# Monitoring procedure for parameter change in causal time series

Jean-Marc Bardet
*SAMM, Université Paris 1 Panthéon-Sorbonne*
*E-mail: Jean-Marc.Bardet@univ-paris1.fr*

and

William Kengne
*THEMA, Université Cergy Pontoise*
*E-mail : William.Kengne@u-cergy.fr*

November 27, 2013

**Abstract** : We propose a new sequential procedure to detect change in the parameters of a process $X = (X_t)_{t\in\mathbb{Z}}$ belonging to a large class of causal models (such as AR($\infty$), ARCH($\infty$), TARCH($\infty$), ARMA-GARCH processes). The procedure is based on a difference between the historical parameter estimator and the updated parameter estimator, where both these estimators are quasi-likelihood estimators. Unlike classical recursive fluctuation test, the updated estimator is computed without the historical observations. The asymptotic behavior of the test is studied and the consistency in power as well as an upper bound of the detection delay are obtained. Some simulation results are reported with comparisons to some other existing procedures exhibiting the accuracy of our new procedure. This procedure coupled with retrospective tests is applied to solve off-line multiple breaks detection in the daily closing values of the FTSE 100 stock index.

*Keywords:* Sequential change detection; Change-point; Causal processes; Quasi-maximum likelihood estimator; Weak convergence.

## 1 Introduction

In statistical inference, many authors have pointed out the danger of omitting the existence of changes in data. Many papers have been devoted to the problem of test for parameter changes in time series models when all data are available, see for instance Horváth (1993), Inclan and Tiao (1994), Kokoszka and Leipus (1999), Aue *et al.* (2009) or Kengne (2012). These papers consider "retrospective" (off-line) changes *i.e.* changes in parameters when all data are available. Another point of view is to progressively detect change when new data arrive; this is the sequential change-point problem.

This problem can be seen as an engineering process control where new data arrive from an information on an industrial system; see for instance Basseville and Nikiforov (1993). In this paper, we will follow the paradigm of Chu *et al.* (1996) which considers this problem of "on-line segmentation" as a classical hypothesis testing with a fixed probability of type I error. Chu *et al.* (1996) studied the sequential change in regression model and pointed out the effects of repeating retrospective tests when new data are observed; this can increase the probability of type I error of the test. They developed two procedures based on cumulative sum (CUSUM) of residuals and recursive parameter fluctuations. Their idea has been generalized and several procedures are now based on this approach. Leisch *et al.* (2000) introduced the generalized fluctuation test based on the recursive moving estimator which contains the test of Chu *et al.* [15] as a special case. Horváth *et al.* (2004) introduced residual CUSUM monitoring procedure where the recursive parameter is based on the historical data. This procedure has been generalized by Aue *et al.* (2006) to the class of linear model with dependent errors. Berkes *et al.* (2004) considered sequential changes in the parameters of GARCH process. According to the fact that the functional limit theorem assumed by Chu *et al.* [15] is not satisfied by the squares of residuals of GARCH process, they developed a procedure based on quasi-likelihood scores. Kirch (2008) and Hušková and Kirch (2012) pointed out that the critical value of the monitoring procedure are usually based on an infinite observation period whereas it is finite in practice. Such procedure leads to a loss of some power. They proposed a bootstrapping methods to obtain critical values (even for a small sample sizes situation) for sequential change-point tests for linear regression models. The monitoring change in linear models has also been carried out by Hušková *et al.* (2009), Černíková *et al.* (2013). Na *et al.* (2011) developed a monitoring procedure for the detection of parameter changes in general time series models. They show that under the null hypothesis of no change, their detector statistic weakly converges to a known distribution. However, the asymptotic behavior of this detector is unknown under the alternative of parameter changes.

In this new contribution, we consider a large class of causal time series and investigate the asymptotic behavior under the null hypothesis of no change but also under the alternative hypothesis of change. More precisely, let $M, f : \mathbb{R}^{\mathbb{N}} \to \mathbb{R}$ be measurable functions, $(\xi_t)_{t\in\mathbf{Z}}$ be a sequence of centered independent and identically distributed (iid) random variables satisfying $\mathrm{var}(\xi_0) = \sigma^2$. We assume that the functions $M$ and $f$ are known up to some unknown parameter $\theta$ belonging to a fixed compact set $\Theta \subset \mathbb{R}^d$. Let $\mathcal{T} \subset \mathbf{Z}$, and for any $\theta \in \Theta$, define

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

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

The existence and properties of this general class of causal and affine processes were studied in Bardet and Wintenberger [4]. Numerous classical time series (such as AR($\infty$), ARCH($\infty$), TARCH($\infty$), ARMA-GARCH or bilinear processes) can be written as a model $\mathcal{M}_{\mathbf{Z}}(M, f)$. The off-line change detection for such class of models has already been studied in Bardet *et al.* [5] and Kengne [25].

