Qin and Liu *Journal of Inequalities and Applications* (2015) 2015:48
DOI 10.1186/s13660-015-0569-8

Journal of Inequalities and Applications
a SpringerOpen Journal

RESEARCH Open Access

A robust test for mean change in dependent observations

Ruibing Qin[1*] and Weiqi Liu[2]

*Correspondence:
rbqin@hotmail.com
[1]School of Mathematical Science, Shanxi University, Taiyuan, Shanxi 030006, P.R. China
Full list of author information is available at the end of the article

Abstract

A robust test based on the indicators of the data minus the sample median is proposed to detect the change in the mean of α-mixing stochastic sequences. The asymptotic distribution of the test is established under the null hypothesis that the mean μ remains as a constant. The consistency of the proposed test is also obtained under the alternative hypothesis that μ changes at some unknown time. Simulations demonstrate that the test behaves well for heavy-tailed sequences.

MSC: Primary 62G08; 62M10

Keywords: change point; median; robust test; consistency

1 Introduction

The problem of a mean change at an unknown location in a sequence of observations has received considerable attention in the literature. For example, Sen and Srivastava [1], Hawkins [2], Worsley [3] proposed tests for a change in the mean of normal series. Yao [4] proposed some estimators of the change point in a sequence of independent variables. For serially correlated data, Bai [5] considered the estimation of the change point in linear processes. Horváth and Kokoszka [6] gave an estimator of the change point in a long-range dependent series.

Most of the existing results in the statistic and econometric literature have concentrated on the case that the innovations are Gaussian. In fact, many economic and financial time series can be very heavy-tailed with infinite variances; see *e.g.* Mittnik and Rachev [7]. Therefore, the series with infinite-variance innovations aroused a great deal of interest of researchers in statistics, such as Phillips [8], Horváth and Kokoskza [9], Han and Tian [10, 11]. It is more efficient to construct robust procedures for heavy-tailed innovations, such as the M procedures in Hušková [12, 13] and the references therein. De Jong *et al.* [14] proposed a robust KPSS test based on the 'sign' of the data minus the sample median, which behaves rather well for heavy-tailed series. In this paper, we shall construct a robust test for the mean change in a sequence.

The rest of this paper is organized as follows: Section 2 introduces the models and necessary assumptions for the asymptotic properties. Section 3 gives the asymptotic distribution and the consistency of the test proposed in the paper. In Section 4, we shall show the statistical behaviors through simulations. All mathematical proofs are collected in the Appendix.

2 Model and assumptions

In the following, we concentrate ourselves on the model as follows:

$$Y_t = \mu(t) + X_t, \quad \mu(t) = \begin{cases} \mu_1, & t \le k_0, \\ \mu_2, & t > k_0, \end{cases} \tag{1}$$

where k_0 is the change point.

In order to obtain the weak convergence and the convergence rate, $X(t)$ satisfies the following.

Assumption 1

1. The X_j are strictly stationary random variables, and $\tilde{\mu}$ is the unique population median of $\{X_t, 1 \le t \le T\}$.
2. The X_j are strong (α-) mixing, and for some finite $r > 2$ and $C > 0$, and for some $\eta > 0$, $\alpha(m) \le Cm^{-r/(r-2)-\eta}$.
3. $X_j - \tilde{\mu}$ has a continuous density $f(x)$ in a neighborhood $[-\eta, \eta]$ of 0 for some $\eta > 0$, and $\inf_{x\in[-\eta,\eta]} f(x) > 0$.
4. $\sigma^2 \in (0, \infty)$, where σ^2 is defined as follows:

$$\sigma^2 = \lim_{T\to\infty} E\left(T^{-1/2}\sum_{t=1}^{T} \operatorname{sgn}(X_t - \tilde{\mu})\right)^2.$$

To derive the CLT of sign-transformed data, we need a kernel estimator, so we make the following assumption on the kernel function.

Assumption 2

1. $k(\cdot)$ satisfies $\int_{-\infty}^{\infty} |\psi(\xi)|\, d\xi < \infty$, where

$$\psi(\xi)\, d\xi = (2\pi)^{-1} \int_{-\infty}^{\infty} k(x) \exp(-it\xi)\, dx.$$

2. $k(x)$ is continuous at all but a finite number of points, $k(x) = k(-x)$, $|k(x)| \le l(x)$ where $l(x)$ is nondecreasing and $\int_0^{\infty} |l(x)|\, dx \le \infty$, and $k(0) = 1$.
3. $\gamma_T/T \to 0$, and $\gamma_T \to \infty$ as $T \to \infty$.

Remark 1 De Jong *et al.* [14] test the stationarity of a sequence under Assumption 1. We detect change in the mean of a sequence, so Assumption 1 holds under the null hypothesis and the alternative one. Since there is no moment condition for X_t in Assumption 1, even Cauchy series are allowed. The α-mixing sequences can include many time series, such as autoregressive or heteroscedastic series under some conditions. Assumption 2 allows some choices such as the Bartlett, quadratic spectral, and Parzen kernel functions.

3 Main results

Let $m_T = \operatorname{med}\{Y_1, \dots, Y_T\}$. Then we transform the data $Y_1, \dots, Y_T$ into the indicator data $\operatorname{sgn}(Y_t - m_T)$, where $\operatorname{sgn}(x) = 1$ if $x > 0$, $\operatorname{sgn}(x) = -1$ if $x < 0$, $\operatorname{sgn}(x) = 0$ if $x = 0$. Based on these indicator data, De Jong *et al.* [14] replace $\hat{\epsilon}_t = Y_t - \bar{Y}_T$ with $\operatorname{sgn}(Y_t - m_T)$ in the usual KPSS test and their simulations show that the new KPSS test exhibits some robustness for the heavy-tailed series.

The popularly used test to detect a mean change is based on the CUSUM type as follows:

$$\Xi_T(\tau) = \frac{[T\tau][T(1-\tau)]}{T^2}\left\{\frac{1}{[T\tau]}\sum_{t=1}^{[T\tau]} Y_t - \frac{1}{[T(1-\tau)]}\sum_{t=[T\tau]+1}^{T} Y_t\right\}. \tag{2}$$

We rewrite $\Xi_T(\tau)$ under H_0 as

$$\Xi_T(\tau) = \frac{[T\tau][T(1-\tau)]}{T^2}\left\{\frac{1}{[T\tau]}\sum_{t=1}^{[T\tau]} (Y_t - \bar{Y}_T) - \frac{1}{[T(1-\tau)]}\sum_{t=[T\tau]+1}^{T} (Y_t - \bar{Y}_T)\right\}, \tag{3}$$

According to the idea of De Jong *et al.* [14], replace $\hat{\epsilon}_t = Y_t - \bar{Y}_T$ with $\operatorname{sgn}(Y_t - m_T)$ in (3); then we get a robust version of CUSUM as follows:

$$\Xi_T = \frac{[T\tau][T(1-\tau)]}{T^2}\left\{\frac{1}{[T\tau]}\sum_{t=1}^{[T\tau]} \operatorname{sgn}(Y_t - m_T) - \frac{1}{[T(1-\tau)]}\sum_{t=[T\tau]+1}^{T} \operatorname{sgn}(Y_t - m_T)\right\}. \tag{4}$$

Then the test statistic proposed in this paper is

$$\Gamma_T = T^{1/2}\sigma^{-1}\max_{\tau\in(0,1)}\left|\Xi_T(\tau)\right|. \tag{5}$$

Under Assumptions 1, 2, we can obtain two asymptotic results as follows.

Theorem 1 *If Assumptions* 1, 2 *hold, then under the null hypothesis* H_0, *we have*

$$T^{1/2}\sigma^{-1}\max_{\tau\in(0,1)}|\Xi_T| \quad\Longrightarrow\quad \sup_{\tau\in(0,1)}\left|W(\tau) - \tau W(1)\right|, \quad \text{as } T\to\infty, \tag{6}$$

where '⟹' stands for the weak convergence.

Under the alternative hypothesis H_1, a change in the mean happens at some time, we denote the time as $[T\tau_0]$. Let $F(\cdot)$ be the common distribution function of X_t and μ^* be the median of

$$F^*(\cdot) = \tau_0 F(\cdot - \mu_1) + (1-\tau_0)F(\cdot - \mu_2). \tag{7}$$

Then we have the following.

