应用概率统计 第二十八卷
第四期 2012年8月

Chinese Journal of Applied Probability
and Statistics Vol.28 No.4 Aug. 2012

Detection of Change Points in Volatility of Non-Parametric Regression by Wavelets *

WANG JINGLE ZHENG MING

(*Department of Statistics, Management School of Fudan University, shanghai, 200433*)

Abstract

This paper studies the detection and estimation of change points in volatility under nonparametric regression models. Wavelet methods are applied to construct the test statistics which can be used to detect change points in volatility. The asymptotic distributions of the test statistics are established. We also utilize the test statistics to construct the estimators for the locations and jump sizes of the change points in volatility. The asymptotic properties of these estimators are derived. Some simulation studies are conducted to assess the finite sample performance of the proposed procedures.

Keywords: Change points in volatility, wavelet coefficient, kernel estimation, local linear smoother, α-mixing.

AMS Subject Classification: 62G05, 91B82.

§1. Introduction

The study of the conditional variance of financial and economic data has draw much attention due to its importance in risk management, portfolio, and derivative pricing. Lamoreux and Lastrapes (1990) gave evidence of jumps in conditional variance. Many attempts have followed since then to test and estimate the change points in volatility. See, for example, Jorion (1988), Vlaar and Palm (1993), Drost et al. (1998) and so on. All of these models are variants of the popular ARCH and GARCH models.

For nonparametric regression models, the change point problem in volatility has also attracted increasing interests in recent years. Müller (1992) used a nonparametric procedure which is based on one sided kernel smoothers to detect the change points in volatility. Chen et al. (2005) pointed out that the one sided kernel procedure has a drawback: the power of the test is weak and the rate of convergence is slow. They proposed a hybrid estimation procedure that combines the least squares and nonparametric methods to estimate the change points of volatility. Moreover, they used the method based on the

*The project supported by the National Natural Science Foundation of China (10971033).
Received December 8, 2011. Revised February 23, 2012.

iterated cumulative sums of squares (ICSS) algorithm proposed by Inclan and Tiao (1994) to detect multiple change points in volatility. The idea of ICSS algorithm is that when the first change point is identified, the whole sample is divided into two sub-samples by the first change point, and then one makes detection of the change point for each sub-sample. This step is repeated until the null hypothesis that volatility is smooth could not be rejected for all sub-samples. However, the major drawback of the ICSS algorithm is that the sub-sample size decreases very fast with the increasing number of change points. They did not propose estimators for the jump sizes of the change points, either.

There have been developments of wavelet applications with respect to financial time series. Many authors used wavelet methods to detect and estimate the change points of the regression function in nonparametric regression models. See, for example, Wang (1995), Wang (1999), Wong et al. (1999), Luan and Xie (2001), Chen et al. (2008), Zhou et al. (2010), among many others. Wong et al. (2001) employed the wavelet methods to identify the jumps for the conditional mean and conditional variance in a heteroscedastic autoregressive model. However, they did not provide estimators for the jump sizes of the change points nor establish the asymptotic distributions for their test statistics.

In this paper, we apply the wavelet methods to detect and estimate the possible change points in volatility of nonparametric regression models with dependent observations. The test statistics we propose can be used to detect change points in volatility. The asymptotic distributions of the test statistics are established. We also use the test statistics to construct estimators of the locations and jump sizes of the change points. The asymptotic properties of these proposed estimators are derived. We carry out some simulation studies to assess the finite sample performance of the proposed methods.

The rest of this paper is organized as follows. Section 2 introduces some notation and the basic model. In Section 3 we develop the test statistics which can be used to detect change points in volatility. The asymptotic distributions of these test statistics are discussed. The estimators of the locations and jump sizes of the change points are constructed in Section 4. Section 5 gives out the setup and results of the simulation studies. All technique details are given in the Appendix.

§2. Model and Notation

We consider the following nonparametric regression model

$$Y_i = T(X_i) + \sigma(X_i)\varepsilon_i, \qquad i = 1, 2, \ldots, n, \tag{2.1}$$

where $T(x) = \mathsf{E}(Y|X = x)$ and $\sigma^2(x) = \mathsf{Var}\,(Y|X = x)$. $\{\varepsilon_i, i = 1, 2, \ldots\}$ is a sequence of independently and identically distributed (i.i.d.) random variables with $\mathsf{E}(\varepsilon_i) = 0$ and $\mathsf{Var}\,(\varepsilon_i) = 1$. $\{(Y_i, X_i), i = 1, 2, \ldots\}$ is a sequence of random vectors. Here, we assume that $\{(Y_i, X_i), i = 1, 2, \ldots\}$ is a strongly mixing sequence, which obviously includes the i.i.d. observations case. It also includes many time series models. For reader's convenience, more detailed assumptions of the model are stated in the Appendix.

Consider the volatility $\sigma^2(x)$ in model (2.1). We assume that $\sigma^2(x)$ has p jump points in the interval $[a, b]$ and $\sigma^2(x)$ can be decomposed as

$$\sigma^2(x) = C(x) + D(x), \tag{2.2}$$

where $D(x) = \sum\limits_{l=1}^{p} d_l I(t_l \leqslant x \leqslant b)$ with $a < t_1 < t_2 < \cdots < t_p < b$, $d_l = \sigma^2(t_l+) - \sigma^2(t_l-)$, and $C(x)$ is twice continuously differentiable on (a, b). The aim in this paper is first to detect the change points and then to estimate p, t_l and d_l, $l = 1, 2, \ldots, p$.

Wavelet methods are applied to detect and estimate the change points in volatility under model (2.1). Before our discussion, we need to introduce some additional notation and the wavelet coefficient of the volatility function. We use similar notation as that in Chen et al. (2008).

Suppose $\{(Y_i, X_i), 1 \leqslant i \leqslant n\}$ is the observed data following model (2.1). Denote

$$I(s, \delta_n) = \Big\{k : \Big|a + \frac{k}{2^J}(b - a) - s\Big| \leqslant \delta_n\Big\},$$

where $\delta_n = 2^{-J}$ and $k \in I_J = \{0, 1, 2, \ldots, 2^J - 1\}$. $J = J(n)$ is a sequence with $J \to \infty$ as $n \to \infty$. We take a wavelet $\psi(x)$ which has compact support on $[-A, A]$ with $A > 1$, and $\psi(x)$ satisfies the conditions (B1)-(B2) listed in the Appendix.

The wavelet coefficient of the volatility function $\sigma^2(x)$ is defined as follows. Let

$$\beta_{J,k} = \int_a^b \sigma^2(x)\psi_{J,k}^{\rm per}(x){\rm d}x,$$

where

$$\psi_{J,k}^{\rm per}(x) = \frac{1}{\sqrt{b-a}}\psi_{J,k}\Big(\frac{x-a}{b-a}\Big)$$

with $\psi_{J,k}(x) = 2^{J/2}\psi(2^J x - k)$ and $k \in I_J = \{0, 1, 2, \ldots, 2^J - 1\}$. It is useful to note that for a given $k \in I_J = \{0, 1, 2, \ldots, 2^J - 1\}$, $\beta_{J,k}$ has relatively large absolute values at fine resolution level J when $a + (k/2^J)(b - a)$ is near to the change points, while $\beta_{J,k}$ has relatively small absolute values at fine resolution level J when $a + (k/2^J)(b - a)$ shifts away from the change points. These properties are important for designing test for change points detection.

§3. Test Statistics and Their Asymptotic Properties

In model (2.1), $\sigma^2(x)$ can be rewritten as

$$\sigma^2(x) = \mathsf{E}(Y^2|X=x) - \mathsf{E}^2(Y|X=x) \triangleq m(x) - T^2(x). \tag{3.1}$$

From the Assumption (A4) in the Appendix, $T(x)$ is smooth on $[a,b]$. Hence, $\sigma^2(x)$ and $m(x)$ have the same change points on $[a,b]$. Based on (3.1), (2.2) can be rewritten as

$$m(x) = \widetilde{C}(x) + D(x), \tag{3.2}$$

where $\widetilde{C}(x) = C(x) + T^2(x)$.

Let $K(x)$ be a kernel function with bounded support $[-C, C]$ for some constant $C > 0$. A nonparametric estimator of $m(x)$ is given by

$$\widehat{m}(x) = \frac{\sum\limits_{i=1}^{n} K_h(X_i - x)Y_i^2}{\sum\limits_{i=1}^{n} K_h(X_i - x)}, \tag{3.3}$$

where $h = h_n$ is a sequence of bandwidths with $h \to 0$ and $nh \to \infty$ as $n \to \infty$. $\widehat{m}(x)$ can be kernel estimation (see Nadaraya, 1964), local linear smoother (see Fan and Gijbels, 1996), or other nonparametric estimators.

Based on $\widehat{m}(x)$, we propose an estimator of wavelet coefficient of the $m(x)$ as

$$U_J(k) = \int_a^b \psi_{J,k}^{\rm per}(x) \frac{\sum\limits_{i=1}^{n} K_h(X_i - x)Y_i^2}{\sum\limits_{i=1}^{n} K_h(X_i - x)} \,\mathrm{d}x. \tag{3.4}$$

This estimator is an integral estimator. An alternative estimator, which is simple in computation, is the following descretized version of $U_J(k)$,

$$W_J(k) = \frac{b-a}{N} \sum_{i=1}^{N} \psi_{J,k}^{\rm per}(w_i) \frac{\sum\limits_{j=1}^{n} K_h(X_j - w_i)Y_j^2}{\sum\limits_{j=1}^{n} K_h(X_j - w_i)}, \tag{3.5}$$

where $N \to \infty$, $w_i = a + (i/N)(b-a)$.

For our discussion, the following assumptions are needed.

(C1) $\lim\limits_{n\to\infty} 2^{2J}(\log n)^3/n = 0$, $\lim\limits_{n\to\infty} 2^{5J}/n = \infty$, $\lim\limits_{n\to\infty} 2^J h \log n = 0$.

(C2) $\lim\limits_{n\to\infty} n2^J/(Nh)^2 = 0$.

Theorem 3.1 Suppose that the assumptions (A1)-(A5) in the Appendix hold. Let t_i, $i = 1, 2, \ldots, p$ be p jump points of $\sigma^2(x)$, and the corresponding jump sizes be denoted as d_i, $i = 1, 2, \ldots, p$.