Suppose now that we observed available historical data $X_1, \cdots, X_n$ with $(X_1, \cdots, X_n) \in \mathcal{M}_{\{1,\cdots,n\}}(M_{\theta_0^*}, f_{\theta_0^*})$ and $\theta_0^* \in \Theta$ is unknown. Then, we observe new data $X_{n+1}, X_{n+2}, \cdots, X_k, \cdots$: the monitoring scheme starts. For each new observation, we would like to know if a change occurs in the parameter $\theta_0^*$. More precisely, we consider the following test problem:

$\mathbf{H_0}$: $\theta_0^*$ is constant over the observation $X_1, \cdots, X_n, X_{n+1}, \cdots$ *i.e.* $(X_n)_{n\in\mathbb{N}} \in \mathcal{M}_{\mathbb{N}}(M_{\theta_0^*}, f_{\theta_0^*})$;

$\mathbf{H_1}$ : there exist $k^* > n$, $(\theta_0^*, \theta_1^*) \in \Theta^2$, with $\theta_0^* \neq \theta_1^*$, such that $(X_1, \cdots, X_{k^*}) \in \mathcal{M}_{\{1,\cdots,k^*\}}(M_{\theta_0^*}, f_{\theta_0^*})$ and $(X_{k^*+n})_{n\in\mathbb{N}} \in \mathcal{M}_{\{k^*+1,\cdots\}}(M_{\theta_1^*}, f_{\theta_1^*})$.

The main contribution of this paper is a new procedure proposed to test $\mathrm{H}_0$ against $\mathrm{H}_1$. For any $k \geq 1$, $\ell, \ell' \in \{1, \cdots, k\}$ (with $\ell \leq \ell'$) let $\widehat{\theta}(X_\ell, \cdots, X_{\ell'})$ be the quasi-maximum likelihood estimator (QMLE in the sequel) of the parameter computed on $\{\ell, \cdots, \ell'\}$ as it is defined in (7). When new data arrive at time $k > n$, we explore the segment $\{\ell, \ell+1, \cdots, k\}$ with $\ell \in \{n - v_n, \cdots, k - v_n\}$ (where $(v_n)_{n\in\mathbb{N}}$ is a fixed sequence of integer numbers) that a distance between $\widehat{\theta}(X_\ell, \cdots, X_k)$ and $\widehat{\theta}(X_1, \cdots, X_n)$ is the largest. We construct a detector taking into account a distance between $\widehat{\theta}(X_\ell, \cdots, X_k)$ and $\widehat{\theta}(X_1, \cdots, X_n)$ and if this distance is larger than a suitable critical value, then $H_0$ is rejected and a model with a new parameter is considered; otherwise, the monitoring scheme continues. We show that this detector is almost surely consistent under $\mathrm{H}_0$ and almost surely diverges to infinity under $\mathrm{H}_1$: the consistency of our procedure follows.

Finally, Monte-Carlo experiments have been done on several processes, comparing our procedure to the ones of Horváth *et al.* [21] (see also Aue *et al.* [2]) and Na *et al.* [31]. It appears that our procedure outperforms these other procedures in terms of test power and detection delay in different frames. An application to financial data (FTSE 100 stock index) allows to detect changes in these data in accordance with historical and economic events.

In the forthcoming Section 2 the assumptions and the definition of the quasi-likelihood estimator are provided. In Section 3 we present the monitoring procedure and the asymptotic results. Section 4 is devoted to a simulation study for AR(1) and GARCH(1, 1) processes. In Section 5 we apply our procedure to famous financial data. The proofs of the main results are provided in Section 6.

# 2 Assumptions and definition of the quasi-likelihood estimator

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

We begin by giving assumptions ensuring the existence and stationarity of a process belonging to a class $\mathcal{M}_{\mathbf{Z}}(f_\theta, M_\theta)$. First, define $\theta \in \mathbb{R}^d$ and

- $M_\theta$ and $f_\theta$ are numerical functions such that for all $(x_i)_{i\in\mathbb{N}} \in \mathbb{R}^{\mathbb{N}}$, $M_\theta\big((x_i)_{i\in\mathbb{N}}\big) \neq 0$.
- $h_\theta := M_\theta^2$.

We will use the following classical notations:

1. $\|\cdot\|$ applied to a vector denotes the Euclidean norm of the vector;
2. for any compact set $\mathcal{K} \subseteq \mathbb{R}^d$ and for any $g : \mathcal{K} \longrightarrow \mathbb{R}^{d'}$, $\|g\|_{\mathcal{K}} = \sup_{\theta\in\mathcal{K}}(\|g(\theta)\|)$;
3. for any set $\mathcal{K} \subseteq \mathbb{R}^d$, $\overset{\circ}{\mathcal{K}}$ denotes the interior of $\mathcal{K}$.

Throughout the sequel, we will assume that the functions $\theta \mapsto M_\theta$ and $\theta \mapsto f_\theta$ are twice continuously differentiable on $\Theta$. Let $\Psi_\theta = f_\theta$, $M_\theta$ and $i = 0,\ 1,\ 2$, then for any compact set $\mathcal{K} \subset \Theta$ define

**Assumption $\mathbf{A}_i(\Psi_\theta, \mathcal{K})$:** *Assume that* $\|\partial^i \Psi_\theta(0)/\partial\theta^i\|_\Theta < \infty$ *and there exists a sequence of non-negative real numbers* $(\alpha_j^{(i)}(\Psi_\theta, \mathcal{K}))_{j\geq 1}$ *such that* $\sum\limits_{j=1}^{\infty} \alpha_j^{(i)}(\Psi_\theta, |) < \infty$ *and*

$$\Big\|\frac{\partial^i \Psi_\theta(x)}{\partial\theta^i} - \frac{\partial^i \Psi_\theta(y)}{\partial\theta^i}\Big\|_{\mathcal{K}} \leq \sum_{j=1}^{\infty} \alpha_j^{(i)}(\Psi_\theta, \mathcal{K})|x_j - y_j| \quad \textit{for all } x, y \in \mathbb{R}^{\mathbb{N}}.$$