Theorem 2 *If Assumptions* 1, 2 *hold, then under the alternative hypothesis* H_1, *we have*

$$\max_{\tau\in(0,1)}\left|\Xi_T(\tau)\right| \xrightarrow{P} \tau_0(1-\tau_0)|\Delta|, \tag{8}$$

where $\Delta = F(\mu^* - \mu_1) - F(\mu^* - \mu_2)$.

Remark 2 By Theorem 1, we reject H_0 if $\Gamma_T > c_p$, where the critic value c_p is the $(1-p)$ quantile of the Kolmogorov-Smirnov distribution. By Theorem 2, Γ_T is consistent asymptotically as the sample size $T \xrightarrow{P} \infty$.

In order to apply the test in (5), we employ the HAC estimator instead of the unknown σ^2 as

$$\hat{\sigma}_T^2 = T^{-1}\sum_{i=1}^{T}\sum_{j=1}^{T} k\bigl((i-j)/\gamma_T\bigr)\operatorname{sgn}(Y_i - m_T)\operatorname{sgn}(Y_j - m_T), \tag{9}$$

then the following theorem proves two results of the estimator $\hat{\sigma}_T^2$ under H_0 and H_A, respectively.

Theorem 3 (i) *Assuming that the conditions of Theorem* 1 *hold, then we have, as* $T \to \infty$,

$$\hat{\sigma}_T^2 \xrightarrow{P} \sigma^2. \tag{10}$$

(ii) *Assuming that the conditions of Theorem* 2 *hold, then we have, as* $T \to \infty$,

$$\hat{\sigma}_T^2 \xrightarrow{P} \sigma_2^2, \tag{11}$$

where σ_2^2 *is defined as follows*:

$$\sigma_2^2 = \lim_{T\to\infty} E\left(T^{-1/2}\sum_{t=1}^{T}\operatorname{sgn}\bigl(Y_t - \mu^*\bigr)\right)^2.$$

4 Simulation and empirical application

4.1 Simulation

In this section, we present Monte Carlo simulations to investigate the size and the power of the robust CUSUM and the ordinary CUSUM tests. Since a lot of information has been lost during the inference by using the indicator data instead of the original data, so we are concerned whether the indicator CUSUM test is robust to the heavy-tailed sequences; moreover, we may ask: how large is the loss in power in using indicators when the data has a nearly normal distribution? The HAC estimator $\hat{\sigma}^2$ in the robust CUSUM test is a kernel estimator, so it is important to analyze whether the performance is affected by the choice of the kernel function $k(\cdot)$ and the bandwidth γ_T.

We consider the model as follows:

$$Y_t = \begin{cases} 0 + X_t, & t \le T\tau_0, \\ \mu_2 + X_t, & t > T\tau_0, \end{cases} \tag{12}$$

X_t is an autoregressive process $X_t = 0.5X_{t-1} + e_t$, where the $\{e_t\}$ are independent noise generated by the program from JP Nolan. We vary the tail thickness of $\{e_t\}$ by the different characteristic indices $\alpha = 1.97, 1.83, 1.41, 1.14$, respectively. Accordingly the break times are $\tau_0 = 0.3, 0.5$, respectively. During the simulations, we adopt 1.358 as the asymptotic critical value of $\sup_{\tau\in(0,1)} |W(\tau) - \tau W(1)|$ at 95% for the various sample sizes $T = 300, 500, 1{,}000$.

First, we consider the size of the tests. Tables 1 and 2 report the results when σ^2 are estimated by the Bartlett kernel and the quadratic spectral kernel with the bandwidth $\gamma_T = [4(T/100)^{1/4}]$ and $\gamma_T = [8(T/100)^{1/4}]$, respectively, in 1,000 repetitions. From Tables 1 and 2, the ordinary CUSUM test based on the Bartlett kernel has better sizes, however,

Table 1 The empirical levels of the robust CUSUM test and the CUSUM test for dependent innovations

	CUSUM			RCUSUM		
	***T* = 300**	***T* = 500**	***T* = 1,000**	***T* = 300**	***T* = 500**	***T* = 1,000**
The tests based on the Bartlett kernel function						
$\alpha = 1.97$	0.045	0.026	0.036	0.042	0.046	0.059
$\alpha = 1.83$	0.028	0.028	0.033	0.037	0.032	0.043
$\alpha = 1.41$	0.010	0.010	0.025	0.030	0.036	0.044
$\alpha = 1.14$	0.005	0.010	0.008	0.045	0.049	0.048
The tests based on the quadratic spectral kernel function						
$\alpha = 1.97$	0.471	0.491	0.489	0.068	0.048	0.050
$\alpha = 1.83$	0.428	0.462	0.478	0.062	0.077	0.063
$\alpha = 1.41$	0.458	0.449	0.486	0.066	0.072	0.053
$\alpha = 1.14$	0.474	0.476	0.507	0.083	0.073	0.055

The values in Table 1 are based on the bandwidth $\gamma_T = [4(T/100)^{1/4}]$.

Table 2 The empirical levels of the robust CUSUM test and the CUSUM test for dependent innovations

	CUSUM			RCUSUM		
	***T* = 300**	***T* = 500**	***T* = 1,000**	***T* = 300**	***T* = 500**	***T* = 1,000**
The tests based on the Bartlett kernel function						
$\alpha = 1.97$	0.028	0.032	0.034	0.034	0.033	0.046
$\alpha = 1.83$	0.019	0.032	0.023	0.034	0.037	0.037
$\alpha = 1.41$	0.009	0.013	0.021	0.035	0.038	0.048
$\alpha = 1.14$	0.004	0.008	0.01	0.038	0.036	0.047
The tests based on the quadratic spectral kernel function						
$\alpha = 1.97$	0.425	0.447	0.470	0.037	0.043	0.040
$\alpha = 1.83$	0.414	0.444	0.456	0.026	0.043	0.048
$\alpha = 1.41$	0.484	0.463	0.483	0.040	0.035	0.041
$\alpha = 1.14$	0.459	0.490	0.454	0.028	0.048	0.042

The values in Table 2 are based on the bandwidth $\gamma_T = [8(T/100)^{1/4}]$.

the one based on the quadratic spectral kernel has a severe problem of overrejection, so we can conclude that the choice of the kernel function has higher impact on the sizes of the two CUSUM tests than the selection of the bandwidth. Comparing the two tests based on the Bartlett kernel, the ordinary CUSUM test becomes underrejecting as the tail index α changes from 2 to 1, and the sizes of the robust test are closer to the nominal size 0.05. Furthermore, the size is closer to 0.05 as the sample size T increases, which is consistent with Theorem 1.

Now we shall show the power of the two tests through empirical powers. The empirical powers are calculated based on the rejection numbers of the null hypothesis H_0 in 1,000 repetitions when the alternative hypothesis H_1 holds. The results are included in Tables 3, 4, 5, 6. On the basis of Tables 3, 4, 5, 6, we can draw some conclusions. (i) The two CUSUM tests based on the Bartlett kernel and the quadratic spectral kernel become more powerful as the sample size T becomes larger. (ii) As the tail of the innovations gets heavier, the ordinary CUSUM test becomes less powerful, especially, the test hardly works, while the CUSUM test based on indicators is rather robust to the heavy-tailed innovations. (iii) The selection of the bandwidth has lower impact on the powers of the two CUSUM tests.

