

RESEARCH

Open Access



Detection of multiple change points for linear processes under negatively super-additive dependence

Yuncaï Yu¹, Xinsheng Liu^{1*}, Ling Liu² and Piao Zhao¹

*Correspondence:
xsliu@nuaa.edu.cn

¹State Key Laboratory of Mechanics and Control of Mechanical Structures, Institute of Nano Science and Department of Mathematics, Nanjing University of Aeronautics and Astronautics, Nanjing, China
Full list of author information is available at the end of the article

Abstract

This paper focuses on the issue of detecting the multiple change points for linear processes under negatively super-additive dependence (NSD). We propose a CUSUM-type method in the multiple variance change model and establish the weak convergence rate of the change points estimation. To carry out this method, we give a multiple variance-change iterative (MVCI) algorithm. Additionally, some simulations are implemented to substantiate the validity of the CUSUM-type method. Comparison with some best methods indicates that the CUSUM-type change point estimation is computationally competitive and superior in terms of the mean squared error (MSE).

MSC: 62F03; 62F05; 62F12

Keywords: Multiple change points; Linear processes under NSD; Variance change; CUSUM-type estimation; Weak convergence rate

1 Introduction

As a common feature of 'big data', change point arises in many areas such as signal processing (Basseville [1]), finance (Chen and Gupta [2]), ecology (Hawkins [3]), disease outbreak watch (Sparks et al. [4]), and neuroscience (Ratnam et al. [5]; Lena et al. [6]) and has been much investigated in the last few decades. To detect change point and estimate its location, there has emerged a number of approaches including least squares (LS, Bai [7]), Bayesian method (Fearnhead [8]), maximum likelihood (Zou et al. [9]), and some non-parametric methods (Matteson and James [10]; Haynes et al. [11]). The cumulative sum (CUSUM) method, based on the LS estimation, is a very attractive one for detecting the variance change in a sequence because it avoids some assumptions about the underlying error distribution function and is computed simply (Gombay et al. [12]). For independent sequences, Gombay et al. [12] constructed the CUSUM statistic to detect and estimate the change of variance. Wang and Wang [13] used the CUSUM test to detect the variance change in a linear process with long memory errors. Zhao et al. [14] considered the ratio test for variance change in a linear process. Qin et al. [15] investigated the strong convergence rate of the CUSUM estimator of the variance change in linear processes.

However, most of the references above assume the change point number in a sequence is one, which is a serious restriction when applied to practical problems. For multiple change

point detection, Inclán and Tiao [16] employed the cumulative sums of squares to detect the multiple changes of variance in the uncorrelated sequences. Lavielle [17] obtained the convergence rate for multiple change detection for strongly mixing and strongly dependent processes. Li and Zhao [18] gave the convergence rate for multiple change-points estimation of moving-average processes. More recently, Haynes et al. [11] proposed a computationally efficient nonparametric approach for change point detection, and Laurentiu et al. [19] offered the Bayesian loss-based approach to analyze change point problem. But both of them require the information of the underlying error distribution function, which may lead to the complexity of computation.

In this contribution, we consider the following multiple variance change model:

$$Y_t = \mu + \sigma_i e_t, \quad t_{i-1}^* \leq t \leq t_i^*, \quad 1 \leq i \leq r, \tag{1}$$

where r is the known number of change points, μ and σ_i ($1 \leq i \leq r$) are parameters, t_i^* , $1 \leq i \leq r$, $t_0^* = 0$, $t_{r+1}^* = n$ are the true change locations with $t_i^* = [\tau_i n]$, where $[x]$ denotes the integer part of x , $\boldsymbol{\tau} = (\tau_1^*, \tau_2^*, \dots, \tau_r^*)$ are the change points, and e_t is linear processes given as follows:

$$e_t = \sum_{j=0}^{\infty} a_j \varepsilon_{t-j}, \tag{2}$$

where a_j is an array of real numbers satisfying $\sum_{j=0}^{\infty} a_j^2 < \infty$, $\{\varepsilon_m, m \in \mathbb{Z}\}$ are stationary random variables.

Under the independent or dependent assumptions of $\{\varepsilon_m, m \in \mathbb{Z}\}$, the convergence rates of the single change point estimators have been established for the linear processes (2). We refer to Bai [7] and Qin et al. [15] for independence case, to Li and Zhao [18] for linear negative quadrant dependence, and to Wang and Wang [13] for long range dependence. In this article, we will consider the multiple variance change model, and simultaneously $\{\varepsilon_m, m \in \mathbb{Z}\}$ are negatively super-additive dependence (NSD) whose definition is based on the super-additive functions.

Definition 1 (Hu [20]) A function ϕ is called super-additive if

$$\phi(\mathbf{x} \vee \mathbf{y}) + \phi(\mathbf{x} \wedge \mathbf{y}) \geq \phi(\mathbf{x}) + \phi(\mathbf{y})$$

for all $\mathbf{x}, \mathbf{y} \in R^n$, where “ \vee ” is componentwise maximum and “ \wedge ” represents componentwise minimum.

Definition 2 (Hu [20]) A random vector (X_1, X_2, \dots, X_n) is said to be NSD if

$$E\phi(X_1, X_2, \dots, X_n) \leq E\phi(X_1^*, X_2^*, \dots, X_n^*), \tag{3}$$

where $\{X_m^*, 1 \leq m \leq n\}$ are independent random variables that have the same marginal distribution with $\{X_m, 1 \leq m \leq n\}$ for each i , and ϕ is a super-additive function such that the expectations in (3) exist.

Definition 3 (Wang et al. [21]) A sequence of random variables $(X_1, X_2, \dots, X_n, \dots)$ is called NSD if, for all $n \geq 1$, (X_1, X_2, \dots, X_n) is NSD.

NSD has received considerable attention since it includes the well-known negative association (see Christofides and Vaggelatos [22]). Eghbal et al. [23] explored the strong law of large numbers and the rate of convergence for NSD sequences with the existence of high order moments. Shen et al. [24] and Wu et al. [25] got the almost sure and complete convergence, respectively, for NSD random variables. Wang et al. [26] investigated the complete convergence, and Yu et al. [27] established the central limit theorem for weighted sums of NSD random variables. Moreover, NSD samples have been introduced to various models; for example, under NSD errors, Yu et al. [27] considered the M-test problem of regression parameters in a linear model; Wang et al. [28] studied the strong consistency and weak consistency of the LS estimators in an EV regression model, and Yu et al. [29] obtained the convergence rates of the wavelet thresholding estimators in a nonparametric regression model.

The aim of this study is to detect the multiple change points for linear processes under NSD. We propose the CUSUM-type change point estimator in model (1) and establish the weak convergence rate of the estimator with the mean parameter μ estimated by its LS estimator. Moreover, some simulations are implemented by *R Software* to compare the CUSUM-type estimator with some methods. The result indicates that the CUSUM-type change point estimator is broadly comparable with those obtained by the typical methods.

The remainder of this paper is organized as follows. In Sect. 2, we describe the CUSUM-type multiple change point estimation and give the weak convergence rate of this estimator. Also, we give a multiple variance-change iterative (MVCI) algorithm to evaluate the estimator. In Sect. 3, some simulations are presented to show the performances of the estimator. Finally, the proofs of the main results are given in Sect. 4.

2 Estimation and main results

Let $\tilde{Y}_j = Y_j - \hat{\mu}_n$, where $\hat{\mu}_n = \frac{1}{n} \sum_{t=1}^n Y_t$ is the LS estimator of the mean μ . Assume that

$$A_{n,r} = \{(t_0, t_1, \dots, t_{r+1}), t_0 = 0 < t_1 < \dots < t_r < t_{r+1} = n\}$$

is a set of allowable r -partitions. We further consider the following set of allowable r -partitions:

$$A_{n,r}^{\delta_n} = \{(t_0, t_1, \dots, t_{r+1}) : t_i - t_{i-1} \geq n\delta_n\},$$

where δ_n is a non-increasing non-negative sequence satisfying $\delta_n \rightarrow 0$ and $n\delta_n \rightarrow \infty$.