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 $\mathbf{A}_i(h_\theta, \mathcal{K})$:** *Assume that* $\|\partial^i h_\theta(0)/\partial\theta^i\|_\Theta < \infty$ *and there exists a sequence of non-negative real numbers* $(\alpha_j^{(i)}(h_\theta, \mathcal{K}))_{j\geq 1}$ *such as* $\sum\limits_{j=1}^{\infty} \alpha_j^{(i)}(h_\theta, \mathcal{K}) < \infty$ *and*

$$\Big\|\frac{\partial^i h_\theta(x)}{\partial\theta^i} - \frac{\partial^i h_\theta(y)}{\partial\theta^i}\Big\|_{\mathcal{K}} \leq \sum_{j=1}^{\infty} \alpha_j^{(i)}(h_\theta, \mathcal{K})|x_j^2 - y_j^2| \quad \textit{for all } x, y \in \mathbb{R}^{\mathbb{N}}.$$

The Lipschitz-type hypothesis $A_i(\Psi_\theta, \mathcal{K})$ are classical when studying the existence of solutions of models (see for instance [17]). Using a result of [4], for $r \geq 1$ and a model $\mathcal{M}_{\mathbf{Z}}(M_\theta, f_\theta)$, it is interesting to define the set:

$$\Theta(r) := \Big\{\theta \in \Theta,\ A_0(f_\theta, \{\theta\}) \text{ and } A_0(M_\theta, \{\theta\}) \text{ hold with } \sum_{j\geq 1} \alpha_j^{(0)}(f_\theta, \{\theta\}) + (E\,|\xi_0|^r)^{1/r} \sum_{j\geq 1} \alpha_j^{(0)}(M_\theta, \{\theta\}) < 1\Big\}$$
$$\bigcup \Big\{\theta \in \Theta,\ f_\theta = 0 \text{ and } A_0(h_\theta, \{\theta\}) \text{ holds with } (E\,|\xi_0|^r)^{2/r} \sum_{j\geq 1} \alpha_j^{(0)}(h_\theta, \{\theta\}) < 1\Big\}.$$

Then, if $\theta \in \Theta(r)$ the existence of a unique causal, stationary and ergodic solution $X = (X_t)_{t\in\mathbf{Z}} \in \mathcal{M}_{\mathbf{Z}}(f_\theta, M_\theta)$ satisfying $E\,|X_t|^r < \infty$ is ensured (see more details in [4]). The subset $\Theta(r)$ is defined as a union to consider accurately general causal models and ARCH-type models simultaneously.

Here there are assumptions required for studying QMLE asymptotic properties:

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

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

**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 and there exists $\ell > 2$ such that $\alpha_j^{(i)}(f_\theta, \Theta) + \alpha_j^{(i)}(M_\theta, \Theta) + \alpha_j^{(i)}(h_\theta, \Theta) = \mathcal{O}(j^{-\ell})$ for $j \in I\!N$.

Note that in this last assumption, as in [4], we use 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 $\alpha_\ell^{(i)}(M_\theta, \Theta) = 0$.

## 2.2 Examples

### 2.2.1 AR($\infty$) processes

Consider a AR($\infty$) process which is a generalization of ARMA($p, q$) processes and is defined by:

$$X_t = \phi_0(\theta_0^*) + \sum_{j\geq 1} \phi_j(\theta_0^*) X_{t-j} + \xi_t\ ,\ t \in \mathbf{Z}. \tag{2}$$

This process belongs to the class $\mathcal{M}_{\mathbf{Z}}(M_{\theta_0^*}, f_{\theta_0^*})$ where:

- $f_\theta(x_1, \cdots) = \phi_0(\theta) + \sum_{j\geq 1} \phi_j(\theta) x_j$ and therefore $\alpha_j^{(0)}(f_\theta, \Theta) = \|\phi_j(\theta)\|_\Theta$ for $j \in I\!N^*$;
- $M_\theta \equiv 1$ and therefore $\alpha_j^{(0)}(M_\theta, \Theta) = 0$ for $j \in I\!N^*$ and $\alpha_0^{(0)}(M_\theta, \Theta) = 1$.

Then, $\theta_0^* \in \overset{\circ}{\Theta}$ where we can chose $\Theta$ as a compact subset of $\Theta(4) \subset I\!R^d$ such as

$$\Theta(4) = \big\{\theta \in I\!R^d;\ \sum_{j\geq 1} |\phi_j(\theta)| < 1\big\}.$$

Then

- Assumption $\mathbf{D}(\Theta)$ holds if $\underline{h} = \inf_{\theta\in\Theta}(|\phi_0(\theta)|) > 0$;
- Assumption $\mathbf{K}(f_\theta, M_\theta, \Theta)$ holds if there exists $\ell > 2$ and and if $\theta \mapsto \phi_j(\theta)$ are twice differentiable functions satisfying $\max\big(\|\psi_j(\theta)\|_\Theta, \|\phi'_j(\theta)\|_\Theta, \|\phi''_j(\theta)\|_\Theta\big) = O(j^{-\ell})$ for $j \in I\!N$.