Finally, we consider the effects of the skewness in the innovations $\{e_t\}$ on the power of the proposed test through simulations. In order to obtain the results reported in Table 7, we take the $e(t)$ in the model (12) as chi square distributions with a freedom degree n =

Table 3 The empirical powers of the robust CUSUM test and the CUSUM test for dependent innovations

	CUSUM			RCUSUM		
	$T = 300$	$T = 500$	$T = 1{,}000$	$T = 300$	$T = 500$	$T = 1{,}000$
The change point $\tau_0 = 0.3$						
$\alpha = 1.97$	0.849	0.991	0.998	0.951	0.999	1.000
$\alpha = 1.83$	0.692	0.919	0.977	0.964	1.000	1.000
$\alpha = 1.41$	0.222	0.361	0.530	0.957	0.995	1.000
$\alpha = 1.14$	0.047	0.065	0.076	0.964	0.998	1.000
The change point $\tau_0 = 0.5$						
$\alpha = 1.97$	0.988	0.997	0.997	0.991	1.000	1.000
$\alpha = 1.83$	0.913	0.966	0.979	0.985	1.000	1.000
$\alpha = 1.41$	0.360	0.531	0.651	0.994	1.000	1.000
$\alpha = 1.14$	0.097	0.108	0.133	0.996	1.000	1.000
The change point $\tau_0 = 0.7$						
$\alpha = 1.97$	0.972	0.995	0.999	0.958	0.999	1.000
$\alpha = 1.83$	0.875	0.944	0.978	0.962	0.997	1.000
$\alpha = 1.41$	0.300	0.446	0.542	0.964	0.999	1.000
$\alpha = 1.14$	0.063	0.080	0.104	0.972	1.000	1.000

The values in Table 3 are based on the Bartlett kernel and the bandwidth $\gamma_T = [4(T/100)^{1/4}]$.

Table 4 The empirical powers of the robust CUSUM test and the CUSUM test for dependent innovations

	CUSUM			RCUSUM		
	$T = 300$	$T = 500$	$T = 1{,}000$	$T = 300$	$T = 500$	$T = 1{,}000$
The change point $\tau_0 = 0.3$						
$\alpha = 1.97$	0.348	0.848	0.995	0.921	1.000	1.000
$\alpha = 1.83$	0.241	0.676	0.953	0.931	0.993	1.000
$\alpha = 1.41$	0.111	0.242	0.409	0.944	0.997	0.997
$\alpha = 1.14$	0.029	0.056	0.080	0.943	1.000	1.000
The change point $\tau_0 = 0.5$						
$\alpha = 1.97$	0.931	0.995	0.997	0.993	1.000	1.000
$\alpha = 1.83$	0.796	0.954	0.985	0.989	1.000	1.000
$\alpha = 1.41$	0.285	0.456	0.605	0.990	1.000	1.000
$\alpha = 1.14$	0.057	0.088	0.106	0.989	1.000	1.000
The change point $\tau_0 = 0.7$						
$\alpha = 1.97$	0.937	0.997	0.997	0.949	1.000	1.000
$\alpha = 1.83$	0.783	0.926	0.969	0.934	1.000	1.000
$\alpha = 1.41$	0.238	0.373	0.553	0.938	0.997	1.000
$\alpha = 1.14$	0.046	0.068	0.094	0.948	0.997	1.000

The values in Table 4 are based on the Bartlett kernel and the bandwidth $\gamma_T = [8(T/100)^{1/4}]$.

1, 2 and 10, respectively. On the basis of the simulations, the skewness of the innovations affects the powers the two CUSUM test significantly.

4.2 Empirical application

In this section, we take an empirical application on a series of daily stock price of LBC (SHANDONG LUBEI CHEMICAL Co., LTD) in the Shanghai Stocks Exchange. The stock prices in the group are observed from July 1st, 2004 to December 30th, 2005 with samples of 367 observations (as shown in Figure 1) and can be found in http://stock.business.sohu.com. As in Figure 2, the logarithm sequence is seen to exhibit a number of 'outliers', which are a manifestation of their heavy-tailed distributions, see Wang *et al.* [15]; the data can be well fitted by stable sequences.

Table 5 The empirical powers of the robust CUSUM test and the CUSUM test for dependent innovations

	CUSUM			RCUSUM		
	***T* = 300**	***T* = 500**	***T* = 1,000**	***T* = 300**	***T* = 500**	***T* = 1,000**
The change point $\tau_0 = 0.3$						
$\alpha = 1.97$	0.979	1.000	1.000	0.869	0.964	0.999
$\alpha = 1.83$	0.957	0.995	0.996	0.847	0.957	0.994
$\alpha = 1.41$	0.824	0.882	0.917	0.729	0.855	0.963
$\alpha = 1.14$	0.644	0.672	0.652	0.574	0.753	0.895
The change point $\tau_0 = 0.5$						
$\alpha = 1.97$	0.998	0.999	1.000	0.939	0.983	1.000
$\alpha = 1.83$	0.982	0.994	0.992	0.915	0.979	0.998
$\alpha = 1.41$	0.802	0.826	0.889	0.805	0.929	0.996
$\alpha = 1.14$	0.604	0.593	0.646	0.670	0.819	0.943
The change point $\tau_0 = 0.7$						
$\alpha = 1.97$	0.993	1.000	1.000	0.873	0.961	0.996
$\alpha = 1.83$	0.736	0.773	0.845	0.820	0.947	0.999
$\alpha = 1.41$	0.736	0.773	0.845	0.717	0.867	0.972
$\alpha = 1.14$	0.570	0.556	0.594	0.577	0.731	0.878

The values in Table 5 are based on the quadratic spectral kernel and the bandwidth $\gamma_T = [4(T/100)^{1/4}]$.

Table 6 The empirical powers of the robust CUSUM test and the CUSUM test for dependent innovations

	CUSUM			RCUSUM		
	***T* = 300**	***T* = 500**	***T* = 1,000**	***T* = 300**	***T* = 500**	***T* = 1,000**
The change point $\tau_0 = 0.3$						
$\alpha = 1.97$	0.467	0.881	1.000	0.808	0.941	0.999
$\alpha = 1.83$	0.521	0.874	0.993	0.764	0.920	0.995
$\alpha = 1.41$	0.658	0.770	0.893	0.582	0.788	0.961
$\alpha = 1.14$	0.565	0.629	0.668	0.440	0.642	0.847
The change point $\tau_0 = 0.5$						
$\alpha = 1.97$	0.974	0.999	1.000	0.891	0.967	0.997
$\alpha = 1.83$	0.958	0.987	0.994	0.866	0.969	0.999
$\alpha = 1.41$	0.792	0.860	0.897	0.726	0.876	0.992
$\alpha = 1.14$	0.594	0.640	0.631	0.568	0.720	0.921
The change point $\tau_0 = 0.7$						
$\alpha = 1.97$	0.992	1.000	1.000	0.782	0.924	0.997
$\alpha = 1.83$	0.974	0.981	0.992	0.802	0.924	0.990
$\alpha = 1.41$	0.749	0.800	0.881	0.604	0.756	0.942
$\alpha = 1.14$	0.544	0.580	0.590	0.448	0.598	0.838

The values in Table 6 are based on the quadratic spectral kernel and the bandwidth $\gamma_T = [8(T/100)^{1/4}]$.

Table 7 The empirical powers of the two CUSUM test for the skewed dependent innovations

	CUSUM			RCUSUM		
	$\chi^2(1)$	$\chi^2(2)$	$\chi^2(10)$	$\chi^2(1)$	$\chi^2(2)$	$\chi^2(10)$
$\tau_0 = 0.3$	0.9400	0.6690	0.3550	0.0	0.6760	0.2090
$\tau_0 = 0.5$	0.9940	0.8130	0.4270	0.0350	0.8280	0.2880
$\tau_0 = 0.7$	0.9900	0.7140	0.3480	0.0150	0.7530	0.2250

The values in Table 7 are based on the Bartlett kernel and the bandwidth $\gamma_T = [4(T/100)^{1/4}]$.

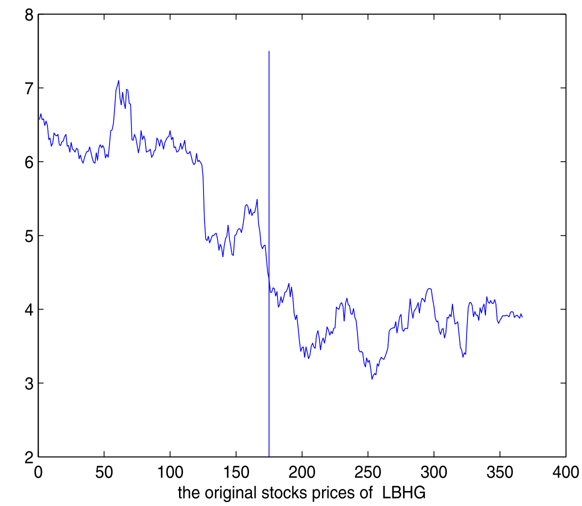

Figure 1 Stock prices of LBC in Shanghai Stock Exchange.

