



HAL
open science

A test for parameter change in general causal time series models

William Charky Kengne

► **To cite this version:**

William Charky Kengne. A test for parameter change in general causal time series models. 2011.
hal-00561015v1

HAL Id: hal-00561015

<https://hal.science/hal-00561015v1>

Preprint submitted on 31 Jan 2011 (v1), last revised 2 Jul 2011 (v2)

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

A test for parameter change in general causal time series models

William Charky Kengne^{1,*}

SAMM, Université Paris 1 Panthéon-Sorbonne, 90 rue de Tolbiac 75634-Paris Cedex 13, France

Abstract

We consider a process $X = (X_t)_{t \in \mathbb{Z}}$ belonging to a large class of causal models. We assume that the model depends on a parameter θ_0 and consider the problem of test for change of the parameter. The test statistic is based on the quasi-maximum likelihood estimator (QLME) of the parameter. Given a significance level $\alpha \in (0, 1)$, it is shown that the asymptotic size of the test less than α . Under the local alternative that there is one change, we show that the test statistic converges almost surely to ∞ . Some simulation results for AR(1), ARCH(1) and GARCH(1,1) models are reported.

Keywords: Semi-parametric test; Change of parameters; Causal processes; Quasi-maximum likelihood estimator; Weak convergence.

1. Introduction

Many statistical data can be represented by models which may change over time, for instance hydraulic flow, climate data. Before any inference on these data, it is crucial to test whether a change has not occurred in the model.

Since Page [20] in 1955, real advances have been done about tests for change detection. Horvath [9] proposed a test for detecting change in parameter of autoregressive processes based on weighted supremum and L_p -functionals of the residual sums. The CUSUM statistic which was introduced by Brown et al. [7] in 1975, was modified by Inclan and Tiao [11] for testing change in variance of independent random variable. Their test has asymptotically correct size but the consistency in power is unknown. Numerous work devoted to the CUSUM-type procedure (see for instance Lee et al. [17] in the regression models with ARCH errors, Kokoszka and Leipus [14] the case of ARCH(∞) or Aue et al. [1] for testing break in covariance). Horváth et al. [10] suggested to compute the ratio of the CUSUM functionals instead of the differences for testing change in the mean of a time series. Berkes et al. [5] used a test based on approximate likelihood scores for testing parameter constancy in GARCH(p,q) models. These procedures are either done in a parametric case, either the asymptotic power is unknown. The present work is a new

*Tel.: +33(0)144078706.

Email address: William-Charky.Kengne@malix.univ-paris1.fr (William Charky Kengne)

¹Supported by AUF (Agence Universitaire de la Francophonie) and Edulink ACP-EU project.

contribution to the problem of test for change detection.

In this paper, we consider a general class $\mathcal{M}_T(M, f)$ of causal (non-anticipative) time series. Let $M, f : \mathbb{R}^N \rightarrow \mathbb{R}$ be measurable functions, $(\xi_t)_{t \in \mathbb{Z}}$ be a sequence of centered independent and identically distributed (iid) random vectors called the innovations and satisfying $\text{var}(\xi_0) = \sigma^2$ and Θ a compact subset of \mathbb{R}^d . Let $T \subset \mathbb{Z}$, for any $\theta \in \Theta$, define

Class $\mathcal{M}_T(M_\theta, f_\theta)$: *The process $X = (X_t)_{t \in \mathbb{Z}}$ belongs to $\mathcal{M}_T(M_\theta, f_\theta)$ if it satisfies the relation:*

$$X_{t+1} = M_\theta((X_{t-i})_{i \in \mathbb{N}})\xi_t + f_\theta((X_{t-i})_{i \in \mathbb{N}}) \quad \text{for all } t \in T. \quad (1.1)$$

The existence and properties of these general affine processes were studied in Bardet and Wintenberger [2]. Numerous classical time series are included in $\mathcal{M}_{\mathbb{Z}}(M, f)$: for instance AR(∞), ARCH(∞), TARARCH(∞), ARMA-GARCH or bilinear processes.

Now, assume that a trajectory (X_1, \dots, X_n) of $X = (X_t)_{t \in \mathbb{Z}}$ is observed and consider the following hypothesis:

H_0 : there exists $\theta_0 \in \Theta$ such that (X_1, \dots, X_n) belong in the class $\mathcal{M}_{\{1, \dots, n\}}(M_{\theta_0}, f_{\theta_0})$;
 H_1 : there exists $K \geq 2$, $\theta_1, \dots, \theta_K \in \Theta$ a partition $\{T_1^n, \dots, T_K^n\}$ of $\{1, \dots, n\}$ such that (X_1, \dots, X_n) belong in $\bigcap_{j=1}^K \mathcal{M}_{T_j^n}(M_{\theta_j}, f_{\theta_j})$.

Thus, it is easy to see that under H_1 all stationary property is lost after the first change. This is not the case in many existing works (for instance Kouamo et al. [12]) where the stationarity or both K-th order stationarity after the change is an essential assumption.

This work is devoted to the test for change detection (see Bardet et al. [3] for procedure of estimation of instants of change). We consider semi-parametric test statistics based on the QLME (see Bardet and Wintenberger [2] for more reference for QLME) which is notably a modification of the statistic introduced by Lee et al. [15]. Then, we obtain the consistency in power under the alternative of one change and simulation results compared to some other approaches show that our procedure is more powerful. In Section 2 we present assumptions and construct test statistics. In Section 3 we give some asymptotic results. The empirical study of AR(1), ARCH(1) and GARCH(1,1) are detailed in Section 4 and the proofs of main results are presented in Section 5.

2. Assumptions and test statistics

2.1. Assumptions on the class of models $\mathcal{M}_{\mathbb{Z}}(f_\theta, M_\theta)$

Let $\theta \in \mathbb{R}^d$ and M_θ and f_θ be numerical functions such that for all $(x_i)_{i \in \mathbb{N}} \in \mathbb{R}^{\mathbb{N}}$, $M_\theta((x_i)_{i \in \mathbb{N}}) \neq 0$ and $f_\theta((x_i)_{i \in \mathbb{N}}) \in \mathbb{R}$. We use the following different norms:

1. $\|\cdot\|$ applied to a vector denotes the Euclidean norm of the vector;
2. for any compact set $\Theta \subseteq \mathbb{R}^d$ and for any $g : \Theta \rightarrow \mathbb{R}^d$; $\|g\|_{\Theta} = \sup_{\theta \in \Theta} (\|g(\theta)\|)$.

Let $\Psi_{\theta} = f_{\theta}$, M_{θ} and $i = 0, 1, 2$, then for any compact set $\Theta \subseteq \mathbb{R}^d$, define

Assumption $\mathbf{A}_i(\Psi_{\theta}, \Theta)$: Assume that $\|\partial^i \Psi_{\theta}(0)/\partial \theta^i\|_{\Theta} < \infty$ and there exists a sequence of non-negative real number $(\alpha_i^{(k)}(\Psi_{\theta}, \Theta))_{i \geq 1}$ such that $\sum_{k=1}^{\infty} \alpha_k^{(i)}(\Psi_{\theta}, \Theta) < \infty$ satisfying

$$\left\| \frac{\partial^i \Psi_{\theta}(x)}{\partial \theta^i} - \frac{\partial^i \Psi_{\theta}(y)}{\partial \theta^i} \right\|_{\Theta} \leq \sum_{k=1}^{\infty} \alpha_k^{(i)}(\Psi_{\theta}, \Theta) |x_k - y_k| \quad \text{for all } x, y \in \mathbb{R}^N.$$

The sequel refer to the particular case called "ARCH-type process" if $f_{\theta} = 0$ and if the following assumption holds on $h_{\theta} = M_{\theta}^2$:

Assumption $\mathbf{A}_i(h_{\theta}, \Theta)$: Assume that $\|\partial^i h_{\theta}(0)/\partial \theta^i\|_{\Theta} < \infty$ and there exists a sequence of non-negative real number $(\alpha_i^{(k)}(h_{\theta}, \Theta))_{i \geq 1}$ such as $\sum_{k=1}^{\infty} \alpha_k^{(i)}(h_{\theta}, \Theta) < \infty$ satisfying

$$\left\| \frac{\partial^i h_{\theta}(x)}{\partial \theta^i} - \frac{\partial^i h_{\theta}(y)}{\partial \theta^i} \right\|_{\Theta} \leq \sum_{k=1}^{\infty} \alpha_k^{(i)}(h_{\theta}, \Theta) |x_k^2 - y_k^2| \quad \text{for all } x, y \in \mathbb{R}^N.$$

Assumption $\mathbf{D}(\Theta)$: $\exists \underline{h} > 0$ such that $\inf_{\theta \in \Theta} (|h_{\theta}(x)|) \geq \underline{h}$ for all $x \in \mathbb{R}^N$.

Assumption $\mathbf{Id}(\Theta)$: For all $\theta, \theta' \in \Theta^2$,

$$\left(f_{\theta}(X_0, X_{-1}, \dots) = f_{\theta'}(X_0, X_{-1}, \dots) \text{ and } h_{\theta}(X_0, X_{-1}, \dots) = h_{\theta'}(X_0, X_{-1}, \dots) \text{ a.s.} \right) \Rightarrow \theta = \theta'.$$

Assumption $\mathbf{Var}(\Theta)$: For all $\theta \in \Theta$, one of the families $\left(\frac{\partial f_{\theta}}{\partial \theta^i}(X_0, X_{-1}, \dots)\right)_{1 \leq i \leq d}$ or $\left(\frac{\partial h_{\theta}}{\partial \theta^i}(X_0, X_{-1}, \dots)\right)_{1 \leq i \leq d}$ is a.s. linearly independent.

