Testing for parameter constancy in general causal time series models

August 11, 2018

by William Charky KENGNE [1]

SAMM, Université Paris 1 Panthéon-Sorbonne, 90 rue de Tolbiac 75634-Paris Cedex 13, France.
E-mail: William-Charky.Kengne@malix.univ-paris1.fr

Abstract : We consider a process $X = (X_t)_{t \in \mathbf{Z}}$ belonging to a large class of causal models including AR(∞), ARCH(∞), TARCH(∞),... models. We assume that the model depends on a parameter $\theta_0 \in I\!\!R^d$ and consider the problem of testing for change in the parameter. Two statistics $\widehat{Q}_n^{(1)}$ and $\widehat{Q}_n^{(2)}$ are constructed using quasi-likelihood estimator (QLME) of the parameter. Under the null hypothesis that there is no change, it is shown that each of these two statistics weakly converges to the supremum of the sum of the squares of independent Brownian bridges. Under the local alternative that there is one change, we show that the test statistic $\widehat{Q}_n = \max\big(\widehat{Q}_n^{(1)}, \widehat{Q}_n^{(2)}\big)$ diverges to infinity. Some simulation results for AR(1), ARCH(1) and GARCH(1,1) models are reported to show the applicability and the performance of our procedure with comparisons to some other approaches.

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 [23] in 1955, real advances have been done about tests for change detection. Horvath [11] proposed a test for detecting a change in the 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.* [9] in 1975, was modified by Inclan and Tiao [13] for testing change in variance of independent random variables. Their test has asymptotically correct size but the asymptotic power is unknown. Numerous works devoted to the CUSUM-type procedure, for instance Kim *et al.* [15] for testing change in parameters of GARCH(1,1), Kokoszka and Leipus [17] in the specific case of ARCH(∞) or Aue *et al.* [1] for testing breaks in covariance.

Kulperger and Yu [18] studied the high moment partial sum process based on residuals and applied it to residual CUSUM test in GARCH model. Horváth *et al.* [12] 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.* [6] used a test based on approximate likelihood scores for testing parameter constancy in GARCH(p,q) models. These procedures are mostly developed in a parametric framework and their asymptotic powers are unknown. The present work is a new contribution to the challenging 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 $M, f :$ $I\!\!R^{I\!\!N} \to I\!\!R$ be measurable functions, $(\xi_t)_{t\in\mathbf{Z}}$ be a sequence of centered independent and identically distributed (iid) random variables called the innovations and satisfying $\mathrm{var}(\xi_0) = \sigma^2$ and Θ a compact subset of $I\!\!R^d$. Let $T \subset \mathbf{Z}$, and for any $\theta \in \Theta$, define

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

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

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

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

H_0 : there exists $\theta_0 \in \Theta$ such that $(X_1, \cdots, X_n)$ belongs to the class $\mathcal{M}_{\{1,\cdots,n\}}(M_{\theta_0}, f_{\theta_0})$;

H_1 : there exist $K \geq 2$, $\theta_1, \cdots, \theta_K \in \Theta$ such that $(X_1, \cdots, X_n)$ belongs to $\bigcap\limits_{j=1}^{K}\mathcal{M}_{T_j^n}(M_{\theta_j}, f_{\theta_j})$ where $T_j^n = \{t_{j-1} + 1, t_{j-1} + 2, \cdots, t_j\}$ with $0 = t_0 < t_1 < \cdots < t_{K-1} < t_K = n$.

Thus, it is easy to see that under H_1 the property of stationary is lost after the first change. This is not the case in many existing works (for instance Kouamo *et al.* [16]) where the stationarity or the K-th order stationarity after the change is an essential assumption.

In this paper we study a new test for change detection (see Bardet *et al.* [3] for the procedure of the estimation of the instants of change). We consider a semi-parametric test statistic based on the QLME which is a modification of the statistic introduced by Lee *et al.* [19]. For $k, k' \in \{2, \cdots, n-1\}$ (with $k \leq k'$) let $\widehat{\theta}_n(X_k, \cdots, X_{k'})$ be the QLME of the parameter computed on $\{k, \cdots, k'\}$. The basic idea of our procedure is that : under H_0, $\widehat{\theta}_n(X_1, \cdots, X_k)$ and $\widehat{\theta}_n(X_{k+1}, \cdots, X_n)$ are close to $\widehat{\theta}_n(X_1, \cdots, X_n)$ and the distances $\|\widehat{\theta}_n(X_1, \cdots, X_k) - \widehat{\theta}_n(X_1, \cdots, X_n)\|$ and $\|\widehat{\theta}_n(X_{k+1}, \cdots, X_n) - \widehat{\theta}_n(X_1, \cdots, X_n)\|$ are not too large. Therefore, we show that the test statistic is finite under the null hypothesis and diverges to infinity under the alternative of one change. Simulation results compared to some other procedures show that our procedure is more powerful. In Section 2 we present assumptions, some examples and construct the test statistic. In Section 3 we give some asymptotic results. The empirical studies of AR(1), ARCH(1) and GARCH(1,1) are detailed in Section 4 and the proofs of the main results are presented in Section 5.

2   Assumptions and test statistics

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

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

1. $\|\cdot\|$ applied to a vector denotes the Euclidean norm of the vector;

2. for any compact set $\Theta \subseteq I\!\!R^d$ and for any $g : \Theta \longrightarrow I\!\!R^{d'}$, $\|g\|_\Theta = \sup_{\theta \in \Theta}(\|g(\theta)\|)$.

Let $\Psi_\theta = f_\theta$, M_θ and $i = 0, 1, 2$, then for any compact set $\Theta \subseteq I\!\!R^d$, 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* $\sum\limits_{k=1}^{\infty} \alpha_k^{(i)}(\Psi_\theta, \Theta) < \infty$ *satisfying*

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

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

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* $\sum\limits_{k=1}^{\infty} \alpha_k^{(i)}(h_\theta, \Theta) < \infty$ *satisfying*

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

Then define the set:

$$\Theta(r) := \{\theta \in \Theta, \ A_0(f_\theta, \{\theta\}) \text{ and } A_0(M_\theta, \{\theta\}) \text{ hold with } \sum_{k \geq 1} \alpha_k^{(0)}(f_\theta, \theta) + (E\,|\xi_0|^r)^{1/r} \sum_{k \geq 1} \alpha_k^{(0)}(M_\theta, \theta) < 1\}$$

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

The Lipschitz-type hypothesis $A_i(\Psi_\theta, \Theta)$ are classical when studying the existence of solutions of the general model. If $\theta \in \Theta(r)$ the existence of a unique causal, stationary and ergodic solution $X = (X_t)_{t \in \mathbf{Z}} \in \mathcal{M}_{\mathbf{Z}}(f_\theta, M_\theta)$ is assured (see [2]). The subset $\Theta(r)$ is defined as a reunion to consider accurately general causal models and ARCH-type models simultaneously.

The following assumptions are needed to study QLME property.