- if $(\xi_t)$ is a sequence of non-degenerate random variables, Assumptions $\mathbf{Id}(\Theta)$ and $\mathbf{Var}(\Theta)$ hold.

**Particular case of AR($p$) processes**
Assume that

$$X_t = \phi_0^* + \sum_{j=1}^{p} \phi_j^* X_{t-j} + \xi_t \quad \text{with } p \in I\!N^*.$$

The true parameter of the model is denoted by $\theta_0^* = (\phi_0^*, \phi_1^*, \cdots, \phi_p^*) \in \Theta$ where $\Theta = \{\theta = (\phi_0, \phi_1, \cdots, \phi_p) \in I\!R^{p+1} \ / \sum_{j=1}^{p} |\phi_j| < 1\}$.

### 2.2.2 ARCH($\infty$) processes

Consider a ARCH($\infty$) process which is generalization of GARCH($p,q$) processes and is defined by:

$$X_t = \sigma_t \, \xi_t \quad \text{and} \quad \sigma_t^2 = \psi_0(\theta_0^*) + \sum_{j=1}^{\infty} \psi_j(\theta_0^*) X_{t-j}^2 \ , \ t \in \mathbf{Z}. \tag{3}$$

This process, introduced by Robinson [34], is a "'ARCH-type process"' belonging to the class $\mathcal{M}_{\mathbf{Z}}(f_{\theta_0^*}, M_{\theta_0^*},)$ where:

- $f_\theta(x_1, \cdots) \equiv 0$ and therefore $\alpha_j^{(0)}(f_\theta, \Theta) = 0$ for $j \in I\!N^*$;
- $h_\theta(x_1, \cdot) = M_\theta^2(x_1, \cdots) = \psi_0(\theta) + \sum_{j \geq 1} \psi_j(\theta) x_j^2$ and therefore $\alpha_j^{(0)}(h_\theta, \Theta) = \|\psi_j(\theta)\|_\Theta$ for $j \in I\!N^*$.

Then $\theta_0^* \in \overset{\circ}{\Theta}$, where we can chose $\Theta$ as a compact subset of $\Theta(4) \subset I\!R^d$ where

$$\Theta(4) = \big\{\theta \in I\!R^d; \ (E\,|\xi_0|^4)^{1/2} \sum_{j=1}^{\infty} |\psi_j(\theta)| < 1\big\}.$$

Then

- Assumption $\mathbf{D}(\Theta)$ holds if $\underline{h} = \inf_{\theta \in \Theta}(\psi_0(\theta)) > 0$;
- Assumption $\mathbf{K}(f_\theta, M_\theta, \Theta)$ holds if there exists $\ell > 2$ such as $\theta \mapsto \phi_j(\theta)$ are twice differentiable functions satisfying $\max\big(\|\psi_j(\theta)\|_\Theta, \|\psi_j'(\theta)\|_\Theta, \|\psi_j''(\theta)\|_\Theta\big) = O(j^{-\ell})$ for $j \in I\!N$.
- if $(\xi_t^2)$ is a sequence of non-degenerate random variables, Assumptions $\mathbf{Id}(\Theta)$ and $\mathbf{Var}(\Theta)$ hold.

**Particular case of GARCH$(1,1)$ processes**
Assume that

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

with $\theta_0^* = (\alpha_0^*, \alpha_1^*, \beta_1^*) \in \Theta \subset ]0, \infty[ \times [0, \infty[^2$ such as $\Theta \subset \Theta(2) = \{\alpha_1^* + \beta_1^* < 1\}$. The ARCH($\infty$) representation of this process is

$$\sigma_t^2 = \alpha_0^*/(1 - \beta_1^*) + \alpha_1^* \sum_{j \geq 1} (\beta_1^*)^{j-1} X_{t-j}^2,$$

and therefore $f_\theta(x_1, \cdots) \equiv 0$ and $h_\theta(x_1, \cdot) = \alpha_0^*/(1 - \beta_1^*) + \alpha_1^* \sum_{j \geq 1} (\beta_1^*)^{j-1} x_{t-j}^2$.

### 2.2.3 TARCH($\infty$) processes

The process $X$ is called Threshold ARCH($\infty$) (TARCH($\infty$) in the sequel) if it satisfies

$$X_t = \sigma_t \xi_t \quad \text{and} \quad \sigma_t = b_0(\theta_0^*) + \sum_{j=1}^{\infty} \Big[ b_j^+(\theta_0^*) \max(X_{t-j}, 0) - b_j^-(\theta_0^*) \min(X_{t-j}, 0) \Big] \ , \ t \in \mathbf{Z} \tag{4}$$

where the parameters $b_0(\theta)$, $b_j^+(\theta)$ and $b_j^-(\theta)$ are assumed to be non negative real numbers and $\theta \in \overset{\circ}{\Theta}$ where $\Theta$ is a compact subset of $\Theta(4)$ defined by

$$\Theta(4) = \Big\{ \theta \in \mathbb{R}^d \ \Big/ \ \big(E\,|\xi_0|^4\big)^{1/4} \sum_{j=1}^{\infty} \max\big(b_j^-(\theta), b_j^+(\theta)\big) < 1 \Big\}$$

since $\alpha_j^{(0)}(M, \{\theta\}) = \max\big(b_j^-(\theta), b_j^+(\theta)\big)$. This class of processes is a generalization of the class of TGARCH($p$,$q$) processes (introduced by Rabemananjara and Zakoïan [33]). Then,