For each t_i , $1 \leq i \leq r + 1$, we define

$$R(t_i) = \frac{(t_i - t_{i-1})(t_{i+1} - t_i)}{(t_{i+1} - t_{i-1})^2} \left| \frac{1}{t_i - t_{i-1}} \sum_{t=t_{i-1}+1}^{t_i} \tilde{Y}_t^2 - \frac{1}{t_{i+1} - t_i} \sum_{t=t_i+1}^{t_{i+1}} \tilde{Y}_t^2 \right|.$$

Denote $\hat{\tau}^{\delta_n} = \hat{\mathbf{t}}^{\delta_n} / n$, the CUSUM-type multiple change point estimator is given by

$$\hat{\tau}^{\delta_n} = \arg \max_{\mathbf{t} \in A_{n,r}^{\delta_n}} \frac{1}{n} \sum_{i=1}^{r+1} R(t_i). \tag{4}$$

To derive our results, we list several conditions as follows.

- (A1) $\{\varepsilon_m, m \in \mathbb{Z}\}$ are stationary NSD random variables with $E\varepsilon_m = 0$ and $\text{Var}(\varepsilon_m) = \sigma^2 < \infty$.
- (A2) For all $l \geq 1$, we have $\sum_{m:|l-m|\geq u} |\text{Cov}(\varepsilon_l, \varepsilon_m)| \rightarrow 0$, as $u \rightarrow \infty$.
- (A3) $E\varepsilon_m^4 < \infty$ holds for all $m \geq 1$.
- (A4) $\sum_{j=0}^\infty |a_j| < \infty$.

Remark 1 Conditions (A1) and (A2) are easily satisfied (see Yu et al. [29]). (A3) is often applied to obtain the convergence rate of change point estimator (e.g., Qin et al. [15]; Shi et al. [30]). Condition (A4) is weaker than Bai [11], which requires $\sum_{j=0}^\infty j|a_j| < \infty$. Furthermore, condition (A4) implies that $\sum_{j=0}^\infty a_j^2 < \infty$ and $\sum_{j=0}^\infty a_j^4 < \infty$.

Theorem 1 *Assume that conditions (A1)–(A4) hold. Then, for all $1 \leq j \leq r$, we have*

$$\hat{\tau}_j^{\delta_n} \rightarrow \tau_j^*, \quad \text{in probability.}$$

When the mean μ is known, Qin et al. [15] established the strong convergence of the CUSUM estimator. It is obvious that Theorem 1 is still true when μ is known, and we will give the following corollary without proof.

Corollary 1 *If the mean is known ($\mu = 0$), conditions (A1)–(A4) hold, then we have the same conclusion of Theorem 1.*

Under assumptions (A1)–(A4), we can further establish the convergence rate of the CUSUM-type multiple change point estimator $\hat{\tau}^{\delta_n}$.

Theorem 2 *Let $M(n)$ be a natural number sequence with $M(n) \rightarrow \infty$. Then, under the conditions of Theorem 1, we further have*

$$\hat{\tau}_j^{\delta_n} - \tau_j^* = o(M(n)/n), \quad \text{in probability.}$$

To implement the CUSUM-type multiple change-point method, we also give the multiple variance-change iterative (MVCI) algorithm based on Qin et al. [15] and Shi et al. [30] as follows:

- Step 1.* Choose $\eta \geq 1$, compute $\hat{\mu}_n$ and $\{\tilde{Y}_t^2\}$.
- Step 2.* Set $i = 1$, $m = 0$, and $l = \lfloor n\delta_n \rfloor$. Divide the sample into L subintervals I_j with the equal interval length l .
- Step 3.* For each subinterval $I_j, j = 1, 2, \dots, L$, find $\hat{t}_j^{(i)} = \arg \max_{t \in (1+m, m+l)} R(t_j)$.
- Step 4.* Compute the set $\Delta = \{R(\hat{t}_j^{(i)})\}$, and select r change locations which correspond to r maximum values of $R(\hat{t}_j^{(i)})$ in the set Δ .
- Step 5.* For the selected r change locations $\hat{t}_j^{(i)}, j = 1, \dots, r$, find $\hat{t}_j^{(i+1)} = \arg \max_{t \in (\hat{t}_j^{(i)} - 2M(l), \hat{t}_j^{(i+1)} + 2M(l))} R(t_j)$.
- Step 6.* Set $l = 4M(l)$ and $m = \hat{t}_j^{(i)} - 2M(l)$.
- Step 7.* If $\|\hat{\mathbf{t}}^{(i+1)} - \hat{\mathbf{t}}^{(i)}\|_\infty < \eta$, then proceed to Step 8, otherwise set $i = i + 1$, go back to Step 3.
- Step 8.* $\hat{\mathbf{t}}_{\text{MVCI}} = \hat{\mathbf{t}}^{(i)}$ and $\hat{\tau}_{\text{MVCI}} = \hat{\mathbf{t}}^{(i)}/n$.

3 Simulation studies

We present a set of simulation studies to illustrate the availability of the CUSUM-type MVCI algorithm via R packages. Additionally, we implement some available competitors including segment neighborhood (SN), pruned exact linear time (PELT), binary segmentation (BS), and wild binary segmentation (WBS) to compare the performance of the MVCI algorithm.

In model (1), we take $r = 4, \mu = 0, \sigma_1 = 2, \sigma_2 = 4, \sigma_3 = 8, \sigma_4 = 4, \sigma_5 = 2$, and suppose the true change locations $t_1^* = 100, t_2^* = 200, t_3^* = 300, t_4^* = 400$, we model the NSD sequence $\{\varepsilon_m, m \in \mathbb{Z}\}$ as a multivariate mixture of normal distribution with joint distribution $N(0, 0, 1, 4; -0.5)$. The sample size is taken to be $n = 500$ and the weight functions are satisfied $a_j = 2^{-j}, j \in \mathbb{Z}$. Figure 1 displays the simulated sequence of $Y_t, 1 \leq t \leq 500$, and the true change locations.

To carry out the SN (Auger and Lawrence [31]), PELT (Killick et al. [32]), BS (Killick and Eckley [33]), we use the penalty likelihood method, which can be implemented by *change-point* package in (Killick [34]). As to WBS (Killick and Eckley [33]), we utilize package *wbsts* (Korkas, Karolos, and Piotr [35]) in our model with the threshold $\lambda_n = C\sqrt{2}\log^{1/2}n$, where $C = 1$. We assume the parameter $\delta_n = n^{-1/2}$ in the MVCI algorithm. The mean squared error (MSE) of the CUSUM-type variance change point estimator of τ^* is defined as $MSE = \frac{1}{r} \sum_{i=1}^r (\hat{\tau}_i - \tau_i^*)^2$, and the performances of the above methods are described in Table 1 (all of the simulations are run for 100 replicates).