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

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

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

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

As in [2], we will 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 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 and there exists $l > 2$ such that $\alpha_j^{(i)}(f_\theta, \Theta) + \alpha_j^{(i)}(M_\theta, \Theta) + \alpha_j^{(i)}(h_\theta, \Theta) = \mathcal{O}(j^{-l})$, for i= 0, 1.
Throughout the sequel, we will assume that the functions $\theta \mapsto M_\theta$ and $\theta \mapsto f_\theta$ are twice continuously differentiable on Θ.

2.2   Examples

1. AR(∞) models.
   Consider the AR(∞) process defined by :

   $$X_t = \sum_{k \geq 1} \phi_k(\theta_0^*) X_{t-k} + \xi_t \ , \ t \in \mathbf{Z}$$

   with $\theta_0^* \in \Theta$, where Θ is a compact subset of $I\!\!R^d$ such that $\sum_{k \geq 1} \|\phi_k(\theta)\|_\Theta < 1$. The process belongs to the class $\mathcal{M}_\mathbf{Z}(M_{\theta_0^*}, f_{\theta_0^*})$ where $f_\theta(x_1, \cdots) = \sum_{k \geq 1} \phi_k(\theta) x_k$ and $M_\theta \equiv 1$ for all $\theta \in \Theta$. Then Assumptions D(Θ) and $A_0(f_\theta, \Theta)$ hold with $\underline{h} = 1$ and $\alpha_k^{(0)}(f_\theta, \Theta) = \|\phi_k(\theta)\|_\Theta$. If there exists $\ell > 2$ and ϕ_k twice differentiable such as $\|\phi_k(\theta)\|_\Theta = \|\phi_k'(\theta)\|_\Theta = \|\phi_k''(\theta)\|_\Theta = O(k^{-\ell})$, then Assumptions $\mathbf{K}(f_\theta, M_\theta, \Theta)$ holds. Moreover, if ξ_0 is a nondegenerate random variable, Id(Θ) and Var(Θ) hold. For any $r \geq 1$ such that $E\,|\xi_0|^r < \infty$, $\Theta(r) = \Theta$.

2. GARCH(p,q) models.
   Consider the GARCH(p,q) process defined by :

   $$X_t = \sigma_t\,\xi_t \ , \ \sigma_t^2 = \alpha_0^* + \sum_{k=1}^{q} \alpha_k^* X_{t-k}^2 + \sum_{k=1}^{p} \beta_k^* \sigma_{t-k}^2 \ , \quad t \in \mathbf{Z}$$

   with $E\,(\xi_0^2) = 1$ and $\theta_0^* := (\alpha_0^*, \cdots, \alpha_q^*, \beta_1^*, \cdots, \beta_p^*) \in \Theta$ where Θ is a compact subset of $]0, \infty[\times [0, \infty[^{p+q}$ such that $\sum_{k=1}^{q} \alpha_k + \sum_{k=1}^{p} \beta_k < 1$ for all $\theta \in \Theta$. Then there exists (see Bollerslev [8] or Nelson and Cao [22]) a nonnegative sequence $(\psi_k(\theta_0^*))_{k \geq 0}$ such that $\sigma_t^2 = \psi_0(\theta_0^*) + \sum_{k \geq 1} \psi_k(\theta_0^*) X_{t-k}^2$ with $\psi_0(\theta_0^*) = \alpha_0^*/(1 - \sum_{k=1}^{p} \beta_k^*)$. This process belongs to a class $\mathcal{M}_\mathbf{Z}(M_{\theta_0^*}, f_{\theta_0^*})$ where $M_\theta(x_1, \cdots) = \sqrt{\psi_0(\theta) + \sum_{k \geq 1} \psi_k(\theta) x_k}$ and $f_\theta \equiv 0 \ \forall \theta \in \Theta$. Assumptions D($\Theta$) holds with $\underline{h} = \inf_{\theta \in \Theta}(\alpha_0)$. If there exists $0 < \rho_0 < 1$ such that for any $\theta \in \Theta$, $\sum_{k=1}^{q} \alpha_k + \sum_{k=1}^{p} \beta_k \leq \rho_0$ then the sequences $(\|\psi_k(\theta)\|_\Theta)_{k \geq 1}$, $(\|\psi_k'(\theta)\|_\Theta)_{k \geq 1}$ and $(\|\psi_k''(\theta)\|_\Theta)_{k \geq 1}$ decay exponentially fast (see Berkes *et al.* [5]), thus Assumption $\mathbf{K}(f_\theta, M_\theta, \Theta)$ holds. Moreover, if ξ_0^2 is a nondegenerate random variable, Id(Θ) and Var(Θ) hold. For $r \geq 2$ denote

   $$\Theta(r) = \big\{\theta \in \Theta \ ; \ (E\,|\xi_0|^r)^{2/r} \sum_{k=1}^{q} \alpha_k + \sum_{k=1}^{p} \beta_k < 1\big\}.$$

2.3 Test statistics

Assume that a trajectory $(X_1, \cdots, X_n)$ is observed. It is clear that if $(X_1, \cdots, X_n) \in \mathcal{M}_{\{1, \cdots, n\}}(M_\theta, f_\theta)$, then for $T \subset \{1, \cdots, n\}$, 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 $f_\theta^t = f_\theta(X_{t-1}, X_{t-2} \ldots)$, $M_\theta^t = M_\theta(X_{t-1}, X_{t-2} \ldots)$ and $h_\theta^t = M_\theta^{t\,2}$. Therefore, we approximate the conditional log-likelihood with :

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

with $\widehat{f}_\theta^t = f_\theta(X_{t-1}, \ldots, X_1, 0, 0, \cdots)$, $\widehat{M}_\theta^t = M_\theta(X_{t-1}, \ldots, X_1, 0, 0, \cdots)$ and $\widehat{h}_\theta^t = (\widehat{M}_\theta^t)^2$.
For $T \subset \{1, \cdots, n\}$, define the estimator $\widehat{\theta}_n(T) := \underset{\theta \in \Theta}{\mathrm{argmax}}(\widehat{L}_n(T, \theta))$. Moreover, for $1 \leq k \leq n$, denote $T_k = \{1, \cdots, k\}$ and $\overline{T}_k = \{k+1, \cdots, n\}$.