- Assumption $\mathbf{D}(\Theta)$ holds if $\underline{h} = \inf_{\theta \in \Theta} b_0(\theta) > 0$;
- Assumption $\mathbf{K}(f_\theta, M_\theta, \Theta)$ holds if there exists $\ell > 2$ and and if $\theta \mapsto b_j^-(\theta)$ and $\theta \mapsto b_j^+(\theta)$ are twice differentiable functions satisfying

$$\max\big(\|b_j^-(\theta)\|_\Theta, \|b_j^+(\theta)\|_\Theta, \|\frac{\partial}{\partial\theta} b_j^-(\theta)\|_\Theta, \|\frac{\partial}{\partial\theta} b_j^+(\theta)\|_\Theta, \|\frac{\partial^2}{\partial\theta^2} b_j^-(\theta)\|_\Theta, \|\frac{\partial^2}{\partial\theta^2} b_j^+(\theta)\|_\Theta\big) = O(j^{-\ell}) \quad \text{for} \quad j \in \mathbb{N}.$$

Unfortunately, for TARCH($\infty$) it is not possible to provide simple conditions for obtaining Assumptions $\mathbf{Id}(\Theta)$ and $\mathbf{Var}(\Theta)$ as for AR($\infty$) or ARCH($\infty$) processes.

## 2.3 Quasi-maximum likelihood estimators

For $1 \le \ell \le \ell'$, denote

$$T_{\ell,\ell'} := \{\ell, \ell+1, \cdots, \ell'\}.$$

Let $k > n \ge 2$. If $(X_1, \cdots, X_k) \in \mathcal{M}_{\{1,\cdots,k\}}(M_\theta, f_\theta)$, then for any subset $T_{\ell,\ell'} \subset \{1, \cdots, k\}$, the conditional quasi-(log)likelihood computed on $T$ is given by:

$$L(T_{\ell,\ell'}, \theta) := -\frac{1}{2} \sum_{t \in T_{\ell,\ell'}} q_t(\theta) \ \ \text{with} \ \ q_t(\theta) = \frac{(X_t - f_\theta^t)^2}{h_\theta^t} + \log(h_\theta^t) \tag{5}$$

where $f_\theta^t = f_\theta\big(X_{t-1}, X_{t-2} \ldots\big)$, $M_\theta^t = M_\theta\big(X_{t-1}, X_{t-2} \ldots\big)$ and $h_\theta^t = {M_\theta^t}^2$. The classical approximation of this conditional log-likelihood (see more details in Bardet and Wintenberger [4]) is given by:

$$\widehat{L}(T_{\ell,\ell'}, \theta) := -\frac{1}{2} \sum_{t \in T_{\ell,\ell'}} \widehat{q}_t(\theta) \quad \text{where} \quad \widehat{q}_t(\theta) := \frac{\big(X_t - \widehat{f}_\theta^t\big)^2}{\widehat{h}_\theta^t} + \log\big(\widehat{h}_\theta^t\big) \tag{6}$$

with $\widehat{f}_\theta^t = f_\theta\big(X_{t-1}, \ldots, X_1, 0, 0, \cdots\big)$, $\widehat{M}_\theta^t = M_\theta\big(X_{t-1}, \ldots, X_1, 0, 0, \cdots\big)$ and $\widehat{h}_\theta^t = (\widehat{M}_\theta^t)^2$.
For $T_{\ell,\ell'} \subset \{1, \cdots, k\}$, define the quasi maximum-likelihood estimator (QMLE) of $\theta$ computed on $T_{\ell,\ell'}$ by

$$\widehat{\theta}(T_{\ell,\ell'}) := \underset{\theta \in \Theta}{\operatorname{argmax}}(\widehat{L}(T_{\ell,\ell'}, \theta)). \tag{7}$$

In Bardet and Wintenberger [4] it was established that if $(X_1, \cdots, X_n)$ is an observed trajectory of $X \subset \mathcal{M}_{\mathbf{Z}}(f_{\theta_0^*}, M_{\theta_0^*})$ with $\theta_0^* \in \overset{\circ}{\Theta}(4)$ and if $\Theta$ is a compact set such as Assumptions $\mathbf{A}_i(f_\theta, M_\theta, \Theta)$ (or $\mathbf{A}_i(h_\theta, \Theta)$) hold for $i = 0, 1, 2$ and under Assumptions $\mathbf{D}(\Theta)$, $\mathbf{Id}(\Theta)$, $\mathbf{Var}(\Theta)$, $\mathbf{K}(f_\theta, M_\theta, \Theta)$, then

$$\sqrt{n}\big(\widehat{\theta}(T_{1,n}) - \theta_0^*\big) \ \xrightarrow[n\to\infty]{\mathcal{D}} \ \mathcal{N}\big(0\,,\, F\,G^{-1}\,F\big), \tag{8}$$

with

$$G := E\Big[\frac{\partial q_0(\theta_0^*)}{\partial\theta}\frac{\partial q_0(\theta_0^*)}{\partial\theta}'\Big] \quad \text{and} \quad F := E\Big[\frac{\partial^2 q_0(\theta_0^*)}{\partial\theta\partial\theta'}\Big], \tag{9}$$