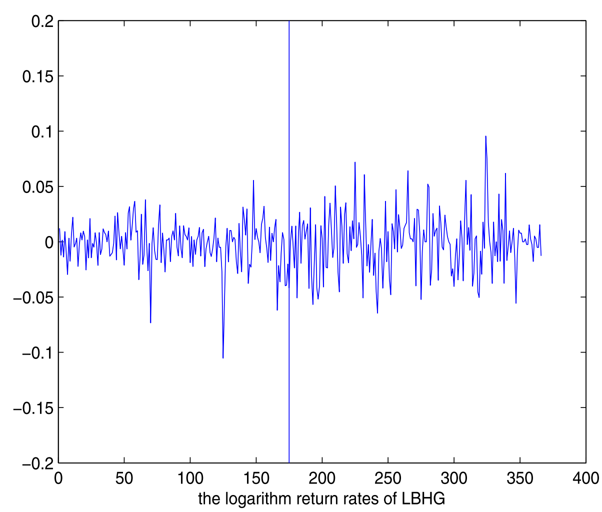

Figure 2 The logarithm return rates of LBC in Shanghai Stock Exchange.

Fitting a mean and computing the test proposed in this paper $\Gamma_1 = 4.2123 > 1.358$, which indicates that a change in mean occurred, and $\Xi_T(k)$ attains its maximum at $k_0 = 175$ (21st, March, 2004) (as shown in Figure 3). Recall that LBC issued an announcement that its net profits in 2005 would decrease to 50% of that in 2004, in the 3rd Session Board of Directors' 17th Meeting on March 8th, 2005 ($k_1 = 166$). The influence of the bad news was so strong that the stock price fell immediately in the following nine days, the mean of the logarithm return rate has a significant change after $k_0 = 175$.

5 Concluding remarks

In this paper, we construct a nonparametric test based on the indicators of the data minus the sample median. When there exists no change in the mean of the data, the test has the usual distribution of the sup of the absolute value of a Brownian bridge. As Bai [5] pointed out, it is a difficult task in applications of autoregressive models. First, the order

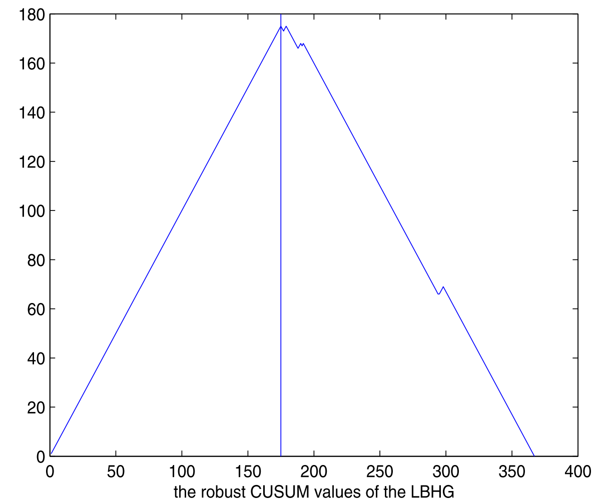

Figure 3 The robust CUSUM values of LBC in Shanghai Stock Exchange.

of an autoregressive model is not assumed to be known a priori and has to be estimated. Second, the often-used way to determine the order via the Akaike information criterion (AIC) and the Bayes information criterion (BIC) tends to overestimate its order if a change exists. However, the proposed test does not rely on the precise autoregressive models and the prior knowledge on the tail index α, so the proposed test is more applicable, although there exists a little distortion in its size for dependent sequences.

Appendix: Proofs of main results

The proof of Theorem 1 is based on the following four lemmas.

Lemma 1 *For L_r-bounded strong (α-) mixing random variables $y_{Tt} \in R$, for which the mixing coefficients satisfy $\alpha(m) \le Cm^{-r/(r-2)-\eta}$ for some $\eta > 0$,*

$$E \max_{1\le i\le T} \left(\sum_{t=1}^{i} (y_{Tt} - Ey_{Tt}) \right)^2 \le C' \sum_{t=1}^{T} \|y_{Tt}\|_r^2 \tag{13}$$

for constants C and C', where $\|X\| = (E|X|^r)^{1/r}$.

This lemma is Lemma 1 in De Jong *et al.* [14]; it is crucial for the proof of the following lemmas and theorems.

Lemma 2 *Let*

$$y_j(\phi) = \operatorname{sgn}(Y_j - \mu_1 - \tilde{\mu} - \phi T^{-1/2}) - \operatorname{sgn}(Y_j - \mu_1 - \tilde{\mu}). \tag{14}$$

If the null hypothesis H_0 holds, then under Assumption 1, for all $K, \varepsilon > 0$,

$$\lim_{\delta\to 0} \limsup_{T\to\infty} P\left(\sup_{\phi,\phi'\in[-K,K]:|\phi-\phi'|<\delta} T^{-1/2} \sum_{t=1}^{T} \left| y_j(\phi) - y_j(\phi') - Ey_j(\phi) + Ey_j(\phi') \right| > \varepsilon \right) = 0. \tag{15}$$

Proof Since the proof is similar to Lemma 2 of De Jong *et al.* [14], we omit it. □

Lemma 3 *Let $y_j(\phi)$ be as in* (14), *and let*

$$G_T(\tau,\phi) = T^{-1/2}\sum_{j=1}^{[T\tau]} y_j(\phi). \tag{16}$$

If the null hypothesis H_0 holds, then under Assumption 1, *for any $K > 0$,*

$$\sup_{\tau\in[0,1]}\sup_{\phi\in[-K,K]}\left|G_T(\tau,\phi) - EG_T(\tau,\phi)\right| \xrightarrow{P} 0. \tag{17}$$

Proof The proof is similar to Lemma 3 of De Jong *et al.* [14], so we omit it. □

Lemma 4 *If the null hypothesis H_0 holds, then under Assumption* 1,

$$T^{1/2}(m_T - \mu_1 - \tilde{\mu}) = 2^{-1}f(0)^{-1}\sigma W_T(1) + o_P(1). \tag{18}$$

Proof The proof is similar to Lemma 4 of De Jong *et al.* [14], so we omit it. □

Proof of Theorem 1 According to Lemma 4, we can find a large K so that $-K \le T^{1/2}(m_T - \mu_1 - \tilde{\mu}) \le K$. Then

$$\begin{aligned} T^{-1/2}S_{T,[T\tau]} &= T^{-1/2}\sum_{j=1}^{[T\tau]}\operatorname{sgn}(Y_j - m_T) = T^{-1/2}\sum_{j=1}^{[T\tau]}\operatorname{sgn}\bigl((Y_j - \mu_1 - \tilde{\mu}) - (m_T - \mu_1 - \tilde{\mu}_1)\bigr) \\ &= G_T\bigl(\tau, T^{1/2}(m_T - \mu_1 - \tilde{\mu})\bigr) - EG_T\bigl(\tau, T^{1/2}(m_T - \mu_1 - \tilde{\mu})\bigr) \\ &\quad + T^{-1/2}\sum_{j=1}^{[T\tau]}\operatorname{sgn}(Y_j - \mu_1 - \tilde{\mu}) - 2T^{-1/2}[T\tau](m_T - \mu_1 - \tilde{\mu})f(\tilde{m}_T - \mu_1 - \tilde{\mu}) \\ &= I_1 + I_2 - I_3, \end{aligned} \tag{19}$$

where $\tilde{m}_T$ is on the line between m_T and $\mu + \tilde{\mu}$ and $\tilde{m}_T - \mu_1 - \tilde{\mu} = o_P(1)$ by Lemma 4. Then $I_1 = o_P(1)$ holds uniformly for all $\tau \in [0,1]$ by Lemmas 3, 4. By definition, $I_2 = \sigma W_T(\tau)$. $I_3 = \tau\sigma W_T(1) + o_P(1)$ by Lemma 4. So we have

$$T^{-1/2}S_{T,[T\tau]} = \sigma\bigl(W_T(\tau) - \tau W_T(1)\bigr) + o_P(1). \tag{20}$$