Table 1 presents the average MSEs of the MVCI, SN, PELT, BS, and WBS methods. Generally, the first change point is overestimated and the rest change points are underestimated. When the sample size is large ($n = 500$), all of the methods can estimate the change points available, but the MVCI method is superior in terms of the average MSE. This also

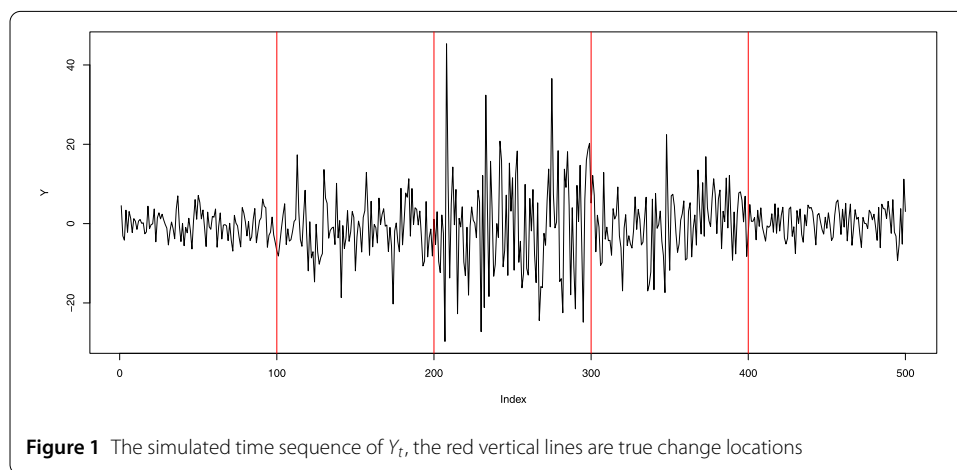


Table 1 Comparison of the MVCI algorithm with SN, PELT, BS, and WBS methods

τ_j^*	n	SN	PELT	BS	WBS	MVCI
0.2	500	0.222	0.222	0.218	0.210	0.210
0.4	500	0.410	0.398	0.392	0.406	0.396
0.6	500	0.590	0.590	0.594	0.590	0.592
0.8	500	0.784	0.792	0.796	0.792	0.790
MSE		23.5×10^{-5}	16.3×10^{-5}	11.0×10^{-5}	7.5×10^{-5}	7.0×10^{-5}

indicates that the CUSUM-type variance-change method is computationally competitive with some of best change point estimation methods.

4 Proof of the theorems

Throughout the proof, let C be a general positive constant and $c_0, c_1, c_2, C_0, C_1, \dots, C_4$ be some positive constants. Denote $x^+ = xI(x \geq 0)$ and $x^- = -xI(x < 0)$. In the following we will state some lemmas which are needed.

Lemma 1 (Hu [20]) *An NSD random sequence $\{X_m, m \geq 1\}$ possesses the following properties.*

(P1) *For any x_1, x_2, \dots, x_n ,*

$$P(X_1 \leq x_1, X_2 \leq x_2, \dots, X_n \leq x_n) \leq \prod_{m=1}^n P(X_m \leq x_m).$$

(P2) *$\{-X_1, -X_2, \dots, -X_n\}$ is also NSD.*

(P3) *Let f_1, f_2, \dots be a sequence of non-decreasing Borel functions, then $\{f_n(X_n), n \geq 1\}$ is still an NSD random sequence.*

Lemma 2 (Wang et al. [21]) *Suppose that $\{X_m, m \geq 1\}$ is an NSD random sequence with $EX_m = 0$ and $E|X_m|^\alpha < \infty$ for some $\alpha \geq 2$, then for all n ,*

$$E\left(\max_{1 \leq k \leq n} \left| \sum_{m=1}^k X_m \right|^\alpha\right) \leq C \left\{ \sum_{m=1}^n E|X_m|^\alpha + \left(\sum_{m=1}^n EX_m^2 \right)^{\alpha/2} \right\}.$$

Lemma 3 *Suppose that $\{X_m, m \geq 1\}$ is an NSD random sequence with conditions (A1)–(A2) hold, $\{a_m, 1 \leq m \leq n, n \geq 1\}$ is a sequence of real numbers satisfying $\sum_{m=1}^\infty a_m^2 < \infty$. Then*

$$\sigma_n^2 = \text{Var}\left(\sum_{m=1}^\infty a_m \varepsilon_{n-m}\right) \leq C\sigma^2.$$

Proof For a pair of NSD random variables X, Y , by property (P1) in Lemma 1, we have

$$H(x, y) = P(X \leq x, Y \leq y) - P(X \leq x)P(Y \leq y) \leq 0.$$

The covariance of X and Y is verified to be negative by

$$\text{Cov}(X, Y) = E(XY) - E(X)E(Y) = \int \int H(x, y) dx dy \leq 0. \tag{5}$$

Then, for $u \geq 1$,

$$\begin{aligned} & \sum_{l,m=1, |l-m| \geq u} |a_l a_m \text{Cov}(X_l, X_m)| \\ & \leq \sum_{l=1} \sum_{m=l+u} (a_l^2 + a_m^2) |\text{Cov}(X_l, X_m)| \end{aligned}$$

$$\begin{aligned} &\leq \sum_{l=1} a_l^2 \sum_{m=l+u} |\text{Cov}(X_l, X_m)| + \sum_{m=u+1} a_m^2 \sum_{l=1}^{m-u} |\text{Cov}(X_l, X_m)| \\ &\leq \sum_{l=1} a_l^2 \sum_{|m-l|\geq u} |\text{Cov}(X_l, X_m)| \\ &\leq \sup_l \sum_{m=1, |l-m|\geq u} |\text{Cov}(X_l, X_m)| \left(\sum_{m=1} a_m^2 \right). \end{aligned}$$

Hence, by condition (A2), for a fixed small $\varepsilon > 0$, there exists a positive integer $u = u_\varepsilon$ such that

$$\sum_{l,m=1, |l-m|\geq u} |a_l a_m \text{Cov}(X_l, X_m)| \leq \varepsilon.$$

Set $K = [1/\varepsilon]$ and $Y_m = \sum_{l=um+1}^{u(m+1)} a_l X_l, m = 0, 1, \dots, n,$

$$\Upsilon_m = \left\{ l : 2Km \leq l \leq 2Km + K, |\text{Cov}(Y_l, Y_{l+1})| \leq \frac{2}{K} \sum_{l=2Km}^{2Km+K} \text{Var}(Y_m) \right\}.$$

Define $h_0 = 0, h_{m+1} = \min\{h : h > h_m, h \in \Upsilon_m\},$ and put

$$\begin{aligned} Z_m &= \sum_{l=h_m+1}^{h_{m+1}} Y_l, \quad m = 0, 1, \dots, n, \\ \Lambda_m &= \{u(h_m + 1) + 1, \dots, u(h_{m+1} + 1)\}. \end{aligned}$$

Note that

$$Z_m = \sum_{l \in \Lambda_m} a_l X_l, \quad m = 0, 1, \dots, n.$$

It is easy to see that $\#\Lambda_m \leq 3Ku,$ where $\#$ stands for the cardinality of a set. From Lemma 2, it follows that

$$\begin{aligned} \sigma_n^2 &= E \sum_{m=1} Z_m^2 + \sum_{1 \leq m < l \leq n} |\text{Cov}(Z_m, Z_l)| \\ &= \sum_{m=1} E \left(\sum_{l \in \Lambda_m} a_l X_l \right)^2 + \sum_{1 \leq m < l \leq n, |m-l|=1} |\text{Cov}(Z_m, Z_l)| \\ &\quad + \sum_{1 \leq m < l \leq n, |m-l|>1} |\text{Cov}(Z_m, Z_l)| \\ &\leq \sum_{m=1} a_m^2 E \left(\sum_{l \in \Lambda_m} X_l \right)^2 + \sum_{m=1} |\text{Cov}(Y_{h_m}, Y_{h_{m+1}})| \\ &\quad + \sum_{1 \leq m < l \leq n, |m-l|\geq u} |a_m a_l| |\text{Cov}(Z_m, Z_l)| \\ &\leq \sum_{m=1} \sum_{l \in \Lambda_m} a_l^2 E(X_l)^2 + \frac{1}{K} \sum_{m=1} \text{Var}(Y_{h_m}) + \varepsilon \end{aligned}$$

$$\begin{aligned} &\leq \sum_{m=1} \sum_{l \in \Lambda_m} a_l^2 E(X_l)^2 + \frac{u}{K} \sum_{m=1} \sum_{l \in \Lambda_m} a_l^2 E(X_l)^2 + \varepsilon \\ &\leq \sum_{m=1} E(a_m X_m)^2 + \frac{Cu}{K} \sum_{m=1} E(a_m X_m)^2 + \varepsilon = C\sigma^2, \end{aligned}$$

which completes the proof of Lemma 3. □

Lemma 4 *Suppose that e_t is linear processes under NSD random sequence with conditions (A1)–(A4) hold, let $\sigma_\varepsilon^2 = E(e_\varepsilon^2)$. Then*

$$E\left(\sum_{t=1}^k (e_t^2 - \sigma_\varepsilon^2)\right)^2 \leq Ck.$$

Proof According to Lemma 3, we obtain

$$\sigma_\varepsilon^2 = E\left(\sum_{j=0}^\infty a_j \varepsilon_{t-j}\right)^2 \leq C\sigma^2 < \infty.$$