Assumption $\mathbf{D}(\Theta)$ will be required to define the QMLE, $\mathbf{Id}(\Theta)$ to the consistence of the QMLE and $\mathbf{Var}(\Theta)$ to asymptotic normality.

As in [2], make the convention that if $\mathbf{A}_i(M_{\theta}, \Theta)$ holds then $\alpha_{\ell}^{(i)}(h_{\theta}, \Theta) = 0$ and if $\mathbf{A}_i(h_{\theta}, \Theta)$ holds then $\alpha_{\ell}^{(i)}(M_{\theta}, \Theta) = 0$. Denote :

Assumption $\mathbf{K}(f_{\theta}, M_{\theta}, \Theta)$: for $i = 0, 1, 2$, $\mathbf{A}_i(f_{\theta}, \Theta)$ and $\mathbf{A}_i(M_{\theta}, \Theta)$ (or $\mathbf{A}_i(h_{\theta}, \Theta)$) hold with $\alpha_j^{(i)}(f, \Theta) + \alpha_j^{(i)}(M, \Theta) + \alpha_j^{(i)}(h, \Theta) = \mathcal{O}(j^{-l})$ for some $l > 2$, for $i = 0, 1$.

Also define the set :

$$\begin{aligned} \Theta(r) := & \{\theta \in \Theta, A_0(f_\theta, \{\theta\}) \text{ and } A_0(M_\theta, \{\theta\}) \text{ hold with } \sum_{k \geq 1} \alpha_k^{(0)}(f_\theta, \theta) + (\mathbb{E}|\xi_0|^r)^{1/r} \sum_{k \geq 1} \alpha_k^{(0)}(M_\theta, \theta) < 1\} \\ & \cup \{\theta \in \Theta, f_\theta = 0 \text{ and } A_0(h_\theta, \{\theta\}) \text{ hold with } (\mathbb{E}|\xi_0|^r)^{2/r} \sum_{k \geq 1} \alpha_k^{(0)}(h_\theta, \theta) < 1\}. \end{aligned}$$

Throughout the following, we assume that the functions $\theta \mapsto M_\theta, f_\theta$ are 2-times continuously differentiable on Θ .

2.2. Test statistics

Assume that a trajectory (X_1, \dots, X_n) is observed. It is clear that if $(X_1, \dots, X_n) \in \mathcal{M}_{\{1, \dots, n\}}(M_\theta, f_\theta)$, then for $T \subset \{1, \dots, n\}$, the conditional quasi-(log)likelihood computed on T is given by :

$$L_n(T, \theta) := -\frac{1}{2} \sum_{t \in T} q_t(\theta) \quad \text{with} \quad q_t(\theta) = \frac{(X_t - f_\theta^t)^2}{h_\theta^t} + \log(h_\theta^t)$$

where $f_\theta^t = f_\theta(X_{t-1}, X_{t-2}, \dots)$, $M_\theta^t = M_\theta(X_{t-1}, X_{t-2}, \dots)$ and $h_\theta^t = M_\theta^{t^2}$. Therefore, we approximate the conditional log-likelihood with :

$$\widehat{L}_n(T, \theta) := -\frac{1}{2} \sum_{t \in T} \widehat{q}_t(\theta) \quad \text{where} \quad \widehat{q}_t(\theta) := \frac{(X_t - \widehat{f}_\theta^t)^2}{\widehat{h}_\theta^t} + \log(\widehat{h}_\theta^t)$$

with $\widehat{f}_\theta^t = f_\theta(X_{t-1}, \dots, X_1, 0, 0, \dots)$, $\widehat{M}_\theta^t = M_\theta(X_{t-1}, \dots, X_1, 0, 0, \dots)$ and $\widehat{h}_\theta^t = (\widehat{M}_\theta^t)^2$.

For $T \subset \{1, \dots, n\}$, define the estimator $\widehat{\theta}_n(T) := \operatorname{argmax}_{\theta \in \Theta} (\widehat{L}_n(T, \theta))$. Moreover, for $1 \leq k \leq n$, denote $T_k = \{1, \dots, k\}$ and $\overline{T}_k = \{k+1, \dots, n\}$.

Now, define

$$\widehat{G}_n(T) := \frac{1}{\operatorname{Card}(T)} \sum_{t \in T} \left(\frac{\partial \widehat{q}_t(\widehat{\theta}_n(T))}{\partial \theta} \right) \left(\frac{\partial \widehat{q}_t(\widehat{\theta}_n(T))}{\partial \theta} \right)' \quad \text{and} \quad \widehat{F}_n(T) := -\frac{2}{\operatorname{Card}(T)} \left(\frac{\partial^2 \widehat{L}_n(T, \widehat{\theta}_n(T))}{\partial \theta \partial \theta'} \right).$$

For $k = 1, \dots, n-1$, denote :

$$\widehat{\Sigma}_{n,k} := \frac{k}{n} \widehat{F}_n(T_k) \widehat{G}_n(T_k)^{-1} \widehat{F}_n(T_k) \mathbf{1}_{\det(\widehat{G}_n(T_k)) \neq 0} + \frac{n-k}{n} \widehat{F}_n(\overline{T}_k) \widehat{G}_n(\overline{T}_k)^{-1} \widehat{F}_n(\overline{T}_k) \mathbf{1}_{\det(\widehat{G}_n(\overline{T}_k)) \neq 0}.$$

Let $(v_n)_{n \in \mathbb{N}}$ be a sequences satisfying $v_n \leq n$, $v_n \rightarrow \infty$ and $v_n/n \rightarrow 0$ (as $n \rightarrow \infty$). Denote $\Pi_n = [v_n, n - v_n] \cap \mathbb{N}$ and define the statistics:

$$\widehat{Q}_n^{(1)} := \max_{k \in \Pi_n} \widehat{Q}_{n,k}^{(1)} \quad \text{where} \quad \widehat{Q}_{n,k}^{(1)} := \frac{k^2}{n} (\widehat{\theta}_n(T_k) - \widehat{\theta}_n(T_n))' \widehat{\Sigma}_{n,k} (\widehat{\theta}_n(T_k) - \widehat{\theta}_n(T_n)).$$

$$\widehat{Q}_n^{(2)} := \max_{k \in \Pi_n} \widehat{Q}_{n,k}^{(2)} \quad \text{where} \quad \widehat{Q}_{n,k}^{(2)} := \frac{(n-k)^2}{n} (\widehat{\theta}_n(\overline{T}_k) - \widehat{\theta}_n(T_n))' \widehat{\Sigma}_{n,k} (\widehat{\theta}_n(\overline{T}_k) - \widehat{\theta}_n(T_n)).$$

$$\widehat{Q}_n := \max(\widehat{Q}_n^{(1)}, \widehat{Q}_n^{(2)}) \quad \text{which is the test statistic.}$$

Remark 2.1. In our simulation, we choice $v_n = (\ln n)^3$. Note that, in practice the computation of $\widehat{Q}_{n,k}^{(j)}$ is not easy if k is too small or very close to n .

3. Asymptotic results

3.1. Asymptotic under null hypothesis

Theorem 3.1. Assume $\mathbf{D}(\Theta)$, $\mathbf{Id}(\Theta)$, \mathbf{Var} and $\mathbf{K}(f_\theta, M_\theta, \Theta)$. Under null hypothesis, if $\theta_0 \in \overset{\circ}{\Theta}(4)$, then for $j = 1, 2$

$$\widehat{Q}_n^{(j)} \xrightarrow[n \rightarrow \infty]{\mathcal{D}} \sup_{0 \leq \tau \leq 1} \|W_d(\tau)\|^2$$

where W_d is a d -dimensional Brownian bridges.

For any $\alpha \in (0, 1)$, C_α denotes the $(1 - \alpha/2)$ -quantile of the distribution of $\sup_{0 \leq \tau \leq 1} \|W_d(\tau)\|^2$. Then, the following corollary is a direct application of Theorem 3.1.

Corollary 3.1. Under assumptions of Theorem 3.1 :

$$\forall \alpha \in (0, 1) \quad \limsup_{n \rightarrow \infty} P(\widehat{Q}_n > C_\alpha) \leq \alpha.$$

Remark 3.1. 1. Theorem 3.1 and Corollary 3.1 imply that a large value of \widehat{Q}_n means there is a change in the model. At a nominal level α , the critical region of the test is $(\widehat{Q}_n > C_\alpha)$.

2. Quantile values of the distribution of $\sup_{0 \leq \tau \leq 1} \|W_d(\tau)\|^2$ are known (see Kieffer [13] for $d \in \{1, \dots, 5\}$ or Lee et al. [15] for $d \in \{1, \dots, 10\}$).

Figure 1 is an illustration of the test procedure for AR(1) process. At a level $\alpha = 0.05$, the empirical $(1 - \alpha/2)$ quantile of $\sup_{0 \leq \tau \leq 1} (W_1(\tau))^2$ is equal to 2.20. 1 a-) and b-) show that, the values of $\widehat{Q}_{n,k}^{(1)}$ and $\widehat{Q}_{n,k}^{(2)}$ are all below the red line which represents the limit of the critical region. Figure 1 c-) and d-) show that $\widehat{Q}_{n,k}^{(1)}$ and $\widehat{Q}_{n,k}^{(2)}$ are larger and increases around the point where the change occurred. We will show in the next section that under the alternative that there is one change, the statistic $\widehat{Q}_n = \max(\widehat{Q}_n^{(1)}; \widehat{Q}_n^{(2)})$ is asymptotically infinite.