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

For $k = 1, \cdots, n-1$, 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 sequence satisfying $v_n \to \infty$ and $v_n/n \to 0$ (as $n \to \infty$). Denote $\Pi_n = [v_n, n - v_n] \cap \mathbb{N}$ and define the statistics:

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

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

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

Remark 2.1 *Note that, the choice of (v_n) is crucial in practice. We evaluated the procedure with $v_n = [\log n]$, $[(\log n)^2]$, $[(\log n)^3]$ and recommend to use $v_n = [(\log n)^2]$ for linear model and $v_n = [(\log n)^{5/2}]$ for GARCH-type model.*

Lee and Song [21] constructed a test for detecting changes in parameters of ARMA-GARCH models. Their test statistic uses a matrix $\widehat{\Sigma}_n$ which depends on the estimator $\widehat{\theta}_n(T_n)$. Under the null hypothesis (the parameter θ_0 does not change), $\widehat{\Sigma}_n$ is a consistent estimator of $FG^{-1}F$ where $G = E\left[(\partial q_0(\theta_0)/\partial\theta)(\partial q_0(\theta_0)/\partial\theta)'\right]$ and $F = E\left[\partial^2 q_0(\theta_0)/\partial\theta\partial\theta'\right]$. Under the alternative, the model depends on several parameters and $\widehat{\theta}_n(T_n)$ may not be a consistent estimator of one of them. Therefore, the convergence of the matrix $\widehat{\Sigma}_n$ is not assured. Thus, the asymptotic behavior of the test statistic may be very difficult to study. To solve this problem, we

introduce the family of matrices $\{\widehat{\Sigma}_{n,k}, \ k \in \Pi_n\}$. It is easy to see that under the null hypothesis, any sequence $(\widehat{\Sigma}_{n,k_n})_{n>1, k_n \in \Pi_n}$ is consistent. We show (see proof of Theorem 3.2) that under the local alternative where there is one change in the model, there exists a sequence $(\widehat{\Sigma}_{n,k_n^*})_{n>1, k_n^* \in \Pi_n}$ which converges.

3   Asymptotic results

3.1   Asymptotic under the null hypothesis

Theorem 3.1 *Assume* $\boldsymbol{D}(\Theta)$, $\boldsymbol{Id}(\Theta)$, $\boldsymbol{Var}$ *and* $\boldsymbol{K}(f_\theta, M_\theta, \Theta)$. *Under the null hypothesis* H_0, *if* $\theta_0 \in \overset{\circ}{\Theta}(4)$, *then for* $j = 1, 2$,

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

where W_d *is a d-dimensional Brownian bridge.*

For any $\alpha \in (0,1)$, let C_α denote 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 :*

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

Remark 3.1 *Quantile values of the distribution of* $\sup_{0 \leq \tau \leq 1} \|W_d(\tau)\|^2$ *are known (see for instance Kieffer [14] for* $d \in \{1, \cdots, 5\}$ *or Lee et al. [19] for* $d \in \{1, \cdots, 10\}$*).*

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 $(\widehat{Q}_n > C_\alpha)$.

Figure 1 is an illustration of the test procedure for AR(1) process. At a level $\alpha = 0.05$, for $d = 1$, $C_\alpha \simeq 2.20$. Figure 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 increase around the point where the change occurs.

As it can be observed on the Figure 1 and Figure 2 , the statistics $\widehat{Q}_n^{(1)}$ and $\widehat{Q}_n^{(1)}$ are not clearly equal. Figure 2 shows the typical example for $ARCH(1)$ with one change where $\widehat{Q}_n^{(1)} < C_\alpha$ and $\widehat{Q}_n^{(2)} > C_\alpha$. In general, we don't know if under the alternative hypothesis each of statistics $\widehat{Q}_n^{(1)}$ and $\widehat{Q}_n^{(2)}$ take large values. But, under the local alternative of one change, the maximum of these two statistics diverges to infinity (see Theorem 3.2). This is the reason why we define the critical region as $\{\max(\widehat{Q}_n^{(1)}, \widehat{Q}_n^{(2)}) > C_\alpha\}$.

3.2   The asymptotic under a local alternative

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

$H_1^{(loc)}$: there exist $\tau^* \in (0,1)$ and $\theta_1^*, \theta_2^* \in \Theta$ with $\theta_1^* \neq \theta_2^*$ such that $X_1, \cdots X_{[n\tau^*]} \in \mathcal{M}_{T_{[n\tau^*]}}(M_{\theta_1^*}, f_{\theta_1^*})$

and $X_{[n\tau^*]+1}, \cdots, X_n \in \mathcal{M}_{\overline{T}_{[n\tau^*]}}(M_{\theta_2^*}, f_{\theta_2^*})$.

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

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

Remark 3.2 *1-) Theorem 3.2 shows that the test with the local alternative $H_1^{(loc)}$ is consistent in power. 2-) This procedure can also be used to test multiple changes using ICSS type algorithm developed by Inclán and Tiao [13].*
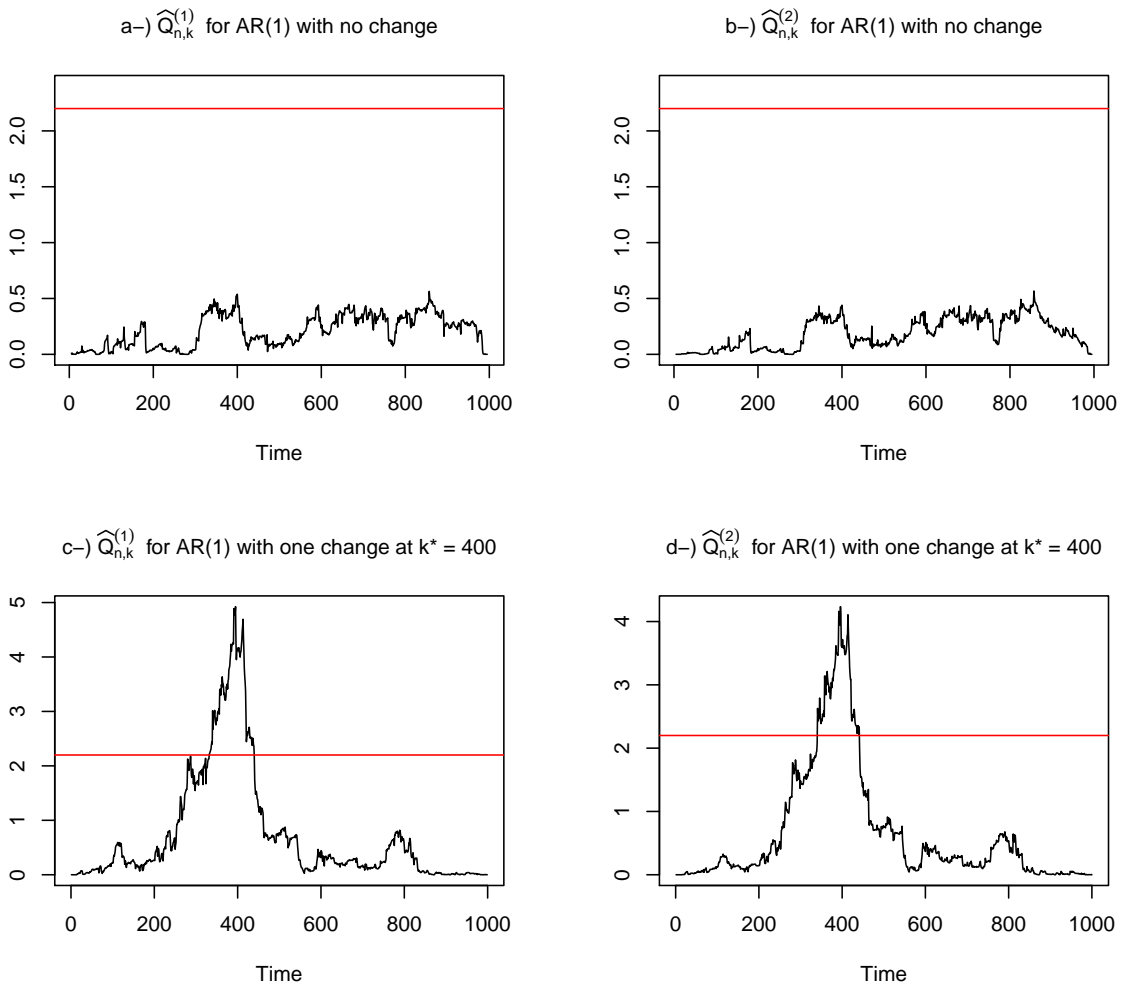
a–) $\widehat{Q}_{n,k}^{(1)}$ for AR(1) with no change