Obviously, there exists a positive number c_0 such that

$$e_t^2 - \sigma_\varepsilon^2 = \sum_{j=0}^\infty a_j^2 (\varepsilon_{t-j}^2 - c_0\sigma^2) + 2 \sum_{0 \leq s < l < \infty} a_s a_l \varepsilon_{t-s} \varepsilon_{t-l}.$$

Hence

$$\begin{aligned} E(e_t^2 - \sigma_\varepsilon^2)^2 &= E\left[\sum_{s=0}^\infty a_s^2 (\varepsilon_{t-s}^2 - c_0\sigma^2)\right]^2 + 4E\left(\sum_{0 \leq s < l < \infty} a_s a_l \varepsilon_{t-s} \varepsilon_{t-l}\right)^2 \\ &\quad + 4E\left(\left[\sum_{l=0}^\infty a_l^2 (\varepsilon_{t-l}^2 - c_0\sigma^2)\right]\left[\sum_{0 \leq s < l < \infty} a_s a_l \varepsilon_{t-s} \varepsilon_{t-l}\right]\right) \\ &= E\left[\sum_{s=0}^\infty a_s^4 (\varepsilon_{t-s}^2 - c_0\sigma^2)^2\right] + 2E\left[\sum_{0 \leq s < l < \infty} a_s^2 a_l^2 (\varepsilon_{t-l}^2 - c_0\sigma^2)\right] (\varepsilon_{t-s}^2 - c_0\sigma^2) \\ &\quad + 4E\left(\left[\sum_{0 \leq s' < l' < \infty} a_{s'} a_{l'} \varepsilon_{t-s'} \varepsilon_{t-l'}\right]\left[\sum_{0 \leq s' < l' < \infty} a_{s'} a_{l'} \varepsilon_{t-s'} \varepsilon_{t-l'}\right]\right) \\ &\quad + 4E\left(\left[\sum_{j=0}^\infty \sum_{0 \leq s < l < \infty} a_j^2 a_s a_l (\varepsilon_{t-s}^2 - c_0\sigma^2) \varepsilon_{t-s} \varepsilon_{t-l}\right]\right) \\ &\leq E(\varepsilon_j^2 - c_0\sigma^2)^2 \sum_{s=0}^\infty a_s^4 + 4 \sum_{0 \leq s < l < \infty} a_s^2 a_l^2 E(\varepsilon_{t-s}^2 \varepsilon_{t-l}^2) \\ &=: H_1 + H_2. \end{aligned}$$

By conditions (A3) and (A4), we obtain

$$H_1 \leq C(E\varepsilon_t^4 - 2c_0\sigma^2 E\varepsilon_t^2 + c_0^2\sigma^4) < +\infty. \tag{6}$$

Decomposing ε_t as $\varepsilon_t = \varepsilon_t^+ - \varepsilon_t^-$, from properties (P2) and (P3) in Lemma 1, one can see that $\varepsilon_t^+, \varepsilon_t^-, (\varepsilon_t^+)^2$ and $(\varepsilon_t^-)^2$ are NSD random sequences. From formula (5), we have $E(XY) \leq E(X)E(Y)$, then

$$\begin{aligned}
 H_2 &= 4 \sum_{0 \leq s < l < \infty} a_s^2 a_l^2 E((\varepsilon_{t-s}^+ - \varepsilon_{t-s}^-)^2 (\varepsilon_{t-l}^+ - \varepsilon_{t-l}^-)^2) \\
 &\leq 4 \sum_{0 \leq s < l < \infty} a_s^2 a_l^2 E([(\varepsilon_{t-s}^+)^2 + (\varepsilon_{t-s}^-)^2] \cdot [(\varepsilon_{t-l}^+)^2 + (\varepsilon_{t-l}^-)^2]) \\
 &\leq 16 \sum_{0 \leq s < l < \infty} a_s^2 a_l^2 \sigma^4 < \infty.
 \end{aligned}
 \tag{7}$$

Combining (6) and (7), we get

$$E(e_t^2 - \sigma_\varepsilon^2)^2 < \infty.
 \tag{8}$$

Now we consider the cross term. For any $t < j$, one can see that

$$\begin{aligned}
 &E[(e_t^2 - \sigma_\varepsilon^2)(e_j^2 - \sigma_\varepsilon^2)] \\
 &= E \left[\sum_{s=0}^{\infty} a_s^2 (\varepsilon_{t-s}^2 - \sigma^2) + 2 \sum_{0 \leq s < l < \infty} a_s a_l \varepsilon_{t-s} \varepsilon_{t-l} \right] \\
 &\quad \cdot \left[\sum_{s'=0}^{\infty} a_{s'}^2 (\varepsilon_{j-s'}^2 - c_0 \sigma^2) + 2 \sum_{0 \leq s' < l' < \infty} a_{s'} a_{l'} \varepsilon_{t-s'} \varepsilon_{j-l'} \right] \\
 &= E \left\{ \sum_{s=0}^{\infty} a_s^2 (\varepsilon_{t-s}^2 - c_0 \sigma^2) \cdot \sum_{\bar{s}=0}^{\infty} a_{\bar{s}}^2 (\varepsilon_{t-\bar{s}}^2 - \sigma^2) \right. \\
 &\quad \left. + 2 \sum_{s=0}^{\infty} a_s^2 (\varepsilon_{t-s}^2 - c_0 \sigma^2) \sum_{0 \leq s' < l' < \infty} a_{s'} a_{l'} \varepsilon_{t-s'} \varepsilon_{j-l'} \right\} \\
 &\quad + 2E \left\{ \sum_{s'=0}^{\infty} a_{s'}^2 (\varepsilon_{t-s'}^2 - c_0 \sigma^2) \sum_{0 \leq s < l < \infty} a_s a_l \varepsilon_{t-s} \varepsilon_{t-l} \right\} \\
 &\quad + 4E \left\{ \sum_{0 \leq s < l < \infty} a_s a_l \varepsilon_{t-s} \varepsilon_{t-l} \sum_{0 \leq s' < l' < \infty} a_{s'} a_{l'} \varepsilon_{j-s'} \varepsilon_{j-l'} \right\} \\
 &= \sum_{s=0}^{\infty} \sum_{s'=0}^{\infty} a_s^2 a_{s'}^2 E\{(\varepsilon_{t-s}^2 - c_0 \sigma^2)(\varepsilon_{t-s'}^2 - c_0 \sigma^2)\} \\
 &\quad + 2 \sum_{s=0}^{\infty} \sum_{0 \leq s' < l' < \infty} a_s^2 a_{s'} a_{l'} E\{(\varepsilon_{t-s}^2 - c_0 \sigma^2) \varepsilon_{j-s'}^2\} \varepsilon_{j-l'}^2 \\
 &\quad + 2 \sum_{s'=0}^{\infty} \sum_{0 \leq s < l < \infty} a_{s'}^2 a_s a_l E\{(\varepsilon_{t-s'}^2 - c_0 \sigma^2) \varepsilon_{t-s} \varepsilon_{t-l}\} \\
 &\quad + 4 \sum_{0 \leq s < l < \infty} \sum_{0 \leq s' < l' < \infty} a_s a_l a_{s'} a_{l'} E\{\varepsilon_{t-s} \varepsilon_{t-l} \varepsilon_{j-s'} \varepsilon_{j-l'}\}.
 \end{aligned}$$