3.2. The asymptotic under some local alternative

In this subsection, we consider a local alternative that there is one change in the model. More precisely, define

$$H_1^{(loc)} : \text{there exists } \tau^* \in (0, 1) \text{ and } \theta_1^*, \theta_2^* \in \Theta \text{ with } \theta_1^* \neq \theta_2^* \text{ such that } X_1, \dots, X_{[n\tau^*]} \in \mathcal{M}_{T_{[n\tau^*]}}(M_{\theta_1^*}, f_{\theta_1^*}) \\ \text{and } X_{[n\tau^*]+1}, \dots, X_n \in \mathcal{M}_{\overline{T}_{[n\tau^*]}}(M_{\theta_2^*}, f_{\theta_2^*}).$$

Theorem 3.2. Assume $\mathbf{D}(\Theta)$, $\mathbf{Id}(\Theta)$, \mathbf{Var} and $\mathbf{K}(f_\theta, M_\theta, \Theta)$. Under $H_1^{(loc)}$, if $\theta_1^*, \theta_2^* \in \overset{\circ}{\Theta}(4)$, then

$$\widehat{Q}_n \xrightarrow[n \rightarrow \infty]{a.s.} \infty.$$

Remark 3.2. 1. Theorem 3.2 shows that the test with local alternative $H_1^{(loc)}$ is consistent in power.
2. This procedure can be used to test multiple change using ICSS type algorithm developed by Inclán and Tiao [11].

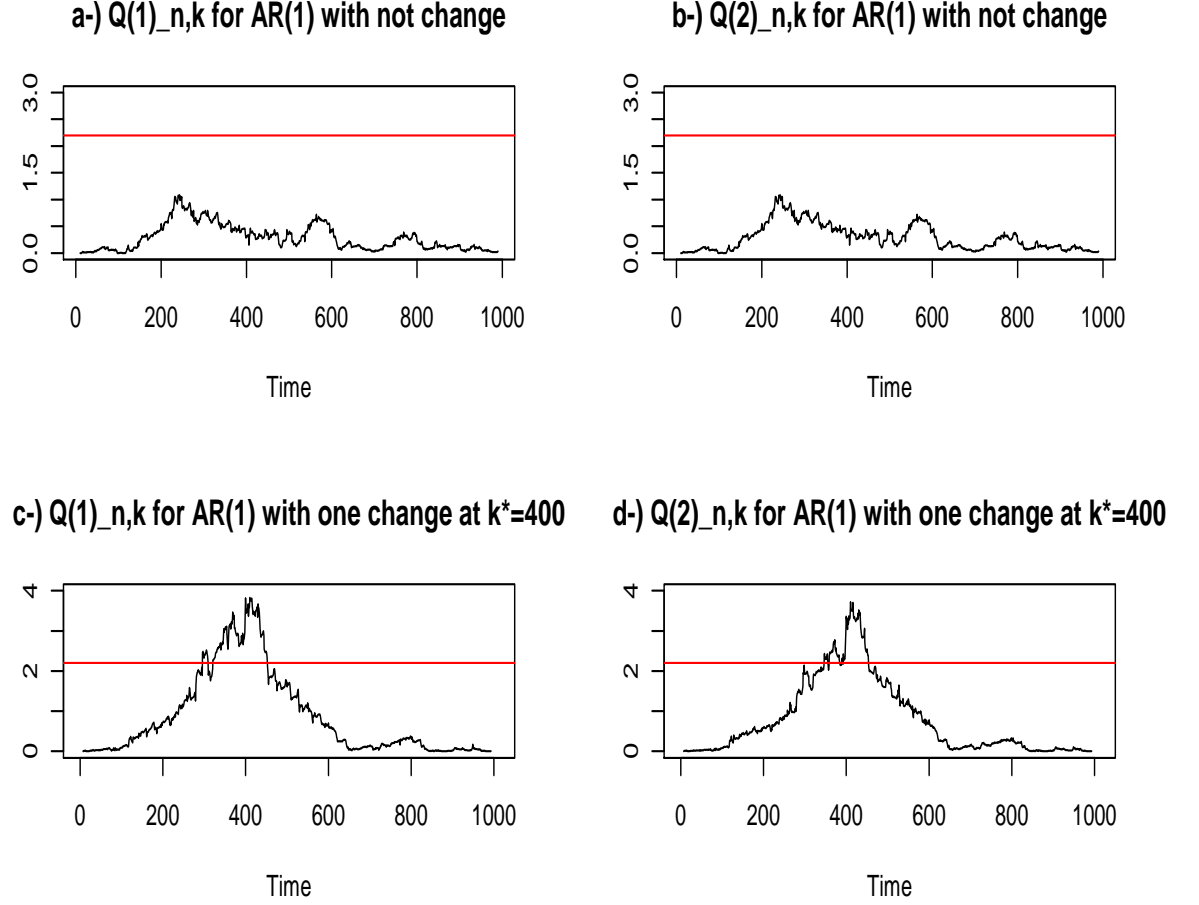


Figure 1: The previous statistics compute for 1000 sample of AR(1). a-) and b-) are respectively $\widehat{Q}_{n,k}^{(1)}$ and $\widehat{Q}_{n,k}^{(2)}$ for AR(1) with parameter $\phi_1 = 0.3$ does not change. c-) and d-) are $\widehat{Q}_{n,k}^{(1)}$ and $\widehat{Q}_{n,k}^{(2)}$ for AR(1) with parameter $\phi_1 = 0.3$ changes to 0.5 at $k^* = 400$

4. Some examples.

In this section, we evaluate the performance of our procedure through empirical study. We compare our results with those obtained by Kouamo et al. [12], Lee and Na [16] and Lee et al. [17]. For a sample size n , \widehat{Q}_n is computed with $v_n = (\ln n)^3$ and is compared to the critical value of the test.

4.1. Test for parameter change in AR(p) models.

Let us consider a AR(p) process : $X_t = \sum_{k=1}^p \phi_k X_{t-k} + \xi_t$ with $p \in \mathbb{N}^*$. The parameter of model is denote by $\theta = (\phi_1, \dots, \phi_p)$ and let $\Theta = \{\theta = (\phi_1, \dots, \phi_p) \in \mathbb{R}^p / \sum_{i=1}^p |\phi_i| < 1\}$. Since $M_\theta \equiv 1$, $\Theta(r) = \Theta$

for any $r \geq 1$. Assume (X_1, \dots, X_n) is observed, we have for any $\theta \in \Theta$, $\widehat{q}_t(\theta) = (X_t - \sum_{k=1}^p \phi_k X_{t-k})^2$,
 $\frac{\partial \widehat{q}_t(\theta)}{\partial \theta} = -2(X_t - \sum_{k=1}^p \phi_k X_{t-k}) \cdot (X_{t-1}, X_{t-2}, \dots, X_{t-p})$ and for $1 \leq i, j \leq n$ $\frac{\partial^2 \widehat{q}_t(\theta)}{\partial \phi_i \partial \phi_j} = 2X_{t-i}X_{t-j}$.

We consider a AR(1) process with one parameter. At level $\alpha = 0.05$, the critical value is $C_\alpha = 2.20$. For $n = 1024, 2048, 4096$; we simulate n observations of AR(1), in the following situations : (i) there is not change, the parameter of model is $\theta_0 = 0.9$ and (ii) there is one change, the parameter $\theta_0 = 0.9$ changes to θ_1 at $n/2$. The following table indicate the proportion of number of rejections of null hypothesis out of 100 repetitions.

| | $n = 1024$ | $n = 2048$ | $n = 4096$ |
|--|-----------------------|-----------------------|-----------------------|
| Empirical levels | 0.080 (0.134 ; 0.092) | 0.070 (0.100 ; 0.062) | 0.050 (0.082 ; 0.040) |
| Empirical powers when $\theta_1 = 0.5$ | 0.980 (0.590 ; 0.530) | 0.990 (0.720 ; 0.680) | 0.990 (0.810 ; 0.790) |

Table 1: Empirical levels and powers at nominal level 0.05 of test for parameter change in AR(1) model. Empirical levels are computed when $\theta_0 = 0.9$; empirical powers are computed when θ_0 changes to θ_1 at $n/2$. Figures in brackets the results obtained by Kouamo et al. [12] at the scale J=4 with KSM and CVM statistic in wavelet domain.

Table 1 shows that the empirical level of the test decreases as n increases and equals to 0.05 when $n = 4096$. These levels are close to those obtained by Kouamo et al. with CVM (Cramér-Von Mises) test statistics. However, our test is more powerful.

4.2. Test for parameter change in GARCH(1,1) models.

Consider the GARCH(1,1) model defined by:

$$\forall t \in \mathbb{Z}, X_t = \sigma_t \xi_t \quad \text{with} \quad \sigma_t^2 = a_0 + a_1 X_{t-1}^2 + b_1 \sigma_{t-1}^2 \quad a_0, a_1, b_1 \geq 0.$$

Assume $a_0 > 0$ and $a_1 + b_1 < 1$. From Nelson and Cao, 1992 (see [19]), we have $\sigma_t^2 = a_0/(1 - b_1) + a_1 \sum_{k \geq 1} b_1^{k-1} X_{t-k}^2$. Thus, the ARCH(∞) representation of model is given by

$$X_t = \sqrt{a_0/(1 - b_1) + a_1 X_{t-1}^2 + a_1 \sum_{k \geq 2} b_1^{k-1} X_{t-k}^2} \cdot \xi_t \quad \text{for all } t \in \mathbb{Z}.$$

Assume $\mathbb{E}\xi^4 = 1$ and denote $\Theta(4) = \Theta = \{\theta = (a_0, a_1, b_1) \in (\mathbb{R}_+)^3; a_0 > 0, a_1 + b_1 < 1\}$. For all $\theta \in \Theta$, denote $f_\theta^t \equiv 0$ and $h_\theta^t = (M_\theta^t)^2$ with $M_\theta^t = \sqrt{a_0/(1 - b_1) + a_1 X_{t-1}^2 + a_1 \sum_{k \geq 2} b_1^{k-1} X_{t-k}^2}$.

