Change point detection based on empirical quantiles

Lajos Horváth, Agnieszka Jach, and Piotr Kokoszka

Abstract. The paper studies the problem of testing the null hypothesis that the observations $X_1, \ldots, X_n$ have the same distribution or follow the same stochastic model versus the alternative that the distribution or the model change at some point in the sample. Using a L^1, L^2 and max type statistics based on empirical quantiles, the validity of the subsampling method for independent obseravtions and GARCH processes is verified. A simulation study assesses the performance of the tests in finite samples.

1. Introduction

This paper focuses on a method proposed by Csörgő and Horváth (1987) who studied it theoretically in the case of independent observations. It is designed to detect a change of an arbitrary nature in the distribution of the observations.

Consider the observations $X_1, \ldots, X_n$, their empirical distribution function

$$F_n(x) = \frac{1}{n}\sum_{i=1}^{n} \mathbf{1}\left\{X_i \le x\right\}, \quad -\infty < x < \infty, \tag{1.1.1}$$

and their empirical quantile function

$$Q_n(t) = \inf\left\{x : F_n(x) \ge t\right\}, \quad 0 < t < 1.$$

Introduce also the score function

$$\psi_t(x) = \begin{cases} -(1-t) & \text{if} \quad x \le 0 \\ t & \text{if} \quad x > 0 \end{cases} \tag{1.1.2}$$

and the random field

$$Y_n(s,t) = n^{-1/2} \sum_{1 \le i \le [ns]} \psi_t(X_i - Q_n(t)), \quad 0 < s,t < 1. \tag{1.1.3}$$

Received by the editors March 5, 2002.
1991 *Mathematics Subject Classification.* Primary 62G30; Secondary 62G09.
Key words and phrases. Change point, empirical quantiles, GARCH, subsampling.
Research partially supported by NATO grant PST.CLG.977607. Agnieszka Jach and Piotr Kokoszka partially supported by USU New Faculty grant.

To understand why statistics designed to detect a change in the distribution of the observations $X_1, \ldots, X_n$ can be based on the field (1.1.3), observe that if there is no change point, then for every s, $X_1 \ldots, X_{[ns]}$ should have empirical quantiles close to those of $X_1, \ldots, X_n$. Thus, for each t we expect tns of the $X_1 \ldots, X_{[ns]}$ to the left of $Q_n(t)$ and $(1-t)ns$ of them to the right of $Q_n(t)$. The expected value of $n^{1/2}Y_n(s,t)$ is thus $(tns)(t-1)+(1-t)ns(t)=0$. If there is a change point, then the field (1.1.3) will deviate from zero.

For ease of reference, we state here the null and alternative hypotheses:

$$H_0: \quad X_1, \ldots, X_n \text{ have the same distribution;}$$

$$\begin{aligned} H_A: \quad & \text{There is } \lambda \in (0,1) \text{ and distribution functions } F \text{and } G \text{ such that} \\ & F(t) \neq G(t) \text{ for some } t \text{ and} \\ & P(X_i \le t) = F(t),\ 1 \le i \le [n\lambda] \text{ and } P(X_i \le t) = G(t),\ [n\lambda] \le i \le n. \end{aligned}$$

Csörgő and Horváth (1987) obtained the weak convergence of $Y_n(s,t)$ under H_0 in the case of independent observations:

Theorem 1.1. *If $X_1, X_2, \ldots$ are independent, identically distributed random variables with continuous distribution function, then*

$$Y_n(s,t) \to Y(s,t) \quad \text{in } D([0,1]\times[0,1]),$$

where $Y(t,s)$ is a Gaussian process with $EY(t,s)=0$ and

$$E[Y(t,s)Y(t',s') = (\min(t,t')-tt')(\min(s,s')-ss').$$

The limiting Gaussian process $Y(t,s)$ appeared earlier in Blum *et al.* (1961) in the context of testing for independence, c.f. also Csörgő (1979).