(a) If (C1) is satisfied, then for any $k \in I(t_i, 2^{-J}(b-a))$, we have

$$U_J(k) = 2^{-J/2}(b-a)^{1/2} d_i \int_1^A \psi(x)\mathrm{d}x + O_p(n^{-1/2}), \tag{3.6}$$

and for any $k \notin \bigcup_{i=1}^{p} I(t_i, 2^{-J}(b-a))$, we have

$$U_J(k) = O_p(n^{-1/2}). \tag{3.7}$$

(b) If (C1) and (C2) are satisfied, (3.6) and (3.7) hold for the discretized estimator $W_J(k)$.

Theorem 3.1 implies that at fine resolution level J, the wavelet coefficients nearby the jump points are significantly larger than those that are a little farther away from the jumps. Therefore, the points $a + (k/2^J)(b-a)$, corresponding to large empirical wavelet coefficients $U_J(k)$, may be potential change points. We want to test whether or not $\sigma^2(x)$ jumps at these potential change points. The following results play a role in detecting the jump points of the volatility $\sigma^2(x)$.

Theorem 3.2 Suppose that the assumptions (A1)-(A5) in the Appendix hold. When there is no change point in $\sigma^2(x)$, we have

(a) if (C1) is satisfied, then for any $k \in I_J$,

$$\sqrt{n}U_J(k) \longrightarrow N(0, \sigma_u^2(x_0))$$

in distribution, where $k/2^J \to \theta \in (0,1)$, $x_0 = a + \theta(b-a)$ and

$$\sigma_u^2(x) = \frac{\int_{-A}^{A} \psi^2(z)\mathrm{d}z}{f(x)} \{\sigma^4(x)\mathsf{E}(\varepsilon^2-1)^2 + 4T^2(x)\sigma^2(x) + 4T(x)\sigma^3(x)\mathsf{E}\varepsilon^3\};$$

(b) if (C1) and (C2) are satisfied, then for any $k \in I_J$,

$$\sqrt{n}W_J(k) \longrightarrow N(0, \sigma_u^2(x_0))$$

in distribution.

Note that $\sigma_u^2(x)$ in Theorem 3.2 includes unknown parameters $\mathsf{E}\varepsilon^r$, $r = 3, 4$. It is difficult to estimate these parameters in practical applications. Therefore, we will introduce another version of $\sigma_u^2(x)$. Denote $Z_i = Y_i^2$, and then model (2.1) can be rewritten as $Z_i = \mu(X_i) + \sigma_Z(X_i)\widetilde{\varepsilon}_i$, where $\mu(x) = \mathsf{E}(Z|X=x)$ and $\sigma_Z^2(x) = \mathsf{Var}\,(Z|X=x)$. $\{\widetilde{\varepsilon}_i, i=1,2,\ldots\}$ is a sequence of i.i.d. random variables with $\mathsf{E}(\widetilde{\varepsilon}_i)=0$, $\mathsf{Var}\,(\widetilde{\varepsilon}_i)=1$. When $\sigma^2(x)$ is smooth, $\mu(x)$ is also smooth by (3.1). Denote

$$\tau^2(x) = \frac{\sigma_Z^2(x)}{f(x)}\int_{-A}^{A}\psi^2(z)\mathrm{d}z.$$

Based on the fact that $\sigma_Z^2(X) = \mathsf{Var}\,(Y^2|X) = \mathsf{E}((Y^2 - \mathsf{E}(Y^2|X))^2|X)$, and some calculations, we obtain that $\sigma_u^2(x) = \tau^2(x)$.

Note that $\tau^2(x)$ includes the unknown density function $f(x)$ and heteroscedastic variance $\sigma_Z^2(x)$. There exist various methods for estimating these quantities. For example, we can adopt the kernel method for estimating $f(x)$ as follows

$$f_n(x) = \frac{1}{nh}\sum_{i=1}^{n}K\Big(\frac{X_i - x}{h}\Big),$$

where h is a bandwidth and $K(x)$ is a kernel function. The variance $\sigma_Z^2(x)$ can be estimated by

$$\widehat{\sigma}_Z^2(x) = \frac{\sum\limits_{i=1}^{n}K_h(X_i - x)(Y_i^2 - \widehat{\mu}(X_i))^2}{\sum\limits_{i=1}^{n}K_h(X_i - x)},$$

where $\widehat{\mu}(x)$ is any consistent estimator for $\mu(x)$.

Suppose that one wants to test whether or not a particular point t is a change point. Denote $d = \sigma^2(t+) - \sigma^2(t-)$. To test whether $\sigma^2(x)$ jumps at t amounts to test:

$$H_0: d = 0 \qquad \text{vs} \qquad H_1: d \neq 0.$$

To develop a test statistic for testing H_0 against H_1, define

$$T^{(U)}(k) = \frac{U_J(k)}{\widehat{\tau}(t)} \qquad \text{or} \qquad T^{(W)}(k) = \frac{W_J(k)}{\widehat{\tau}(t)},$$

where $\widehat{\tau}(t)$ is any estimator for $\tau(t)$, k is chosen sufficiently close to $2^J(t-a)/(b-a)$. Theorem 3.2 gives out approximate critical values for the tests under the null hypothesis H_0. The decision rule is to reject H_0 if $\sqrt{n}|T^{(U)}(k)|$ or $\sqrt{n}|T^{(W)}(k)|$ is larger than $C_{1-\alpha}$, the $1-\alpha/2$ quartile of the $N(0,1)$ distribution.

§4. Estimation of Jump Sizes and Change Points

We first consider the case that $\sigma^2(x)$ has only one change point t_0 in $[a,b]$, the corresponding jump size is denoted by d_0. We propose the following estimators of t_0 and

d_0:

$$t_0^U = a + \frac{k_1^U(b-a)}{2^J} \quad \text{or} \quad t_0^W = a + \frac{k_1^W(b-a)}{2^J},$$

$$d_0^U = \frac{2^{J/2}U_J(k_1^U)}{(b-a)^{1/2}\displaystyle\int_1^A \psi(x)\mathrm{d}x} \quad \text{or} \quad d_0^W = \frac{2^{J/2}W_J(k_1^W)}{(b-a)^{1/2}\displaystyle\int_1^A \psi(x)\mathrm{d}x},$$

where

$$k_1^U = \arg\max_k |T^{(U)}(k)|, \qquad k_1^W = \arg\max_k |T^{(W)}(k)|.$$

The following theorems establish the convergence rates of the proposed estimators.

Theorem 4.1 Under the assumptions (A1)-(A5) in the Appendix, if (C1) and (C2) are satisfied, we have

$$t_0^U = t_0 + O_p(2^{-J}), \qquad t_0^W = t_0 + O_p(2^{-J}).$$

Theorem 4.2 Under the assumptions (A1)-(A5) in the Appendix, if (C1) and (C2) are satisfied, we have

$$d_0^U = d_0 + O_p((2^{-J}n)^{-1/2}), \qquad d_0^W = d_0 + O_p((2^{-J}n)^{-1/2}).$$

Next, we consider the case that $\sigma^2(x)$ has p jump points, where p is unknown. Theorem 3.2 implies that $\sqrt{n}|T^{(U)}(k)| \leqslant C_{1-\alpha}$ with approximate probability $1-\alpha$ when $\sigma^2(x)$ is smooth. So we can take $C_{1-\alpha}$ as a threshold and use it to determine which points in $[a,b]$ are change points.

From Theorem 3.1, there may be several large wavelet coefficients near a single jump, so how to combine those wavelet coefficients which have large absolute values into suitable subgroups is a key point. Each subgroup should be regarded as the result from a single jump. Our analysis uses the ρ-division method introduced by Li and Xie (1999). Let $\rho = 2A$. Denote

$$E^{(U)} = \{k : |T^{(U)}(k)| \geqslant C_{1-\alpha}n^{-1/2}\} \triangleq \{k_1, k_2, \ldots, k_m\},$$

where m is a finite number. Put $m_1 = \max\{i : 1 \leqslant i \leqslant m, k_i \leqslant k_1+\rho\}$; if $m_1 < m$, then put $m_2 = \max\{i : m_1 \leqslant i \leqslant m, k_i \leqslant k_{m_1+1} + \rho\}$; if $m_2 < m$, then define m_3 in a similar way, and so no. At the end, we can get a series of integers $\{m_j : 1 \leqslant m_1 < m_2 < \cdots < m_q = m\}$.

Let $E_1^{(U)} = \{k_i : 1 \leqslant i \leqslant m_1\}, E_2^{(U)} = \{k_i : m_1 < i \leqslant m_2\}, \ldots, E_q^{(U)} = \{k_i : m_{q-1} < i \leqslant m_q\}$, then $E^{(U)} = \bigcup\limits_{j=1}^{q} E_j^{(U)}$. The estimator of the number p of change points is constructed as follows

$$\widehat{p}^U = \begin{cases} q, & E^{(U)} \neq \varnothing, \\ 0, & E^{(U)} = \varnothing. \end{cases}$$

The estimators of locations t_i and jump sizes d_i for change points are constructed as follows

$$t_i^U = a + \frac{k_i(b-a)}{2^J}, \qquad d_i^U = \frac{2^{J/2}U_J(k_i)}{(b-a)^{1/2}\displaystyle\int_1^A \psi(x)\mathrm{d}x}, \qquad i=1,2,\ldots,\widehat{p}^U,$$

where

$$k_i = \arg\max_k\{|T^{(U)}(k)|, k\in E_i^{(U)}\}, \qquad i=1,2,\ldots,\widehat{p}^U.$$

These procedures also hold for the test statistic $T^{(W)}(k)$. The corresponding estimators which are constructed by $T^{(W)}(k)$ in a similar way are denoted as $\widehat{p}^W, t_i^W, d_i^W$, $i=1,2,\ldots,\widehat{p}^W$.

The following theorem establishes the convergence rates of the proposed estimators.