b–) $\widehat{Q}_{n,k}^{(2)}$ for AR(1) with no change

c–) $\widehat{Q}_{n,k}^{(1)}$ for AR(1) with one change at k* = 400

d–) $\widehat{Q}_{n,k}^{(2)}$ for AR(1) with one change at k* = 400

Figure 1 : Typical realization of the statistics $\widehat{Q}_{n,k}^{(1)}$ and $\widehat{Q}_{n,k}^{(2)}$ for 1000 sample of AR(1) with $v_n = [\log n]$. a-) and b-) are the case of AR(1) where parameter $\phi_1 = 0.3$ remains constant. c-) and d-) are the case of AR(1) with parameter $\phi_1 = 0.3$ changing to 0.5 at $k^* = 400$.
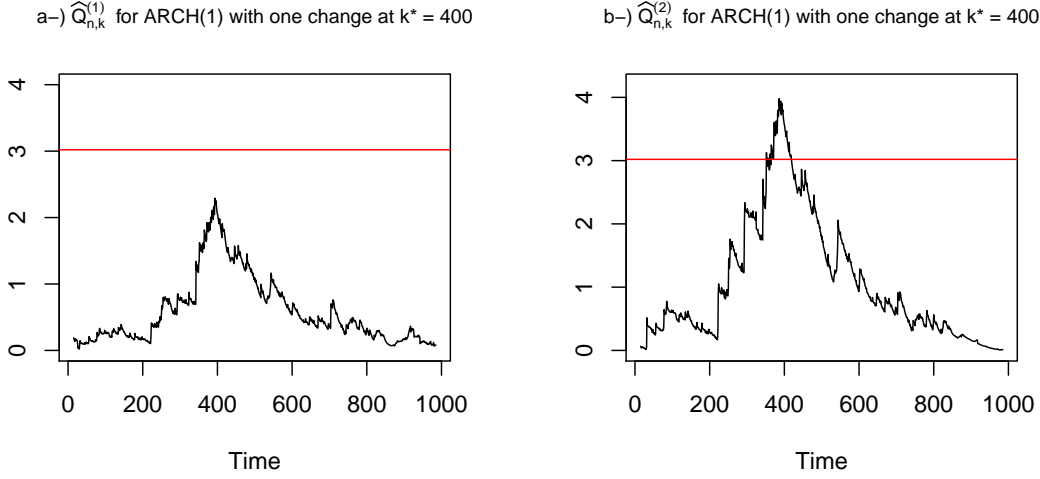
a-) $\widehat{Q}_{n,k}^{(1)}$ for ARCH(1) with one change at k* = 400   b-) $\widehat{Q}_{n,k}^{(2)}$ for ARCH(1) with one change at k* = 400

Figure 2 : Typical realization of the statistics $\widehat{Q}_{n,k}^{(1)}$ and $\widehat{Q}_{n,k}^{(2)}$ for 1000 sample of ARCH(1) with $v_n = [(\log n)^2]$. a-) and b-) are the case of ARCH(1) with parameter $\theta_1 = (1, 0.3)$ changing to $(1, 0.15)$ at $k^* = 400$.

4   Some simulations results

In this section, we evaluate the performance of the procedure through empirical study. We compare our results with those obtained by Kouamo *et al.* [16], Lee and Na [20] and the results obtained from the residual CUSUM test by using the statistics defined by Kulperger and Yu [18]. For a sample size n, $\widehat{Q}_n$ is computed with $v_n = [(\ln n)^2]$ for AR model and $v_n = [(\ln n)^{5/2}]$ for GARCH model and is compared to the critical value of the test. In the following models, $(\xi_t)_{t \in \mathbf{Z}}$ are iid standard Gaussian random variables.

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 $p \in I\!N^*$. The true parameter of the model is denoted by $\theta_0^* = (\phi_1^*, \cdots, \phi_p^*) \in \Theta$ where $\Theta = \{\theta = (\phi_1, \cdots, \phi_p) \in I\!R^p \ / \sum_{i=1}^{p} |\phi_i| < 1\}$. Since $M_\theta \equiv 1$, $\Theta(r) = \Theta$ for any $r \geq 1$. Assume $(X_1, \cdots, X_n)$ is observed, we have for any $\theta \in \Theta$, $\widehat{q}_t(\theta) = \left(X_t - \sum_{k=1}^{p} \phi_k X_{t-k}\right)^2$, $\frac{\partial \widehat{q}_t(\theta)}{\partial \theta} = -2\left(X_t - \sum_{k=1}^{p} \phi_k X_{t-k}\right) \cdot (X_{t-1}, X_{t-2}, \cdots, X_{t-p})$ 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 $C_\alpha \simeq 2.20$. For $n = 1024, 2048, 4096$; we generate a sample $(X_1, \cdots, X_n)$ in the following situations : (i) there is no change and the parameter of the model is $\theta_0 = 0.9$ and (ii) there is one change and the parameter $\theta_0 = 0.9$ changes to

$\theta_1 = 0.5$ at $n/2$. The following table indicates the proportion of the number of rejections of the 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 the test for parameter change in AR(1) model. The empirical levels are computed when $\theta_0 = 0.9$; the empirical powers are computed when θ_0 changes to $\theta_1 = 0.5$ at $n/2$. Figures in brackets are the results obtained by Kouamo *et al.* [16] 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 statistic. The results obtained with our test statistic $\widehat{Q}_n$ are clearly more accurate.