From Theorem 1.1 a number of test statistics can be constructed. For example

$$\sup_{0<s<1}\sup_{0<t<1}|Y_n(s,t)| \xrightarrow{d} \sup_{0<s<1}\sup_{0<t<1}|Y(s,t)|, \tag{1.1.4}$$

$$\int_0^1\int_0^1 Y_n^2(s,t)dsdt \xrightarrow{d} \int_0^1\int_0^1 Y^2(s,t)dsdt \tag{1.1.5}$$

and

$$\int_0^1\int_0^1 |Y_n(s,t)|dsdt \xrightarrow{d} \int_0^1\int_0^1 |Y(s,t)|dsdt. \tag{1.1.6}$$

The critical values for the right-hand sides of (1.1.4) and (1.1.5) were computed, respectively, in Picard (1985) and Blum *et al.* (1961). Cotteril and Csörgő (1985) and Martynov (1992) obtained the characteristic function of $\int\int Y^2(s,t)dsdt$. Inverting the characteristic function, the distribution function at points of interest can be calculated. Nothing is known about the distribution function of the limit in (1.1.6). For easy reference, we provide some critical values in Table B.1. The values in last row of this Table were obtained by Monte Carlo simulation.

Assuming independent observations, Csörgő and Horváth (1997, p. 103) studied $Y_n(s,t)$ under H_A. Suppose F and G are two continuous distrbution functions. Then under H_A

$$n^{-1/2} Y_n(s,t) \overset{a.s.}{\to} g(s,t) \tag{1.1.7}$$

for any $0 < s < t < 1$, where

$$g(s,t) = \begin{cases} s(t - F(Q_\lambda(t))), & 0 \le s \le \lambda, \\ (1-s)\lambda(1-\lambda)^{-1}(t - F(Q_\lambda(t))), & \lambda \le s \le 1 \end{cases} \tag{1.1.8}$$

and

$$Q_\lambda(t) = \inf\{x : \lambda F(x) + (1-\lambda)G(x) \ge t\}.$$

If $F(t) \ne G(t)$ for some t, by the continuity of F and G, there is an interval on which F and G are different. Hence, under H_A

$$\sup_{0<s<1} \sup_{0<t<1} |Y_n(s,t)| \overset{P}{\to} \infty, \tag{1.1.9}$$

$$\int_0^1 \int_0^1 Y_n^2(s,t) ds dt \overset{P}{\to} \infty \tag{1.1.10}$$

and

$$\int_0^1 \int_0^1 |Y_n(s,t)| ds dt \overset{P}{\to} \infty. \tag{1.1.11}$$

2. Subsampling independent observations

Our first objective is to check if the tests based on (1.1.4)–(1.1.6), referred to in the following respectively as the Kolmogorov-Smirnov (KS), the Cramér-von Mises (CVM) and L_1 tests, have correct size. For this purpose we generate samples $X_1, \ldots, X_n$ of iid random variables with the same distribution and calculate nominal rejection rates based on $R = 1000$ replications. Note that the suprema and integrals in the left-hand sides of (1.1.4)–(1.1.6) can be calculated exactly as maxima and sums over finite sets. The results are presented in Table B.3.

Subsampling with ordered blocks: To focus attention we describe the proceduret for the CVM test. Denote

$$I_n = \int_0^1 \int_0^1 Y_n^2(s,t) ds dt.$$

Fix an integer $b < n$ and define the blocks $X_k, \ldots X_{k+b-1}$, $1 \le k \le n - b + 1$. For each of these block we compute $I_k(b)$ in the same way that I_n is computed from $X_1, \ldots, X_n$. Let

$$H_{n,b}(x) = \frac{1}{n-b+1} \sum_{k=1}^{n-b+1} \mathbf{1}\{I_k(b) \le x\}$$

and

$$c_b(\alpha) = \sup\{x : H_{n,b}(x) \le 1 - \alpha\}.$$

The null hypothesis of no change in distribution is rejected if $I_n \ge c_b(\alpha)$

In the remainder of this section we will verify that that the above test has correct asymptotic size and is consistent i.e.

$$\lim_{n\to\infty} P(I_n \ge c_b(\alpha)) = \alpha \tag{2.2.1}$$

under H_0 and

$$\lim_{n\to\infty} P(I_n \ge c_b(\alpha)) = 1 \tag{2.2.2}$$

under H_1 provided

$$b \to \infty \quad \text{and} \quad b/n \to 0. \tag{2.2.3}$$

Denote by $H(\cdot)$ the distribution function of the limit in (1.1.5). Since H is continuous, (2.2.1) follows from Theorem 3.5.1(i) of Politis *et al.* (1999). To verify (2.2.2) we show that under H_A for every x

$$\frac{1}{n-b+1}\sum_{k=1}^{n-b+1} \mathbf{1}\{I_k(b) \le x\} \xrightarrow{P} H(x). \tag{2.2.4}$$

Relation (2.2.2) will then follow follow immediately from (1.1.10) and (2.2.4).