Noting that $|T^{-1/2}\sum_{j=1}^{T}\operatorname{sgn}(Y_j - m_T)| \le T^{-1/2}$, we have

$$\begin{aligned} \frac{1}{\sqrt{T}}S_{T,[T(1-\tau)]} &= T^{-1/2}\sum_{j=[T\tau]+1}^{T}\operatorname{sgn}(Y_j - m_T) \\ &= T^{-1/2}\sum_{j=1}^{T}\operatorname{sgn}(Y_j - m_T) - T^{-1/2}\sum_{j=1}^{[T\tau]}\operatorname{sgn}(Y_j - m_T) \\ &= O\bigl(T^{-1/2}\bigr) - \Bigl\{G_T\bigl(\tau, T^{1/2}(m_T - \mu_1 - \tilde{\mu})\bigr) - EG_T\bigl(\tau, T^{1/2}(m_T - \mu_1 - \tilde{\mu})\bigr) \end{aligned}$$

$$+ T^{-1/2}\sum_{j=1}^{[T\tau]}\operatorname{sgn}(Y_j - \mu_1 - \tilde{\mu})$$

$$\left. - 2T^{-1/2}[T\tau](m_T - \mu_1 - \tilde{\mu})f(\tilde{m}_T - \mu_1 - \tilde{\mu})\right\}$$

$$= O\bigl(T^{-1/2}\bigr) - o_P(1) - \bigl\{\sigma W_T(\tau) - \tau\sigma W_T(1)\bigr\}. \tag{21}$$

Based on (20), (21), by the functional central limit theorem,

$$T^{1/2}\sigma^{-1}\max_{\tau\in(0,1)}|\Xi_T| \quad\Longrightarrow\quad \sup_{\tau\in(0,1)}\bigl|W(\tau) - \tau W(1)\bigr|, \quad \text{as } T\to\infty. \tag{22}$$

If we can show $\hat{\sigma}^2 \xrightarrow{P} \sigma^2$, the proof of Theorem 1 is completed. Under the null hypothesis H_0, μ_1 remains as a constant, so we can prove the consistency of $\hat{\sigma}^2$ just as De Jong *et al.* [14]. □

The proof of Theorem 2 is based on Lemmas 1, 5, 6, 7 as follows.

Lemma 5 *If the alternative hypothesis H_1 holds and $k_0 = [T\tau_0]$ is the change point, let $y_j(\phi)$ be as follows*:

$$y_j(\phi) = \operatorname{sgn}\bigl(Y_j - \mu^* - \phi T^{-1/2}\bigr) - \operatorname{sgn}\bigl(Y_j - \mu^*\bigr), \tag{23}$$

then under Assumption 1, *for all $K, \varepsilon > 0$*,

$$\lim_{\delta\to 0}\limsup_{T\to\infty} P\left(\sup_{\phi,\phi'\in[-K,K]:|\phi-\phi'|<\delta} T^{-1/2}\sum_{j=1}^{T}\bigl|y_j(\phi) - y_j\bigl(\phi'\bigr) - Ey_j(\phi) + Ey_j\bigl(\phi'\bigr)\bigr| > \varepsilon\right) = 0. \tag{24}$$

Proof For $y_j(\phi)$ as in (23), we have

$$Ey_j(\phi) = \begin{cases} 2F(\mu^* - \mu_1) - 2F(\mu^* + \phi T^{-1/2} - \mu_1), & j \le k_0, \\ 2F(\mu^* - \mu_2) - 2F(\mu^* + \phi T^{-1/2} - \mu_2), & t > k_0. \end{cases} \tag{25}$$

Then for T large enough such that $KT^{-1/2} \le \eta$, under the alternative hypothesis H_1,

$$\begin{aligned} &\sup_{\phi,\phi':|\phi-\phi'|<\delta} T^{-1/2}\sum_{j=1}^{T}\bigl|Ey_j(\phi) - Ey_j\bigl(\phi'\bigr)\bigr| \\ &\quad\le \sup_{\phi,\phi':|\phi-\phi'|<\delta} T^{-1/2}\sum_{j=1}^{k_0}\bigl|Ey_j(\phi) - Ey_j\bigl(\phi'\bigr)\bigr| + \sup_{\phi,\phi':|\phi-\phi'|<\delta} T^{-1/2}\sum_{j=k_0+1}^{T}\bigl|Ey_j(\phi) - Ey_j\bigl(\phi'\bigr)\bigr| \\ &\quad= I_{11} + I_{12}, \end{aligned} \tag{26}$$

$$I_{11} = 2\sup_{\phi,\phi':|\phi-\phi'|<\delta} T^{-1/2}\sum_{j=1}^{k_0}\bigl|Ey_j(\phi) - Ey_j\bigl(\phi'\bigr)\bigr|$$

$$
\begin{aligned}
&= 2 \sup_{\phi,\phi':|\phi-\phi'|<\delta} T^{-1/2} \sum_{j=1}^{k_0} \bigl|F\bigl(\mu^* + \phi T^{-1/2} - \mu_1\bigr) - F\bigl(\mu^* + \phi' T^{-1/2} - \mu_1\bigr)\bigr| \\
&\le 2 \sup_{\phi,\phi':|\phi-\phi'|<\delta} T^{-1/2} \sum_{j=1}^{k_0} \sup_{x\in[-\eta_1,\eta_1]} f(x) T^{-1/2} \bigl|\phi - \phi'\bigr| \le 2\delta \sup_{x\in[-\eta_1,\eta_1]} f(x),
\end{aligned} \tag{27}
$$

$$
\begin{aligned}
I_{12} &= 2 \sup_{\phi,\phi':|\phi-\phi'|<\delta} T^{-1/2} \sum_{j=k_0+1}^{T} \bigl|Ey_j(\phi) - Ey_j\bigl(\phi'\bigr)\bigr| \\
&= 2 \sup_{\phi,\phi':|\phi-\phi'|<\delta} T^{-1/2} \sum_{j=k_0+1}^{T} \bigl|F\bigl(\mu^* + \phi T^{-1/2} - \mu_2\bigr) - F\bigl(\mu^* + \phi' T^{-1/2} - \mu_2\bigr)\bigr| \\
&\le 2 \sup_{\phi,\phi':|\phi-\phi'|<\delta} T^{-1/2} \sum_{j=k_0+1}^{T} \sup_{x\in[-\eta_2,\eta_2]} f(x) T^{-1/2} \bigl|\phi - \phi'\bigr| \le 2\delta \sup_{x\in[-\eta_2,\eta_2]} f(x),
\end{aligned} \tag{28}
$$

where η stands for different constants at different equations. This establishes equicontinuity of I_{11} and I_{12}. Similar to (26), we have

$$
\begin{aligned}
&\sup_{\phi,\phi'\in[-K,K]:|\phi-\phi'|<\delta} T^{-1/2} \sum_{j=1}^{T} \bigl|y_j(\phi) - y_j\bigl(\phi'\bigr)\bigr| \\
&\quad\le \sup_{\phi,\phi'\in[-K,K]:|\phi-\phi'|<\delta} T^{-1/2} \sum_{j=1}^{k_0} \bigl|y_j(\phi) - y_j\bigl(\phi'\bigr)\bigr| \\
&\qquad + \sup_{\phi,\phi'\in[-K,K]:|\phi-\phi'|<\delta} T^{-1/2} \sum_{j=k_0+1}^{T} \bigl|y_j(\phi) - y_j\bigl(\phi'\bigr)\bigr| \\
&\quad= I_{21} + I_{22}.
\end{aligned} \tag{29}
$$