Theorem 4.3 Under the assumptions (A1)-(A5) in the Appendix, if (C1) and (C2) are satisfied, we have

$$\begin{aligned}
&\lim_{n\to\infty}\mathsf{P}(\widehat{p}^U=p)=1, && \lim_{n\to\infty}\mathsf{P}(\widehat{p}^W=p)=1,\\
&t_i^U = t_i + O_p(2^{-J}), && d_i^U = d_i + O_p((2^{-J}n)^{-1/2}), && i=1,2,\ldots,\widehat{p}^U,\\
&t_i^W = t_i + O_p(2^{-J}), && d_i^W = d_i + O_p((2^{-J}n)^{-1/2}), && i=1,2,\ldots,\widehat{p}^W.
\end{aligned}$$

§5. Simulations

In this section, some simulations are carried out to verify the theoretical findings of the proposed change point analysis methods. Throughout our simulation studies, $m(x)$, $f(x)$ and $\sigma_Z^2(x)$ are estimated by kernel estimation. The Epanechnikov kernel $K(x) = 3(1-x^2/5)I(x^2\leqslant 5)/4\sqrt{5}$ is used to estimate $m(x)$, and the Gaussian kernel $K(x) = (1/\sqrt{2\pi})e^{-x^2/2}$ is used to estimate $f(x)$ and $\sigma_Z^2(x)$. The bandwidths in the estimators of $f(x)$ and $\sigma_Z^2(x)$ are 0.1. The wavelet is chosen to be

$$\psi(x) = 5(x-1)^4 I(1\leqslant x\leqslant 2) + \Big\{\frac{20}{3}(x+1)^3 + 2(x+1)^2\Big\}I(-2\leqslant x\leqslant -1).$$

Example 1 This example concerns the detection and estimation of a unique jump point in volatility. The simulation model is as follows

$$Y_i = T(X_i) + \sigma(X_i)\varepsilon_i, \tag{5.1}$$

where $\{\varepsilon_i\}$ is a sequence of i.i.d. random variables with the standard normal distribution. The volatility function is

$$\sigma(x) = \begin{cases}(x+4)^{-1}, & x<t,\\ (x+4)^{-1}+d, & x\geqslant t,\end{cases} \tag{5.2}$$

and $T(x)=0.2x$. When $\sigma(x)$ has a jump size d at point t, the corresponding jump size in $\sigma^2(x)$ is denoted as $d_0=\sigma^2(t+)-\sigma^2(t-)$. The following designs for X_i are considered:

(a) $\{X_i\}$ is generated from the uniform distribution on $[0,1]$.

(b) $X_{t+1}=0.4+0.3X_t+\xi_{t+1}+0.2\xi_t$ is truncated in $[0,1]$, where $\xi_t\sim N(0,0.25^2)$, $X_1\sim U[0,1]$.

The aim in this example is to study the sizes and powers of the test statistic $T^{(W)}(k)$ at known change point t, and to estimate the location t and the jump size d_0 in $\sigma^2(x)$. When $d_0=0$, without loss of generality, we only present the sizes of the test statistic $T^{(W)}(k)$ at $t=0.5$.

We generate 500 series for two lengths ($n=128$ and $n=256$) from this model. Test size is set to 0.01, 0.05, 0.10. Estimates of change points are calculated for $d=0.5,1,1.5$ at different locations $t=0.3,0.5,0.7$, and the true values of the corresponding jump size d_0 in $\sigma^2(x)$ are to see Tables 1 and 2. The bandwidths in the wavelet coefficients are $h=0.05$ ($n=256$) and $h=0.06$ ($n=128$). The resolution level in simulations is $J=4$.

The simulated results are listed in Table 1 and 2. The design for X_i in Table 1 is from (a), and the design for X_i in Table 2 is from (b). From Table 1 and 2, the location of change point seems to have an effect on the power of the tests, and the bias and variance of the estimates decrease as the sample size increases.

Example 2 This example is used to make comparison with Example 5.1. We consider model (5.1) with AR(1) process error term. Assume that error $\{\varepsilon_i\}$ is a sequence of stationary random variables satisfying the AR(1) model

$$\varepsilon_i=(0.4\varepsilon_{i-1}+e_i)/\sigma,\qquad \sigma=1.09,\ \varepsilon_1\sim N(0,1),\ e_i\sim N(0,1).$$

$\{X_i\}$ is a sequence of i.i.d. uniform random variables on $[0,1]$. The other parameters are all the same as in Example 1. The simulation results are listed in Table 3.

The results for this example are similar to those listed in Tables 1 and 2, which imply that our methods may be used in the case of correlated errors.

Example 3 We consider model (5.1) with multiple jumps in volatility. The error term is the same as in Example 1. Assume that the regression function is $T(x)=0.1x$, and the volatility is

$$\sigma(x)=\begin{cases}0.4-0.1\sqrt{x+4}, & x<t_1,\\ 0.4-0.1\sqrt{x+4}+d_1, & t_1\leqslant x<t_2,\\ 0.4-0.1\sqrt{x+4}+d_1-d_2, & x\geqslant t_2,\end{cases}\tag{5.3}$$

Table 1 Sizes and powers of statistic $T^{(W)}(k)$ for Example 1 with uniform design (a) for X_i for various sample sizes n

			Test			Estimation			
t	d	d_0	$\alpha=10\%$	5%	1%	$\widehat{t}$	se	$\widehat{d}_0$	se
$n=256$									
	0	0	13%	6.8%	2%				
0.3	0.5	0.48	99.8%	99%	96.8%	0.314	0.060	0.493	0.100
	1	1.47	100%	100%	99%	0.308	0.046	1.488	0.290
	1.5	2.94	100%	100%	99%	0.309	0.057	3.007	0.582
0.5	0.5	0.47	100%	99.8%	98.8%	0.515	0.058	0.484	0.098
	1	1.44	100%	100%	99.4%	0.511	0.049	1.490	0.288
	1.5	2.92	100%	100%	99.2%	0.508	0.047	2.986	0.567
0.7	0.5	0.46	100%	100%	98.4%	0.705	0.041	0.482	0.107
	1	1.43	100%	100%	99.8%	0.706	0.040	1.466	0.306
	1.5	2.89	100%	100%	99.8%	0.706	0.036	2.968	0.583
$n=128$									
	0	0	13%	8.8%	1.6%				
0.3	0.5	0.48	98.8%	96.6%	82.2%	0.362	0.142	0.501	0.137
	1	1.47	98.4%	96%	86.2%	0.344	0.119	1.451	0.382
	1.5	2.94	98.8%	96%	87.6%	0.348	0.126	2.976	0.734
0.5	0.5	0.47	99.4%	98.2%	86.2%	0.529	0.080	0.485	0.132
	1	1.44	99.8%	98.4%	92.2%	0.527	0.078	1.438	0.361
	1.5	2.92	99.2%	98.6%	92.4%	0.527	0.076	3.032	0.850
0.7	0.5	0.46	99.6%	98.8%	89.4%	0.712	0.045	0.477	0.146
	1	1.43	100%	99%	95.8%	0.710	0.043	1.416	0.369
	1.5	2.89	100%	99.2%	96.4%	0.711	0.043	2.919	0.812

$\{X_i\}$ is a sequence of i.i.d. uniform random variables on $[0,1]$. The volatility has two different change points at t_1 and t_2 ($t_1 < t_2$). The jump sizes for two change points are d_1 and d_2 ($d_1 \geqslant 0$, $d_2 \geqslant 0$). The corresponding jump sizes in $\sigma^2(x)$ are denoted as: $d_{10}=\sigma^2(t_1+)-\sigma^2(t_1-)$, $d_{20}=\sigma^2(t_2+)-\sigma^2(t_2-)$.

The aim of this example is to study the sizes and powers of the test $T^{(W)}(k)$ at known change points t_1 and t_2, and to estimate the locations t_1,t_2 and jump sizes d_{10},d_{20} in $\sigma^2(x)$. Without loss of generality, let $|d_{10}|\geqslant|d_{20}|$ in this example.

In this example, we set $t_1=0.3$, $t_2=0.7$. The jump sizes in $\sigma(x)$ are set to $d_1=$

Table 2　Sizes and powers of statistic $T^{(W)}(k)$ for Example 1 with mixing design (b) for X_i for various sample sizes n

t	d	d_0	Test $\alpha=10\%$	5%	1%	Estimation $\widehat{t}$	se	$\widehat{d}_0$	se
$n=256$									
	0	0	12.2%	5.2%	2.2%				
0.3	0.5	0.48	99.6%	98.8%	94.8%	0.304	0.067	0.493	0.100
	1	1.47	100%	99.6%	96.8%	0.301	0.048	1.472	0.285
	1.5	2.94	99.8%	99.4%	98.6%	0.303	0.052	2.994	0.524
0.5	0.5	0.47	100%	100%	99.8%	0.508	0.053	0.480	0.090
	1	1.44	100%	100%	99.8%	0.507	0.042	1.464	0.257
	1.5	2.92	100%	100%	100%	0.504	0.033	2.931	0.512
0.7	0.5	0.46	100%	100%	100%	0.712	0.040	0.482	0.103
	1	1.43	100%	100%	100%	0.710	0.038	1.477	0.293
	1.5	2.89	100%	100%	100%	0.709	0.038	2.930	0.584
$n=128$									
	0	0	15%	8.2%	1.2%				
0.3	0.5	0.48	95.8%	91.4%	73%	0.333	0.117	0.488	0.126
	1	1.47	97.6%	94.6%	80.6%	0.322	0.105	1.463	0.356
	1.5	2.94	96.4%	92.8%	79.8%	0.323	0.112	2.890	0.704
0.5	0.5	0.47	99.8%	99.4%	92.4%	0.522	0.072	0.481	0.121
	1	1.44	99.8%	99.6%	97%	0.514	0.060	1.456	0.336
	1.5	2.92	100%	100%	98%	0.519	0.066	2.893	0.652
0.7	0.5	0.46	99.2%	99.2%	95.6%	0.719	0.048	0.488	0.148
	1	1.43	100%	100%	98.2%	0.719	0.045	1.408	0.392
	1.5	2.89	100%	100%	99%	0.717	0.064	2.891	0.800

$0, 0.4, 0.8, 1.2, 1.6$, $d_2 = 0, 0.4, 0.8$. We generate 500 series for two lengths ($n = 128$ and $n = 256$) from this model. Test size is set to 0.01, 0.05, 0.10. The bandwidths in the wavelet coefficients are $h = 0.05$ ($n = 256$) and $h = 0.06$ ($n = 128$). The resolution level in simulations is $J = 4$.