Observe that

$$\frac{1}{n-b+1}\sum_{k=1}^{n-b+1} \mathbf{1}\{I_k(b) \le x\} = \frac{1}{n-b+1}\left\{\sum_{1\le k\le[n\lambda]-b} \mathbf{1}\{I_k(b) \le x\}\right. \tag{2.2.5}$$

$$\left. + \sum_{[n\lambda]-b<k\le[n\lambda]} \mathbf{1}\{I_k(b) \le x\} + \sum_{[n\lambda]<k\le n-b+1} \mathbf{1}\{I_k(b) \le x\}\right\}.$$

Since $I_k(b),\ 1 \le k \le [n\lambda]-b$ is computed from independent, identically distributed random variables

$$\frac{1}{n-b+1}\sum_{1\le k\le[n\lambda]-b} \mathbf{1}\{I_k(b) \le x\} \xrightarrow{P} \lambda H(x). \tag{2.2.6}$$

The same argument gives

$$\frac{1}{n-b+1}\sum_{[n\lambda]<k\le n-b+1} \mathbf{1}\{I_k(b) \le x\} \xrightarrow{P} (1-\lambda) H(x). \tag{2.2.7}$$

Since $b/n \to 0$,

$$\frac{1}{n-b+1}\sum_{[n\lambda]-b<k\le[n\lambda]} \mathbf{1}\{I_k(b) \le x\} \xrightarrow{P} 0. \tag{2.2.8}$$

Putting together (2.2.5)–(2.2.8), we see that (2.2.4) holds under H_A.

The consistency of the subsampling tests based on the KS and L_1 statistics is established in the same way.

3. Extension to GARCH processes

The observations $X_1, \dots, X_n$ are said to follow a GARCH(p,q) model if they satisfy the equations:

$$X_t = \sigma_t \varepsilon_t, \tag{3.3.1}$$

$$\sigma_t^2 = \omega + \sum_{j=1}^{p} \alpha_j X_{t-j}^2 + \sum_{j=1}^{q} \beta_j \sigma_{t-j}^2. \tag{3.3.2}$$

The *innovations* ε_k in (3.3.1) are independent and identically distributed and $\omega, \alpha_j, \beta_j$ are nonnegative parameters. A necessary and sufficient condition for the existence of a unique stationary solution to equations (3.3.1), (3.3.2) was given by Bourgerol and Picard (1992a, 1992b). For ease of reference we formulate this condition in the Appendix as relations (A.A.1) and (A.A.2).

We impose the following conditions on the distribution of the innovations ε_t in (3.3.1):

Assumption 3.1. We assume that for some $0 < \theta \le 1$, $C > 0$ and any $t < t'$

$$P\{t < \varepsilon_0 \le t'\} \le C|t - t'|^{\theta} \tag{3.3.3}$$

and

$$E(\ln^+ |\varepsilon_0|)^{2+16/\theta} < \infty. \tag{3.3.4}$$

The common distribution function of the X_t's is denoted by F and its inverse (quantile function) by Q. Let

$$\gamma(s, s') = EJ_0(s)J_0(s') + \sum_{1 \le k < \infty} EJ_0(s)J_k(s') + \sum_{1 \le k < \infty} EJ_0(s')J_k(s), \tag{3.3.5}$$

where $J_k(s) = \mathbf{1}\{X_k \le s\} - F(s)$.

Theorem 3.1. *If conditions* (A.A.1) *and* (A.A.2) *and Assumption 3.1 hold, then there are Gaussian processes* $\Gamma_n(t, s)$ *such that*

$$E\Gamma_n(t, s) = 0 \text{ and } E[\Gamma_n(t, s)\Gamma_n(t', s')] = \gamma(Q(t), Q(t'))(\min(s, s') - ss') \tag{3.3.6}$$

and

$$\sup_{0 \le t \le 1} \sup_{1 \le s \le 1} \left| n^{-1/2} \sum_{1 \le i \le ns} \psi_t(X_i - Q_n(t)) - \Gamma_n(t, s) \right| = o_P(1).$$

The proof of Theorem 3.1 is given in Section A.2 of the Appendix.