Since $y_j(\phi)$ is non-increasing in ϕ,

$$
\begin{aligned}
I_{21} &= \sup_{\phi,\phi'\in[-K,K]:|\phi-\phi'|<\delta} T^{-1/2} \sum_{j=1}^{k_0} \bigl|y_j(\phi) - y_j\bigl(\phi'\bigr)\bigr| \\
&= \sup_{-[K/\delta]-1\le i\le[K/\delta]} \sup_{\phi,\phi'\in[-K,K]\cap[i\delta,(i+2)\delta]} T^{-1/2} \sum_{j=1}^{k_0} \bigl|y_j(\phi) - y_j\bigl(\phi'\bigr)\bigr| \\
&\le \sup_{-[K/\delta]-1\le i\le[K/\delta]} T^{-1/2} \sum_{j=1}^{k_0} \bigl|y_j(i\delta) - y_j\bigl((i+2)\delta\bigr)\bigr| \\
&= o_P(1) + \sup_{-[K/\delta]-1\le i\le[K/\delta]} T^{-1/2} \sum_{j=1}^{k_0} E\bigl|y_j(i\delta) - y_j\bigl((i+2)\delta\bigr)\bigr| \\
&\le o_P(1) + \sup_{\phi,\phi'\in[-K,K]:|\phi-\phi'|<2\delta} T^{-1/2} \sum_{j=1}^{k_0} E\bigl|y_j(\phi) - y_j\bigl(\phi'\bigr)\bigr|,
\end{aligned} \tag{30}
$$

and the last term has been proved earlier to be equicontinuous. Similarly $y_j(\phi)$ is non-increasing in ϕ, we have

$$
\begin{aligned}
I_{22} &= \sup_{\phi,\phi'\in[-K,K]:|\phi-\phi'|<\delta} T^{-1/2}\sum_{j=k_0+1}^{T}\bigl|y_j(\phi)-y_j(\phi')\bigr| \\
&= \sup_{-[K/\delta]-1\le i\le[K/\delta]}\ \sup_{\phi,\phi'\in[-K,K]\cap[i\delta,(i+2)\delta]} T^{-1/2}\sum_{j=k_0+1}^{T}\bigl|y_j(\phi)-y_j(\phi')\bigr| \\
&\le \sup_{-[K/\delta]-1\le i\le[K/\delta]} T^{-1/2}\sum_{j=k_0+1}^{T}\bigl|y_j(i\delta)-y_j\bigl((i+2)\delta\bigr)\bigr| \\
&= o_P(1)+\sup_{-[K/\delta]-1\le i\le[K/\delta]} T^{-1/2}\sum_{j=1}^{k_0}E\bigl|y_j(i\delta)-y_j\bigl((i+2)\delta\bigr)\bigr| \\
&\le o_P(1)+\sup_{\phi,\phi'\in[-K,K]:|\phi-\phi'|<2\delta} T^{-1/2}\sum_{j=k_0+1}^{T}E\bigl|y_j(\phi)-y_j(\phi')\bigr|,
\end{aligned}
\tag{31}
$$

and the last term has been proved earlier to be equicontinuous too. By the triangle inequality, for all $\varepsilon > 0$,

$$
\begin{aligned}
&P\Biggl(\sup_{\phi,\phi'\in[-K,K]:|\phi-\phi'|<\delta} T^{-1/2}\sum_{j=1}^{T}\bigl|y_j(\phi)-y_j(\phi')-Ey_j(\phi)+Ey_j(\phi')\bigr|>\varepsilon\Biggr) \\
&\quad\le P\Biggl(\sup_{\phi,\phi'\in[-K,K]:|\phi-\phi'|<\delta} T^{-1/2}\sum_{j=1}^{k_0}\bigl|y_j(\phi)-y_j(\phi')-Ey_j(\phi)+Ey_j(\phi')\bigr|>\varepsilon/2\Biggr) \\
&\qquad+P\Biggl(\sup_{\phi,\phi'\in[-K,K]:|\phi-\phi'|<\delta} T^{-1/2}\sum_{j=k_0+1}^{T}\bigl|y_j(\phi)-y_j(\phi')-Ey_j(\phi)+Ey_j(\phi')\bigr|>\varepsilon/2\Biggr) \\
&\quad=I_{31}+I_{32},
\end{aligned}
\tag{32}
$$

$$
\begin{aligned}
I_{31} &= P\Biggl(\sup_{\phi,\phi'\in[-K,K]:|\phi-\phi'|<\delta} T^{-1/2}\sum_{j=1}^{k_0}\bigl|y_j(\phi)-y_j(\phi')-Ey_j(\phi)+Ey_j(\phi')\bigr|>\varepsilon/2\Biggr) \\
&\le P\Biggl(\sup_{\phi,\phi'\in[-K,K]:|\phi-\phi'|<\delta} T^{-1/2}\sum_{j=1}^{k_0}\bigl|y_j(\phi)-y_j(\phi')\bigr|>\varepsilon/4\Biggr) \\
&\quad+P\Biggl(\sup_{\phi,\phi'\in[-K,K]:|\phi-\phi'|<\delta} T^{-1/2}\sum_{j=1}^{k_0}\bigl|Ey_j(\phi)-Ey_j(\phi')\bigr|>\varepsilon/4\Biggr) \\
&\le o_P(1)+P\Biggl(\sup_{\phi,\phi'\in[-K,K]:|\phi-\phi'|<\delta} T^{-1/2}\sum_{j=1}^{k_0}\bigl|Ey_j(\phi)-Ey_j(\phi')\bigr|>\varepsilon/4\Biggr),
\end{aligned}
\tag{33}
$$

the last term converges to 0 as $\delta \to 0$ by the equicontinuity of (27). Similarly, we can show

$$
\begin{aligned}
I_{32} &= P\Biggl(\sup_{\phi,\phi'\in[-K,K]:|\phi-\phi'|<\delta} T^{-1/2}\sum_{j=k_0+1}^{T}\bigl|y_j(\phi)-y_j(\phi')-Ey_j(\phi)+Ey_j(\phi')\bigr|>\varepsilon/2\Biggr) \\
&\le P\Biggl(\sup_{\phi,\phi'\in[-K,K]:|\phi-\phi'|<\delta} T^{-1/2}\sum_{j=k_0+1}^{T}\bigl|y_j(\phi)-y_j(\phi')\bigr|>\varepsilon/4\Biggr)
\end{aligned}
$$

$$+ I\left(\sup_{\phi,\phi'\in[-K,K]:|\phi-\phi'|<\delta} T^{-1/2}\sum_{j=k_0+1}^{T}\left|Ey_j(\phi)-Ey_j(\phi')\right| > \varepsilon/4\right)$$

$$\le o_P(1) + P\left(\sup_{\phi,\phi'\in[-K,K]:|\phi-\phi'|<\delta} T^{-1/2}\sum_{j=k_0+1}^{T}\left|Ey_j(\phi)-Ey_j(\phi')\right| > \varepsilon/4\right), \tag{34}$$

the last term converges to 0 as $\delta \to 0$ by the equicontinuity of (28) too. Now, we have completed the proof of Lemma 5. □

Lemma 6 *If the alternative hypothesis H_1 holds, let $y_j(\phi)$ be as in* (23), *and*

$$G_T(\tau,\phi) = T^{-1/2}\sum_{j=1}^{[T\tau]} y_j(\phi), \tag{35}$$

then under Assumption 1, *for any $K > 0$*,

$$\sup_{\tau\in[0,1]}\sup_{\phi\in[-K,K]}\left|G_T(\tau,\phi)-EG_T(\tau,\phi)\right| \xrightarrow{P} 0. \tag{36}$$

Proof Just as De Jong *et al.* [14], we can obtain from Kim and Pollard [16, Theorem 2.1]

$$\sup_{\tau\in[0,1]}\sup_{\phi\in[-K,K]}\left|G_T(\tau,\phi)-EG_T(\tau,\phi)\right| \xrightarrow{P} 0 \tag{37}$$

through the arguments for the finite-dimensional convergence for each $\phi \in [-K,K]$ and the stochastic equicontinuity of $\sup_{\tau\in[0,1]}|G_T(\tau,\phi)-EG_T(\tau,\phi)|$. For every $\phi\in[-K,K]$, by Lemma 1, for T large enough such that $KT^{-1/2}\le\eta$, we have