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

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

$$\forall t \in \mathbf{Z}, \quad X_t = \sigma_t \xi_t \quad \text{with} \quad \sigma_t^2 = \alpha_0^* + \alpha_1^* X_{t-1}^2 + \beta_1^* \sigma_{t-1}^2$$

with $\theta_0^* = (\alpha_0^*, \alpha_1^*, \beta_1^*) \in \Theta \subset]0, \infty[\times [0, \infty[^2$ and satisfying $\alpha_1^* + \beta_1^* < 1$. The ARCH($\infty$) representation is $\sigma_t^2 = \alpha_0^*/(1 - \beta_1^*) + \alpha_1^* \sum_{k \geq 1} (\beta_1^*)^{k-1} X_{t-k}^2$. For any $\theta \in \Theta$ and $t = 2, \cdots, n$, we have

$$\widehat{h}_\theta^t = \alpha_0/(1 - \beta_1) + \alpha_1 X_{t-1}^2 + \alpha_1 \sum_{k=2}^t \beta_1^{k-1} X_{t-k}^2 \quad \text{and} \quad \widehat{q}_t(\theta) = X_t^2 / \widehat{h}_\theta^t + \log(\widehat{h}_\theta^t).$$

Therefore, it follows that $\dfrac{\partial \widehat{q}_t(\theta)}{\partial \theta} = \dfrac{1}{\widehat{h}_\theta^t}\Big(1 - \dfrac{X_t^2}{\widehat{h}_\theta^t}\Big)\Big(\dfrac{\partial \widehat{h}_\theta^t}{\partial \alpha_0}, \dfrac{\partial \widehat{h}_\theta^t}{\partial \alpha_1}, \dfrac{\partial \widehat{h}_\theta^t}{\partial \beta_1}\Big)$ with $\partial \widehat{h}_\theta^t / \partial \alpha_1 = X_{t-1}^2 + \sum_{k=2}^t \beta_1^{k-1} X_{t-k}^2$

$\partial \widehat{h}_\theta^t / \partial \alpha_0 = 1/(1 - \beta_1)$, and $\partial \widehat{h}_\theta^t / \partial \beta_1 = \alpha_0/(1 - \beta_1)^2 + \alpha_1 X_{t-2}^2 + \alpha_1 \sum_{k=3}^t (k - 1)\beta_1^{k-2} X_{t-k}^2$.

Let $\theta = (\alpha_0, \alpha_1, \beta_1) = (\theta_1, \theta_2, \theta_3) \in \Theta$, for $1 \leq i, j \leq 3$, we have

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

with $\quad \partial^2 \widehat{h}_\theta^t / \partial \alpha_0^2 = 0, \ \partial^2 \widehat{h}_\theta^t / \partial \alpha_0 \partial \alpha_1 = 0, \ \partial^2 \widehat{h}_\theta^t / \partial \alpha_1^2 = 0, \ \partial^2 \widehat{h}_\theta^t / \partial \alpha_1 \partial \beta_1 = X_{t-2}^2 + \sum_{k=3}^t (k - 1)\beta_1^{k-2} X_{t-k}^2,$

$\partial^2 \widehat{h}_\theta^t / \partial \alpha_0 \partial \beta_1 = 1/(1 - \beta_1)^2$ and $\quad \partial \widehat{h}_\theta^t / \partial \beta_1^2 = 2\alpha_0/(1 - \beta_1)^3 + 2\alpha_1 X_{t-3}^2 + \alpha_1 \sum_{k=4}^t (k - 1)(k - 2)\beta_1^{k-3} X_{t-k}^2$.

1. Case of ARCH(1). Assume $\beta_1 = 0$ and $\theta = (\alpha_0, \alpha_1)$. At level $\alpha = 0.05$, the critical value is $C_\alpha \simeq 3.02$. For $n = 500, 800, 1000$; we generate a sample $(X_1, \cdots, X_n)$ in the following situations : (i) there is no change, the parameter of the model is $\theta_0 = (1, 0.3)$ and (ii) there is one change, the parameter $\theta_0 = (1, 0.3)$ changes to two different values of $\theta_1 = (0.5, 0.3)$ and $\theta_1 = (0.5, 0.6)$ at $n/2$. The following table indicates the proportion of the number of rejections of the 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 $\theta_1 = (0.5, 0.6)$	0.876	(0.498)	0.976	(0.589)	0.984	(0.606)

Table 2: Empirical levels and powers at nominal level 0.05 of the test for parameter change in ARCH(1) model. The empirical levels are computed when $\theta_0 = (1, 0.3)$; the empirical powers are computed when θ_0 changes to θ_1 at $n/2$. Figures in brackets are the results obtained by Lee and Na [20].

2. Case of GARCH(1,1). Now $\theta = (\alpha_0, \alpha_1, \beta_1)$. At level $\alpha = 0.05$, the critical value is $C_\alpha \simeq 3.47$. For $n = 500, 1000, 1500$; we generate a sample $(X_1, \cdots, X_n)$ in the following situations : (i) there is no change, the parameter of the model is $\theta_0 = (1, 0.4, 0.1)$ and (ii) there is one change, the parameter $\theta_0 = (1, 0.4, 0.1)$ changes two different values of $\theta_1 = (0.7, 0.4, 0.1)$ and $\theta_1 = (1, 0.4, 0.3)$ at $n/2$. The following table indicates the proportion of the number of rejections of the null hypothesis out of 500 repetitions.

	$n = 500$		$n = 1000$		$n = 1500$	
Empirical levels	0.100	(0.030)	0.078	(0.032)	0.052	(0.042)
Empirical powers when $\theta_1 = (0.7, 0.4, 0.1)$	0.498	(0.334)	0.752	(0.658)	0.934	(0.848)
Empirical powers when $\theta_1 = (1, 0.4, 0.3)$	0.654	(0.404)	0.968	(0.772)	0.976	(0.922)

Table 3: Empirical levels and powers at nominal level 0.05 of test the for parameter change in GARCH(1,1) model. The empirical levels are computed when $\theta_0 = (1, 0.4, 0.1)$; the empirical powers are computed when θ_0 changes to θ_1 at $n/2$. Figures in brackets are the results of the residual CUSUM test using $CUSUM^{(2)}$ statistic defined by Kulperger and Yu [18].

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 empirical level is less than 0.05 when $n = 800$. It is not very surprising because the asymptotic size of the test is less than $\alpha = 0.05$. This is not the case for GARCH model. It 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. In Table 2, figures in brackets are the results obtained by Lee and Na [20] using the CUSUM test based on conditional least-squares estimator. In Table 3, figures in brackets are the results of the residual CUSUM test that we obtained by using $CUSUM^{(2)}$ statistic studied by Kulperger and Yu [18]. Once again, our test statistic $\widehat{Q}_n$ provides best results.

5   Proofs of the main results

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 $\varepsilon > 0$, $P(|\psi_n| \geq \varepsilon |r_n|) \to 0$ as $n \to \infty$. Write $\psi_n = O_P(r_n)$ to mean : for all $\varepsilon > 0$, there exists $C > 0$ such that $P(|\psi_n| \geq C |r_n|) \leq \varepsilon$ for n large enough.

5.1   Some preliminary results

First, let us prove useful technical lemmas.