Remark 3.1. Theorem 3.1 implies the convegence in distribution of all continuous functionals, in particular counterparts of relations (1.1.4), (1.1.5) and (1.1.6), with Y replaced by Γ, hold.

Next we consider the alternative when the parameters change after $X_{[n\lambda]}$, for some $\lambda \in (0,1)$. We thus assume that (3.3.1) holds, but

$$\sigma_k^2 = \begin{cases} \omega + \sum_{j=1}^p \alpha_j X_{k-j}^2 + \sum_{j=1}^q \beta_j \sigma_{k-1}^2, & -\infty < k \le [n\lambda] \\ \omega' + \sum_{j=1}^p \alpha_j' X_{k-j}^2 + \sum_{j=1}^q \beta_j' \sigma_{k-1}^2, & [n\lambda] < k < \infty, \end{cases} \tag{3.3.7}$$

where

$$\boldsymbol{\theta} := (\omega, \alpha_1, \dots, \alpha_p, \beta_1, \dots, \beta_q) \neq (\omega', \alpha_1', \dots, \alpha_p', \beta_1', \dots, \beta_q') =: \boldsymbol{\theta}'.$$

Replacing $\boldsymbol{\theta}$ by $\boldsymbol{\theta}'$ in Section A.1, we define the matrices $\mathbf{A}_n(\boldsymbol{\theta}')$ and the top Lyapunov exponent $\gamma_L(\boldsymbol{\theta}')$. In the following we assume that conditions (A.A.1) and (A.A.2) hold for the parameter $\boldsymbol{\theta}'$ as well.

Recall that the common distribution function of the X_i, $1 \le i \le [n\lambda]$ is F. In Section A.3 we explain that $\lim_{n\to\infty} P(X_n \le t) =: G(t)$ exists. Adapting the proof in Csörgő and Horváth (1997), p. 103, to the present case, it can be verified that (1.1.7) holds. If we assume that the parameters changes in such a way that $F(t) \neq G(t)$ for some t, then (1.1.9)–(1.1.11) also hold. We can then use the subsampling method as described in Section 2. To justify its asymptotic validity, it is enough to show that (2.2.4) holds both under the null and the alternative. The proof is rather technical and only an outline is given in Section A.3.

Appendix A. Mathematical appendix

A.1. Existence and uniqueness of the GARCH(p,q) process

We begin by introducing the following notation:

$$\boldsymbol{\tau}_n = (\beta_1 + \alpha_1 \varepsilon_n^2, \beta_2, \dots, \beta_{q-1}) \in R^{q-1},$$
$$\boldsymbol{\xi}_n = (\varepsilon_n^2, 0, \dots, 0) \in R^{q-1}$$

and

$$\boldsymbol{\alpha} = (\alpha_2, \dots, \alpha_{p-1}) \in R^{p-2}.$$

Without loss of generality we may and shall assume that $\min(p,q) \ge 2$. Define the $(p+q-1) \times (p+q-1)$ matrix A_n, written in block form, by

$$\mathbf{A}_n = \begin{bmatrix} \boldsymbol{\tau}_n & \beta_q & \boldsymbol{\alpha} & \alpha_p \\ \boldsymbol{I}_{q-1} & 0 & 0 & 0 \\ \boldsymbol{\xi}_n & 0 & 0 & 0 \\ 0 & 0 & \boldsymbol{I}_{p-2} & 0 \end{bmatrix},$$

where $\boldsymbol{I}_d$ is a $d \times d$ identity matrix. The norm of any $d \times d$ matrix $\mathbf{M}$ is defined by

$$||\mathbf{M}|| = \sup\{||\mathbf{Mx}||_d / ||\mathbf{x}||_b : \mathbf{x} \in R^d, \mathbf{x} \neq 0\},$$

where $||\cdot||_d$ is the usual (Euclidean) norm in R^d. The top Lyapunov exponent γ_L associated with the sequence of matrices A_n, $n \ge 0$, is defined as

$$\gamma_L = \gamma_L(\varepsilon) = \inf_{0 \le n < \infty} \frac{1}{n+1} E \ln ||\mathbf{A}_0 \mathbf{A}_1 \dots \mathbf{A}_n||,$$

assuming that

(A.A.1) $$E(\ln ||\mathbf{A}_0||) < \infty.$$

Bourgerol and Picard (1992a, 1992b) showed that if (A.A.1) holds, then equations (3.3.1) and (3.3.2) have a unique stationary solution if and only if

(A.A.2) $$\gamma_L < 0.$$

A.2. Proof of Theorem 3.1

The proof of Theorem 3.1 id based on the following stong approximation due to Berkes and Horváth (2001).