$$\begin{aligned}
&E\sup_{\tau\in[0,1]}\left|G_T(\tau,\phi)-EG_T(\tau,\phi)\right|^2\\
&\quad\le CT^{-1}\sum_{j=1}^{T}\left\|\operatorname{sgn}\left(Y_j-\mu^*-\phi T^{-1/2}\right)-\operatorname{sgn}\left(Y_j-\mu^*\right)\right\|_r^2\\
&\quad= CT^{-1}\sum_{j=1}^{k_0}\left\|\operatorname{sgn}\left(X_j+\mu_1-\mu^*-\phi T^{-1/2}\right)-\operatorname{sgn}\left(X_j+\mu_1-\mu^*\right)\right\|_r^2\\
&\qquad+ CT^{-1}\sum_{j=k_0+1}^{T}\left\|\operatorname{sgn}\left(X_j+\mu_2-\mu^*-\phi T^{-1/2}\right)-\operatorname{sgn}\left(X_j+\mu_2-\mu^*\right)\right\|_r^2\\
&\quad\le C'\tau_0\left|F\left(\mu^*-\mu_1+KT^{-1/2}\right)-F\left(\mu^*-\mu_1-KT^{-1/2}\right)\right|^{2/r}\\
&\qquad+ C'(1-\tau_0)\left|F\left(\mu^*-\mu_2+KT^{-1/2}\right)-F\left(\mu^*-\mu_2-KT^{-1/2}\right)\right|^{2/r}\\
&\quad\le C''\tau_0\left(\sup_{x\in[-\eta_1,\eta_1]} f(x)2KT^{-1/2}\right)^{2/r}\\
&\qquad+ C''(1-\tau_0)\left(\sup_{x\in[-\eta_2,\eta_2]} f(x)2KT^{-1/2}\right)^{2/r},
\end{aligned} \tag{38}$$

where constants $C, C', C'' > 0$. Now we have shown the finite-dimensional convergence for each $\phi \in [-K, K]$. By the triangle inequality,

$$\begin{aligned}
&\Big| \sup_{\tau\in[0,1]} \big| G_T(\tau,\phi) - EG_T(\tau,\phi) \big| - \sup_{\tau\in[0,1]} \big| G_T(\tau,\phi') - EG_T(\tau,\phi') \big| \Big| \\
&\quad \le \sup_{\tau\in[0,1]} \big| G_T(\tau,\phi) - EG_T(\tau,\phi) - G_T(\tau,\phi') + EG_T(\tau,\phi') \big| \\
&\quad \le T^{-1/2} \sum_{j=1}^{T} \big| y_j(\phi) - y_j(\phi') - Ey_j(\phi) + Ey_j(\phi') \big|. \qquad (39)
\end{aligned}$$

Now stochastic equicontinuity follows from Lemma 5. □

Lemma 7 *If the alternative hypothesis H_1 holds, then under Assumption* 1,

$$T^{1/2}(m_T - \mu^*) = O_P(1), \qquad (40)$$

where μ^* is defined as the median of (7).

Proof For T large enough such that $T \ge K^2\eta^{-2}$,

$$\begin{aligned}
&\sup_{\phi>K} T^{-1/2} \sum_{j=1}^{T} \operatorname{sgn}\big(Y_j - \mu^* - \phi T^{-1/2}\big) \\
&\quad \le T^{-1/2} \sum_{j=1}^{T} \operatorname{sgn}\big(Y_j - \mu^* - KT^{-1/2}\big) \\
&\quad = T^{-1/2} \sum_{j=1}^{k_0} \operatorname{sgn}\big(X_j + \mu_1 - \mu^* - KT^{-1/2}\big) \\
&\qquad + T^{-1/2} \sum_{j=k_0+1}^{T} \operatorname{sgn}\big(X_j + \mu_2 - \mu^* - KT^{-1/2}\big) \\
&\quad \le o_P(1) + T^{-1/2} \sum_{j=1}^{k_0} \big(1 - 2F\big(KT^{-1/2} + \mu^* - \mu_1\big)\big) \\
&\qquad + o_P(1) + T^{-1/2} \sum_{j=k_0+1}^{T} \big(1 - 2F\big(KT^{-1/2} + \mu^* - \mu_2\big)\big) \\
&\quad = o_P(1) + T^{1/2}\big(1 - 2F^*\big(\mu^* + KT^{-1/2}\big)\big) \\
&\quad = o_P(1) - 2K \inf_{-\eta<(x-\mu^*)<\eta} f^*(x), \qquad (41)
\end{aligned}$$

which implies that $\limsup_{T\to\infty} P(T^{1/2}(m_T - \mu^*) > K)$ can be made arbitrarily small by choosing K large enough under the alternative hypothesis H_1. For $P(T^{1/2}(m_T - \mu^*) < -K)$, a similar result can be derived, which proves that $m_T - \mu^* = O_P(T^{-1/2})$ under the alternative hypothesis H_1. □

Proof of Theorem 2 We just consider the case $k = [T\tau] \ge k_0 = [T\tau_0]$, the case $k \le k_0$ can be analyzed similarly. According to Lemma 7, we can find K large enough so that $-K \le$

$T^{-1/2}(m_T - \mu^*) \leq K$ will happen with arbitrarily large probability. Then

$$\begin{aligned}
T^{-1}S_{T,[T\tau]} &= T^{-1}\sum_{j=1}^{[T\tau]}\operatorname{sgn}(Y_j - m_T) \\
&= T^{-1}\sum_{j=1}^{[T\tau_0]}\operatorname{sgn}\bigl((X_j + \mu_1 - \mu^*) - (m_T - \mu^*)\bigr) \\
&\quad + T^{-1}\sum_{j=[T\tau_0]+1}^{[T\tau]}\operatorname{sgn}\bigl((X_j + \mu_2 - \mu^*) - (m_T - \mu^*)\bigr) \\
&= T^{-1/2}\bigl(G_T\bigl(\tau, T^{1/2}(m_T - \mu^*)\bigr) - EG_T\bigl(\tau, T^{1/2}(m_T - \mu^*)\bigr)\bigr) \\
&\quad + T^{-1}\sum_{j=1}^{[T\tau]}\operatorname{sgn}(Y_j - \mu^*) + T^{-1/2}EG_T\bigl(\tau, T^{1/2}(m_T - \mu^*)\bigr) \\
&= I_1 + I_2 + I_3. \qquad (42)
\end{aligned}$$

Then $I_1 = o_P(1)$ by Lemma 6, $I_3 = o_P(1)$ by (25), with Proposition 2.8 of Fan and Yao [17], we have

$$T^{-1}S_{T,[T\tau]} \xrightarrow{P} \tau_0 F(\mu^* - \mu_1) + (\tau - \tau_0)F(\mu^* - \mu_2).$$

Similarly, we can obtain

$$\begin{aligned}
T^{-1}S^2_{T,[T\tau]} &= T^{-1}\sum_{j=[T\tau]+1}^{T}\operatorname{sgn}(Y_j - m_T) \\
&\xrightarrow{P} (1 - \tau)F(\mu^* - \mu_2). \qquad (43)
\end{aligned}$$

By the definition of $\Xi_T(\tau)$,

$$\Xi_T(\tau) \xrightarrow{P} (1 - \tau)\tau_0\bigl|F(\mu^* - \mu_1) - F(\mu^* - \mu_2)\bigr| \qquad (44)$$

as $T \to \infty$, and $\tau \geq \tau_0$, so

$$\sup_{\tau\in(0,1)}\bigl|\Xi_T(\tau)\bigr| \xrightarrow{P} \tau_0(1 - \tau_0)\bigl|F(\mu^* - \mu_1) - F(\mu^* - \mu_2)\bigr|. \qquad (45)$$

□

Proof of Theorem 3 Under the hypothesis H_0, there is no shift in the mean, so the proof of (i) is nearly similar to the proof of the consistency of $\hat{\sigma}_T$, so we just gave the details of the proof of (ii).