Assume (X_1, \dots, X_n) is observed, for any $\theta \in \Theta$ and $t = 2, \dots, n$, we have

$$\widehat{h}_\theta^t = a_0/(1 - b_1) + a_1 X_{t-1}^2 + a_1 \sum_{k=2}^t b_1^{k-1} X_{t-k}^2 \quad \text{and} \quad \widehat{q}_t(\theta) = X_t^2 / \widehat{h}_\theta^t + \log(\widehat{h}_\theta^t).$$

Therefore, it follows that $\frac{\partial \widehat{q}_t(\theta)}{\partial \theta} = \frac{1}{\widehat{h}_\theta^t} \left(1 - \frac{X_t^2}{\widehat{h}_\theta^t}\right) \left(\frac{\partial \widehat{h}_\theta^t}{\partial a_0}, \frac{\partial \widehat{h}_\theta^t}{\partial a_1}, \frac{\partial \widehat{h}_\theta^t}{\partial b_1}\right)$ with $\partial \widehat{h}_\theta^t / \partial a_0 = 1/(1 - b_1)$, $\partial \widehat{h}_\theta^t / \partial a_1 = X_{t-1}^2 + \sum_{k=2}^t b_1^{k-1} X_{t-k}^2$ and $\partial \widehat{h}_\theta^t / \partial b_1 = a_0/(1 - b_1)^2 + a_1 X_{t-2}^2 + a_1 \sum_{k=3}^t (k-1) b_1^{k-2} X_{t-k}^2$.

Let $\theta = (a_0, a_1, b_1) = (\theta_1, \theta_2, \theta_3) \in \Theta$, for $1 \leq i, j \leq 3$, we have

$$\frac{\partial^2 \widehat{q}_t(\theta)}{\partial \theta_i \partial \theta_j} = \frac{1}{(\widehat{h}_\theta^t)^2} \left(\frac{2X_t^2}{\widehat{h}_\theta^t} - 1 \right) \frac{\partial \widehat{h}_\theta^t}{\partial \theta_i} \frac{\partial \widehat{h}_\theta^t}{\partial \theta_j} + \frac{1}{\widehat{h}_\theta^t} \left(1 - \frac{X_t^2}{\widehat{h}_\theta^t} \right) \frac{\partial^2 \widehat{h}_\theta^t}{\partial \theta_i \partial \theta_j}$$

with $\partial^2 \widehat{h}_\theta^t / \partial a_0^2 = 0$, $\partial^2 \widehat{h}_\theta^t / \partial a_0 \partial a_1 = 0$, $\partial^2 \widehat{h}_\theta^t / \partial a_0 \partial b_1 = 1/(1 - b_1)^2$, $\partial^2 \widehat{h}_\theta^t / \partial a_1^2 = 0$, $\partial^2 \widehat{h}_\theta^t / \partial a_1 \partial b_1 = X_{t-2}^2 + \sum_{k=3}^t (k-1)b_1^{k-2} X_{t-k}^2$, and $\partial \widehat{h}_\theta^t / \partial b_1^2 = 2a_0/(1 - b_1)^3 + 2a_1 X_{t-3}^2 + a_1 \sum_{k=4}^t (k-1)(k-2)b_1^{k-3} X_{t-k}^2$.

1. Case of ARCH(1). Takes $b_1 = 0$ and $\theta = (a_0, a_1)$. At level $\alpha = 0.05$, the critical value is $C_\alpha = 3.02$. For $n = 500, 800, 1000$; we simulate n observations of ARCH(1), in the following situations : (i) there is not change, the parameter of model is $\theta_0 = (1, 0.3)$ and (ii) there is one change, the parameter $\theta_0 = (1, 0.3)$ changes to θ_1 at $n/2$. The following table indicate the proportion of number of rejections of null hypothesis out of 500 repetitions.

| | $n = 500$ | $n = 800$ | $n = 1000$ |
|---|---------------|---------------|---------------|
| Empirical levels | 0.068 (0.088) | 0.048 (0.080) | 0.036 (0.074) |
| Empirical powers when $\theta_1 = (0.5, 0.3)$ | 0.948 (0.922) | 0.972 (0.987) | 0.998 (0.995) |
| Empirical powers when $\theta_1 = (0.5, 0.9)$ | 0.862 (0.626) | 0.968 (0.687) | 0.994 (0.716) |

Table 2: Empirical levels and powers at nominal level 0.05 of test for parameter change in ARCH(1) model. Empirical levels are computed when $\theta_0 = (1, 0.3)$; empirical powers are computed when θ_0 changes to θ_1 at $n/2$. Figures in brackets the results obtained by Lee and Na [16].

2. Case of GARCH(1,1). Now $\theta = (a_0, a_1, b_1)$. At level $\alpha = 0.05$, the critical value is $C_\alpha = 3.47$. For $n = 500, 800, 1000$; we simulate n observations of GARCH(1,1), in the following situations : (i) there is not change, the parameter of model is $\theta_0 = (0.5, 0.2, 0.2)$ and (ii) there is one change, the parameter $\theta_0 = (0.5, 0.2, 0.2)$ changes to θ_1 at $n/2$. The following table indicate the proportion of number of rejections of null hypothesis out of 500 repetitions.

| | $n = 500$ | $n = 800$ | $n = 1000$ |
|--|---------------|---------------|---------------|
| Empirical levels | 0.080 (0.026) | 0.070 (0.033) | 0.066 (0.049) |
| Empirical powers when $\theta_1 = (3.0, 0.2, 0.2)$ | 0.592 (0.306) | 0.716 (0.866) | 0.870 (0.990) |
| Empirical powers when $\theta_1 = (0.5, 0.2, 0.6)$ | 0.740 (0.537) | 0.942 (0.806) | 0.960 (0.902) |

Table 3: Empirical levels and powers at nominal level 0.05 of test for parameter change in GARCH(1,1) model. Empirical levels are computed when $\theta_0 = (0.5, 0.2, 0.2)$; empirical powers are computed when θ_0 changes to θ_1 at $n/2$. Figures in brackets the results obtained by Lee et al. [17].

Table 2 and Table 3 show that the empirical level of the test decreases and the empirical power increases as n increases. For ARCH model, we can see that the level less than 0.05 when $n = 800$ it is not the case

for GARCH model. This is explained by the fact that the application of the procedure to GARCH model requires ARCH(∞) representation. Thus, the information contained in all the past of the process is not used because it is not observed. The comparison with results of Lee et al. and Lee and Na shows that in general, our procedure is more powerful.

5. Proofs of the main result

Let $(\psi_n)_n$ and $(r_n)_n$ be sequences of random variables. Throughout this section, we use the notation $\psi_n = o_P(r_n)$ to mean : for all $\varepsilon > 0$, $P(|\psi_n| \geq \varepsilon |r_n|) \rightarrow 0$ as $n \rightarrow \infty$. Write $\psi_n = O_P(r_n)$ to mean : for all $\varepsilon > 0$, there exists $C > 0$ such that $P(|\psi_n| \geq C|r_n|) < \varepsilon$ for n large enough.

5.1. Some preliminary result

First, let us prove a useful technical lemma.

Under the null hypothesis that the observation (X_1, \dots, X_n) belong in the class $\mathcal{M}_{\{1, \dots, n\}}(M_{\theta_0}, f_{\theta_0})$, define the matrix $G := \mathbb{E} \left[\frac{\partial q_0(\theta_0)}{\partial \theta} \frac{\partial q_0(\theta_0)'}{\partial \theta} \right]$ (where ' denotes the transpose) and $F := \mathbb{E} \left[\frac{\partial^2 q_0(\theta_0)}{\partial \theta \partial \theta'} \right]$. Under assumption **Var**, F is a non-singular matrix (see [2]).

Lemma 5.1. *Assume the functions $\theta \mapsto M_\theta$ and $\theta \mapsto f_\theta$ are 2-times continuously differentiable on Θ . Under null hypothesis $\mathbf{D}(\Theta)$ and **Var**, G is a symmetric, positive definite matrix.*

Proof It is clear that G is symmetric. Moreover, for $1 \leq i \leq d$, we have :

$$\frac{\partial q_0(\theta_0)}{\partial \theta_i} = -2 \frac{\xi_0}{\sqrt{h_{\theta_0}^0}} \frac{\partial f_{\theta_0}^0}{\partial \theta_i} - \frac{\xi_0^2}{h_{\theta_0}^0} \frac{\partial h_{\theta_0}^0}{\partial \theta_i} + \frac{1}{h_{\theta_0}^0} \frac{\partial h_{\theta_0}^0}{\partial \theta_i}$$
Thus, using independence of ξ_0 and X_{-1}, X_{-2}, \dots we obtain :

$$\mathbb{E} \left[\frac{\partial q_0(\theta_0)'}{\partial \theta} \frac{\partial q_0(\theta_0)}{\partial \theta} \right] = 4\mathbb{E} \left[\frac{1}{h_{\theta_0}^0} \frac{\partial f_{\theta_0}^0}{\partial \theta} \frac{\partial f_{\theta_0}^0}{\partial \theta} \right] + \mathbb{E} \left((\xi_0^2 - 1)^2 \right) \mathbb{E} \left[\frac{1}{(h_{\theta_0}^0)^2} \frac{\partial h_{\theta_0}^0}{\partial \theta} \frac{\partial h_{\theta_0}^0}{\partial \theta} \right]. \quad (5.1)$$

Since $\mathbb{E}\xi_0^2 = 1$, it is easy to see that $\mathbb{E} \left((\xi_0^2 - 1)^2 \right) > 0$.