where $'$ denotes the transpose and with $q_0$ defined in (5). Note that under assumptions $\mathbf{D}(\Theta)$ and $\mathbf{Var}(\Theta)$, $G$ is symmetric positive definite (see [25]) and $F$ is non-singular (see [4]). Also define the matrix

$$\widehat{G}(T_{\ell,\ell'}) := \frac{1}{\ell'-\ell+1}\sum_{t\in T_{\ell,\ell'}}\Big(\frac{\partial\widehat{q}_t(\widehat{\theta}(T_{\ell,\ell'}))}{\partial\theta}\Big)\Big(\frac{\partial\widehat{q}_t(\widehat{\theta}(T_{\ell,\ell'}))}{\partial\theta}\Big)' \quad \text{and} \quad \widehat{F}(T_{\ell,\ell'}) := -\frac{2}{\ell'-\ell+1}\Big(\frac{\partial^2\widehat{L}(T_{\ell,\ell'},\widehat{\theta}(T_{\ell,\ell'}))}{\partial\theta\partial\theta'}\Big). \tag{10}$$

Under the previous assumptions, $\widehat{G}(T_{1,n})$ and $\widehat{F}(T_{1,n})$ converge almost surely to $G$ and $F$ respectively. Hence,

$$\sqrt{n}\,\widehat{G}(T_{1,n})^{-1/2}\widehat{F}(T_{1,n})\big(\widehat{\theta}(T_{1,n})-\theta_0^*\big) \xrightarrow[n\to\infty]{\mathcal{D}} \mathcal{N}\big(0\,,\,I_d\big) \tag{11}$$

with $I_d$ the identity matrix. This result will be the starting point of the following monitoring procedure.

# 3 The monitoring procedure and asymptotic results

## 3.1 The monitoring procedure

In the sequel, $(X_1,\cdots,X_n)$ is supposed to be the historical available observations belonging to the class $\mathcal{M}_{\{1,\cdots,n\}}(f_{\theta_0^*},M_{\theta_0^*})$. At a monitoring instant $k$, our procedure evaluates the difference between $\widehat{\theta}(T_{\ell,k})$ and $\widehat{\theta}(T_{1,n})$ for any $\ell=n,\cdots,k$. More precisely, from (11), for any $k>n$ define the statistic (called the detector)

$$\widehat{C}_{k,\ell} := \sqrt{n}\,\frac{k-\ell}{k}\,\big\|\widehat{G}(T_{1,n})^{-1/2}\widehat{F}(T_{1,n})\big(\widehat{\theta}(T_{\ell,k})-\widehat{\theta}(T_{1,n})\big)\big\|$$

for $\ell=n,\cdots,k$. Since the matrix $\widehat{G}(T_{1,n})$ is asymptotically symmetric and positive definite (see [25]), $\widehat{G}(T_{1,n})^{-1/2}$ exists for $n$ large enough and $\widehat{C}_{k,\ell}$ is well defined.
At the beginning of the monitoring scheme and when $\ell$ is close to $k$, the length of $T_{\ell,k}$ is too small, therefore the numerical algorithm used to compute $\widehat{\theta}(T_{\ell,k})$ cannot converge. This can introduce a large distortion in the procedure. To avoid this, we introduce a sequence of integer numbers $(v_n)_{n\in\mathbb{N}}$ with $v_n << n$ and compute $\widehat{C}_{k,\ell}$ for $\ell\in\{n-v_n,n-v_n+1,\cdots,k-v_n\}$. Thus, for any $k>n$ denote

$$\Pi_{n,k} := \{n-v_n,n-v_n+1,\cdots,k-v_n\}.$$

For technical reasons, assume that,

$$v_n\to\infty \quad \text{and} \quad v_n/\sqrt{n}\to 0 \ \ (n\to\infty).$$

According to Remark 1 of [25], we can choose $v_n=[(\log n)^\delta]$ with $\delta>1$.

Note that, if change does not occur at time $k>n$, for any $\ell\in\Pi_{n,k}$, both the estimators $\widehat{\theta}(T_{\ell,k})$ and $\widehat{\theta}(T_{1,n})$ are close and the detector $\widehat{C}_{k,\ell}$ is not too large.
Let $T>1$ ($T$ can be equal to infinity). The monitoring scheme rejects $H_0$ at the first time $k$ satisfying $n<k\le[Tn]+1$ and there exists $\ell\in\Pi_{n,k}$ such that $\widehat{C}_{k,\ell}>c$ for a suitably chosen constant $c>0$, where $[x]$ denote the integer part of $x$. The procedure is called **closed-end** method when $T<\infty$ and **open-end** method when $T=\infty$.

To be more general, we will use a function $b:(0,\infty)\mapsto(0,\infty)$, called a boundary function satisfying:

**Assumption B:** $b:(0,\infty)\mapsto(0,\infty)$ is a non-increasing and continuous function such as $\underset{0<t<\infty}{\text{Inf}}\, b(t)>0$.