The simulation results for Example 3 are listed in Table 4. As the sample size increases, the bias and variance of the estimates decrease. The sizes and powers of the test statistic $T^{(W)}(k)$ at $t = 0.3$ and the estimates of $t = 0.3$ are similar to those in the Table 1, 2 and 3. As to the second change point t_2, with the increasing difference of jump sizes

Table 3 Sizes and powers of statistic $T^{(W)}(k)$ for Example 2 with correlate error term for various sample sizes n

			Test			Estimation			
t	d	d_0	$\alpha=10\%$	5%	1%	$\widehat{t}$	se	$\widehat{d}_0$	se
$n=256$									
	0	0	14%	7.6%	1.4%				
0.3	0.5	0.48	99.8%	99.6%	97.4%	0.318	0.076	0.493	0.112
	1	1.47	100%	100%	98.8%	0.314	0.064	1.497	0.324
	1.5	2.94	100%	99.8%	98.6%	0.309	0.041	2.983	0.612
0.5	0.5	0.47	100%	100%	99%	0.508	0.049	0.497	0.104
	1	1.44	99.8%	99.8%	99.4%	0.509	0.045	1.502	0.310
	1.5	2.92	100%	99.8%	99.2%	0.509	0.047	2.990	0.620
0.7	0.5	0.46	100%	99.8%	99.6%	0.704	0.040	0.480	0.109
	1	1.43	100%	99.6%	99.4%	0.702	0.036	1.452	0.298
	1.5	2.89	100%	100%	99.8%	0.703	0.036	2.975	0.566
$n=128$									
	0	0	12.8%	6.8%	0.8%				
0.3	0.5	0.48	98.4%	95%	80.6%	0.351	0.121	0.504	0.148
	1	1.47	98.6%	96.8%	87.4%	0.338	0.105	1.475	0.406
	1.5	2.94	98.2%	97%	88.8%	0.342	0.112	2.995	0.783
0.5	0.5	0.47	99%	96.6%	87%	0.532	0.080	0.492	0.140
	1	1.44	99.8%	99.2%	93.6%	0.522	0.064	1.459	0.393
	1.5	2.92	100%	99.8%	95.4%	0.530	0.078	2.979	0.831
0.7	0.5	0.46	99.8%	98%	91%	0.713	0.047	0.490	0.147
	1	1.43	99.8%	99.4%	94.8%	0.713	0.044	1.404	0.427
	1.5	2.89	99.6%	99%	95.6%	0.711	0.045	2.884	0.805

between t_1 and t_2 ($|d_{10}| \geqslant |d_{20}|$), the bias of the estimates for jump sizes of the second change point t_2 becomes larger.

§6. Appendix

For our discussion, the following assumptions are needed for model (2.1):

(A1) $\{X_i, i=1,2,\ldots\}$ is a sequence of stationary and strongly mixing random vectors, with mixing coefficient $\alpha(\cdot)$. Let $\{l_n\}$ be a sequence of positive integers such that

Table 4 Sizes and powers of statistic $T^{(W)}(k)$ for Example 3 with multiple change points for various sample sizes n

		Test 1			Test 2			Estimation 1		Estimation 2	
d_{10}	d_{20}	$\alpha=10\%$	5%	1%	$\alpha=10\%$	5%	1%	$\hat{t}_1$(se)	$\hat{d}_{10}$(se)	$\hat{t}_2$(se)	$\hat{d}_{20}$(se)
$n=256$, $t_1=0.3$, $t_2=0.7$											
0	0	10%	5%	1.8%	8.8%	5%	0.8%				
0.31	0	100%	99.8%	97%	10.4%	5.4%	1.2%	0.293(0.048)	0.311(0.083)		
0.31	0.31	100%	99.4%	97.2%	100%	99.6%	96.6%	0.309(0.044)	0.317(0.068)	0.692(0.042)	0.316(0.067)
0.95	0	100%	99.8%	98.8%	13%	5.8%	0.8%	0.291(0.047)	0.925(0.240)		
0.95	0.31	100%	99.8%	98%	94%	87%	62.6%	0.306(0.039)	0.962(0.198)	0.686(0.062)	0.687(0.198)
0.95	0.93	100%	100%	98.8%	100%	100%	99.4%	0.303(0.039)	0.958(0.193)	0.692(0.038)	0.944(0.192)
1.90	0.31	100%	100%	99.6%	73.8%	61.6%	36.8%	0.306(0.037)	1.957(0.413)	0.681(0.075)	1.170(0.410)
1.90	0.93	100%	99.8%	97.6%	99.8%	99.2%	94.4%	0.307(0.040)	1.948(0.400)	0.695(0.046)	1.706(0.405)
3.18	0.93	100%	100%	99.4%	97%	91.8%	74.8%	0.309(0.041)	3.203(0.665)	0.689(0.058)	2.392(0.634)
$n=128$, $t_1=0.3$, $t_2=0.7$											
0	0	16.2%	10%	2.4%	14.2%	8.2%	3%				
0.31	0	98.8%	96%	82%	12.8%	5.4%	0.6%	0.312(0.063)	0.309(0.098)		
0.31	0.31	97.4%	95.8%	84.6%	98.2%	95.4%	81.8%	0.317(0.052)	0.326(0.094)	0.680(0.055)	0.322(0.096)
0.95	0	98.4%	96.8%	89.2%	11.6%	6.6%	1.2%	0.303(0.060)	0.917(0.281)		
0.95	0.31	99.6%	97.4%	86%	75.2%	61.8%	33.4%	0.311(0.052)	0.947(0.245)	0.678(0.073)	0.679(0.248)
0.95	0.93	99.4%	98%	89.2%	98.8%	96.2%	86.6%	0.317(0.050)	0.948(0.234)	0.686(0.051)	0.953(0.240)
1.90	0.31	98.6%	96.6%	88.2%	52.8%	40.2%	16.6%	0.311(0.052)	1.877(0.495)	0.673(0.091)	1.206(0.478)
1.90	0.93	98.8%	98%	89.8%	93.8%	89%	71.6%	0.313(0.046)	1.892(0.485)	0.682(0.061)	1.644(0.469)
3.18	0.93	99.2%	97.8%	88.2%	83.8%	72.2%	46.2%	0.315(0.051)	3.146(0.810)	0.677(0.068)	2.391(0.840)

$l_n \to \infty$ and $l_n = o((n2^{-J})^{1/2})$.

(i) $(n2^J)^{1/2}\alpha(l_n) \to \infty$ as $n \to \infty$.

(ii) There exist $e > 2$ and $a > 1 - 2/e$ such that $\sum\limits_{l=1}^{\infty} l^a(\alpha(l))^{1-2/e} < \infty$.

(A2) The density $f(x)$ of X_1 is bounded away from zero and infinity on some open subset U, where U is a non-empty open neighborhood of the origin of R and $[a,b] \subset U$.

(A3) The conditional density $f(x_i|x)$ of X_i, given $X_1 = x$, is also bounded away from zero and infinity on U.

(A4) $f(x)$ is a twice bounded derivative function, and $T(x)$ and $C(x)$ are third continuously differentiable on U.

(A5) Let $\{\varepsilon_i, i = 1, 2, \ldots\}$ be a sequence of i.i.d. random variables and for each i, ε_i is independent of $\{(X_j, Y_{j-1}), j \leqslant i\}$. Also, $\mathsf{E}|X|^\lambda < \infty$, $\mathsf{E}|\varepsilon|^\lambda < \infty$ and $\mathsf{E}|Y|^\lambda < \infty$ for some $\lambda > 4$.

These assumptions are satisfied by most time series models. (A2) and (A3) are necessary for the kernel estimation with dependent data. (A1) and (A5) are to simplify proofs. (A4) is necessary for kernel smoother. For a detailed discussion of these assumptions, see Chen et al. (2008), Zhou et al. (2010) and Xia (1998).

The wavelet $\psi(x)$ in this paper satisfies the following properties:

(B1) $\psi(x)$ has finite support, say, $[-A, A]$, $A > 1$, and $\psi(x) = 0$, $x \in [-1, 1]$. $\psi(x)$ has bounded variation on $[-A, A]$.

(B2) $\psi(x)$ is a twice continuously differentiable function on $[-A, A]$. Furthermore

$$\int_{-A}^{A} \psi(x)\mathrm{d}x = 0, \quad \int_{-A}^{A} x\psi(x)\mathrm{d}x = 0, \quad \int_{1}^{A} \psi(x)\mathrm{d}x \neq 0, \quad \int_{1}^{A} x\psi(x)\mathrm{d}x \neq 0.$$

$$0 < \Big|\int_{y}^{A} \psi(x)\mathrm{d}x\Big| < \Big|\int_{1}^{A} \psi(x)\mathrm{d}x\Big|, \qquad 0 < \Big|\int_{-A}^{-y} \psi(x)\mathrm{d}x\Big| < \Big|\int_{-A}^{-1} \psi(x)\mathrm{d}x\Big|,$$

for all $1 < y < A$.

We only provide proofs for the theorems based on the integral estimator $U_J(k)$. The proofs based on the discretized estimator $W_J(k)$ can be took in the similar way. The theorems are proved based on kernel estimation. Let C denote a generic constant which may vary depending on the context.