Under **Var**, one of the two matrix of the right-hand side of relation (5.1) is positive definite and the other is semi-positive definite. Thus, G is positive definite. \blacksquare

Now, recall that $F := \mathbb{E} \left[\frac{\partial^2 q_0(\theta_0)}{\partial \theta \partial \theta'} \right]$. Let $T \subset \{1, \dots, n\}$. For any $\theta \in \Theta$ and $i = 1, \dots, d$, by Taylor expansion of $\partial L_n(T, \theta_0) / \partial \theta_i$, there exist $\bar{\theta}_{n,i} \in [\theta_0, \theta]$ such that:

$$\frac{\partial L_n(T, \theta)}{\partial \theta_i} = \frac{\partial L_n(T, \theta_0)}{\partial \theta_i} + \frac{\partial^2 L_n(T, \bar{\theta}_{n,i})}{\partial \theta \partial \theta_i} (\theta - \theta_0) \quad (5.2)$$

where $[a, b] = \{\lambda a + (1 - \lambda)b ; \lambda \in [0, 1]\}$. Denote $\bar{F}_n(T, \theta) = -2 \left(\frac{1}{\text{card}(T)} \frac{\partial^2 L_n(T, \bar{\theta}_{n,i})}{\partial \theta \partial \theta_i} \right)_{1 \leq i \leq d}$. Then, (5.2) implies,

$$\text{Card}(T) \bar{F}_n(T, \theta) (\theta - \theta_0) = -2 \left(\frac{\partial L_n(T, \theta)}{\partial \theta} - \frac{\partial L_n(T, \theta_0)}{\partial \theta} \right). \quad (5.3)$$

Similarly, for any $\theta \in \Theta$ we can find a matrix $\tilde{F}_n(T, \theta)$ such that

$$\text{Card}(T)\tilde{F}_n(T, \theta)(\theta - \theta_0) = -2\left(\frac{\partial \hat{L}_n(T, \theta)}{\partial \theta} - \frac{\partial \hat{L}_n(T, \theta_0)}{\partial \theta}\right). \quad (5.4)$$

Takes in (5.4) $\theta = \hat{\theta}_n(T)$ and uses the fact that $\partial \hat{L}_n(T, \hat{\theta}_n(T))/\partial \theta = 0$ (because $\hat{\theta}_n(T)$ is a local extremum of $\hat{L}_n(T, \cdot)$), it comes

$$\text{Card}(T)\tilde{F}_n(T, \hat{\theta}_n(T))(\hat{\theta}_n(T) - \theta_0) = 2\frac{\partial \hat{L}_n(T, \theta_0)}{\partial \theta}. \quad (5.5)$$

Remark 5.1. If $\text{Card}(T) \xrightarrow[n \rightarrow \infty]{} \infty$ and $\theta = \theta(n) \xrightarrow[n \rightarrow \infty]{} \theta_0$, then $\bar{F}_n(T, \theta) \xrightarrow[n \rightarrow \infty]{a.s.} F$ and $\tilde{F}_n(T, \theta) \xrightarrow[n \rightarrow \infty]{a.s.} F$ (see [2] and [3]). In particular, if $\text{Card}(T) \xrightarrow[n \rightarrow \infty]{} \infty$, then $\bar{F}_n(T, \hat{\theta}_n(T)) \xrightarrow[n \rightarrow \infty]{a.s.} F$ and $\tilde{F}_n(T, \hat{\theta}_n(T)) \xrightarrow[n \rightarrow \infty]{a.s.} F$.

Lemma 5.2. Under assumptions of Theorem 3.1

$$\frac{1}{\sqrt{n}} \max_{k \in \Pi_n} \|k(\tilde{F}_n(T_k, \hat{\theta}_n(T_k)) - F)(\hat{\theta}_n(T_k) - \theta_0)\| = o_P(1).$$

Proof For $k \in \Pi_n$, we know that $\sqrt{k}(\hat{\theta}_n(T_k) - \theta_0)$ converges in distribution to the Gaussian law as $n \rightarrow \infty$ (see Theorem 2 of [2]). Therefore, $\max_{k \in \Pi_n} \|\sqrt{k}(\hat{\theta}_n(T_k) - \theta_0)\| = O_P(1)$. Remark 5.1 implies that $\max_{k \in \Pi_n} \|\tilde{F}_n(T_k, \hat{\theta}_n(T_k)) - F\| = o(1)$ a.s. Thus

$$\begin{aligned} \frac{1}{\sqrt{n}} \max_{k \in \Pi_n} \|k(\tilde{F}_n(T_k, \hat{\theta}_n(T_k)) - F)(\hat{\theta}_n(T_k) - \theta_0)\| &\leq \max_{k \in \Pi_n} \|\tilde{F}_n(T_k, \hat{\theta}_n(T_k)) - F\| \times \max_{k \in \Pi_n} \|\sqrt{k}(\hat{\theta}_n(T_k) - \theta_0)\| \\ &= o(1)O_P(1) \text{ a.s.} \\ &= o_P(1). \end{aligned}$$

■

Under assumptions of Theorem 3.1, the matrix G is invertible. Denote $\Sigma = FG^{-1}F$

$$\begin{aligned} Q_n^{(1)} &:= \max_{k \in \Pi_n} Q_{n,k}^{(1)} \quad \text{where} \quad Q_{n,k}^{(1)} := \frac{k^2}{n} (\hat{\theta}_n(T_k) - \hat{\theta}_n(T_n))' \Sigma (\hat{\theta}_n(T_k) - \hat{\theta}_n(T_n)) \quad \text{and} \\ Q_n^{(2)} &:= \max_{k \in \Pi_n} Q_{n,k}^{(2)} \quad \text{where} \quad Q_{n,k}^{(2)} := \frac{(n-k)^2}{n} (\hat{\theta}_n(\bar{T}_k) - \hat{\theta}_n(T_n))' \Sigma (\hat{\theta}_n(\bar{T}_k) - \hat{\theta}_n(T_n)). \end{aligned}$$

Lemma 5.3. Under assumptions of Theorem 3.1

$$\max_{k \in \Pi_n} |\hat{Q}_{n,k}^{(j)} - Q_{n,k}^{(j)}| = o_P(1) \quad \text{for } j = 1, 2.$$

Proof The proof is made for $j = 1$, proceed the same for $j = 2$. For any $k \in \Pi_n$, we have

$$\begin{aligned} |\hat{Q}_{n,k}^{(1)} - Q_{n,k}^{(1)}| &\leq \frac{k^2}{n} \|\hat{\theta}_n(T_k) - \hat{\theta}_n(T_n)\|^2 \|\hat{\Sigma}_{n,k} - \Sigma\| \\ &\leq 2\frac{k^2}{n} (\|\hat{\theta}_n(T_k) - \theta_0\|^2 + \|\hat{\theta}_n(T_n) - \theta_0\|^2) \|\hat{\Sigma}_{n,k} - \Sigma\| \\ &\leq 2(\|\sqrt{k}(\hat{\theta}_n(T_k) - \theta_0)\|^2 + \|\sqrt{n}(\hat{\theta}_n(T_n) - \theta_0)\|^2) \|\hat{\Sigma}_{n,k} - \Sigma\|. \end{aligned} \quad (5.6)$$

Since $k \in \Pi_n$, $k, n - k \rightarrow \infty$ as $n \rightarrow \infty$. Therefore, $\sqrt{k}(\hat{\theta}_n(T_k) - \theta_0) = O_P(1)$ as $n \rightarrow \infty$, $\sqrt{n}(\hat{\theta}_n(T_n) - \theta_0) = O_P(1)$, $\hat{F}_n(T_k) \xrightarrow[n \rightarrow \infty]{a.s.} F$, $\hat{F}_n(\bar{T}_k) \xrightarrow[n \rightarrow \infty]{a.s.} F$, $\hat{G}_n(T_k) \xrightarrow[n \rightarrow \infty]{a.s.} G$ and $\hat{G}_n(\bar{T}_k) \xrightarrow[n \rightarrow \infty]{a.s.} G$ which is invertible.