Let $s' = j - t + s, l' = j - t + l$, similar to the proof of inequality (8), we have

$$\begin{aligned}
 & E[(e_t^2 - \sigma_\varepsilon^2)(e_j^2 - \sigma_\varepsilon^2)] \\
 &= \sum_{s=0}^{\infty} a_s^2 a_{j-t+s}^2 E(\varepsilon_{t-s}^2 - c_0 \sigma^2)^2 + 4 \sum_{0 \leq s < l < \infty} a_s a_l a_{j-t+s} a_{j-t+l} E(\varepsilon_{t-s} \varepsilon_{t-l})^2 < \infty.
 \end{aligned}$$

Hence

$$\begin{aligned}
 & \sum_{1 \leq t < j \leq k} |E[(e_t^2 - \sigma_\varepsilon^2)(e_j^2 - \sigma_\varepsilon^2)]| \\
 &= \sum_{1 \leq t < j \leq k} \left| E(\varepsilon_t^2 - c_0 \sigma^2)^2 \sum_{s=0}^{\infty} a_s^2 a_{j-t+s}^2 + 4 \sigma^2 \sum_{0 \leq s < l < \infty} a_s a_l a_{j-t+s} a_{j-t+l} \right| \\
 &\leq C \sigma^4 \sum_{t=1}^{k-1} \sum_{j=t+1}^k \sum_{s=0}^{\infty} a_s^2 a_{j-t+s}^2 + 4 \sigma^4 \sum_{t=1}^{k-1} \sum_{j=t+1}^k \sum_{0 \leq s < l < \infty} |a_s a_l a_{j-t+s} a_{j-t+l}| \\
 &= C \left\{ \sum_{t=1}^{k-1} \sum_{s=0}^{\infty} a_s^2 \sum_{j=t+1}^k a_{j-t+s}^2 + \sum_{t=1}^{k-1} \sum_{0 \leq s < l < \infty} |a_s a_l| \sum_{j=t+1}^k |a_{j-t+s} a_{j-t+l}| \right\} \\
 &\leq C \left\{ k \left(\sum_{s=0}^{\infty} a_s^2 \right)^2 + k \left(\sum_{s=0}^{\infty} a_s \right)^2 \left(\sum_{u=0}^{\infty} a_u^2 \right) \right\} \leq Ck. \quad \square
 \end{aligned}$$

Lemma 5 Let Y_1, Y_2, \dots, Y_n be a sample from model (1), assume that $\tilde{Y}_t = Y_t - \hat{\mu}_n, \hat{\mu}_n = \frac{1}{n} \sum_{t=1}^n Y_t$, if assumptions (A1)–(A4) hold, then for any $\varepsilon > 0$,

$$P\left(\max_{1 \leq k \leq n} \frac{1}{k} \left| \sum_{t=1}^k (\tilde{Y}_t^2 - E\tilde{Y}_t^2) \right| > \varepsilon \right) \leq \frac{C}{\sqrt{n}}.$$

Proof Note that

$$\sum_{t=1}^k \tilde{Y}_t^2 = \sum_{t=1}^k \sigma_i^2 e_t^2 - 2(\hat{\mu}_n - \mu) \sum_{t=1}^k \sigma_i^2 e_t^2 + k(\hat{\mu}_n - \mu)^2,$$

where $e_t = (Y_t - \mu)/\sigma_i$, then

$$\begin{aligned}
 & P\left(\max_{1 \leq k \leq n} \frac{1}{k} \left| \sum_{t=1}^k (\tilde{Y}_t^2 - E\tilde{Y}_t^2) \right| > \varepsilon \right) \\
 &\leq P\left(\max_{1 \leq k \leq n} \frac{\sigma_i^2}{k} \left| \sum_{t=1}^k (e_t^2 - Ee_t^2) \right| > \frac{\varepsilon}{3} \right) + P\left(|(\hat{\mu}_n - \mu)^2 - E(\hat{\mu}_n - \mu)^2| > \frac{\varepsilon}{3} \right) \\
 &\quad + P\left(\max_{1 \leq k \leq n} \frac{\sigma_i^2}{k} \left| 2(\hat{\mu}_n - \mu) \sum_{t=1}^k e_t^2 - 2E(\hat{\mu}_n - \mu) \sum_{t=1}^k e_t^2 \right| > \frac{\varepsilon}{3} \right) \\
 &=: J_1 + J_2 + J_3.
 \end{aligned}$$

Applying Lemma 4, it is easy to see that $J_1 \leq C/\sqrt{n}$. From Markov's inequality, J_2 is bounded by

$$\begin{aligned} J_2 &= P\left(\sigma_i^2 \left| \left(\frac{1}{n} \sum_{t=1}^n e_t\right)^2 - E\left(\frac{1}{n} \sum_{t=1}^n e_t\right)^2 \right| > \frac{\varepsilon}{3}\right) \\ &\leq \frac{3\sigma_i^2}{\varepsilon} E\left| \left(\frac{1}{n} \sum_{t=1}^n e_t\right)^2 - E\left(\frac{1}{n} \sum_{t=1}^n e_t\right)^2 \right| \\ &\leq \frac{6\sigma_i^2}{\varepsilon} E\left(\frac{1}{n} \sum_{t=1}^n e_t\right)^2 \leq \frac{C}{n}. \end{aligned}$$

Now, we will show that $J_3 \leq C/\sqrt{n}$. By Cauchy–Schwarz's inequality, it follows

$$\begin{aligned} &\max_{1 \leq k \leq n} E\left| (\hat{\mu}_n - \mu) \frac{1}{k} \sum_{t=1}^k \sigma_i e_t \right| \\ &\leq \sigma_i^2 \max_{1 \leq k \leq n} \left(E\left(\frac{1}{n} \sum_{t=1}^n e_t\right)^2 \right)^{1/2} \max_{1 \leq k \leq n} \left(E\left(\frac{1}{k} \sum_{t=1}^k e_t\right)^2 \right)^{1/2} \leq \frac{C}{\sqrt{n}}. \end{aligned}$$

Therefore

$$\begin{aligned} &P\left(\max_{1 \leq k \leq n} \frac{1}{k} \left| 2(\hat{\mu}_n - \mu) \sum_{t=1}^k \sigma_i e_t \right| > \frac{\varepsilon}{6}\right) \\ &\leq P\left(\max_{1 \leq k \leq n} \left| \left(\sum_{t=1}^k e_t\right)^2 + \frac{1}{k} \left(\sum_{t=1}^k e_t^2\right) \right| > \frac{\varepsilon}{6\sigma_i^2}\right) \\ &\leq P\left(\max_{1 \leq k \leq n} \left(\sum_{t=1}^k e_t\right)^2 > \frac{\varepsilon}{12\sigma_i^2}\right) + P\left(\max_{1 \leq k \leq n} \frac{1}{k} \left(\sum_{t=1}^k e_t^2\right) > \frac{\varepsilon}{12\sigma_i^2}\right) \\ &\leq \frac{c_1}{\sqrt{n}} + \frac{c_2}{n} \leq \frac{C}{\sqrt{n}}. \end{aligned}$$