Lemma 6.1 Under the assumptions (A1)-(A5). $K(x)$ is a continuously differentiable kernel function with finite support $[-C, C]$, and $\displaystyle\int K(x)\mathrm{d}x = 1$, $\displaystyle\int xK(x)\mathrm{d}x = 0$. Let $h \to 0$, $nh \to \infty$ as $n \to \infty$, then for any positive integer i, we have

$$\sum_{t=1}^{n} K_h(X_t - x)(X_t - x)^i = nh^{i+1}\Phi_i f(x) + nh^{i+2}\Phi_{i+1} f'(x) + O(nh^{i+1}c_n) \quad \text{a.s.},$$

$$\sum_{t=1}^{n} K_h(X_t - x)\varepsilon_t = O_p((nh)^{1/2}),$$

$$\sum_{t=1}^{n} K_h(X_t - x)\Big(\frac{X_t - x}{h}\Big)\varepsilon_t = O_p((nh)^{1/2}),$$

uniformly for $x \in [a, b]$, where

$$K_h(x) = K\Big(\frac{x}{h}\Big), \qquad c_n = h^2 + \Big(\frac{\log n}{nh}\Big)^{1/2}, \qquad \Phi_i = \int x^i K(x)\mathrm{d}x.$$

Lemma 6.2 Assume that $\psi(x)$ satisfies the assumptions (B1)-(B2), and that $C(x)$ is a continuously differentiable function in the order of two. Then, uniformly for $k \in I_J$,

$$\int_a^b \psi_{J,k}^{\mathrm{per}}(x) \frac{\sum\limits_{j=1}^{n} K_h(X_j - x)(C(X_j) - C(x))}{\sum\limits_{j=1}^{n} K_h(X_j - x)}\mathrm{d}x = O_p(2^{-J/2}hc_n),$$

$$\int_a^b \psi_{J,k}^{\mathrm{per}}(x)C(x)\mathrm{d}x = O(2^{-5J/2}).$$

Lemma 6.3 Assume that $\psi(x)$ satisfies the assumptions (B1)-(B2), and that $K(x)$ is a kernel function with finite support $[-C,C]$. Let $h\to 0$ as $n\to\infty$. We have

$$\int_a^b \psi^{\rm per}_{J,k}(x)\frac{\sum\limits_{j=1}^n K_h(X_j-x)\Big(\sum\limits_{i=1}^p d_i I(t_i\leqslant X_j\leqslant b)\Big)}{\sum\limits_{j=1}^n K_h(X_j-x)}{\rm d}x=2^{-J/2}(b-a)^{1/2}d_i\int_1^A\psi(x){\rm d}x,$$

uniformly for $k\in I(t_i,2^{-J}(b-a))$, and

$$\int_a^b \psi^{\rm per}_{J,k}(x)\frac{\sum\limits_{j=1}^n K_h(X_j-x)\Big(\sum\limits_{i=1}^p d_i I(t_i\leqslant X_j\leqslant b)\Big)}{\sum\limits_{j=1}^n K_h(X_j-x)}{\rm d}x=0,$$

uniformly for $k\notin\bigcup\limits_{i=1}^p I(t_i,2^{-J}A(b-a))$.

Lemma 6.1, Lemma 6.2 and Lemma 6.3 come from Chen et al. (2008).

Proof of Theorem 3.1 Note that $U_J(k)$ can be decomposed into three parts:

$$U_J(k)=U_J^{(1)}(k)+U_J^{(2)}(k)+U_J^{(3)}(k),$$

where

$$U_J^{(1)}(k)=\int_a^b \psi^{\rm per}_{J,k}(x)\frac{\sum\limits_{j=1}^n K_h(X_j-x)T^2(X_j)}{\sum\limits_{j=1}^n K_h(X_j-x)}{\rm d}x,$$

$$U_J^{(2)}(k)=\int_a^b \psi^{\rm per}_{J,k}(x)\frac{\sum\limits_{j=1}^n K_h(X_j-x)\sigma^2(X_j)\varepsilon_j^2}{\sum\limits_{j=1}^n K_h(X_j-x)}{\rm d}x,$$

$$U_J^{(3)}(k)=\int_a^b \psi^{\rm per}_{J,k}(x)\frac{\sum\limits_{j=1}^n K_h(X_j-x)2T(X_j)\sigma(X_j)\varepsilon_j}{\sum\limits_{j=1}^n K_h(X_j-x)}{\rm d}x.$$

We first consider $U_J^{(1)}(k)$. From Lemma 6.2 and assumption (C1), we have

$$U_J^{(1)}(k)=O_p(2^{-5J/2}+2^{-J/2}hc_n)=O_p(n^{-1/2}),$$

for all $k\in I_J$.

As to $U_J^{(3)}(k)$, based on (2.2), $\sigma(x)$ can be decomposed into $\sigma(x)=\widetilde{C}(x)+\widetilde{D}(x)$, where $\widetilde{C}(x)$ is twice continuously differentiable on (a,b), $\widetilde{D}(x)=\sum\limits_{l=1}^{p}\widetilde{d}_l I(t_l\leqslant x\leqslant b)$, $\widetilde{d}_l=\sigma(t_l+)-\sigma(t_l-)$. Then

$$\begin{aligned}U_J^{(3)}(k) &= \int_a^b \psi_{J,k}^{\mathrm{per}}(x)\frac{\sum\limits_{j=1}^{n}K_h(X_j-x)2T(X_j)\widetilde{C}(X_j)\varepsilon_j}{\sum\limits_{j=1}^{n}K_h(X_j-x)}\mathrm{d}x\\ &\quad+\int_a^b \psi_{J,k}^{\mathrm{per}}(x)\frac{\sum\limits_{j=1}^{n}K_h(X_j-x)2T(X_j)\widetilde{D}(X_j)\varepsilon_j}{\sum\limits_{j=1}^{n}K_h(X_j-x)}\mathrm{d}x\\ &= U_J^{(3)(C)}(k)+U_J^{(3)(D)}(k).\end{aligned}$$

From the same arguments of Lemma B.1 in Chen et al. (2008), we obtain

$$\frac{\sum\limits_{j=1}^{n}K_h(X_j-x)T(X_j)I(t_l\leqslant X_j\leqslant b)\varepsilon_j}{\sum\limits_{j=1}^{n}K_h(X_j-x)}=I(t_l\leqslant x\leqslant b)\frac{\sum\limits_{j=1}^{n}K_h(X_j-x)T(X_j)\varepsilon_j}{\sum\limits_{j=1}^{n}K_h(X_j-x)}.$$

According to the Lemma A.4 of Chen et al. (2008), we have

$$U_J^{(3)(C)}(k)=O_p(n^{-1/2}).$$

Therefore

$$\begin{aligned}U_J^{(3)}(k) &= 2\sum_{l=1}^{p}\widetilde{d}_l\int_a^b \psi_{J,k}^{\mathrm{per}}(x)I(t_l\leqslant x\leqslant b)\frac{\sum\limits_{j=1}^{n}K_h(X_j-x)T(X_j)\varepsilon_j}{\sum\limits_{j=1}^{n}K_h(X_j-x)}\mathrm{d}x+O_p(n^{-1/2})\\ &= 2\sum_{l=1}^{p}\widetilde{d}_l\int_{t_l}^b \psi_{J,k}^{\mathrm{per}}(x)\frac{\sum\limits_{j=1}^{n}K_h(X_j-x)T(X_j)\varepsilon_j}{\sum\limits_{j=1}^{n}K_h(X_j-x)}\mathrm{d}x+O_p(n^{-1/2}).\end{aligned}$$

In fact, Taylor's formula implies that

$$\begin{aligned}\sum_{j=1}^{n}K_h(X_j-x)T(X_j)\varepsilon_j &= T(x)\sum_{j=1}^{n}K_h(X_j-x)\varepsilon_j+T'(x)\sum_{j=1}^{n}K_h(X_j-x)(X_j-x)\varepsilon_j\\ &\quad+\frac{1}{2}\sum_{j=1}^{n}T''(\xi_j)K_h(X_j-x)(X_j-x)^2\varepsilon_j,\end{aligned}$$

where ξ_j lies between X_j and x. By Lemma 6.1, we have

$$\sup_{x\in\Lambda}\Big|\sum_{j=1}^{n} K_h(X_j-x)\varepsilon_j\Big| = O_p((nh)^{-1/2}),$$
$$\sup_{x\in\Lambda}\Big|\sum_{j=1}^{n} K_h(X_j-x)(X_j-x)\varepsilon_j\Big| = O_p((nh)^{-1/2}h),$$
$$\sup_{x\in\Lambda}\Big|\sum_{j=1}^{n} K_h(X_j-x)(X_j-x)^2\varepsilon_j\Big| = O_p((nh)^{-1/2}h^2),$$

where $\Lambda=[a-\Delta_0, b+\Delta_0]$ for some $\Delta_0>0$. Therefore

$$U_J^{(3)}(k) = O_p((2^J nh)^{-1/2}) + O_p(n^{-1/2}) = O_p(n^{-1/2}).$$

Now, we consider $U_J^{(2)}(k)$

$$\begin{aligned} U_J^{(2)}(k) &= \int_a^b \psi_{J,k}^{\mathrm{per}}(x)\frac{\sum_{j=1}^{n}K_h(X_j-x)\sigma^2(X_j)}{\sum_{j=1}^{n}K_h(X_j-x)}\mathrm{d}x+\int_a^b \psi_{J,k}^{\mathrm{per}}(x)\frac{\sum_{j=1}^{n}K_h(X_j-x)\sigma^2(X_j)(\varepsilon_j^2-1)}{\sum_{j=1}^{n}K_h(X_j-x)}\mathrm{d}x \\ &= \int_a^b \psi_{J,k}^{\mathrm{per}}(x)\frac{\sum_{j=1}^{n}K_h(X_j-x)\sigma^2(X_j)}{\sum_{j=1}^{n}K_h(X_j-x)}\mathrm{d}x + O_p(n^{-1/2}). \end{aligned}$$

From Lemma 6.2 and Lemma 6.3, we have

$$U_J^{(2)}(k) = 2^{-J/2}(b-a)^{1/2}d_l\int_1^A \psi(x)\mathrm{d}x + O_p(2^{-5J/2}+2^{-J/2}hc_n) + O_p(n^{-1/2}),$$

uniformly for $k\in I(t_l, 2^{-J}A(b-a))$, and

$$U_J^{(2)}(k) = O_p(2^{-5J/2}+2^{-J/2}hc_n) + O_p(n^{-1/2}) = O_p(n^{-1/2}),$$

for $k\notin \bigcup_{l=1}^{p} I(t_l, 2^{-J}A(b-a))$. This completes the proof of Theorem 3.1. □

