). The increasing intensity of the strongest tropical cyclones. *Nature* **455** 9295.
- EUBANK, R. L. and SPECKMAN, P. L. (1993). Confidence bands in nonparametric regression. *J. Amer. Statist. Assoc.* **88** 1287–1301. [MR1245362](#)
- FAN, J. and GIJBELS, I. (1996). *Local Polynomial Modelling and Its Applications*. Chapman and Hall, London. [MR1383587](#)
- FAN, J. and JIANG, J. (2005). Nonparametric inference for additive models. *J. Amer. Statist. Assoc.* **100** 890–907. [MR2201017](#)
- FAN, J. and JIANG, J. (2007). Nonparametric inference with generalized likelihood ratio tests. *TEST* **16** 409–444. [MR2365172](#)
- FAN, J. and ZHANG, W. Y. (2000). Simultaneous confidence bands and hypothesis testing in varying-coefficient models. *Scand. J. Statist.* **27** 715–731. [MR1804172](#)
- FAN, J., ZHANG, C. and ZHANG, J. (2001). Generalized likelihood ratio statistics and Wilks phenomenon. *Ann. Statist.* **29** 153–193. [MR1833962](#)
- FARAWAY, J. J. and JHUN, M. (1990). Bootstrap choice of bandwidth for density estimation. *J. Amer. Statist. Assoc.* **85** 1119–1122. [MR1134508](#)
- FITZENBERGER, B. (1998). The moving blocks bootstrap and robust inference for linear least squares and quantile regressions. *J. Econometrics* **82** 235–287. [MR1613422](#)
- GHOSH, S. and DRAGHICESCU, D. (2002). An algorithm for optimal bandwidth selection for smooth nonparametric quantiles and distribution functions. In *Statistical Data Analysis Based on the  $L_1$ -Norm and Related Methods* (Y. Dodge, ed.) 161–168. Birkhäuser, Basel. [MR2001312](#)
- HALL, P. and TITTERINGTON, D. M. (1988). On confidence bands in nonparametric density estimation and regression. *J. Multivariate Anal.* **27** 228–254. [MR0971184](#)
- HALL, P. (1993). On Edgeworth expansion and bootstrap confidence bands in nonparametric curve estimation. *J. R. Stat. Soc. Ser. B Stat. Methodol.* **55** 291–304.
- HÄRDLE, W. and MAMMEN, E. (1993). Comparing nonparametric versus parametric regression fits. *Ann. Statist.* **21** 1926–1947. [MR1245774](#)
- HART, J. D. (1997). *Nonparametric Smoothing and Lack-of-Fit Tests*. Springer, New York. [MR1461272](#)
- HE, X. and ZHU, L.-X. (2003). A lack-of-fit test for quantile regression. *J. Amer. Statist. Assoc.* **98** 1013–1022. [MR2041489](#)