Thus, for n large enough, $\hat{G}_n(T_k)$ and $\hat{G}_n(\bar{T}_k)$ are invertible. It follows that as $n \rightarrow \infty$,

$$\begin{aligned} \|\hat{\Sigma}_{n,k} - \Sigma\| &= \left\| \frac{k}{n} \hat{F}_n(T_k) \hat{G}_n(T_k)^{-1} \hat{F}_n(T_k) + \frac{n-k}{n} \hat{F}_n(\bar{T}_k) \hat{G}_n(\bar{T}_k)^{-1} \hat{F}_n(\bar{T}_k) - FG^{-1}F \right\| \\ &= \left\| \frac{k}{n} \left(\frac{k}{n} \hat{F}_n(T_k) \hat{G}_n(T_k)^{-1} \hat{F}_n(T_k) - FG^{-1}F \right) + \frac{n-k}{n} \left(\hat{F}_n(\bar{T}_k) \hat{G}_n(\bar{T}_k)^{-1} \hat{F}_n(\bar{T}_k) - FG^{-1}F \right) \right\| \\ &\leq \|\hat{F}_n(T_k) \hat{G}_n(T_k)^{-1} \hat{F}_n(T_k) - FG^{-1}F\| + \|\hat{F}_n(\bar{T}_k) \hat{G}_n(\bar{T}_k)^{-1} \hat{F}_n(\bar{T}_k) - FG^{-1}F\| = o(1) \text{ a.s.} \end{aligned}$$

Therefore, (5.6) implies $\max_{k \in \Pi_n} |\hat{Q}_{n,k}^{(1)} - Q_{n,k}^{(1)}| = o_P(1)$. \blacksquare

Lemma 5.4. *Under assumptions of Theorem 3.1*

$$\frac{-2}{\sqrt{n}} \frac{\partial L_n(T_{[n\tau]}, \theta_0)}{\partial \theta} \xrightarrow{\mathcal{D}} W_G(\tau) \text{ in } D([0, 1], \mathbb{R}^d)$$

where W_G is a d -dimensional Gaussian process with zero mean and covariance matrix $\min(\tau, s)G$.

Proof Recall that $-2 \frac{\partial L_n(T_{[n\tau]}, \theta_0)}{\partial \theta} = \sum_{t=1}^{[n\tau]} \frac{\partial q_t(\theta_0)}{\partial \theta}$. Denote $\mathcal{F}_t = \sigma(X_{t-1}, \dots)$. Since X is stationary and ergodic, the same for the process $(\frac{\partial q_t(\theta_0)}{\partial \theta})_{t \in \mathbb{Z}}$. Moreover, $(\frac{\partial q_t(\theta_0)}{\partial \theta}, \mathcal{F}_t)$ is a square integrable martingale difference process (see [2]) with covariance matrix G . Then, the result follow by using Theorem 23.1 Billingsley (1968) (see [6] page 206). \blacksquare

Lemma 5.5. *Under assumptions of Theorem 3.1*

$$\frac{-2}{\sqrt{n}} G^{-1/2} \left(\frac{\partial L_n(T_{[n\tau]}, \theta_0)}{\partial \theta} - \frac{[n\tau]}{n} \frac{\partial L_n(T_n, \theta_0)}{\partial \theta} \right) \xrightarrow{\mathcal{D}} W_d(\tau) \text{ in } D([0, 1], \mathbb{R}^d)$$

where W_d is a d -dimensional Brownian bridges.

Proof By Lemma 5.4, it comes

$$\frac{-2}{\sqrt{n}} \left(\frac{\partial L_n(T_{[n\tau]}, \theta_0)}{\partial \theta} - \frac{[n\tau]}{n} \frac{\partial L_n(T_n, \theta_0)}{\partial \theta} \right) \xrightarrow{\mathcal{D}} W_G(\tau) - \tau W_G(1) \text{ in } D([0, 1], \mathbb{R}^d).$$

Since the process $\{W_G(\tau) - \tau W_G(1), 0 \leq \tau \leq 1\}$ have covariance matrix $(\min(\tau, s) - \tau s)G$, the covariance matrix of the process $\{G^{-1/2}(W_G(\tau) - \tau W_G(1)), 0 \leq \tau \leq 1\}$ is $(\min(\tau, s) - \tau s)I_d$ (where I_d is d -dimensional identity matrix). Therefore, the process is equal (in distribution) to a d -dimensional Brownian bridge. Thus, the result follows. \blacksquare

Lemma 5.6. *Under assumptions of Theorem 3.1*

$$\frac{-2}{\sqrt{n}} G^{-1/2} \frac{\partial \hat{L}_n(T_{[n\tau]}, \hat{\theta}_n(T_n))}{\partial \theta} \xrightarrow{\mathcal{D}} W_d(\tau) \text{ in } D([0, 1], \mathbb{R}^d).$$

Proof From [2], we have $\frac{1}{\sqrt{n}} \left\| \frac{\partial L_n(T_n, \cdot)}{\partial \theta} - \frac{\partial \widehat{L}_n(T_n, \cdot)}{\partial \theta} \right\|_{\Theta} = o_P(1)$, it implies,

$$\frac{1}{\sqrt{n}} \max_{k \in \Pi_n} \left\| \frac{\partial L_n(T_k, \cdot)}{\partial \theta} - \frac{\partial \widehat{L}_n(T_k, \cdot)}{\partial \theta} \right\|_{\Theta} = o_P(1). \quad (5.7)$$

Let $k \in \Pi_n$. Apply (5.3) with $T = T_k$ and $\theta = \widehat{\theta}_n(T_n)$, we have

$$k \overline{F}_n(T_k, \widehat{\theta}_n(T_n)) (\widehat{\theta}_n(T_n) - \theta_0) = -2 \left(\frac{\partial L_n(T_k, \widehat{\theta}_n(T_n))}{\partial \theta} - \frac{\partial L_n(T_k, \theta_0)}{\partial \theta} \right).$$

By plugging it in (5.7), we have

$$\frac{1}{\sqrt{n}} \max_{k \in \Pi_n} \left\| \frac{\partial \widehat{L}_n(T_k, \widehat{\theta}_n(T_n))}{\partial \theta} - \frac{\partial L_n(T_k, \theta_0)}{\partial \theta} + \frac{1}{2} k \overline{F}_n(T_k, \widehat{\theta}_n(T_n)) (\widehat{\theta}_n(T_n) - \theta_0) \right\| = o_P(1). \quad (5.8)$$

But, by Remark 5.1, it comes that

$$\begin{aligned} \frac{1}{\sqrt{n}} \max_{k \in \Pi_n} \left\| k (\overline{F}_n(T_k, \widehat{\theta}_n(T_n)) - \overline{F}_n(T_n, \widehat{\theta}_n(T_n))) (\widehat{\theta}_n(T_n) - \theta_0) \right\| \\ \leq \frac{1}{\sqrt{n}} \max_{k \in \Pi_n} \left\| k (\overline{F}_n(T_k, \widehat{\theta}_n(T_n)) - \overline{F}_n(T_n, \widehat{\theta}_n(T_n))) \right\| \times \|\sqrt{n}(\widehat{\theta}_n(T_n) - \theta_0)\| \\ = o(1) O_P(1) \text{ a.s.} \\ = o_P(1). \end{aligned}$$

Thus, (5.8) becomes

$$\frac{1}{\sqrt{n}} \max_{k \in \Pi_n} \left\| \frac{\partial \widehat{L}_n(T_k, \widehat{\theta}_n(T_n))}{\partial \theta} - \frac{\partial L_n(T_k, \theta_0)}{\partial \theta} + \frac{1}{2} k \overline{F}_n(T_n, \widehat{\theta}_n(T_n)) (\widehat{\theta}_n(T_n) - \theta_0) \right\| = o_P(1). \quad (5.9)$$

Applying (5.3) with $T = T_n$, $\theta = \widehat{\theta}_n(T_n)$, and using $(1/\sqrt{n})(\partial L_n(T_n, \widehat{\theta}_n(T_n))/\partial \theta) = o_P(1)$ (see [2]), it follows

$$\overline{F}_n(T_n, \widehat{\theta}_n(T_n)) (\widehat{\theta}_n(T_n) - \theta_0) = \frac{2}{n} \frac{\partial L_n(T_n, \theta_0)}{\partial \theta} + o_P\left(\frac{1}{\sqrt{n}}\right). \quad (5.10)$$

Therefore, (5.9) becomes

$$\frac{1}{\sqrt{n}} \max_{k \in \Pi_n} \left\| \frac{\partial \widehat{L}_n(T_k, \widehat{\theta}_n(T_n))}{\partial \theta} - \frac{\partial L_n(T_k, \theta_0)}{\partial \theta} + \frac{k}{n} \frac{\partial L_n(T_n, \theta_0)}{\partial \theta} \right\| = o_P(1). \quad (5.11)$$

Now, let $0 < \tau < 1$, for large value of n , we have $[\tau n] \in \Pi_n$; write

$$\begin{aligned} \frac{-2}{\sqrt{n}} G^{-1/2} \frac{\partial \widehat{L}_n(T_{[\tau n]}, \widehat{\theta}_n(T_n))}{\partial \theta} &= \frac{-2}{\sqrt{n}} G^{-1/2} \left[\frac{\partial \widehat{L}_n(T_{[\tau n]}, \widehat{\theta}_n(T_n))}{\partial \theta} - \left(\frac{\partial L_n(T_{[\tau n]}, \theta_0)}{\partial \theta} - \frac{[\tau n]}{n} \frac{\partial L_n(T_n, \theta_0)}{\partial \theta} \right) \right. \\ &\quad \left. + \left(\frac{\partial L_n(T_{[\tau n]}, \theta_0)}{\partial \theta} - \frac{[\tau n]}{n} \frac{\partial L_n(T_n, \theta_0)}{\partial \theta} \right) \right] \end{aligned}$$

and the result follows by using (5.11) and Lemma 5.5. ■

5.2. Proof Theorem 3.1 and Theorem 3.2

Proof of Theorem 3.1 .