Theorem A.1. *In the notation and under the assumptions of Theorem 3.1, there is a Gaussian process process $\hat{\Gamma}(x,y)$ with*

$$E\hat{\Gamma}(x,y) = 0, \quad E[\hat{\Gamma}(x,y)\hat{\Gamma}(x',y')] = \gamma(x,x')\min(y,y')$$

such that for some $\kappa > 0$

$$\sup_{1\le y\le T}\ \sup_{-\infty<x<\infty}\left|\sum_{1\le k\le y} J_k(x) - \hat{\Gamma}(x,y)\right| = o(T^{1/2}(\log T)^{-\kappa})\ a.s.$$

Lemma A.2. *Under the assumptions of Theorem 3.1,*

$$\sup_{0\le t\le 1}|F(Q_n(t)) - t| = O_P(n^{-1/2}).$$

Proof. By Lemma 1.1 on p. 24 of Csörgő and Horváth (1993),

$$\begin{aligned}\sup_{0\le t\le 1}|F(Q_n(t)) - t| &\le \sup_{0\le t\le F(Q_n(1))}|F_n(Q(t)) - t| \\ &\le \sup_{0\le t\le 1}|F_n(Q(t)) - t| \\ &\le \sup_{-\infty<x<\infty}|F_n(x) - F(x)| = O_P(n^{-1/2}),\end{aligned}$$

on account of Theorem A.1.

Lemma A.3. *Under the assumptions of Theorem 3.1,*

$$\sup_{0\le t\le 1}|t - F(Q_n(t)) - (F_n(Q_n(t)) - F(Q_n(t)))| = o_P(n^{-1/2}(\log n)^{-\kappa}),$$

for $\kappa > 0$ as in Theorem A.1.

Proof. By Lemma 1.2 on p. 25 of Csörgő and Horváth (1993), we have

$$t - F(Q_n(t)) = F_n(Q_n(t)) - F(Q_n(t)) - (F_n(Q_n(t)) - t)$$

and

$$\sup_{0\le t\le 1}|F_n(Q_n(t)) - t| \le \sup_{0\le t\le 1}|F_n(Q(t)) - F_n(Q(t-))|.$$

Using Theorem A.1, we get

$$\sup_{0\le t\le 1}|F_n(Q(t)) - F_n(Q(t-))| \overset{a.s.}{=} o(n^{-1/2}(\log n)^{-\kappa})$$

$$+n^{-1/2} \sup_{-\infty<x<\infty} \left|\hat{\Gamma}(x,n) - \hat{\Gamma}(x-,n)\right| \overset{a.s.}{=} o(n^{-1/2}(\log n)^{-\kappa})$$

by the almost sure continuity of $n^{-1/2}\hat{\Gamma}(x,n)$ (c.f. Berkes and Horváth (2001), p. 803).

Lemma A.4. *Under the assumptions of Theorem 3.1,*

$$\sup_{0\le t\le 1} \left|n^{1/2}(t - F(Q_n(t))) - n^{-1/2}\hat{\Gamma}(Q(t),n)\right| = o_P(1)$$

Proof. Combining Theorem A.1 and Lemma A.3, we conclude

$$\sup_{0\le t\le 1} \left|n^{1/2}(t - F(Q_n(t))) - n^{-1/2}\hat{\Gamma}(Q_n(t),n)\right| \overset{a.s.}{=} o(n^{-1/2}(\log n)^{-\kappa}).$$

Lemma A.2 and the almost sure continuity of $n^{-1/2}\hat{\Gamma}(Q(t),n)$ yield

$$n^{-1/2} \sup_{0\le t\le 1} \left|\hat{\Gamma}(Q_n(t),n) - \hat{\Gamma}(Q(t),n)\right|$$

$$\le n^{-1/2} \sup_{0\le t\le 1-h} \sup_{|x|\le h} \left|\hat{\Gamma}(Q(t+x),n) - \hat{\Gamma}(Q(t),n)\right| = o_P(1),$$

where

$$h = \sup_{0\le t\le 1} |F(Q_n(t)) - t| = O_P(n^{-1/2}). \tag{A.A.3}$$

Proof of Theorem 3.1: We note that

$$\sum_{1\le i\le [ns]} \psi_t(X_i - Q_n(t)) = [ns]\left\{t - F_{[ns]}(Q_n(t))\right\}$$

$$= [ns]\left\{F(Q_n(t)) - F_{[ns]}(Q_n(t))\right\} + [ns]\left\{t - F(Q_n(t))\right\}.$$