- HOROWITZ, J. L. and SPOKOINY, V. G. (2001). An adaptive, rate-optimal test of a parametric model against a nonparametric alternative. *Econometrica* **69** 599–631. [MR1828537](#)
- HOROWITZ, J. L. and SPOKOINY, V. G. (2002). An adaptive, rate-optimal test of linearity for median regression models. *J. Amer. Statist. Assoc.* **97** 822–835. [MR1941412](#)
- INGSTER, Y. I. (1993). Asymptotically minimax hypothesis testing for nonparametric alternatives. I–III. *Math. Methods Statist.* **2** 85–114; **3** 171–189; **4** 249–268.
- JHUN, M. (1988). Bootstrapping density estimates. *Comm. Statist. Theory Methods* **17** 61–78. [MR0940967](#)
- KATZ, R. W. and BROWN, B. G. (1992). Extreme events in a changing climate: Variability is more important than averages. *Climatic Change* **21** 289–302.
- KIM, M. O. (2007). Quantile regression with varying coefficients. *Ann. Statist.* **35** 92–108. [MR2332270](#)
- KOENKER, R. (2005). *Quantile Regression*. Cambridge Univ. Press, Cambridge. [MR2268657](#)
- LEPSKI, O. V. and SPOKOINY, V. G. (1999). Minimax nonparametric hypothesis testing: The case of an inhomogeneous alternative. *Bernoulli* **5** 333–358. [MR1681702](#)
- MAMMEN, E. (1993). Bootstrap and wild bootstrap for high dimensional linear models. *Ann. Statist.* **21** 255–285. [MR1212176](#)
- NEUMANN, M. H. and KREISS, J. P. (1998). Regression-type inference in nonparametric autoregression. *Ann. Statist.* **26** 1570–1613. [MR1647701](#)
- POLITIS, D. N. and ROMANO, J. P. (1994). The stationary bootstrap. *J. Amer. Statist. Assoc.* **89** 1303–1313. [MR1310224](#)
- POLLARD, D. (1991). Asymptotics for least absolute deviation regression estimators. *Econometric Theory* **7** 186–199. [MR1128411](#)
- ROSENKRANTZ, W. (2000). Confidence bands for quantile functions: A parametric and graphic alternative for testing goodness-of-fit. *Amer. Statist.* **54** 185–190.
- SHAO, J. and TU, D. (1995). *The Jackknife and Bootstrap*. Springer, New York. [MR1351010](#)
- STUTE, W. (1997). Nonparametric model checks for regression. *Ann. Statist.* **25** 613–641. [MR1439316](#)
- STUTE, W., GONZALEZ MANTEIGA, W. and PRESEDO QUINDIMIL, M. (1998). Bootstrap approximations in model checks for regression. *J. Amer. Statist. Assoc.* **93** 141–149. [MR1614600](#)
- SUN, J. and LOADER, C. R. (1994). Simultaneous confidence bands for linear regression and smoothing. *Ann. Statist.* **22** 1328–1345. [MR1311978](#)
- WANG, L. (2007). A simple nonparametric test for diagnosing nonlinearity in Tobit median regression model. *Statist. Probab. Lett.* **77** 1034–1042. [MR2370429](#)
- WANG, L. (2008). Nonparametric test for checking lack-of-fit of quantile regression model under random censoring. *Canad. J. Statist.* **36** 321–336. [MR2431684](#)
- WU, W. B. (2005). Nonlinear system theory: Another look at dependence. *Proc. Natl. Acad. Sci. USA* **102** 14150–14154. [MR2172215](#)
- WU, W. B. (2007). M-estimation of linear models with dependent errors. *Ann. Statist.* **35** 495–521. [MR2336857](#)
- WU, W. B. and ZHAO, Z. (2007). Inference of trends in time series. *J. R. Stat. Soc. Ser. B Stat. Methodol.* **69** 429–446. [MR2323759](#)
- WU, W. B. and ZHOU, Z. (2010). Gaussian approximations for non-stationary multiple time series. *Statist. Sinica*. To appear.
- XIA, Y. (1998). Bias-corrected confidence bands in nonparametric regression. *J. R. Stat. Soc. Ser. B Stat. Methodol.* **60** 797–811. [MR1649488](#)
- YU, K. and JONES, M. C. (1998). Local linear quantile regression. *J. Amer. Statist. Assoc.* **93** 228–237. [MR1614628](#)
- ZHANG, C. and DETTE, H. (2004). A power comparison between nonparametric regression tests. *Statist. Probab. Lett.* **66** 289–301. [MR2045474](#)

- ZHAO, Z. and WU, W. B. (2008). Confidence bands in nonparametric time series regression. *Ann. Statist.* **36** 1854–1878. [MR2435458](#)
- ZHENG, J. X. (1996). A consistent test of functional form via nonparametric estimation techniques. *J. Econometrics* **75** 263–289. [MR1413644](#)
- ZHENG, J. X. (1998). A consistent nonparametric test of parametric regression models under conditional quantile restrictions. *Econometric Theory* **14** 123–138. [MR1613710](#)
- ZHOU, Z. and WU, W. B. (2009). Local linear quantile estimation for non-stationary time series. *Ann. Statist.* **37** 2696–2729. [MR2541444](#)
- ZHOU, Z. and WU, W. B. (2010). Simultaneous inference of linear models with time-varying coefficients. *J. Roy. Statist. Soc. Ser. B*. To appear.

DEPARTMENT OF STATISTICS  
UNIVERSITY OF TORONTO  
100 ST. GEORGE ST.  
TORONTO, ONTARIO M5S 3G3  
CANADA  
E-MAIL: [zhou@utstat.toronto.edu](mailto:zhou@utstat.toronto.edu)