Then the monitoring scheme rejects $H_0$ at the first time $k$ ( with $n<k\le[Tn]+1$ ) such as there exists $\ell\in\Pi_{n,k}$ satisfying $\widehat{C}_{k,\ell}>b((k-\ell)/n)$. Hence define the stopping time:

$$\tau(n) := \text{Inf}\Big\{n<k\le[Tn]+1 \,/\, \exists\ell\in\Pi_{n,k},\ \widehat{C}_{k,\ell}>b((k-\ell)/n)\Big\} = \text{Inf}\Big\{n<k\le[Tn]+1 \,/\, \max_{\ell\in\Pi_{n,k}}\frac{\widehat{C}_{k,\ell}}{b((k-\ell)/n)}>1\Big\}$$

with the convention that $\text{Inf}\{\emptyset\} = \infty$. Therefore, we have

$$P\{\tau(n) < \infty\} = P\Big\{ \max_{\ell \in \Pi_{n,k}} \frac{\widehat{C}_{k,\ell}}{b((k-\ell)/n)} > 1 \ \text{ for some } k \ \text{ between } n \text{ and } [Tn]+1 \Big\}$$
$$= P\Big\{ \sup_{n<k\leq [Tn]+1} \max_{\ell \in \Pi_{n,k}} \frac{\widehat{C}_{k,\ell}}{b((k-\ell)/n)} > 1 \Big\}. \tag{12}$$

The challenge is to choose a suitable boundary function $b(\cdot)$ such as for some given $\alpha \in (0,1)$,

$$\lim_{n\to\infty} P_{H_0}\{\tau(n) < \infty\} = \alpha$$

and

$$\lim_{n\to\infty} P_{H_1}\{\tau(n) < \infty\} = 1$$

where the hypothesis $H_0$ and $H_1$ are specified in Section 1.
In the case where $b(\cdot)$ is a constant positive value, $b \equiv c$ with $c > 0$, these conditions lead to compute a threshold $c = c_\alpha$ depending on $\alpha$. If change is detected under $H_1$ *i.e.* $\tau(n) < \infty$ and $\tau(n) > k^*$, then the detection delay is defined by

$$\widehat{d}_n = \tau(n) - k^*.$$

Using the previous notations, Na *et al.* [31] defined the following detector

$$\widehat{D}_k := \sqrt{n} \big\| \widehat{G}(T_{1,n})^{-1/2} \widehat{F}(T_{1,n}) \big( \widehat{\theta}(T_{1,k}) - \widehat{\theta}(T_{1,n}) \big) \big\|. \tag{13}$$

At the step $k$ of the monitoring scheme, their recursive estimator is based on $X_1, \cdots, X_n, \cdots, X_k$. One can see that this estimator is highly influenced by historical data. Assume that a change occurs at time $k^* \leq k$, in the sequel of the procedure, the recursive estimator contents the observations $X_1, \cdots, X_n, \cdots, X_{k^*-1}$ which depends on $\theta_0^*$. Then, one must wait longer before the difference between $\widehat{\theta}(X_1, \cdots, X_n)$ and $\widehat{\theta}(X_1, \cdots, X_n, \cdots, X_k)$ becomes significant at a step $k > k^*$. Therefore, their procedure cannot be effective in terms of detection delay (see also the simulations study Section 4).

Berkes *et al.* (2004) considered an estimator based on historical data to compute the quasi-likelihood scores. They used the fact that the partial derivatives applied to a vector $\mathbf{u}$ is equal to 0 if and only if $\mathbf{u}$ is the true parameter of the model. So, when change occurs, their detector growths asymptotically to infinity. Therefore, their procedure is consistent. They proved this result for GARCH$(p,q)$ models.

## 3.2 Asymptotic behaviour under the null hypothesis

Under $H_0$, the parameter $\theta_0^*$ does not change over the new observations. Thus we have the result

**Theorem 3.1.** *Assume* $\boldsymbol{D}(\Theta)$, $\boldsymbol{Id}(\Theta)$, $\boldsymbol{Var}(\Theta)$, $\boldsymbol{K}(f_\theta, M_\theta, \Theta)$, $\mathbf{B}$ *and* $\theta_0^* \in \overset{\circ}{\Theta}(4)$. *Under* $H_0$ *and with* $W_d$ *a d-dimensional standard Brownian motion,*

- *(i) if* $T = \infty$ *(open-end procedure), then*

$$\lim_{n\to\infty} P\{\tau(n) < \infty\} = P\Big\{ \sup_{t>1} \sup_{1<s<t} \frac{\|W_d(s) - sW_d(1))\|}{t\ b(s)} > 1 \Big\},$$

- *(ii) if* $T < \infty$ *(closed-end procedure), then*

$$\lim_{n\to\infty} P\{\tau(n) < \infty\} = P\Big\{ \sup_{1<t\leq T} \sup_{1<s<t} \frac{\|W_d(s) - sW_d(1))\|}{t\ b(s)} > 1 \Big\}.$$

For simulations, we will use the most "natural" boundary function $b(\cdot) = c$ with $c$ a positive constant since it satisfies the above assumptions imposed to $b(\cdot)$. Hence, Corollary 3.1 follows directly from Theorem 3.1.

**Corollary 3.1.** *Assume $b(t) = c > 0$ for all $t \geq 0$. Under the assumptions of Theorem 3.1, and with $T \in (1, \infty)$,*

$$\lim_{n\to\infty} P\{\tau(n) < \infty\} = P\{U_{d,T} > c\}$$

*where, using the convention $U_{d,T} = U_{d,\infty}$ when $T = \infty$,*

$$U_{d,T} = \sup_{1<t\leq T} \sup_{1<s<t} \frac{1}{t}\|W_d(s) - sW_d(1))\|. \tag{14}$$

Hence, at a nominal level $\alpha \in (0, 1)$, we take $c = c(\alpha)$ the $(1 - \alpha)$-quantile of the distribution of $U_{d,T}$ which can be computed through Monte-Carlo simulations (see Section 4).