Using Theorem A.1, Lemma A.2 and the almost sure continuity of $\hat{\Gamma}(x,y)$, we get

$$n^{-1/2} \sup_{0\le s\le 1} \sup_{0\le t\le 1} \left|[ns]\left\{F_{[ns]}(Q_n(t) - F(Q_n(t))\right\} - \hat{\Gamma}(Q(t),ns)\right|$$

$$\le n^{-1/2} \sup_{0\le s\le 1} \sup_{0\le t\le 1-h} \sup_{|x|\le h} \left|\hat{\Gamma}(Q(t+x),ns) - \hat{\Gamma}(Q(t),ns)\right|$$

$$+n^{-1/2} \sup_{0\le s\le 1} \sup_{0\le t\le 1} \left|[ns]\left\{F_{[ns]}(Q_n(t)) - F(Q_n(t))\right\} - \hat{\Gamma}(Q_n(t),ns)\right| = o_P(1),$$

where h is defined by (A.A.3). Hence Lemma A.4 yields

$$n^{-1/2} \sup_{0\le s\le 1} \sup_{0\le t\le 1} \left|\sum_{1\le i\le [ns]} \psi_t(X_i - Q_n(t)) - \Gamma_n(s,t)\right| = o_P(1),$$

where

$$\Gamma_n(s,t) = n^{-1/2}\left(s\hat{\Gamma}(Q(t),n) - n^{-1/2}\hat{\Gamma}(Q(t),ns)\right).$$

Direct verification shows that $\Gamma_n(s,t)$ is a Gaussian process satisfying (3.3.6).

A.3. Consistency of the subsampling test for GARCH processes

The proof is rather technical and we only outline the main steps. Berkes and Horváth (2001) constructed random variables $X_i' = g_i(\varepsilon_i, \ldots, \varepsilon_{i-[i\rho]})$ with some $0 < \rho < 1$ such that X_i and X_i' are near each other. Let $I_k'(b)$ be defined as $I_k(b)$ but using X_i' rather that X_i, c.f. Section 2. By Lemma 2.5 of Berkes and Horváth (2001)

$$\sup_x \left| \frac{1}{n-b+1} \sum_{k=1}^{n-b+1} \left(\mathbf{1}\left\{I_k'(b) \le x\right\} - \mathbf{1}\left\{I_k(b) \le x\right\} \right) \right| \xrightarrow{P} 0.$$

Using the definition of X_i', we get

$$\operatorname{Var}\left(\frac{1}{n-b+1} \sum_{k=1}^{n-b+1} \mathbf{1}\left\{I_k'(b) \le x\right\} \right) = O\left(n^{1+\rho}(n-b+1)^{-2}\right) = o(1)$$

and

$$\lim_{n\to\infty} \frac{1}{n-b+1} \sum_{k=1}^{n-b+1} P\left\{I_k'(b) \le x\right\} = H(x),$$

completing the proof of (2.2.4) under H_0.

Next we assume that H_A holds. Let $\bar{\varepsilon}_i,\ -\infty < i \le [n\lambda]$ be independent, identically distributed random variables independent of $\{\varepsilon_i,\ -\infty < i < infty\}$ distributed as ε_0. Let $\bar{\varepsilon}_i = \varepsilon_i$ if $i > [n\lambda]$.