Thus the proof of Lemma 5 is completed. □

Proof of Theorem 1 Let $\delta_0 = (\sigma_i^2 - \sigma_{i-1}^2) \sum_{j=0}^{\infty} a_j^2$, for $t_i \leq t_i^*$, we have

$$\begin{aligned} ER(t_i) &= \frac{(t_i - t_{i-1})(t_{i+1} - t_i)}{(t_{i+1} - t_{i-1})^2} \left| E\left\{ \frac{1}{t_i - t_{i-1}} \sum_{t=t_{i-1}+1}^{t_i} \tilde{Y}_t^2 - \frac{1}{t_{i+1} - t_i} \sum_{t=t_i+1}^{t_{i+1}} \tilde{Y}_t^2 \right\} \right| \\ &= \frac{(t_i - t_{i-1})(t_{i+1} - t_i)}{(t_{i+1} - t_{i-1})^2} \left| E\left\{ \frac{1}{t_i - t_{i-1}} \sum_{t=t_{i-1}+1}^{t_i} \tilde{Y}_t^2 - \frac{1}{t_{i+1} - t_i} \sum_{t=t_i+1}^{t_i^*} \tilde{Y}_t^2 - \frac{1}{t_{i+1} - t_i} \sum_{t=t_i^*+1}^{t_{i+1}} \tilde{Y}_t^2 \right\} \right| \\ &= \frac{(t_i - t_{i-1})(t_{i+1} - t_i)}{(t_{i+1} - t_{i-1})^2} \left| \sigma_{i-1}^2 \sum_{j=0}^{\infty} a_j^2 - \frac{t_i^* - t_i}{t_{i+1} - t_i} \sigma_{i-1}^2 \sum_{j=0}^{\infty} a_j^2 - \frac{t_{i+1} - t_i^*}{t_{i+1} - t_i} \sigma_i^2 \sum_{j=0}^{\infty} a_j^2 \right| \end{aligned}$$

$$\begin{aligned}
 &= \frac{(t_i - t_{i-1})(t_{i+1} - t_i)}{(t_{i+1} - t_{i-1})^2} \frac{(t_{i+1} - t_i^*)}{t_{i+1} - t_i} |\delta_0| \\
 &= \frac{(t_i - t_{i-1})(t_{i+1} - t_i^*)}{(t_{i+1} - t_{i-1})^2} |\delta_0|.
 \end{aligned}$$

Similarly, for $t_i \geq t_i^*$,

$$\begin{aligned}
 &ER(t_i) \\
 &= \frac{(t_i - t_{i-1})(t_{i+1} - t_i)}{(t_{i+1} - t_{i-1})^2} \left| E \left\{ \frac{1}{t_i - t_{i-1}} \sum_{t=t_{i-1}+1}^{t_i^*} \tilde{Y}_t^2 + \frac{1}{t_i - t_{i-1}} \sum_{t=t_i^*+1}^{t_i} \tilde{Y}_t^2 - \frac{1}{t_{i+1} - t_i} \sum_{t=t_i^*+1}^{t_{i+1}} \tilde{Y}_t^2 \right\} \right| \\
 &= \frac{(t_i - t_{i-1})(t_{i+1} - t_i)}{(t_{i+1} - t_{i-1})^2} \left| \frac{t_i^* - t_i}{t_{i+1} - t_i} \sigma_{i-1}^2 \sum_{j=0}^{\infty} a_j^2 + \frac{t_i - t_i^*}{t_i - t_{i-1}} \sigma_i^2 \sum_{j=0}^{\infty} a_j^2 - \sigma_i^2 \sum_{j=0}^{\infty} a_j^2 \right| \\
 &= \frac{(t_{i+1} - t_i)(t_i^* - t_{i-1})}{(t_{i+1} - t_{i-1})^2} |\delta_0|.
 \end{aligned}$$

Note that $ER(t_i)$ is increasing for $t_i \leq t_i^*$ decreasing while $t_i \geq t_i^*$, thus the maximum of $ER(t_i)$ is

$$|ER(t_i^*)| = \frac{(t_{i+1} - t_i^*)(t_i^* - t_{i-1})}{(t_{i+1} - t_{i-1})^2} |\delta_0|.$$

By direct calculation, it follows

$$|ER(t_i^*)| - |ER(t_i)| \geq \frac{(t_i^* - t_{i-1}) \wedge (t_{i+1} - t_i^*)}{(t_{i+1} - t_{i-1})^2} |t_i - t_i^*| |\delta_0| \geq \frac{C_0}{n} |t_i - t_i^*| = C_0 |\tau_i - \tau_i^*|.$$

In order to prove Theorem 1, it is desired to show that, for any $\varepsilon > 0$,

$$P(\|\hat{\boldsymbol{\tau}} - \boldsymbol{\tau}^*\|_{\infty} \geq \varepsilon) \rightarrow 0.$$

Since $R(t_i) = R(t_i) - ER(t_i) + (ER(t_i) - R(t_i^*)) + R(t_i^*)$ and $|R(t_i^*) - ER(t_i^*)| \leq \max_{t \in A_{n,r}^{\delta_n}} |R(t_i) - ER(t_i)|$, then

$$\begin{aligned}
 |R(t_i)| - |R(t_i^*)| &\leq |R(t_i) - ER(t_i)| + |R(t_i^*) - ER(t_i^*)| + ER(t_i) - ER(t_i^*) \\
 &\leq 2 \max_{t \in A_{n,r}^{\delta_n}} |R(t_i) - ER(t_i)| + ER(t_i) - ER(t_i^*).
 \end{aligned}$$

Define $\Lambda_{n,r} = \{\mathbf{t} \in A_{n,r}^{\delta_n}, \|\mathbf{t} - \mathbf{t}^*\|_{\infty} \geq n\varepsilon\}$, then

$$\begin{aligned}
 P(\|\hat{\boldsymbol{\tau}} - \boldsymbol{\tau}^*\|_{\infty} \geq \varepsilon) &\leq P\left(\max_{\mathbf{t} \in \Lambda_{n,r}} \sum_{i=1}^r \{|R(t_i)| - |R(t_i^*)|\} \geq 0\right) \\
 &\leq P\left(2 \max_{\mathbf{t} \in \Lambda_{n,r}} \sum_{i=1}^r \{|R(t_i) - ER(t_i)|\} - \sum_{i=1}^r C_0 \|\hat{\boldsymbol{\tau}} - \boldsymbol{\tau}^*\|_{\infty} \geq 0\right) \\
 &\leq P\left(\max_{1 \leq i \leq r} |R(t_i) - ER(t_i)| \geq \delta\right), \tag{9}
 \end{aligned}$$

where $\delta = C_0\varepsilon/2$ is an arbitrarily small positive number. According to the definition of $ER(t_i)$, one can see that

$$\begin{aligned} & \max_{1 \leq i \leq r} |R(t_i) - ER(t_i)| \\ &= \max_{1 \leq t_{i-1} < t_i < t_{i+1} \leq n} \frac{(t_i - t_{i-1})(t_{i+1} - t_i)}{(t_{i+1} - t_{i-1})^2} \left| \left\{ \frac{1}{t_i - t_{i-1}} \sum_{t=t_{i-1}+1}^{t_i} (\tilde{Y}_t^2 - E\tilde{Y}_t^2) \right. \right. \\ & \quad \left. \left. - \frac{1}{t_{i+1} - t_i} \sum_{t=t_i+1}^{t_{i+1}} (\tilde{Y}_t^2 - E\tilde{Y}_t^2) \right\} \right| \\ &\leq \max_{1 \leq t_{i-1} < t_i \leq n} \frac{1}{t_i - t_{i-1}} \left| \sum_{t=t_{i-1}+1}^{t_i} (\tilde{Y}_t^2 - E\tilde{Y}_t^2) \right| \\ & \quad + \max_{1 \leq t_i < t_{i+1} \leq n} \frac{1}{t_{i+1} - t_i} \left| \sum_{t=t_i+1}^{t_{i+1}} (\tilde{Y}_t^2 - E\tilde{Y}_t^2) \right|. \end{aligned} \tag{10}$$

From (9) and (10), the proof of Theorem 1 will be completed by showing

$$P\left(\max_{1 \leq t_{i-1} < t_i \leq n} \frac{1}{t_i - t_{i-1}} \left| \sum_{t=t_{i-1}+1}^{t_i} (\tilde{Y}_t^2 - E\tilde{Y}_t^2) \right| > \delta\right) \rightarrow 0, \quad n \rightarrow \infty, \tag{11}$$

and

$$P\left(\max_{1 \leq t_i < t_{i+1} \leq n} \frac{1}{t_{i+1} - t_i} \left| \sum_{t=t_i+1}^{t_{i+1}} (\tilde{Y}_t^2 - E\tilde{Y}_t^2) \right| > \delta\right) \rightarrow 0, \quad n \rightarrow \infty. \tag{12}$$