We make the proof for $j = 1$, proceed the same for $j = 2$. By Lemma 5.3, it suffices to show that $Q_n^{(1)} \xrightarrow{\mathcal{D}} \sup_{0 \leq \tau \leq 1} \|W_d(\tau)\|^2$. Uses (5.7), (5.5) with $T = T_k$ and Lemma 5.2 it follows

$$\begin{aligned} \frac{1}{\sqrt{n}} \max_{k \in \Pi_n} \left\| \frac{\partial L_n(T_k, \theta_0)}{\partial \theta} - \frac{1}{2} k F (\hat{\theta}_n(T_k) - \theta_0) \right\| &= \frac{1}{\sqrt{n}} \max_{k \in \Pi_n} \left\| \frac{\partial \hat{L}_n(T_k, \theta_0)}{\partial \theta} - \frac{1}{2} k F (\hat{\theta}_n(T_k) - \theta_0) \right\| + o_P(1) \\ &= \frac{1}{\sqrt{n}} \max_{k \in \Pi_n} \left\| \frac{1}{2} k \tilde{F}_n(T_k, \hat{\theta}_n(T_k)) (\hat{\theta}_n(T_n) - \theta_0) - \frac{1}{2} k F (\hat{\theta}_n(T_k) - \theta_0) \right\| + o_P(1) \\ &= \frac{1}{\sqrt{n}} \max_{k \in \Pi_n} \left\| \frac{1}{2} k (\tilde{F}_n(T_k, \hat{\theta}_n(T_k)) - F) (\hat{\theta}_n(T_n) - \theta_0) \right\| + o_P(1) = o_P(1). \end{aligned} \quad (5.12)$$

Using (5.11) and 5.12, we have

$$\begin{aligned} \frac{1}{\sqrt{n}} \max_{k \in \Pi_n} \left\| \frac{\partial L_n(T_k, \hat{\theta}_n(T_n))}{\partial \theta} - \frac{1}{2} k F (\hat{\theta}_n(T_k) - \hat{\theta}_n(T_n)) \right\| \\ &= \frac{1}{\sqrt{n}} \max_{k \in \Pi_n} \left\| \frac{\partial L_n(T_k, \theta_0)}{\partial \theta} - \frac{k}{n} \frac{\partial L_n(T_n, \theta_0)}{\partial \theta} - \frac{1}{2} k F (\hat{\theta}_n(T_k) - \hat{\theta}_n(T_n)) \right\| + o_P(1) \\ &= \frac{1}{\sqrt{n}} \max_{k \in \Pi_n} \left\| \frac{1}{2} k F (\hat{\theta}_n(T_k) - \theta_0) - \frac{k}{n} \frac{\partial L_n(T_n, \theta_0)}{\partial \theta} - \frac{1}{2} k F (\hat{\theta}_n(T_k) - \hat{\theta}_n(T_n)) \right\| + o_P(1) \\ &= \frac{1}{\sqrt{n}} \max_{k \in \Pi_n} \left\| \frac{1}{2} k F (\hat{\theta}_n(T_n) - \theta_0) - \frac{k}{n} \frac{\partial L_n(T_n, \theta_0)}{\partial \theta} \right\| + o_P(1) \\ &\leq \sqrt{n} \left\| \frac{1}{2} F (\hat{\theta}_n(T_n) - \theta_0) - \frac{1}{n} \frac{\partial L_n(T_n, \theta_0)}{\partial \theta} \right\| + o_P(1). \end{aligned} \quad (5.13)$$

Note that

$$\begin{aligned} \left\| \sqrt{n} (F - \bar{F}_n(T_n, \hat{\theta}_n(T_n))) (\hat{\theta}_n(T_n) - \theta_0) \right\| &\leq \|F - \bar{F}_n(T_n, \hat{\theta}_n(T_n))\| \left\| \sqrt{n} (\hat{\theta}_n(T_n) - \theta_0) \right\| \\ &= o(1) O_P(1) \quad \text{a.s.} \\ &= o_P(1). \end{aligned}$$

By plugging it in (5.13) and applying (5.3) with $T = T_n$ and $\theta = \hat{\theta}_n(T_n)$, we have

$$\begin{aligned} \frac{1}{\sqrt{n}} \max_{k \in \Pi_n} \left\| \frac{\partial L_n(T_k, \hat{\theta}_n(T_n))}{\partial \theta} - \frac{1}{2} k F (\hat{\theta}_n(T_k) - \hat{\theta}_n(T_n)) \right\| &\leq \sqrt{n} \left\| \frac{1}{2} \bar{F}_n(T_n, \hat{\theta}_n(T_n)) (\hat{\theta}_n(T_n) - \theta_0) - \frac{1}{n} \frac{\partial L_n(T_n, \theta_0)}{\partial \theta} \right\| \\ &\quad + o_P(1). \end{aligned} \quad (5.14)$$

Therefore, using (5.10), (5.14) implies

$$\frac{1}{\sqrt{n}} \max_{k \in \Pi_n} \left\| \frac{\partial L_n(T_k, \hat{\theta}_n(T_n))}{\partial \theta} - \frac{1}{2} k F (\hat{\theta}_n(T_k) - \hat{\theta}_n(T_n)) \right\| = o_P(1). \quad (5.15)$$

Now, let $0 < \tau < 1$, for large value of n , we have $[\tau n] \in \Pi_n$; write

$$\begin{aligned} \frac{-2}{\sqrt{n}} G^{-1/2} \frac{\partial \widehat{L}_n(T_{[n\tau]}, \widehat{\theta}_n(T_n))}{\partial \theta} &= -\frac{[n\tau]}{\sqrt{n}} G^{-1/2} F(\widehat{\theta}_n(T_{[n\tau]}) - \widehat{\theta}_n(T_n)) \\ &\quad - 2G^{-1/2} \frac{1}{\sqrt{n}} \left[\frac{\partial \widehat{L}_n(T_{[n\tau]}, \widehat{\theta}_n(T_n))}{\partial \theta} - \frac{1}{2} [n\tau] F(\widehat{\theta}_n(T_{[n\tau]}) - \widehat{\theta}_n(T_n)) \right]. \end{aligned}$$

Therefore, using (5.15) we have

$$-\frac{[n\tau]}{\sqrt{n}} G^{-1/2} F(\widehat{\theta}_n(T_{[n\tau]}) - \widehat{\theta}_n(T_n)) = \frac{-2}{\sqrt{n}} G^{-1/2} \frac{\partial \widehat{L}_n(T_{[n\tau]}, \widehat{\theta}_n(T_n))}{\partial \theta} + o_P(1)$$

and the result follows by using Lemma 5.6. ■

Proof of Theorem 3.2 .

Let $\tau^* \in (0, 1)$ the true value of break. Denote $k^* = [n\tau^*]$. For n large enough , $k^* \in \Pi_n$. Therefore, we have for $j = 1, 2$, $\widehat{Q}_n^{(j)} = \max_{k \in \Pi_n} \widehat{Q}_{n,k}^{(j)} \geq \widehat{Q}_{n,k^*}^{(j)}$. Thus, it follows that

$$\widehat{Q}_n = \max(\widehat{Q}_n^{(1)}, \widehat{Q}_n^{(2)}) \geq \max(\widehat{Q}_{n,k^*}^{(1)}, \widehat{Q}_{n,k^*}^{(2)}). \quad (5.16)$$

Since $\theta_1^*, \theta_2^* \in \overset{\circ}{\Theta}(4)$, it comes from [2] that the model $\mathcal{M}_{\mathbb{Z}}(M_{\theta_1^*}, f_{\theta_1^*})$ and $\mathcal{M}_{\mathbb{Z}}(M_{\theta_2^*}, f_{\theta_2^*})$ have a 4-order stationary solution which we denote $(X_{t,j})_{t \in \mathbb{Z}}$ for $j = 1, 2$.

For $j = 1, 2$ denote for any $t \in \mathbb{Z}$, $q_{t,j}(\theta) := (X_{t,j} - f_{\theta}^{t,j})^2 / (h_{\theta}^{t,j}) + \log(h_{\theta}^{t,j})$ with $f_{\theta}^{t,j} := f_{\theta}(X_{t-1,j}, X_{t-2,j}, \dots)$, $h_{\theta}^{t,j} := (M_{\theta}^{t,j})^2$ where $M_{\theta}^{t,j} := M_{\theta}(X_{t-1,j}, X_{t-2,j}, \dots)$. Also denote for $j = 1, 2$

$$F^{(j)} = \mathbb{E} \left[\frac{\partial^2 q_{0,j}(\theta_j^*)}{\partial \theta \partial \theta'} \right] \quad \text{and} \quad G^{(j)} = \mathbb{E} \left[\left(\frac{\partial q_{0,j}(\theta_j^*)}{\partial \theta} \right) \left(\frac{\partial q_{0,j}(\theta_j^*)}{\partial \theta} \right)' \right].$$

For $j = 1, 2$, Lemma 5.1 implies that the matrix $G^{(j)}$ is symmetric positive definite and Corollary 5.1 of [3] implies $\widehat{G}_n(T_{k^*}) \xrightarrow[n \rightarrow \infty]{a.s.} G^{(1)}$ and $\widehat{G}_n(\overline{T}_{k^*}) \xrightarrow[n \rightarrow \infty]{a.s.} G^{(2)}$. Lemma 4 of [2] implies $\widehat{F}_n(T_{k^*}) \xrightarrow[n \rightarrow \infty]{a.s.} F^{(1)}$ and $\widehat{F}_n(\overline{T}_{k^*}) \xrightarrow[n \rightarrow \infty]{a.s.} F^{(2)}$. Therefore, it follows that

$$\begin{aligned} \widehat{\Sigma}_{n,k^*} &:= \frac{k^*}{n} \widehat{F}_n(T_{k^*}) \widehat{G}_n(T_{k^*})^{-1} \widehat{F}_n(T_{k^*}) \mathbf{1}_{\det(\widehat{G}_n(T_{k^*})) \neq 0} + \frac{n - k^*}{n} \widehat{F}_n(\overline{T}_{k^*}) \widehat{G}_n(\overline{T}_{k^*})^{-1} \widehat{F}_n(\overline{T}_{k^*}) \mathbf{1}_{\det(\widehat{G}_n(\overline{T}_{k^*})) \neq 0} \\ &\xrightarrow[n \rightarrow \infty]{a.s.} \tau^* F^{(1)} (G^{(1)})^{-1} F^{(1)} + (1 - \tau^*) F^{(2)} (G^{(2)})^{-1} F^{(2)}. \quad (5.17) \end{aligned}$$

Denote $\Sigma = \tau^* F^{(1)} (G^{(1)})^{-1} F^{(1)} + (1 - \tau^*) F^{(2)} (G^{(2)})^{-1} F^{(2)}$. It is easy to see that Σ is symmetric positive definite.