Proof of Theorem 3.2 From the proof of Theorem 3.1, we have

$$\begin{aligned} U_J(k) &= \int_a^b \psi_{J,k}^{\mathrm{per}}(x)\frac{\sum_{j=1}^{n}K_h(X_j-x)\sigma^2(X_j)(\varepsilon_j^2-1)}{\sum_{j=1}^{n}K_h(X_j-x)}\mathrm{d}x \\ &\quad + \int_a^b \psi_{J,k}^{\mathrm{per}}(x)\frac{\sum_{j=1}^{n}K_h(X_j-x)2T(X_j)\sigma(X_j)\varepsilon_j}{\sum_{j=1}^{n}K_h(X_j-x)}\mathrm{d}x + o_p(n^{-1/2}). \end{aligned}$$

Similar to the proof of Lemma A.2 in Zhou et al. (2010), we obtain

$$\begin{aligned} U_J(k) &= \int_a^b \psi_{J,k}^{\text{per}}(x) \frac{\frac{1}{nh}\sum\limits_{j=1}^n K_h(X_j-x)\sigma^2(X_j)(\varepsilon_j^2-1)}{f(x)}\mathrm{d}x \\ &\quad + \int_a^b \psi_{J,k}^{\text{per}}(x) \frac{\frac{1}{nh}\sum\limits_{j=1}^n K_h(X_j-x)2T(X_j)\sigma(X_j)\varepsilon_j}{f(x)}\mathrm{d}x + o_p(n^{-1/2}) \\ &= \frac{1}{nh}\sum_{j=1}^n (\sigma^2(X_j)(\varepsilon_j^2-1)+2T(X_j)\sigma(X_j)\varepsilon_j)\int_a^b \psi_{J,k}^{\text{per}}(x)\frac{K_h(X_j-x)}{f(x)}\mathrm{d}x + o_p(n^{-1/2}) \\ &\triangleq \frac{1}{nh}\sum_{j=1}^n Z_j + o_p(n^{-1/2}), \end{aligned}$$

where $Z_j = Z_j^{(1)} + Z_j^{(2)}$ with

$$Z_j^{(1)} = \sigma^2(X_j)(\varepsilon_j^2-1)\int_a^b \psi_{J,k}^{\text{per}}(x)\frac{K_h(X_j-x)}{f(x)}\mathrm{d}x,$$
$$Z_j^{(2)} = 2T(X_j)\sigma(X_j)\varepsilon_j\int_a^b \psi_{J,k}^{\text{per}}(x)\frac{K_h(X_j-x)}{f(x)}\mathrm{d}x.$$

We first consider $\mathsf{Var}\,(Z_j)$. Obviously, $\{Z_j, j=1,2,\ldots,n\}$ is a stationary sequence, and $\mathsf{E}(Z_j)=0$, $j=1,2,\ldots,n$. Denote $\mathscr{F}_j=\sigma(X_j,X_{j-1},\ldots)$, then $\mathsf{E}(Z_j|\mathscr{F}_j)=0$. We know that

$$\mathsf{Var}\,(Z_j) = \mathsf{Var}\,(Z_j^{(1)}) + \mathsf{Var}\,(Z_j^{(2)}) + 2\mathsf{Cov}\,(Z_j^{(1)}, Z_j^{(2)}).$$

As to $\mathsf{Var}\,(Z_j^{(1)})$,

$$\mathsf{Var}\,(Z_j^{(1)}) = \mathsf{E}(\mathsf{Var}\,(Z_j^{(1)}|\mathscr{F}_j)) + \mathsf{Var}\,(\mathsf{E}(Z_j^{(1)}|\mathscr{F}_j)) = \mathsf{E}(\mathsf{Var}\,(Z_j^{(1)}|\mathscr{F}_j)),$$

where

$$\begin{aligned} \mathsf{Var}\,(Z_j^{(1)}|\mathscr{F}_j) &= \sigma^4(X_j)\Big(\int_a^b \psi_{J,k}^{\text{per}}(x)\frac{K_h(X_j-x)}{f(x)}\mathrm{d}x\Big)^2 \mathsf{E}((\varepsilon_j^2-1)^2|\mathscr{F}_j) \\ &= \frac{\sigma^4(X_j)}{f^2(X_j)}\Big(\int_a^b \psi_{J,k}^{\text{per}}(x)K_h(X_j-x)\mathrm{d}x\Big)^2 \mathsf{E}((\varepsilon_j^2-1)^2) + o(h^2). \end{aligned}$$

Hence

$$\begin{aligned} \mathsf{Var}\,(Z_j^{(1)}) &= \mathsf{E}(\mathsf{Var}\,(Z_j^{(1)}|\mathscr{F}_j)) \\ &= \mathsf{E}((\varepsilon_j^2-1)^2)\int_a^b \frac{\sigma^4(X_j)}{f^2(X_j)}\Big(\int_a^b \psi_{J,k}^{\text{per}}(x)K_h(X_j-x)\mathrm{d}x\Big)^2 f(X_j)\mathrm{d}X_j + o(h^2). \end{aligned}$$

Based on the fact that $|a+2^{-J}(b-a)(y+k)-x_0| = o(1)$ for any $y\in[-A,A]$, we can show that

$$\mathsf{Var}\,(Z_j^{(1)}) = h^2\mathsf{E}((\varepsilon^2-1)^2)\frac{\sigma^4(x_0)}{f(x_0)}\gamma(J,h) + o(h^2),$$

where

$$\gamma(J,h)=\int_a^b\Big(\frac{1}{h}\int_a^b\psi_{J,k}^{\mathrm{per}}(x)K_h(y-x)\mathrm{d}x\Big)^2\mathrm{d}y.$$

Using the condition $\lim\limits_{n\to\infty}2^Jh\log n=0$ in the assumption (C1), we can show that

$$\gamma(J,h)=\int_{-A}^{A}\psi^2(z)\mathrm{d}z+o(1).$$

Hence

$$\mathsf{Var}\,(Z_j^{(1)})=h^2\mathsf{E}((\varepsilon^2-1)^2)\frac{\sigma^4(x_0)}{f(x_0)}\int_{-A}^{A}\psi^2(z)\mathrm{d}z+o(1).$$

Similarly, we can show that

$$\mathsf{Var}\,(Z_j^{(2)})=\frac{4h^2\sigma^2(x_0)T^2(x_0)}{f(x_0)}\int_{-A}^{A}\psi^2(z)\mathrm{d}z+o(1),$$

$$\mathsf{Cov}\,(Z_j^{(1)},Z_j^{(2)})=\frac{4h^2\mathsf{E}(\varepsilon^3)\sigma^3(x_0)T(x_0)}{f(x_0)}\int_{-A}^{A}\psi^2(z)\mathrm{d}z+o(1).$$

Hence

$$\begin{aligned}\mathsf{Var}\,(Z_j)&=\frac{\int_{-A}^{A}\psi^2(z)\mathrm{d}zh^2}{f(x_0)}\{\sigma^4(x_0)\mathsf{E}(\varepsilon^2-1)^2+4T^2(x_0)\sigma^2(x_0)+4T(x_0)\sigma^3(x_0)\mathsf{E}\varepsilon^3\}\\&=\sigma_u^2(x_0)h^2.\end{aligned}$$

Denote

$$V_j=\frac{Z_j}{h\sigma_u(x_0)}=\frac{Z_j^{(1)}+Z_j^{(2)}}{h\sigma_u(x_0)}=V_j^{(1)}+V_j^{(2)}.$$

We have

$$U_J(k)=\frac{1}{nh}\sum_{j=1}^{n}Z_j+o_p(n^{-1/2})=\Big(\frac{\sigma_u^2(x_0)}{n}\Big)^{1/2}\frac{1}{\sqrt{n}}\sum_{j=1}^{n}V_j+o_p(n^{-1/2}).$$

Obviously, $\mathsf{E}(V_j)=0$ and $\mathsf{Var}\,(V_j)=1+o(1)\longrightarrow 1$.

Now, we show that

$$\sum_{j=2}^{n}|\mathsf{Cov}\,(V_1,V_j)|\longrightarrow 0. \tag{A.1}$$

For $j\neq 1$,

$$\begin{aligned}\Big|\mathsf{Cov}\,(V_1,V_j)|&=\Big|\frac{\mathsf{E}(Z_1^{(1)}+Z_1^{(2)})(Z_j^{(1)}+Z_j^{(2)})}{h^2\sigma_u^2(x_0)}\Big|\\&\leqslant\frac{1}{h^2\sigma_u^2(x_0)}\{|\mathsf{E}(Z_1^{(1)}Z_j^{(1)})|+|\mathsf{E}(Z_1^{(1)}Z_j^{(2)})|+|\mathsf{E}(Z_1^{(2)}Z_j^{(1)})|+|\mathsf{E}(Z_1^{(2)}Z_j^{(2)})|\}\\&\triangleq\frac{1}{h^2\sigma_u^2(x_0)}\{R_1+R_2+R_3+R_4\}.\end{aligned}$$

As to R_1, we have

$$\begin{aligned} R_1 &= |\mathsf{E}\{\mathsf{E}(Z_1^{(1)}Z_j^{(1)}|\mathscr{F}_j)\}| \\ &\leqslant C\Big|\mathsf{E}\Big\{h^2\sigma^2(X_1)\sigma^2(X_j)\frac{\psi_{J,k}^{\rm per}(X_1)}{f(X_1)}\frac{\psi_{J,k}^{\rm per}(X_j)}{f(X_j)}\Big\}\Big| \\ &\leqslant C2^{-J}h^2. \end{aligned}$$

Similarly, $R_2 \leqslant C2^{-J}h^2$, $R_3 \leqslant C2^{-J}h^2$, $R_4 \leqslant C2^{-J}h^2$. Hence

$$|\mathsf{Cov}\,(V_1, V_j)| = O(2^{-J}). \tag{A.2}$$

Let C_n be a sequence of integers such that $C_n \to \infty$, $C_n/2^J \to 0$. Denote

$$\sum_{j=2}^{n}|\mathsf{Cov}\,(V_1,V_j)| = \sum_{j=2}^{C_n}|\mathsf{Cov}\,(V_1,V_j)| + \sum_{j=C_n+1}^{n}|\mathsf{Cov}\,(V_1,V_j)| \triangleq F_1 + F_2.$$