We define a new GARCH process

$$\bar{X}_k = \bar{\sigma}_k \bar{\varepsilon}_k,$$

$$\bar{\sigma}_k^2 = \omega' + \sum_{j=1}^{p} \alpha_j' \bar{X}_{k-j}^2 + \sum_{i=1}^{q} \beta_i' \bar{\sigma}_{k-i}^2.$$

Using Bougerol and Picard (1992a) and Lemmas 2.1 and 2.2 of Berkes and Horváth (2001) we get

$$E\left[\frac{1}{n-b+1} \sum_{[n\lambda]<k\le n-b+1} \left(\mathbf{1}\left\{I_k(b) \le x\right\} - \mathbf{1}\left\{\bar{I}_k(b) \le x\right\} \right) \right]^2 = o(1),$$

where in $\bar{I}_k(b)$ we used $\bar{X}_i$ in place of X_i. Observe that the parameters of the X_i remain constant for $-\infty < i \le [n\lambda]$ and there is no change in the parameters of the $\bar{X}_i$.

Arguing as under H_0, we get

$$\frac{1}{n-b+1} \sum_{1\le k\le [n\lambda]-b} \mathbf{1}\left\{I_k(b) \le x\right\} \xrightarrow{P} \lambda H(x)$$

and

$$\frac{1}{n-b+1} \sum_{[n\lambda]<k\le n-b+1} \mathbf{1}\left\{\bar{I}_k(b) \le x\right\} \xrightarrow{P} (1-\lambda) H(x),$$

completing the proof of (2.2.4) under H_A.

Appendix B. Simulation results

We performed a large number of numerical experiments in order to assess the finite sample performance of the tests studied in this paper. To save space we report only some results and describe the general picture that has emerged. Our first finding was that asymptotic CVM and L_1 tests perform very well; they have nearly correct size and good power. The empirical size of the KS test is noticeably smaller than the nominal size and power is also low. In Table B.2 we report empirical sizes based on 500 replications. The standard error of these size estimates is therefore about 0.01. Notice the very good performance of the L^1 test for heavy–tailed stable observations. We recall that the critical values for this test, reported in Table B.1 were obtained using standard normal observations.

Test	α	0.10	0.05	0.02	0.01
KS	$c(\alpha)$	0.715	0.772	0.823	0.874
CVM	$c(\alpha)$	2.286	2.844	3.622	4.230
L^1	$c(\alpha)$	0.155	0.172	0.197	0.214

TABLE B.1. Values $c(\alpha)$ such that $P[\sup_{0<s<1}\sup_{0<t<1}|Y(s,t)| > c(\alpha)] = \alpha$ (KS test), $P[\int_{0<s<1}\int_{0<t<1} Y^2(s,t)dsdt > 2c(\alpha)/\pi^4] = \alpha$ (CVM test) and $P[\int_{0<s<1}\int_{0<t<1}|Y(s,t)|dsdt > c(\alpha)] = \alpha$ (L^1 test). The values $c(\alpha)$ for the L^1 test were obtained by simulating $R = 10000$ times $n = 2000$ standard normal iid observations.

		KS Test			CVM Test			L^1 Test		
	$n \setminus \alpha$	0.10	0.05	0.01	0.10	0.05	0.01	0.10	0.05	0.01
(N)	100	0.025	0.014	0.002	0.082	0.045	0.007	0.092	0.047	0.010
	500	0.049	0.015	0.006	0.093	0.048	0.009	0.093	0.049	0.012
(E)	100	0.022	0.017	0.002	0.096	0.045	0.004	0.109	0.064	0.011
	500	0.055	0.018	0.002	0.105	0.055	0.014	0.098	0.055	0.011
(S)	100	0.034	0.021	0.002	0.099	0.045	0.008	0.125	0.062	0.006
	500	0.045	0.011	0.003	0.083	0.041	0.009	0.091	0.052	0.011

TABLE B.2. Empirical size of the tests based on (1.1.4)–(1.1.6) for standard normal (N), exponential (E) and symmetric α-stable observations with $\alpha = 1.5$ (S).