For all $\rho > 0$ and $\theta \in \Theta$, denote $B_o(\theta, \rho)$ (rep. $B_c(\theta, \rho)$) the open (resp. closed) ball centered at θ of radius ρ in Θ . i.e.

$$B_o(\theta, \rho) = \{x \in \Theta ; \|\theta - x\| < \rho\} \quad \text{and} \quad B_c(\theta, \rho) = \{x \in \Theta ; \|\theta - x\| \leq \rho\}.$$

For $A \subset \Theta$, we denote $A^c = \{x \in \Theta ; x \notin A\}$.

Since $\theta_1^* \neq \theta_2^*$ and $\theta_1^*, \theta_2^* \in \overset{\circ}{\Theta}(4) \subset \overset{\circ}{\Theta}$, then there exists $\rho_1, \rho_2 > 0$ such that $B_o(\theta_1^*, \rho_1) \cap B_o(\theta_2^*, \rho_2) = \emptyset$.

For all $n \in \mathcal{N}$, denote

$$\delta_n^{(j)} = \inf_{x \in B_c(\theta_j^*, \rho_j/2); y \in B_o^c(\theta_j^*, \rho_j)} ((x - y)' \widehat{\Sigma}_{n, k^*} (x - y)) \quad \text{for } j = 1, 2.$$

Also denote $\delta^{(j)} = \inf_{x \in B_c(\theta_j^*, \rho_j/2); y \in B_o^c(\theta_j^*, \rho_j)} ((x - y)' \Sigma (x - y))$ and it is easy to see that $\delta^{(j)} > 0$ for $j = 1, 2$.

Using (5.17), we have

$$\delta_n^{(j)} \xrightarrow[n \rightarrow \infty]{a.s.} \delta^{(j)} \quad \text{for } j = 1, 2. \quad (5.18)$$

From [2] and [3], we have $\widehat{\theta}_n(T_{k^*}) \xrightarrow[n \rightarrow \infty]{a.s.} \theta_1^*$ and $\widehat{\theta}_n(\overline{T}_{k^*}) \xrightarrow[n \rightarrow \infty]{a.s.} \theta_2^*$. Therefore, for n large enough, $\widehat{\theta}_n(T_{k^*}) \in B_o(\theta_1^*, \rho_1/2)$ and $\widehat{\theta}_n(\overline{T}_{k^*}) \in B_o(\theta_2^*, \rho_2/2)$. Thus, two situations may occur

- if $\widehat{\theta}_n(T_n) \in B_o(\theta_2^*, \rho_2)$ i.e. $\widehat{\theta}_n(T_n) \in B_o^c(\theta_1^*, \rho_1)$ then $(\widehat{\theta}_n(T_{k^*}) - \widehat{\theta}_n(T_n))' \widehat{\Sigma}_{n, k^*} (\widehat{\theta}_n(T_{k^*}) - \widehat{\theta}_n(T_n)) \geq \delta_n^{(1)}$.
Therefore,

$$\widehat{Q}_{n, k^*}^{(1)} := \frac{(k^*)^2}{n} (\widehat{\theta}_n(T_{k^*}) - \widehat{\theta}_n(T_n))' \widehat{\Sigma}_{n, k^*} (\widehat{\theta}_n(T_{k^*}) - \widehat{\theta}_n(T_n)) \geq \frac{(k^*)^2}{n} \delta_n^{(1)} \simeq n(\tau^*)^2 \delta_n^{(1)}.$$

- else $\widehat{\theta}_n(T_n) \in B_o^c(\theta_2^*, \rho_2)$ and $(\widehat{\theta}_n(\overline{T}_{k^*}) - \widehat{\theta}_n(T_n))' \widehat{\Sigma}_{n, k^*} (\widehat{\theta}_n(\overline{T}_{k^*}) - \widehat{\theta}_n(T_n)) \geq \delta_n^{(2)}$. Therefore,

$$\widehat{Q}_{n, k^*}^{(2)} = \frac{(n - k^*)^2}{n} (\widehat{\theta}_n(\overline{T}_{k^*}) - \widehat{\theta}_n(T_n))' \widehat{\Sigma}_{n, k^*} (\widehat{\theta}_n(\overline{T}_{k^*}) - \widehat{\theta}_n(T_n)) \geq \frac{(n - k^*)^2}{n} \delta_n^{(2)} \simeq n(1 - \tau^*)^2 \delta_n^{(2)}.$$

In all cases, we have $\widehat{Q}_n \geq \max(\widehat{Q}_{n, k^*}^{(1)}, \widehat{Q}_{n, k^*}^{(2)}) \geq \min(n(\tau^*)^2 \delta_n^{(1)}, n(1 - \tau^*)^2 \delta_n^{(2)})$.

Thus the result follows by using (5.18). ■

Acknowledgements

The author thank Jean-Marc Bardet and Olivier Wintenberger for many helpful discussions which helped to improve this work.

References

- [1] AUE, A., HÖRMANN, S., HORVÁTH, L. AND REIMHERR, M. Break detection in the covariance structure of multivariate time series models. *Ann. Statist.* 37(6B), (2009), 4046-4087.
- [2] BARDET, J.-M. AND WINTENBERGER, O. Asymptotic normality of the quasi-maximum likelihood estimator for multidimensional causal processes. *Ann. Statist.* 37, (2009), 2730-2759.
- [3] BARDET, J.-M. , KENGNE, W. AND WINTENBERGER, O. Detecting multiple change-points in general causal time series using penalized quasi-likelihood. Preprint available on <http://arxiv.org/pdf/1008.0054>.
- [4] BASSEVILLE, M. AND NIKIFOROV, I. Detection of Abrupt Changes: Theory and Applications. Prentice Hall, Englewood Cliffs, NJ, 1993.

- [5] BERKES, I., HORVÁTH, L., AND KOKOSZKA, P. Testing for parameter constancy in GARCH(p; q) models. *Statistics & Probability Letters* 70, (2004), 263–273.
- [6] BILLINGSLEY. *Convergence of Probability Measures*. John Wiley & Sons Inc., New York, 1968.
- [7] BROWN, R.L., DURBIN, J., EVANS, J.M.,. *Techniques for testing the constancy of regression relationships over time*. *Journal of Royal Statistical Society B*, 37, (1975), 149-192.
- [8] FRANCO, C., AND ZAKOÏAN, J.-M. Maximum likelihood estimation of pure garch and arma-garch processes. *Bernoulli* 10 (2004), 605–637.
- [9] HORVÁTH, L. Change in autoregressive processes. *Stochastic Processes. Appl.* 44, (1993), 221-242.
- [10] HORVÁTH, L., HORVÁTH, Z. AND HUSKOVÁ, M. Ratio tests for change point detection. *Inst. Math. Stat.* 1, (2008), 293-304.
- [11] INCLAN, C., TIAO, G. C. Use of cumulative sums of squares for retrospective detection of changes of variance. *Journal of the American Statistical Association* 89, (1994), 913–923.
- [12] KIERFER, J. K-sample analogues of the Kolmogorov-Smirnov and Cramér-v.Mises tests . *Ann. Math. Statist* 30, (1959), 420–447.
- [13] KOUAMO, O., MOULINES, E. AND ROUEFF, F. Testing for homogeneity of variance in the wavelet domain. *In Dependence in Probability and Statistics*, P. Doukhan, G. Lang, D. Surgailis and G. Teysiere. Lecture Notes in Statistic 200 (2010), Springer-Verlag, pp. 420–447.
- [14] KOKOSZKA, P. AND R. LEIPUS Testing for parameter changes in ARCH models. *Lithuanian Mathematical Journal* 39, (1999), 182-195.
- [15] LEE, S. , HA, J. , AND NA, O. The Cusum Test for Parameter Change in Time Series Models . *Scand. J. Statist.* 30, (2003), 781–796.
- [16] LEE, S. AND NA, O. Test for parameter change in stochastic processes based on conditional least-squares estimator. *J. Multivariate Anal.* 93, (2005), 375-393.
- [17] LEE, S. , TOKUTSU, Y., MAEKAWA, K., The CUSUM test for parameter change in regression models with ARCH errors. *Journal of the Japanese Statistical Society* 34, (2004), 173–188.
- [18] LEE, S., SONG, J. Test for parameter change in ARMA models with GARCH innovations . *Statistics & Probability Letters* 78, (2008), 1990–1998.
- [19] NELSON, D. B. AND CAO, C. Q. Inequality Constraints in the Univariate GARCH Model. *Journal of Business & Economic Statistics* 10, (1992), 229–235.
- [20] PAGE, E. S. A test for a change in a parameter occurring at an unknown point. *Biometrika* 42, (1955), 523–526.