Based on (A.2), we have

$$F_1 = O(C_n 2^{-J}) \to 0.$$

Considering F_2, we have

$$|\mathsf{Cov}\,(V_1,V_j)| \leqslant |\mathsf{Cov}\,(V_1^{(1)},V_j^{(1)})| + |\mathsf{Cov}\,(V_1^{(1)},V_j^{(2)})| + |\mathsf{Cov}\,(V_1^{(2)},V_j^{(1)})| + |\mathsf{Cov}\,(V_1^{(2)},V_j^{(2)})|.$$

From Lemma 2.1 of Davydor (1968), we have

$$|\mathsf{Cov}\,(V_1^{(1)},V_j^{(1)})| \leqslant C[\alpha(j-1)]^{1-2/e}(\mathsf{E}|V_1^{(1)}|^e)^{2/e},$$

where $e > 2$. We can show that

$$\mathsf{E}|V_1^{(1)}|^e \leqslant C\frac{1}{h^e}\mathsf{E}|Z_1^{(1)}|^e \leqslant C\mathsf{E}\Big|\psi_{J,k}^{\rm per}(X_1)\frac{\sigma^2(X_1)(\varepsilon_1^2-1)}{f(X_1)}\Big|^e \leqslant C2^{J(e/2-1)}.$$

Therefore

$$|\mathsf{Cov}\,(V_1^{(1)},V_j^{(1)})| \leqslant C[\alpha(j-1)]^{1-2/e}2^{J(1-2/e)}.$$

Similarly,

$$\begin{aligned} |\mathsf{Cov}\,(V_1^{(1)},V_j^{(2)})| &\leqslant C[\alpha(j-1)]^{1-2/e}2^{J(1-2/e)}. \\ |\mathsf{Cov}\,(V_1^{(2)},V_j^{(1)})| &\leqslant C[\alpha(j-1)]^{1-2/e}2^{J(1-2/e)}. \\ |\mathsf{Cov}\,(V_1^{(2)},V_j^{(2)})| &\leqslant C[\alpha(j-1)]^{1-2/e}2^{J(1-2/e)}. \end{aligned}$$

Hence

$$|\mathsf{Cov}\,(V_1,V_j)| \leqslant C[\alpha(j-1)]^{1-2/e}2^{J(1-2/e)}.$$

It then follows that

$$F_2 \leqslant \sum_{j=C_n+1}^{n} C[\alpha(j-1)]^{1-2/e} 2^{J(1-2/e)} \leqslant C 2^{J(1-2/e)} C_n^{-a} \sum_{j=C_n+1}^{n} (j-1)^a [\alpha(j-1)]^{1-2/e}.$$

We choose $C_n = 2^{J(1-2/e)/a}$, so that $C_n/2^J \to 0$ as $a > 1-2/e$ and $e > 2$. According to the assumption (A1), we have $F_2 \to 0$. This implies that (A.1) is true.

To prove Theorem 3.2, it suffices to show that

$$\frac{1}{\sqrt{n}} \sum_{j=1}^{n} V_j \to N(0,1). \tag{A.3}$$

Now, partition the set $\{1, 2, \ldots, n\}$ into $2q_n + 1$ subsets with large blocks $u = u_n$ and small blocks $v = v_n$. Let $q = q_n = [n/(u+v)]$. Denote

$$\eta_j = \sum_{i=j(u+v)+1}^{j(u+v)+u} V_i, \qquad \xi_j = \sum_{i=j(u+v)+u+1}^{(j+1)(u+v)} V_i, \qquad \eta_q = \sum_{i=q(u+v)+1}^{n} V_i.$$

Then, we have

$$\sum_{j=1}^{n} V_j = \sum_{j=0}^{q-1} \eta_j + \sum_{j=0}^{q-1} \xi_j + \eta_q \triangleq S_1 + S_2 + S_3.$$

According to Theorem 18.4 of Ibragimov and Linnik (1971), to prove (A.3), it suffices to show that

$$\frac{1}{n}\mathsf{E}S_2^2 \to 0, \qquad \frac{1}{n}\mathsf{E}S_3^2 \to 0, \tag{A.4}$$

$$\Big|\mathsf{E}\exp(itS_1) - \prod_{j=0}^{q-1} \mathsf{E}\exp(it\eta_j)\Big| \to 0, \tag{A.5}$$

$$\frac{1}{n}\sum_{j=0}^{q-1} \mathsf{E}\eta_j^2 \to 1, \qquad \frac{1}{n}\sum_{j=0}^{q-1} \mathsf{E}(\eta_j^2 I(|\eta_j| > \varepsilon\sqrt{n})) \to 0. \tag{A.6}$$

To prove (A.4)-(A.6), we choose $v = l_n$ in accordance with the assumption (A1). So there exists a constant $p_n \to \infty$ such that $p_n v = o((n2^{-J})^{1/2})$ and $p_n(n2^J)^{1/2}\alpha(v) \to 0$. Let $u = [(n2^{-J})^{1/2}/p_n]$, we obtain

$$\frac{v}{u} \to 0, \qquad \frac{u}{n} \to 0, \qquad \frac{u}{(n2^{-J})^{1/2}} \to 0, \tag{A.7}$$

$$\frac{n}{u}\alpha(v) \to 0. \tag{A.8}$$

We first consider (A.4). Note that

$$\frac{1}{n}\mathsf{E}S_2^2 = \frac{1}{n}\sum_{j=0}^{q-1} \mathsf{Var}\,(\xi_j) + \frac{1}{n}\sum_{i=0}^{q-1}\sum_{j=0,j\neq i}^{q-1} \mathsf{Cov}\,(\xi_i, \xi_j) \triangleq G_1 + G_2.$$

Based on the proof of (A.1) and the stationary of $\{V_j\}$, we have

$$\begin{aligned}\mathsf{Var}\,(\xi_j) &= v\mathsf{Var}\,(V_1)+\sum_{i=1}^{v}\sum_{j=1,j\neq i}^{v}\mathsf{Cov}\,(V_i,V_j)\\ &= v(1+o(1))+2v\sum_{j=2}^{v}\Big(1-\frac{j+1}{v}\Big)\mathsf{Cov}\,(V_1,V_j)\\ &= v+o(v).\end{aligned}$$

Hence

$$G_1=\frac{1}{n}\sum_{j=0}^{q-1}\mathsf{Var}\,(\xi_j)\leqslant\frac{qv}{n}\to 0,$$

by using (A.7). As to G_2, denote $m_j=j(u+v)+u$. Since $i\neq j$, $|m_i+l_1-(m_j+l_2)|>u$, then

$$\begin{aligned}G_2 &= \frac{1}{n}\sum_{i=0}^{q-1}\sum_{j=0,j\neq i}^{q-1}\sum_{l_1=1}^{v}\sum_{l_2=1}^{v}\mathsf{Cov}\,(V_{m_i+l_1},V_{m_j+l_2})\\ &\leqslant \frac{2}{n}\sum_{l_1=1}^{n}\sum_{l_2=l_1+u}^{n-u}|\mathsf{Cov}\,(V_{l_1},V_{l_2})|\\ &\leqslant 2\sum_{j=u}^{n-1}|\mathsf{Cov}\,(V_1,V_j)|=o(1).\end{aligned}$$

Based on the above arguments, the first result in (A.4) has been verified. As to the second result in (A.4), we have

$$\frac{1}{n}\mathsf{E}S_3^2\leqslant\frac{1}{n}(n-q(n-v))\mathsf{Var}\,(V_1)+2\sum_{j=2}^{n}|\mathsf{Cov}\,(V_1,V_j)|\leqslant\frac{v+u}{n}(1+o(1))+o(1)\to 0.$$

Therefore, (A.4) holds.

Note that η_j is a function of random variables $\{V_{j(u+v)+1},\dots,V_{j(u+v)+u}\}$, so η_j is $\mathscr{F}_{i_j}^{j_j}$-measurable, with $i_j=j(u+v)$ and $j_j=j(u+v)+u-1$. Applying Lemma 1.1 of Volkonskii and Rozanov (1959), we have

$$\Big|\mathsf{E}\Big(\prod_{j=0}^{q-1}\exp(it\eta_j)-\prod_{j=0}^{q-1}\mathsf{E}\exp(it\eta_j)\Big)\Big|\leqslant Cq\alpha(v+1)\leqslant C\alpha(v+1)\frac{n}{u+v}=o(1).$$

Hence, (A.5) holds.

As to the first result of (A.6), we have

$$\mathsf{Var}\,(\eta_j)=u\mathsf{Var}\,(V_1)+2\sum_{j=2}^{u}(u-j+1)\mathsf{Cov}\,(V_1,V_j)=u(1+o(1)).$$

Hence

$$\frac{1}{n}\sum_{j=0}^{q-1}\mathsf{E}\eta_j^2=\frac{qu}{n}(1+o(1))=\frac{u}{u+v}\to 1.$$

Now, we consider the second result of (A.6). For a fixed $L > 0$, denote $W_i^L = V_i I(\varepsilon_i^2 \leqslant L)$, $\overline{W}_i^L = \mathsf{E}(W_i^L)$ and $V_i^L = W_i^L - \overline{W}_i^L$. Define

$$S^L = \sum_{j=1}^{n} V_j^L, \qquad \widetilde{S}^L = \sum_{j=1}^{n} (V_j - V_j^L).$$

By the boundedness of $K(x), \sigma(x)$ and $T(x)$, we know that

$$\begin{aligned} |V_i^L| &= |W_i^L - \overline{W}_i^L| \leqslant |W_i^L| + |\overline{W}_i^L| = |W_i^L| + o(2^{-J/2}) \leqslant |V_i I(\varepsilon_i^2 \leqslant L)| \\ &\leqslant \Big|\frac{Z_i^{(1)}}{h\sigma_u(x_0)} I(\varepsilon_i^2 \leqslant L)\Big| + \Big|\frac{Z_i^{(2)}}{h\sigma_u(x_0)} I(\varepsilon_i^2 \leqslant L)\Big| \leqslant C2^{J/2}. \end{aligned}$$

Hence

$$\max_{0 \leqslant j \leqslant q-1} |\eta_j^L| \leqslant C2^{J/2} u,$$

where

$$\eta_j^L = \sum_{i=j(u+v)+1}^{j(u+v)+u} V_i^L.$$

According to (A.7), we know that the set $\{|\eta_j^L| \geqslant \varepsilon\sqrt{n}\}$ is empty for large enough n. Then

$$\max_{0 \leqslant j \leqslant q-1} \mathsf{E}\{(\eta_j^L)^2 I(|\eta_j^L| \geqslant \varepsilon\sqrt{n})\} \to 0.$$

Therefore, the second part of (A.6) holds for the truncated variables η_j^L. Hence

$$\frac{1}{\sqrt{n}} S^L = \frac{1}{\sqrt{n}} \sum_{j=1}^{n} V_j^L \to N(0, \sigma_L^2), \tag{A.9}$$

where $\sigma_L^2 = \mathsf{Var}\,(V_1 I(\varepsilon^2 \leqslant L))$. To prove the second result of (A.6), it suffices to show that as $n \to \infty$ and $L \to \infty$, we have

$$\frac{1}{n} \mathsf{Var}\,(\widetilde{S}^L) \to 0. \tag{A.10}$$

In fact, we know that

$$\widetilde{S}^L = \sum_{j=1}^{n} (V_j I(\varepsilon^2 \geqslant L) - \mathsf{E}(V_j I(\varepsilon^2 \geqslant L))).$$

Hence

$$\frac{1}{n} \mathsf{Var}\,(\widetilde{S}^L) = \mathsf{Var}\,(V_j I(\varepsilon^2 \geqslant L)) \to 0.$$

Therefore, (A.6) is true.