		CVM Test			L^1 Test		
n	$b \setminus \alpha$	0.10	0.05	0.01	0.10	0.05	0.01
100	5	0.016	0.013	0.007	0.025	0.020	0.013
(E)	10	0.036	0.016	0.002	0.040	0.024	0.009
	15	0.047	0.022	0.008	0.030	0.018	0.007
	20	0.129	0.079	0.025	0.104	0.068	0.024
	25	0.150	0.099	0.053	0.147	0.100	0.044
	30	0.157	0.099	0.059	0.142	0.097	0.062
500	15	0.120	0.084	0.022	0.124	0.072	0.026
(E)	20	0.094	0.048	0.018	0.091	0.048	0.014
	25	0.104	0.054	0.022	0.102	0.062	0.022
	50	0.112	0.066	0.028	0.108	0.051	0.019
	75	0.112	0.068	0.024	0.115	0.067	0.027
	100	0.110	0.058	0.018	0.120	0.067	0.022
100	5	0.024	0.020	0.012	0.024	0.020	0.010
(G)	10	0.040	0.022	0.004	0.042	0.014	0.004
	15	0.044	0.024	0.004	0.042	0.020	0.004
	20	0.128	0.078	0.038	0.164	0.100	0.046
	25	0.094	0.056	0.030	0.084	0.054	0.030
	30	0.135	0.082	0.035	0.094	0.054	0.027
500	15	0.128	0.078	0.020	0.168	0.112	0.050
(G)	20	0.126	0.070	0.026	0.128	0.078	0.034
	25	0.134	0.092	0.048	0.150	0.092	0.044
	50	0.122	0.074	0.036	0.132	0.090	0.030
	75	0.146	0.092	0.052	0.118	0.066	0.030
	100	0.150	0.100	0.046	0.144	0.088	0.034

TABLE B.3. Empirical size of the CVM and L^1 subsampling tests for iid exponential observations (E) and observations from GARCH(1,1) model $\omega = 0.1, \alpha = 0.3, \beta = 0.3$ (G).

Subsampling tests require the choice of the block size b. This is a difficult problem and is beyond the intended scope of this paper. Table B.3 shows that values of b between 15 and 20 might be suitable for $n = 100$ and those between 20 and 15 for $n = 500$. This accords with the theoretical requirement that $b \to \infty$ as $n \to \infty$. The more extensive simulations for various GARCH models show that the probability of type I error increases with increasing parameters α and β which control the tails of the obseravtions and the dependence of current variability on the past variability. For models with β greater than 0.5 the probability of type I error may exceed 0.8 for the values of n and b as in Table B.3. This can be explained heuristically by noting that for relatively short samples such processes may appear to have a change-point because there may be one or two periods of very high variability. The performance of subsampling CVM and L^1 tests is comparable.

References

Berkes, I. and Horváth, L. (2001). Strong approximation of the empirical process of GARCH sequences. *The Annals of Applied Probability*, 11, 789–809.

Blum, J. R., Kiefer, J. and Rossenblatt, M. (1961). Distribution free tests of independence based on the sample distribution function. *The Annals of Mathematical Statistics*, 32, 485–498.

Bougerol, P. and Picard, N. (1992a). Strict stationarity of generalized autoregressive processes. *Annals of Probability*, 20, 1714–1730.

Bougerol, P. and Picard, N. (1992b). Stationarity of GARCH processes and of some nonnegative time series. *Journal of Econometrics*, 52, 115–127.

Cotteril, D. S. and Csörgő, M. (1985). On the limiting distribution of and the critical values for the Hoeffding, Blum, Kiefer and Rosenblatt independence criterion. *Statistics & Decisions*, 3, 1–48.

Csörgő, M. (1979). Strong approximations of the Hoeffding, Blum, Kiefer, Rosenblatt multivariate empirical process. *Journal of Multivariate Analysis*, 9, 84–100.

Csörgő, M. and Horváth, L. (1987). Nonparametric tests for the change point problem. *Journal of Statistical Planning and Inference*, 17, 1–9.

Csörgő, M. and Horváth, L. (1993). *Weighted Approximations in Probability and Statistics.* Wiley, New York.

Csörgő, M. and Horváth, L. (1997). *Limit Theorems in Change-Point Analysis.* Wiley, New York.

Martynov, G. V. (1992). Statistical tests based on empirical processes and related questions. *J. Sov. Math.*, 61, 2195–2273.

Picard, D. (1985). Testing and estimating change–point in time series. *Adv. in Appl. Probab.*, 17, 841–867.

Politis, D. N., Romano, J. P. and Wolf, M. (1999). *Subsampling.* Springer.

Department of Mathematics, University of Utah, 155 S 1400 E, Salt Lake City, UT 84341, USA
E-mail address: horvath@math.utah.edu

Department of Mathematics and Statistics, Utah State University, 3900 Old Main Hill, Logan, UT 84322, USA
E-mail address: ajach@math.usu.edu

Department of Mathematics and Statistics, Utah State University, 3900 Old Main Hill, Logan, UT 84322, USA
E-mail address: piotr@math.usu.edu