Since Eq. (12) can be proved similarly as (11), we only consider Eq. (11), thus the proof of Theorem 1 is finished by taking $k = t_i - t_{i-1}$ in Lemma 5. \square

Proof of Theorem 2 Let θ be a constant in the interval $(0, 1)$. Denote $D_{n,r}^{M(n)} = \{t \in A_{n,r}^{\delta_n}, n\theta > \|t - t^*\|_\infty > M(n)\}$. By Theorem 1, we have

$$\begin{aligned} P(\|\hat{\tau} - \tau^*\|_\infty > M(n)/n) &\leq P(\|\hat{\tau} - \tau^*\|_\infty \geq \theta) + P(\theta > \|\hat{\tau} - \tau^*\|_\infty > M(n)/n) \\ &\leq \varepsilon + P\left(\max_{t \in D_{n,r}^{M(n)}} \sum_{i=1}^r \{|R(t_i)| - |R(t_i^*)|\} \geq 0\right). \end{aligned}$$

Without loss of generality, we assume that $\delta_0 < 0$. In view of the fact that $|x| \geq |y|$ is equivalent to (i) $x - y \geq 0$ and $x + y \geq 0$, or (ii) $x - y \leq 0$ and $x + y \leq 0$, then

$$\begin{aligned} & P\left(\max_{t \in D_{n,r}^{M(n)}} \sum_{i=1}^r \{|R(t_i)| - |R(t_i^*)|\} \geq 0\right) \\ &\leq P\left(\max_{t \in D_{n,r}^{M(n)}} \sum_{i=1}^r \{R(t_i) - R(t_i^*)\} \geq 0\right) + P\left(\max_{t \in D_{n,r}^{M(n)}} \sum_{i=1}^r \{R(t_i) + R(t_i^*)\} < 0\right) \\ &\leq P\left(\max_{t \in D_{n,r}^{M(n)}, t_i < t_i^*} \sum_{i=1}^r \{R(t_i) - R(t_i^*)\} \geq 0\right) + P\left(\max_{t \in D_{n,r}^{M(n)}, t_i \geq t_i^*} \sum_{i=1}^r \{R(t_i) - R(t_i^*)\} \geq 0\right) \end{aligned}$$

$$\begin{aligned}
 &+ P\left(\max_{\mathbf{t} \in D_{n,r}^{M(n)}} \sum_{i=1}^r \{R(t_i) + R(t_i^*)\} < 0\right) \\
 &=: T_1 + T_2 + T_3.
 \end{aligned}$$

For $t_i < t_i^*$, we have

$$\begin{aligned}
 &|R(t_i) - ER(t_i) - (R(t_i^*) - ER(t_i^*))| \\
 &\leq \frac{C_1 |t_i - t_i^*|}{(t_{i+1} - t_{i-1})^2} \left| \sum_{t=t_{i-1}+1}^{t_i} (\tilde{Y}_t^2 - E\tilde{Y}_t^2) \right| + \frac{C_2 |t_i - t_i^*|}{(t_{i+1} - t_{i-1})^2} \left| \sum_{t=t_i^*+1}^{t_{i+1}} (\tilde{Y}_t^2 - E\tilde{Y}_t^2) \right| \\
 &+ \frac{C_3}{t_{i+1} - t_{i-1}} \left| \sum_{t=t_{i-1}+1}^{t_i^*} (\tilde{Y}_t^2 - E\tilde{Y}_t^2) \right| + \frac{C_4}{t_{i+1} - t_{i-1}} \left| \sum_{t=t_{i-1}+1}^{t_i^*} (\tilde{Y}_t^2 - E\tilde{Y}_t^2) \right|.
 \end{aligned}$$

Since $\delta_0 < 0$, $ER(t_i) \geq 0$, then for $1 \leq i \leq r$,

$$\begin{aligned}
 T_1 &\leq P\left(\bigcup_{\mathbf{t} \in D_{n,r}^{M(n)}, t_i < t_i^*} \left\{ \sum_{i=1}^r |R(t_i) - ER(t_i) - (R(t_i^*) - ER(t_i^*))| \geq \sum_{i=1}^r (ER(t_i^*) - ER(t_i)) \right\}\right) \\
 &\leq P\left(\bigcup_{\mathbf{t} \in D_{n,r}^{M(n)}, t_i < t_i^*} \left\{ \sum_{i=1}^r |R(t_i) - ER(t_i) - (R(t_i^*) - ER(t_i^*))| \right. \right. \\
 &\quad \left. \left. \geq \sum_{i=1}^r C_0 (t_i^* - t_i) / (t_{i+1} - t_{i-1}) \right\}\right) \\
 &\leq P\left(\max_{n\delta_n \leq t_i - t_{i-1} \leq n} \frac{1}{t_{i+1} - t_{i-1}} \left| \sum_{t=t_{i-1}+1}^{t_i} (\tilde{Y}_t^2 - E\tilde{Y}_t^2) \right| \geq C_1\right) \\
 &+ P\left(\max_{n\delta_n \leq t_i - t_{i-1} \leq n} \frac{1}{t_{i+1} - t_{i-1}} \left| \sum_{t=t_i^*+1}^{t_{i+1}} (\tilde{Y}_t^2 - E\tilde{Y}_t^2) \right| \geq C_2\right) \\
 &+ P\left(\max_{M(n) \leq t_i^* - t_i \leq \theta n} \frac{1}{t_i^* - t_i} \left| \sum_{t=t_{i-1}+1}^{t_i^*} (\tilde{Y}_t^2 - E\tilde{Y}_t^2) \right| \geq C_3\right) \\
 &+ P\left(\max_{M(n) \leq t_i^* - t_i \leq \theta n} \frac{1}{t_i^* - t_i} \left| \sum_{t=t_{i-1}+1}^{t_i^*} (\tilde{Y}_t^2 - E\tilde{Y}_t^2) \right| \geq C_4\right) \\
 &=: Q_1 + Q_2 + Q_3 + Q_4.
 \end{aligned}$$

In the view of $n\delta_n \rightarrow \infty$ and $M(n) \rightarrow \infty$, Lemma 5 yields

$$Q_i \rightarrow 0, \quad i = 1, 2, 3, 4.$$

Thus $T_1 \rightarrow 0$. We can treat T_2 analogously as T_1 , hence $T_2 \rightarrow 0$.

To complete the proof of Theorem 2, it is sufficient to show $T_3 \rightarrow 0$. Since $R(t_i) + R(t_i^*) \leq 0$ implies that $R(t_i) - ER(t_i) + R(t_i^*) - ER(t_i^*) \leq -ER(t_i) - ER(t_i^*) \leq -ER(t_i^*)$, we obtain

$$R(t_i) - ER(t_i) \leq -ER(t_i^*)/2 \quad \text{or} \quad R(t_i^*) - ER(t_i^*) \leq -ER(t_i^*)/2. \tag{13}$$

According to $ER(t_i^*) \geq 0$ ($\delta_0 < 0$), inequality (13) implies that

$$|R(t_i) - ER(t_i)| \geq ER(t_i^*)/2 \quad \text{or} \quad |R(t_i^*) - ER(t_i^*)| \geq ER(t_i^*)/2.$$

Hence

$$\begin{aligned} T_3 &\leq P\left(\bigcup_{t \in D_{n,r}^{M(n)}} \sum_{i=1}^r \{|R(t_i) - ER(t_i)| \geq ER(t_i^*)/2\}\right) \\ &\quad + P\left(\bigcup_{t \in D_{n,r}^{M(n)}} \sum_{i=1}^r \{|R(t_i^*) - ER(t_i^*)| \geq ER(t_i^*)/2\}\right) \\ &\leq 2P\left(\bigcup_{t \in D_{n,r}^{M(n)}} \sum_{i=1}^r |R(t_i) - ER(t_i)| \geq ER(t_i^*)/2\right) \\ &\leq 2rP\left(\max_{1 \leq i \leq r} |R(t_i) - ER(t_i)| \geq ER(t_i^*)/2\right). \end{aligned}$$

Combining (4), (5), (6), and (7), we get

$$P\left(\max_{1 \leq i \leq r} |R(t_i) - ER(t_i)| \geq ER(t_i^*)/2\right) \rightarrow 0.$$

Thus $T_3 \rightarrow 0$. This completes the proof of Theorem 2. □

5 Conclusions

In this study, we consider the multiple variance change model and develop a CUSUM-type methodology for change points estimation. We assume the errors from linear processes under NSD. The weak convergence rate of the change points estimation has been established. Recently, Qin et al. [15] and Shi et al. [30] concentrated on the strong convergence of the CUSUM-type estimator, we believe that the proposed estimation in this paper also has the strong convergent property. Additionally, investigating the change points estimation with the unknown number of the change points is an interesting topic, and this is our next work.