Noting that for y_j defined in (23),

$$\begin{aligned}
\operatorname{sgn}(Y_j - m_T) &= \bigl(y_j\bigl(T^{1/2}(m_T - \mu^*)\bigr) - Ey_j\bigl(T^{1/2}(m_T - \mu^*)\bigr)\bigr) \\
&\quad + Ey_j\bigl(T^{1/2}(m_T - \mu^*)\bigr) + \operatorname{sgn}(Y_j - \mu^*) \\
&= a_{Tj} + b_{Tj} + c_j, \qquad (46)
\end{aligned}$$

so

$$\hat{\sigma}^2 = T^{-1}\sum_{i=1}^{T}\sum_{j=1}^{T}k\big((i-j)/\gamma_T\big)(a_{Ti}+b_{Ti}+c_i)(a_{Tj}+b_{Tj}+c_j). \tag{47}$$

Under the assumption that $\gamma_T \to \infty$, $\gamma_T/T \to 0$, $\gamma_T^2/T \to 0$ as $T \to \infty$, for $y_j(\phi)$ defined in (23), by Lemma 7, we have $b_{Ti} = O_P(T^{-1/2})$, then

$$T^{-1}\sum_{i=1}^{T}\sum_{j=1}^{T}k\big((i-j)/\gamma_T\big)b_{Ti}a_{Tj} \le T^{-3/2}\sum_{j=1}^{T}|a_{Tj}|\sum_{t=-T}^{T}k(t/\gamma_T)\times O_P(1) = O_P(\gamma_T/T),$$

$$T^{-1}\sum_{i=1}^{T}\sum_{j=1}^{T}k\big((i-j)/\gamma_T\big)b_{Ti}b_{Tj} \le C^2T^{-1}\sum_{i=1}^{T}\sum_{j=1}^{T}k\big((i-j)/\gamma_T\big) = O_P\big(T^{-1}\gamma_T\big),$$

and that

$$T^{-1}\sum_{i=1}^{T}\sum_{j=1}^{T}k\big((i-j)/\gamma_T\big)b_{Ti}c_j \le T^{-3/2}\sum_{j=1}^{T}c_j\sum_{i=1}^{T}k\big((i-j)/\gamma_T\big)\times O_P(1)$$

and

$$\begin{aligned}
E\left(T^{-3/2}\sum_{j=1}^{T}c_j\sum_{s=1}^{T}k\big((j-s)/\gamma_T\big)\right)^2 &\le CT^{-3}\sum_{j=1}^{T}\left\|c_t\sum_{s=1}^{T}k\big((j-s)/\gamma_T\big)\right\|_r^2\\
&\le C'T^{-3}\sum_{j=1}^{T}\left(\sum_{s=1}^{T}k\big((j-s)/\gamma_T\big)\right)^2\\
&\le C''T^{-3}\sum_{j=1}^{T}\left(\sum_{t=-T}^{T}k(t/\gamma_T)\right)^2\\
&\le C''T^{-2}\gamma_T^2.
\end{aligned}$$

Therefore under Assumption 1 and the alternative hypothesis H_1, $\hat{\sigma}^2$ is asymptotically equivalent to

$$T^{-1}\sum_{t=1}^{T}\sum_{s=1}^{T}k\big((t-s)/\gamma_T\big)(a_{Tt}+c_s)(a_{Ts}+c_s).$$

Furthermore,

$$\begin{aligned}
&T^{-1}\sum_{t=1}^{T}\sum_{s=1}^{T}k\big((t-s)/\gamma_T\big)a_{Tt}c_s\\
&\quad= \int_{-\infty}^{\infty}\left(T^{-1}\sum_{t=1}^{T}\sum_{s=1}^{T}a_{Tt}c_s\exp\big(\mathrm{i}\xi(t-s)\big)\right)\psi(\xi)\,d\xi\\
&\quad\le T^{-1/2}\sum_{t=1}^{T}|a_{Tt}|\int_{-\infty}^{\infty}\left(T^{-1/2}\sum_{s=1}^{T}c_s\exp(-\mathrm{i}s\xi/\gamma_T)\right)\psi(\xi)\,d\xi,
\end{aligned}$$

and $T^{-1/2}\sum_{t=1}^{T}|a_{Tt}| = o_P(1)$ by Lemma 6 under Assumption 1, and the second term is $O_P(1)$ because

$$E\left|\int_{-\infty}^{\infty}\left(T^{-1/2}\sum_{s=1}^{T}c_s\exp(-\mathrm{i}s\xi/\gamma_T)\right)\psi(\xi)\,d\xi\right| \le \int_{-\infty}^{\infty}\left|\psi(\xi)\right|d\xi\sup_{\xi\in R}\left\|T^{-1/2}\sum_{s=1}^{T}c_s\exp(-\mathrm{i}s\xi/\gamma_T)\right\|_2 < \infty.$$

Finally,

$$T^{-1}\sum_{t=1}^{T}\sum_{s=1}^{T}k\big((t-s)/\gamma_T\big)a_{Tt}a_{Ts} \le \int_{-\infty}^{\infty}\left|\psi(\xi)\right|d\xi\left(T^{-1/2}\sum_{t=1}^{T}|a_{Tt}|\right)^2$$

by the last term is $o_P(1)$ by Lemma 6 and Assumptions 1, 2, so we have shown that $\hat{\sigma}^2$ asymptotically equals

$$T^{-1}\sum_{i=1}^{T}\sum_{j=1}^{T}k\big((i-j)/\gamma_T\big)c_ic_j. \tag{48}$$

Under Assumptions 1, 2 and the alternative hypothesis H_1, $\{\mathrm{sgn}(Y_j-\mu^*)\}$ satisfies the assumptions of Theorem 2.1 in De Jong and Davidson [18], so

$$T^{-1}\sum_{i=1}^{T}\sum_{j=1}^{T}k\big((i-j)/\gamma_T\big)c_ic_j \xrightarrow{P} \sigma_2^2, \tag{49}$$

so the proof of Theorem 2 has been completed. □

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
[1]School of Mathematical Science, Shanxi University, Taiyuan, Shanxi 030006, P.R. China. [2]Institute of Management and Decision, Shanxi University, Taiyuan, Shanxi 030006, P.R. China.

Acknowledgements
The authors are grateful to the two anonymous referees for their careful reading of the manuscript and many helpful comments which improve the manuscript greatly. We would like to thank Prof. John Nolan who provided the software for generating the stable innovations in Section 4. By financial support this work is supported by the National Natural Science Foundation of China (No. 11226217) and the Postdoctoral Science Special Foundation (Nos. 2012M510772, 2013T60266), which is gratefully acknowledged.

Received: 12 May 2014 Accepted: 19 January 2015 Published online: 06 February 2015

References

1. Sen, A, Srivastava, MS: On tests for detecting change in mean. Ann. Stat. **3**, 96-103 (1975)
2. Hawkins, DM: Testing a sequence of observations for a shift in location. J. Am. Stat. Assoc. **72**, 180-186 (1977)
3. Worsley, KJ: On the likelihood ratio test for a shift in locations of normal populations. J. Am. Stat. Assoc. **74**, 365-367 (1979)
4. Yao, YC: Approximating the distribution of the ML estimate of the change point in a sequence of independent random variables. Ann. Stat. **3**, 1321-1328 (1987)
5. Bai, J: Least squares estimation of a shift in linear processes. J. Time Ser. Anal. **15**, 453-472 (1991)

6. Horváth, L, Kokoszka, P: The effect of long-range dependence on change-point estimators. J. Stat. Plan. Inference **64**, 57-81 (1997)
7. Mittnik, S, Rachev, S: Stable Paretian Models in Finance. Wiley, New York (2000)
8. Phillips, PCB: Time series regression with a unit root and infinite-variance errors. Econom. Theory **6**, 44-62 (1990)
9. Horváth, L, Kokoszka, P: A bootstrap approximation to a unit root test statistics for heavy-tailed observations. Stat. Probab. Lett. **62**, 163-173 (2003)
10. Han, S, Tian, Z: Truncating estimator for the mean change-point in heavy-tailed dependent observations. Commun. Stat., Theory Methods **35**, 43-52 (2006)
11. Han, S, Tian, Z: Bootstrap testing for changes in persistence with heavy-tailed innovations. Commun. Stat., Theory Methods **36**, 2289-2299 (2007)
12. Hušková, M: Tests and estimators for change point problem based on *M*-statistics. Stat. Decis. **14**, 115-136 (1996)
13. Hušková, M: *M*-Procedures for detection of changes for dependent observations. Commun. Stat., Simul. Comput. **41**, 1032-1050 (2012)
14. De Jong, RM, Amsler, C, Schimidt, P: A robust version of the kpss test based on indicators. J. Econom. **137**, 311-333 (2007)
15. Wang, JH, Wang, YL, Ke, AM: The stable distributions fitting and testing in stock returns of China. J. Wuhan Univ. Technol. **25**, 99-102 (2003)
16. Kim, J, Pollard, D: Cube root asymptotics. Ann. Stat. **18**, 191-219 (1990)
17. Fan, J, Yao, Q: Nonlinear Time Series: Nonparametric and Parametric Methods. Springer, Berlin (2003)
18. De Jong, RM, Davidson, J: Consistency of kernel estimators of heteroscedastic and autocorrelated covariance matrices. Econometrica **68**, 407-423 (2000)