Under the null hypothesis H_0 the observations $(X_1, \cdots, X_n)$ belong in the class $\mathcal{M}_{\{1,\cdots,n\}}(M_{\theta_0}, f_{\theta_0})$, define the matrix $G := E\left[\dfrac{\partial q_0(\theta_0)}{\partial \theta}\dfrac{\partial q_0(\theta_0)}{\partial \theta}'\right]$ (where $'$ denotes the transpose) and $F := E\left[\dfrac{\partial^2 q_0(\theta_0)}{\partial \theta \partial \theta'}\right]$. 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 the null hypothesis* $\boldsymbol{D}(\Theta)$ *and* $\boldsymbol{Var}$, *G is a symmetric, positive definite matrix.*

Proof.   It is clear that G is symmetric. Moreover, for $1 \le i \le d$, we have :
$\dfrac{\partial q_0(\theta_0)}{\partial \theta_i} = -2\dfrac{\xi_0}{\sqrt{h_{\theta_0}^0}}\dfrac{\partial f_{\theta_0}^0}{\partial \theta_i} - \dfrac{\xi_0^2}{h_{\theta_0}^0}\dfrac{\partial h_{\theta_0}^0}{\partial \theta_i} + \dfrac{1}{h_{\theta_0}^0}\dfrac{\partial h_{\theta_0}^0}{\partial \theta_i}$. Thus, using independence of ξ_0 and $X_{-1},\ X_{-2}, \cdots$ we obtain
:
$$E\left[\frac{\partial q_0(\theta_0)}{\partial \theta}'\frac{\partial q_0(\theta_0)}{\partial \theta}\right] = 4E\left[\frac{1}{h_{\theta_0}^0}\frac{\partial f_{\theta_0}^0}{\partial \theta}'\frac{\partial f_{\theta_0}^0}{\partial \theta}\right] + E\left((\xi_0^2 - 1)^2\right)E\left[\frac{1}{(h_{\theta_0}^0)^2}\frac{\partial h_{\theta_0}^0}{\partial \theta}'\frac{\partial h_{\theta_0}^0}{\partial \theta}\right]. \tag{2}$$

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

Under **Var**, one of the two matrix of the right-hand side of relation (2) is positive definite and the other is semi-positive definite. Thus, G is positive definite. ∎

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

$$\frac{\partial L_n(T, \theta)}{\partial \theta_i} = \frac{\partial L_n(T, \theta_0)}{\partial \theta_i} + \frac{\partial^2 L_n(T, \overline{\theta}_{n,i})}{\partial \theta \partial \theta_i}(\theta - \theta_0) \tag{3}$$

where $[a, b] = \{\lambda a + (1 - \lambda)b\ ;\ \lambda \in [0, 1]\}$. Denote $\overline{F}_n(T, \theta) = -2\left(\dfrac{1}{card(T)}\dfrac{\partial^2 L_n(T, \overline{\theta}_{n,i})}{\partial \theta \partial \theta_i}\right)_{1 \le i \le d}$. Then, (3) implies,
$$Card(T)\overline{F}_n(T, \theta)(\theta - \theta_0) = -2\left(\frac{\partial L_n(T, \theta)}{\partial \theta} - \frac{\partial L_n(T, \theta_0)}{\partial \theta}\right). \tag{4}$$

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

$$Card(T)\widetilde{F}_n(T, \theta)(\theta - \theta_0) = -2\left(\frac{\partial \widehat{L}_n(T, \theta)}{\partial \theta} - \frac{\partial \widehat{L}_n(T, \theta_0)}{\partial \theta}\right). \tag{5}$$

With $\theta = \widehat{\theta}_n(T)$ in (5) and using the fact that $\partial \widehat{L}_n(T, \widehat{\theta}_n(T))/\partial \theta = 0$ (because $\widehat{\theta}_n(T)$ is a local extremum of $\widehat{L}_n(T, \cdot)$), it comes
$$Card(T)\widetilde{F}_n(T, \widehat{\theta}_n(T))(\widehat{\theta}_n(T) - \theta_0) = 2\frac{\partial \widehat{L}_n(T, \theta_0)}{\partial \theta}. \tag{6}$$

Remark 5.1 *If* $Card(T) \underset{n\to\infty}{\longrightarrow} \infty$ *and* $\theta = \theta(n) \underset{n\to\infty}{\longrightarrow} \theta_0$, *then* $\overline{F}_n(T, \theta) \overset{a.s.}{\underset{n\to\infty}{\longrightarrow}} F$ *and* $\widetilde{F}_n(T, \theta) \overset{a.s.}{\underset{n\to\infty}{\longrightarrow}} F$ *(see [2] and [3]). In particular, if* $Card(T) \underset{n\to\infty}{\longrightarrow} \infty$, *then* $\overline{F}_n(T, \widehat{\theta}_n(T)) \overset{a.s.}{\underset{n\to\infty}{\longrightarrow}} F$ *and* $\widetilde{F}_n(T, \widehat{\theta}_n(T)) \overset{a.s.}{\underset{n\to\infty}{\longrightarrow}} F$.

Lemma 5.2 *Under assumptions of Theorem 3.1*

$$\frac{1}{\sqrt{n}} \max_{k \in \Pi_n} \left\| k\big(\widetilde{F}_n(T_k, \widehat{\theta}_n(T_k)) - F\big)(\widehat{\theta}_n(T_k) - \theta_0)\right\| = o_P(1).$$

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

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

∎

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

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

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

Lemma 5.3 *Under assumptions of Theorem 3.1*

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

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

$$
\begin{aligned}
\left|\widehat{Q}_{n,k}^{(1)} - Q_{n,k}^{(1)}\right| &\leq \frac{k^2}{n}\|\widehat{\theta}_n(T_k) - \widehat{\theta}_n(T_n)\|^2\|\widehat{\Sigma}_{n,k} - \Sigma\| \\
&\leq 2\frac{k^2}{n}\big(\|\widehat{\theta}_n(T_k) - \theta_0\|^2 + \|\widehat{\theta}_n(T_n) - \theta_0\|^2\big)\|\widehat{\Sigma}_{n,k} - \Sigma\| \\
&\leq 2\big(\|\sqrt{k}(\widehat{\theta}_n(T_k) - \theta_0)\|^2 + \|\sqrt{n}(\widehat{\theta}_n(T_n) - \theta_0)\|^2\big)\|\widehat{\Sigma}_{n,k} - \Sigma\|. \tag{7}
\end{aligned}
$$

Since $k \in \Pi_n$, $k, n-k \longrightarrow \infty$ as $n \longrightarrow \infty$. Therefore, $\sqrt{k}(\widehat{\theta}_n(T_k) - \theta_0) = O_P(1)$ as $n \longrightarrow \infty$, $\sqrt{n}(\widehat{\theta}_n(T_n) - \theta_0) = O_P(1)$, $\widehat{F}_n(T_k) \xrightarrow[n \to \infty]{a.s.} F$, $\widehat{F}_n(\overline{T}_k) \xrightarrow[n \to \infty]{a.s.} F$, $\widehat{G}_n(T_k) \xrightarrow[n \to \infty]{a.s.} G$ and $\widehat{G}_n(\overline{T}_k) \xrightarrow[n \to \infty]{a.s.} G$ which is invertible. Thus, for n large enough, $\widehat{G}_n(T_k)$ and $\widehat{G}_n(\overline{T}_k)$ are invertible. It follows that as $n \longrightarrow \infty$,

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

Therefore, (7) implies $\max_{k \in \Pi_n} \left|\widehat{Q}_{n,k}^{(1)} - Q_{n,k}^{(1)}\right| = o_P(1)$.