Acknowledgements

The authors would like to thank everyone for help.

Funding

This paper is supported by the Natural Science Foundation of China (No. 61374183); the Postgraduate Research & Practice Innovation Program of Jiangsu Province (No. KYCX19_0149).

Availability of data and materials

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

All authors contributed equally to the writing of this paper. All authors read and approved the final manuscript.

Author details

¹State Key Laboratory of Mechanics and Control of Mechanical Structures, Institute of Nano Science and Department of Mathematics, Nanjing University of Aeronautics and Astronautics, Nanjing, China. ²Department of Information Science and Technology, Donghua University, Shanghai, China.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Received: 15 April 2019 Accepted: 6 August 2019 Published online: 15 August 2019

References

1. Basseville, M.: Detecting changes in signals and systems—a survey. *Automatica* **24**, 309–326 (1988)
2. Chen, J., Gupta, A.: Testing and locating variance change points with application to stock prices. *J. Am. Stat. Assoc.* **92**, 739–747 (1997)
3. Hawkins, S.J., Southward, A.J., Genner, M.J.: Detection of environmental change in a marine ecosystem—evidence from the western English channel. *Sci. Total Environ.* **310**, 245–256 (2003)
4. Sparks, R., Keighley, T., Muscatello, D.: Early warning CUSUM plans for surveillance of negative binomial daily disease counts. *J. Appl. Stat.* **37**, 1911–1930 (2010)
5. Ratnam, R., Goense, J.B., Nelson, M.E.: Change-point detection in neuronal spike train activity. *Neurocomputing* **52**, 849–855 (2003)
6. Lena, K., Go, A., Jutta, K.: Single and multiple change point detection in spike trains: comparison of different CUSUM methods. *Front. Syst. Neurosci.* **10**, 51 (2016)
7. Bai, J.: Least squares estimation of a shift in linear processes. *J. Time Ser. Anal.* **15**(5), 453–472 (1994)
8. Fearnhead, P.: Exact and efficient Bayesian inference for multiple change point problems. *Stat. Comput.* **16**(2), 203–213 (2006)
9. Zou, C., Yin, G., Feng, L., Wang, Z.: Nonparametric maximum likelihood approach to multiple change-point problems. *Ann. Stat.* **42**(3), 970–1002 (2014)
10. Matteson, D.S., James, N.A.: A nonparametric approach for multiple change point analysis of multivariate data. *J. Am. Stat. Assoc.* **109**(505), 334–345 (2014)
11. Haynes, K., Fearnhead, P., Eckley, I.A.: A computationally efficient nonparametric approach for change point detection. *Stat. Comput.* **27**(5), 1293–1305 (2017)
12. Gombay, E., Horvath, L., Huskova, M.: Estimators and tests for change in variances. *Stat. Decis.* **14**, 145–159 (1996)
13. Wang, L.H., Wang, J.D.: Change of variance problem for linear processes with long memory. *Stat. Pap.* **47**, 279–298 (2006)
14. Zhao, W., Xia, Z., Tian, Z.: Ratio test to detect change in the variance of linear process. *Statistics* **45**, 189–198 (2011)
15. Qin, R., Liu, W., Tian, Z.: A strong convergence rate of estimator of variance change in linear processes and its applications. *Statistics* **51**(2), 314–330 (2016)
16. Inclan, C., Tiao, G.C.: Use of cumulative of sums squares for retrospective detection of changes of variance. *J. Am. Stat. Assoc.* **89**, 913–923 (1994)
17. Lavielle, M.: Detection of multiple changes in a sequence of dependent variables. *Stoch. Process. Appl.* **83**, 79–102 (1999)
18. Li, Y.X., Zhao, L.X.: Rate of convergence for multiple change-points estimation of moving-average processes. *Appl. Math. J. Chin. Univ. Ser. B* **20**(4), 416–422 (2005)
19. Hinoveanu, L.C., Leisen, F., Villa, C.: Bayesian loss-based approach to change point analysis. *Comput. Stat. Data Anal.* **129**, 61–78 (2019)
20. Hu, T.Z.: Negatively superadditive dependence of random variables with applications. *Chinese J. Appl. Probab. Statist.* **16**(2), 133–144 (2000)
21. Wang, X.J., Shen, A.T., Chen, Z.Y.: Complete convergence for weighted sums of NSD random variables and its application in the EV regression model. *Test* **24**, 166–184 (2015)
22. Christofides, T.C., Vaggelatos, E.: A connection between super-modular ordering and positive/negative association. *J. Multivar. Anal.* **88**(1), 138–151 (2004)
23. Eghbal, N., Amini, M., Bozorgnia, A.: Some maximal inequalities for quadratic forms of negative super-additive dependence random variables. *Stat. Probab. Lett.* **80**, 587–591 (2010)
24. Shen, Y., Wang, X.J., Yang, W.Z., Hu, S.H.: Almost sure convergence theorem and strong stability for weighted sums of NSD random variables. *Acta Math. Sin.* **29**, 743–756 (2013)
25. Wu, Y., Wang, X.J., Hu, S.H.: Complete convergence for arrays of rowwise negatively super-additive dependent random variables and its applications. *Appl. Math. J. Chin. Univ. Ser. A* **31**, 439–457 (2016)
26. Wang, X.J., Deng, X., Zheng, L.L., Hu, S.H.: Complete convergence for arrays of rowwise negatively super-additive dependent random variables and its applications. *Statistics* **48**, 834–850 (2014)
27. Yu, Y.C., Hu, H.C., Liu, L., Huang, S.Y.: M-test in linear models with negatively super-additive dependent errors. *J. Inequal. Appl.* **2017**, 235 (2017)
28. Wang, X.J., Wu, Y., Hu, S.H.: Strong and weak consistency of LS estimators in the EV regression model with negatively super-additive dependent errors. *AStA Adv. Stat. Anal.* **102**, 41–65 (2018)
29. Yu, Y.C., Liu, X.S., Liu, L., Liu, W.S.: On adaptivity of wavelet thresholding estimators with negatively super-additive dependent noise. *Math. Slovaca* (2019, Accepted)
30. Shi, X., Wu, Y., Miao, B.: Strong convergence rate of estimators of change point and its application. *Comput. Stat. Data Anal.* **53**, 990–998 (2009)
31. Auger, I.E., Lawrence, C.E.: Algorithms for the optimal identification of segment neighborhoods. *Bull. Math. Biol.* **51**(1), 39–54 (1989)
32. Killick, R., Fearnhead, P., Eckley, I.A.: Optimal detection of change points with a linear computational cost. *J. Am. Stat. Assoc.* **107**, 1590–1598 (2012)
33. Fryzlewicz, P.: Wild binary segmentation for multiple change-point detection. *Ann. Stat.* **42**, 2243–2281 (2014)
34. Killick, R., Eckley, I.A.: Changepoint: an R package for change-point analysis. *J. Stat. Softw.* **58**(3), 1–19 (2014)
35. Korkas, K., Fryzlewicz, P.: Multiple change-point detection for non-stationary time series using wild binary segmentation. *Stat. Sin.* **27**, 287–311 (2017)