To complete the proof of Theorem 3.2, we can show that

$$\begin{aligned}
&\Big|\mathsf{E}\exp\Big(it\frac{1}{\sqrt{n}}\sum_{j=1}^{n}V_j\Big)-\exp\Big(-\frac{t^2\sigma^2}{2}\Big)\Big| \\
\leqslant\ &\Big|\mathsf{E}\exp\Big(it\frac{1}{\sqrt{n}}\sum_{j=1}^{n}V_j\Big)\Big\{\exp\Big(it\frac{1}{\sqrt{n}}\widetilde{S}^L\Big)-1\Big\}\Big| \\
&+\Big|\mathsf{E}\exp\Big(it\frac{1}{\sqrt{n}}\widetilde{S}^L\Big)-\exp\Big(-\frac{t^2\sigma_L^2}{2}\Big)\Big|+\Big|\exp\Big(-\frac{t^2\sigma_L^2}{2}\Big)-\exp\Big(-\frac{t^2\sigma^2}{2}\Big)\Big| \\
\triangleq\ &\widetilde{R}_1+\widetilde{R}_2+\widetilde{R}_3,
\end{aligned}$$

where σ^2 is the asymptotic variance of $(1/\sqrt{n})\sum_{j=1}^{n}V_j$. According to (A.10), $\widetilde{R}_1\to 0$, for every $L>0$, as $n\to\infty$. $\widetilde{R}_2\to 0$ as $n\to\infty$ and $L\to\infty$ by using (A.9). $\widetilde{R}_3\to 0$ as $n\to\infty$ and $L\to\infty$ by dominated convergence. Hence, Theorem 3.2 is true. □

Proof of Theorem 4.1 Our proof is similar to the proof of Theorem 2.4 in Chen et al. (2008), so we only provide outline of proof. We know that

$$Y_j^2=\sigma^2(X_j)+(T^2(X_j)+\sigma^2(X_j)(\varepsilon_j^2-1)+2T(X_j)\sigma(X_j)\varepsilon_j).$$

From the proof of Theorem 2.4 in Chen et al. (2008), we know that for all $k\in I_J$

$$\Big|\int_a^b\psi_{J,k}^{\mathrm{per}}(x)\frac{\sum\limits_{j=1}^{n}K_h(X_j-x)T^2(X_j)}{\sum\limits_{j=1}^{n}K_h(X_j-x)}\mathrm{d}x\Big|\leqslant C2^{-3J/2},$$

and for all $k\notin I(t_0,2^{-J}A(b-a))$,

$$\Big|\int_a^b\psi_{J,k}^{\mathrm{per}}(x)\frac{\sum\limits_{j=1}^{n}K_h(X_j-x)\sigma^2(X_j)}{\sum\limits_{j=1}^{n}K_h(X_j-x)}\mathrm{d}x\Big|\leqslant C2^{-3J/2}.$$

According to Lemma 6.3, we know that for $k\in I(t_0,2^{-J}A(b-a))$

$$\Big|\int_a^b\psi_{J,k}^{\mathrm{per}}(x)\frac{\sum\limits_{j=1}^{n}K_h(X_j-x)\sigma^2(X_j)}{\sum\limits_{j=1}^{n}K_h(X_j-x)}\mathrm{d}x\Big|\geqslant C2^{-J/2}.$$

Similar to the proof of Theorem 3.1, we can easily show that

$$\int_a^b \psi_{J,k}^{\mathrm{per}}(x)\frac{\sum\limits_{j=1}^n K_h(X_j-x)\sigma^2(X_j)(\varepsilon_j^2-1)}{\sum\limits_{j=1}^n K_h(X_j-x)}\mathrm{d}x = O_p((2^Jnh)^{-1/2}),$$

$$\int_a^b \psi_{J,k}^{\mathrm{per}}(x)\frac{\sum\limits_{j=1}^n K_h(X_j-x)2T(X_j)\sigma(X_j)\varepsilon_j}{\sum\limits_{j=1}^n K_h(X_j-x)}\mathrm{d}x = O_p((2^Jnh)^{-1/2}).$$

Based on

$$\frac{2^{-J/2}-(2^Jnh)^{-1/2}}{2^{-3J/2}+(2^Jnh)^{-1/2}}\to\infty,$$

we have

$$\max\{|U_J(k)|, k\in I_J\} = \max\{|U_J(k)|, k\in I(t_0, 2^{-J}A(b-a))\}.$$

This implies that

$$|t_0^U - t_0| = \Big|a+\frac{k_1^U(b-a)}{2^J}-t_0\Big| < 2^{-J}A(b-a).$$

This completes the proof of Theorem 4.1. □

Proof of Theorem 4.2 The proof is straightforward from Theorem 3.1. □

Proof of Theorem 4.3 The proof is similar to the proof of Theorem 4.1 in Zhou et al. (2010). □

References

[1] Lamoreux, C.G. and Lastrapes, W.D., Persistence in variance, structural change and the GARCH model, *Journal of Business and Economic Statistics*, **48**(1990), 225–234.

[2] Jorion, P., On jump processes in the foreign exchange and stock markets, *Review of Financial Studies*, **1**(1988), 427–445.

[3] Vlaar, P. and Palm, F., The message in weekly exchange rates in the European monetary system: mean revision, conditional heteroscedasticity, and jumps, *Journal of Business and Economic Statistics*, **11**(1993), 351–360.

[4] Drost, F., Nijman, T. and Werker, B., Estimation and testing in models containing both jumps and conditional heteroscedasticity, *Journal of Business and Economic Statistics*, **16**(1998), 237–243.

[5] Müller, T.G., Change points in nonparametric regression analysis, *The Annals of Statistics*, **20**(1992), 737–761.

[6] Chen, G., Choi, Y. and Zhou, Y., Nonparametric estimation of structural change points in volatility models for time series, *Journal of Econometrics*, **126**(2005), 79–114.

[7] Inclan, C. and Tiao, G.C., Use of cumulative sums of squares for retrospective detection of change of variance, *Journal of the American Statistical Association*, **89**(1994), 913–923.

[8] Wang, Y., Jump and Sharp cusp detection by wavelets, *Biometrika*, **82**(1995), 385–397.

[9] Wang, Y., Change-points via wavelets for indirect data, *Statistica Sinica*, **9**(1999), 103–117.

[10] Wong, H., Ip, W., Li, Y. and Xie, Z., Threshold variable selection by wavelets in open-loop threshold autoregressive models, *Statistics and Probability Letters*, **42**(1999), 375–392.

[11] Luan, Y. and Xie, Z., The wavelet identification for jump points of derivative in regression model, *Statistics and Probability Letters*, **53**(2001), 167–180.

[12] Chen, G., Choi, Y. and Zhou, Y., Detections of changes in return by a wavelet smoother with conditional heteroscedastic volatility, *Journal of Econometrics*, **143**(2008), 227–262.

[13] Zhou, Y., Wan, A.T.K., Xie, S. and Wang, X., Wavelet analysis of change-points in a non-parametric regression with heteroscedastic variance, *Journal of Econometrics*, **159**(2010), 183–201.

[14] Wong, H., Ip, W. and Li, Y., Detection of jumps by wavelets in a heteroscedastic autoregressive model, *Statistics and Probability Letters*, **52**(2001), 365–372.

[15] Nadaraya, E.A., On estimating regression, *Theory of Probability and Its Applications*, **9**(1964), 141–142.

[16] Fan, J. and Gijbels, I., *Local Polynomial Modeling and Its Applications*, Chapman and Hall, London, 1996.

[17] Li, Y. and Xie, Z., The wavelet identification of thresholds and time delay of threshold autoregressive models, *Statistica Sinica*, **9**(1999), 153–166.

[18] Xia, Y., Bias-corrected confidence bands in nonparametric regression, *Journal of Royal Statistical Society Series B*, **60**(1998), 797–811.

[19] Davydov, Y., Convergence of distributions generated by stationary stochastic process, *Theory of Probability and Its Applications*, **13**(1968), 691–696.

[20] Ibragimov, I.A. and Linnik, Yu.A., *Independent and Stationary Sequences of Random Variables*, Wolters-Noordohoff, Groningen, Germany, 1971.

[21] Volkonskii, V.A. and Rozanov, Y.A., Some limit theorems for random functions, I, *Theory of Probability and Its Applications*, **4**(1959), 178–197.

非参数回归中方差变点的小波检测

王景乐　　郑　明

(复旦大学管理学院统计学系, 上海, 200433)

本文主要研究了非参数回归模型中方差函数的变点, 利用小波方法构造的检验量来检测方差中的变点, 建立了这些检验量的渐近分布, 并且运用这些检验量构造了方差变点的位置和跳跃幅度的估计, 给出了这些估计的渐近性质, 并进一步通过随机模拟验证了本文方法在有限样本下的性质.

关键词: 方差变点, 小波系数, 核估计, 局部线性平滑, α-混合.

学科分类号: O212.7.