∎

Lemma 5.4 *Under assumptions of Theorem 3.1*

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

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

Proof. Recall that $-2\frac{\partial L_n(T_{[n\tau]}, \theta_0)}{\partial \theta} = \sum\limits_{t=1}^{[n\tau]} \frac{\partial q_t(\theta_0)}{\partial \theta}$. Denote $\mathcal{F}_t = \sigma(X_{t-1}, \cdots)$. Since X is stationary and ergodic, it is the same for the process $(\frac{\partial q_t(\theta_0)}{\partial \theta})_{t\in\mathbf{Z}}$. 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 [7] page 206). ∎

Lemma 5.5 *Under assumptions of Theorem 3.1*

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

where W_d is a d-dimensional Brownian bridge.

Proof. By Lemma 5.4, it comes

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

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

Lemma 5.6 *Under assumptions of Theorem 3.1*

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

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

$$\frac{1}{\sqrt{n}}\max_{k\in\Pi_n}\Big\|\frac{\partial L_n(T_k, \cdot)}{\partial \theta} - \frac{\partial \widehat{L}_n(T_k, \cdot)}{\partial \theta}\Big\|_\Theta = o_P(1). \tag{8}$$

Let $k \in \Pi_n$. Applying (4) with $T = T_k$ and $\theta = \widehat{\theta}_n(T_n)$, we have

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

By plugging it in (8), we have

$$\frac{1}{\sqrt{n}}\max_{k\in\Pi_n}\Big\|\frac{\partial \widehat{L}_n(T_k, \widehat{\theta}_n(T_n))}{\partial \theta} - \frac{\partial L_n(T_k, \theta_0)}{\partial \theta} + \frac{1}{2}k\overline{F}_n(T_k, \widehat{\theta}_n(T_n))(\widehat{\theta}_n(T_n) - \theta_0)\Big\| = o_P(1). \tag{9}$$

But, by Remark 5.1, it comes that

$$\frac{1}{\sqrt{n}} \max_{k \in \Pi_n} \left\| k\big(\overline{F}_n(T_k, \widehat{\theta}_n(T_n)) - \overline{F}_n(T_n, \widehat{\theta}_n(T_n))\big)(\widehat{\theta}_n(T_n) - \theta_0) \right\|$$

$$\leq \frac{1}{\sqrt{n}} \max_{k \in \Pi_n} \left\| k\big(\overline{F}_n(T_k, \widehat{\theta}_n(T_n)) - \overline{F}_n(T_n, \widehat{\theta}_n(T_n))\big) \right\| \times \left\| \sqrt{n}(\widehat{\theta}_n(T_n) - \theta_0) \right\|$$

$$= o(1)O_P(1) \quad \text{a.s.}$$

$$= o_P(1).$$

Thus, (9) becomes

$$\frac{1}{\sqrt{n}} \max_{k \in \Pi_n} \left\| \frac{\partial \widehat{L}_n(T_k, \widehat{\theta}_n(T_n))}{\partial \theta} - \frac{\partial L_n(T_k, \theta_0)}{\partial \theta} + \frac{1}{2} k \overline{F}_n(T_n, \widehat{\theta}_n(T_n))(\widehat{\theta}_n(T_n) - \theta_0) \right\| = o_P(1). \tag{10}$$

Applying (4) 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

$$\overline{F}_n(T_n, \widehat{\theta}_n(T_n))(\widehat{\theta}_n(T_n) - \theta_0) = \frac{2}{n} \frac{\partial L_n(T_n, \theta_0)}{\partial \theta} + o_P(\frac{1}{\sqrt{n}}). \tag{11}$$

Therefore, (10) becomes

$$\frac{1}{\sqrt{n}} \max_{k \in \Pi_n} \left\| \frac{\partial \widehat{L}_n(T_k, \widehat{\theta}_n(T_n))}{\partial \theta} - \frac{\partial L_n(T_k, \theta_0)}{\partial \theta} + \frac{k}{n} \frac{\partial L_n(T_n, \theta_0)}{\partial \theta} \right\| = o_P(1). \tag{12}$$

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

$$\frac{-2}{\sqrt{n}} G^{-1/2} \frac{\partial \widehat{L}_n(T_{[n\tau]}, \widehat{\theta}_n(T_n))}{\partial \theta} = \frac{-2}{\sqrt{n}} G^{-1/2} \Big[\frac{\partial \widehat{L}_n(T_{[n\tau]}, \widehat{\theta}_n(T_n))}{\partial \theta} - \Big(\frac{\partial L_n(T_{[n\tau]}, \theta_0)}{\partial \theta} - \frac{[n\tau]}{n} \frac{\partial L_n(T_n, \theta_0)}{\partial \theta} \Big)$$

$$+ \Big(\frac{\partial L_n(T_{[n\tau]}, \theta_0)}{\partial \theta} - \frac{[n\tau]}{n} \frac{\partial L_n(T_n, \theta_0)}{\partial \theta} \Big) \Big]$$

and the result follows by using (12) and Lemma 5.5.   ∎

5.2   Proof of Theorem 3.1 and Theorem 3.2

Proof of Theorem 3.1 .
We give the proof for $j = 1$, proceed similarly for $j = 2$. By Lemma 5.3, Theorem 3.1 is established if $Q_n^{(1)} \xrightarrow[n \to \infty]{\mathcal{D}} \sup_{0 \leq \tau \leq 1} \|W_d(\tau)\|^2$. Using (8), (6) with $T = T_k$ and Lemma 5.2 it follows

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

$$= \frac{1}{\sqrt{n}} \max_{k \in \Pi_n} \left\| \frac{1}{2} k \widetilde{F}_n(T_k, \widehat{\theta}_n(T_k))(\widehat{\theta}_n(T_n) - \theta_0) - \frac{1}{2} k F \ (\widehat{\theta}_n(T_k) - \theta_0) \right\| + o_P(1)$$

$$= \frac{1}{\sqrt{n}} \max_{k \in \Pi_n} \left\| \frac{1}{2} k \big(\widetilde{F}_n(T_k, \widehat{\theta}_n(T_k)) - F \big)(\widehat{\theta}_n(T_n) - \theta_0) \right\| + o_P(1) = o_P(1). \tag{13}$$

Using (12) and 13, we have

$$\frac{1}{\sqrt{n}} \max_{k \in \Pi_n} \Big\| \frac{\partial L_n(T_k, \widehat{\theta}_n(T_n))}{\partial \theta} - \frac{1}{2}kF\left(\widehat{\theta}_n(T_k) - \widehat{\theta}_n(T_n)\right) \Big\|$$

$$= \frac{1}{\sqrt{n}} \max_{k \in \Pi_n} \Big\| \frac{\partial L_n(T_k, \theta_0)}{\partial \theta} - \frac{k}{n}\frac{\partial L_n(T_n, \theta_0)}{\partial \theta} - \frac{1}{2}kF\left(\widehat{\theta}_n(T_k) - \widehat{\theta}_n(T_n)\right) \Big\| + o_P(1)$$

$$= \frac{1}{\sqrt{n}} \max_{k \in \Pi_n} \Big\| \frac{1}{2}kF\left(\widehat{\theta}_n(T_k) - \theta_0\right) - \frac{k}{n}\frac{\partial L_n(T_n, \theta_0)}{\partial \theta} - \frac{1}{2}kF\left(\widehat{\theta}_n(T_k) - \widehat{\theta}_n(T_n)\right) \Big\| + o_P(1)$$

$$= \frac{1}{\sqrt{n}} \max_{k \in \Pi_n} \Big\| \frac{1}{2}kF\left(\widehat{\theta}_n(T_n) - \theta_0\right) - \frac{k}{n}\frac{\partial L_n(T_n, \theta_0)}{\partial \theta} \Big\| + o_P(1)$$

$$\leq \sqrt{n}\Big\| \frac{1}{2}F\left(\widehat{\theta}_n(T_n) - \theta_0\right) - \frac{1}{n}\frac{\partial L_n(T_n, \theta_0)}{\partial \theta} \Big\| + o_P(1). \tag{14}$$

Note that

$$\big\| \sqrt{n}(F - \overline{F}_n(T_n, \widehat{\theta}_n(T_n)))\left(\widehat{\theta}_n(T_n) - \theta_0\right) \big\| \leq \big\| F - \overline{F}_n(T_n, \widehat{\theta}_n(T_n)) \big\| \, \big\| \sqrt{n}(\widehat{\theta}_n(T_n) - \theta_0) \big\|$$

$$= o(1)O_P(1) \quad \text{a.s.}$$

$$= o_P(1).$$

By plugging it in (14) and applying (4) with $T = T_n$ and $\theta = \widehat{\theta}_n(T_n)$, we have

$$\frac{1}{\sqrt{n}} \max_{k \in \Pi_n} \Big\| \frac{\partial L_n(T_k, \widehat{\theta}_n(T_n))}{\partial \theta} - \frac{1}{2}kF\left(\widehat{\theta}_n(T_k) - \widehat{\theta}_n(T_n)\right) \Big\| \leq \sqrt{n}\Big\| \frac{1}{2}\overline{F}_n(T_n, \widehat{\theta}_n(T_n))(\widehat{\theta}_n(T_n) - \theta_0) - \frac{1}{n}\frac{\partial L_n(T_n, \theta_0)}{\partial \theta} \Big\|$$

$$+ o_P(1). \tag{15}$$

Therefore, using (11), (15) implies

$$\frac{1}{\sqrt{n}} \max_{k \in \Pi_n} \Big\| \frac{\partial L_n(T_k, \widehat{\theta}_n(T_n))}{\partial \theta} - \frac{1}{2}kF\left(\widehat{\theta}_n(T_k) - \widehat{\theta}_n(T_n)\right) \Big\| = o_P(1). \tag{16}$$

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

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

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

Therefore, using (16) 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

$$\widehat{Q}_n = \max(\widehat{Q}_n^{(1)}, \widehat{Q}_n^{(2)}) \geq \max(\widehat{Q}_{n,k^*}^{(1)}, \widehat{Q}_{n,k^*}^{(2)}). \tag{17}$$

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

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

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

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

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

$$\xrightarrow[n\to\infty]{a.s.} \tau^* F^{(1)}(G^{(1)})^{-1}F^{(1)} + (1-\tau^*)F^{(2)}(G^{(2)})^{-1}F^{(2)}. \quad (18)$$

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 a symmetric positive definite matrix.

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.

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

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 exist $\rho_1 > 0$ and $\rho_2 > 0$ such as $B_o(\theta_1^*, \rho_1) \cap B_o(\theta_2^*, \rho_2) = \emptyset$.

For all $n \in I\!N$, denote

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

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

Using (18), we have

$$\delta_n^{(j)} \xrightarrow[n\to\infty]{a.s.} \delta^{(j)} \text{ for } j = 1, 2. \quad (19)$$

From [2] and [3], we have $\widehat{\theta}_n(T_{k^*}) \xrightarrow[n\to\infty]{a.s.} \theta_1^*$ and $\widehat{\theta}_n(\overline{T}_{k^*}) \xrightarrow[n\to\infty]{a.s.} \theta_2^*$. 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 $(\widehat{\theta}_n(T_{k^*}) - \widehat{\theta}_n(T_n))'\widehat{\Sigma}_{n,k^*}(\widehat{\theta}_n(T_{k^*}) - \widehat{\theta}_n(T_n)) \ge \delta_n^{(1)}$. Therefore,

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

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

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

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

Thus the result follows by using (19).   ∎

Acknowledgements

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. GARCH processes: structure and estimation. *Bernoulli 9* (2003), 201–227.

[6] BERKES, I., HORVÁTH, L., AND KOKOSZKA, P. Testing for parameter constancy in GARCH(p; q) models. *Statistics & Probability Letters 70*, (2004), 263–273.

[7] BILLINGSLEY, P. *Convergence of Probability Measures*. John Wiley & Sons Inc., New York, 1968.

[8] BOLLERSLEV, T. *Generalized autoregressive conditional heteroskedasticity. Journal of Econometrics, 31*, (1986), 307-327.

[9] 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.

[10] FRANCQ, C., AND ZAKOÏAN, J.-M. Maximum likelihood estimation of pure garch and arma-garch processes. *Bernoulli 10* (2004), 605–637.

[11] HORVÁTH, L. Change in autoregressive processes. *Stochastic Processes. Appl. 44*, (1993), 221-242.

[12] HORVÁTH, L., HORVÁTH, Z. AND HUSKOVÁ, M. Ratio tests for change point detection. *Inst. Math. Stat. 1*, (2008), 293-304.

[13] 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.

[14] KIERFER, J. K-sample analogues of the Kolmogorov-Smirnov and Cramér-v.Mises tests . *Ann. Math. Statist 30*, (1959), 420–447.

[15] KIM, S., CHO, S. AND LEE, S. On the CUSUM test for parameter changes in GARCH(1, 1) models. *Comm. Statist. Theory Methods 29*, (2000), 445-462.

[16] 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. Teyssiere. Lecture Notes in Statistic 200 (2010), Springer-Verlag, pp. 420–447.

[17] KOKOSZKA, P. AND R. LEIPUS   Testing for parameter changes in ARCH models. *Lithuanian Mathematical Journal 39*, (1999), 182-195.

[18] KULPERGER R. AND YU, H. High moment partial sum processes of residuals in GARCH models and their applications. *Ann. Statist. 33*, (2005), 2395-2422.

[19] LEE, S. , HA, J. , AND NA, O.   The Cusum Test for Parameter Change in Time Series Models . *Scand. J. Statist. 30*, (2003), 781–796.

[20] 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.

[21] LEE, S., SONG, J. Test for parameter change in ARMA models with GARCH innovations . *Statistics & Probability Letters 78*, (2008), 1990–1998.

[22] NELSON, D. B. AND CAO, C. Q. Inequality Constraints in the Univariate GARCH Model. *Journal of Business & Economic Statistics 10*, (1992), 229–235.

[23] PAGE, E. S. A test for a change in a parameter occurring at an unknown point. *Biometrika 42*, (1955), 523–526.