**Remark 3.1.** *Under $H_0$, we have $\widehat{\theta}(T_{1,n}) \xrightarrow[n\to\infty]{a.s.} \theta_0^*$ (see [4]). Thus denote*

$$\widehat{C}_{k,\ell}^{(0)} := \sqrt{n}\,\frac{k-\ell}{k}\,\big\|\widehat{G}(T_{1,n})^{-1/2}\widehat{F}(T_{1,n})\big(\widehat{\theta}(T_{\ell,k}) - \theta_0^*\big)\big\|.$$

*Under the assumptions of Theorem 3.1, one can easily show that*

$$\sup_{k>n} \max_{\ell\in\Pi_{n,k}} \frac{1}{b((k-\ell)/n)}\,\big|\widehat{C}_{k,\ell} - \widehat{C}_{k,\ell}^{(0)}\big| = o_P(1) \quad \text{as } n\to\infty$$

*also in the case of closed-end situation. Thus, Theorem 3.1 still holds if the stopping time $\tau(n)$ is computed by using the detector $\widehat{C}_{k,\ell}^{(0)}$. Hence, if the parameter $\theta_0^*$ of the historical observations is known, then use the detector $\widehat{C}_{k,\ell}^{(0)}$ instead of $\widehat{C}_{k,\ell}$. But let us note that this situation is infrequent in practice.*

### 3.3 Asymptotic behaviour under the alternative hypothesis

Under $H_1$, the parameter changes from $\theta_0^*$ to $\theta_1^*$ at $k^* > n$, where $\theta_1^* \in \Theta$ and $\theta_0^* \neq \theta_1^*$. The following theorem shows that the detector tends to infinity both in the open-end and closed-end methods.

**Theorem 3.2.** *Assume $\boldsymbol{D}(\Theta)$, $\boldsymbol{Id}(\Theta)$, $\boldsymbol{Var}(\Theta)$, $\boldsymbol{K}(f_\theta, M_\theta, \Theta)$ and $\mathbf{B}$. Under the alternative $H_1$, if $\theta_1^* \neq \theta_0^*$; $\theta_0^*, \theta_1^* \in \overset{\circ}{\Theta}(4)$ and there exists $T^* \in (1, T)$ such as $k^* = k^*(n) = [T^*n]$ then for $k_n = k^*(n) + n^\delta$ with $\delta \in (1/2, 1)$,*

$$\max_{\ell\in\Pi_{n,k_n}} \frac{\widehat{C}_{k_n,\ell}}{b((k_n-\ell)/n)} \xrightarrow[n\to\infty]{a.s.} \infty.$$

**Remark 3.2.** *The Theorem 3.2 still holds if $k^*$ belongs in a set $\{n+1, n+2, \cdots, [T^*n]\}$ for any $T^* \in (1, T)$.*

The forthcoming Corollary 3.2 can be immediately deduced from the relation (12).

**Corollary 3.2.** *Under assumptions of Theorem 3.2,*

$$\lim_{n\to\infty} P\{\tau(n) < \infty\} = 1.$$

**Remark 3.3.** *We know that the monitoring scheme rejects $H_0$ at the first time $k$ where*

$$\max_{\ell\in\Pi_{n,k}} \frac{\widehat{C}_{k,\ell}}{b((k-\ell)/n)} > 1.$$

*Therefore, it follows from Theorem 3.2 that with probability one, the change is asymptotically detected both for open-end and closed-end (when $T^* < T$) procedures and the detection delay $\widehat{d}_n$ can be bounded by $\mathcal{O}_P(n^{1/2+\varepsilon})$ for any $\varepsilon > 0$ (or even by $\mathcal{O}_P\big(\sqrt{n}(\log n)^a\big)$ with $a > 0$ using the same kind of proof).*

# 4 Some simulation and numerical experiments

We present now several simulation and numerical experiments. As a consequence, we assume here that $T < \infty$ (we can not monitoring until infinity) *i.e.* a closed-end procedure.

## 4.1 Concrete procedure of monitoring

At a time $k$ ($n < k \leq [Tn] + 1$), we compute $\widehat{C}_{k,\ell}$ for all $\ell \in \Pi_{n,k}$ to test whether change occurs or not. We have chosen the following parameters:

- $T = 1.5$. Hence the procedure is monitored from $k = n+1$ to $k = 1.5 \times n$. The set $\{n+1, \cdots, 1.5 \times n\}$ is called the monitoring period.
- according to the Remark 1 of [25], $v_n = [(\log n)^{\delta}]$ (with $1 \leq \delta \leq 3$) is chosen. We evaluated the performance of the procedure with $v_n = [\log n]$, $[(\log n)^{3/2}]$, $[(\log n)^2]$, $[(\log n)^3]$ and we recommend to use $v_n = [(\log n)^{3/2}]$ for linear model and $v_n = [(\log n)^2]$ for ARCH-type model.
- the nominal level used in the sequel is $\alpha = 0.05$.

We also have to evaluate the different thresholds of the tests. The forthcoming proposition provides a way for obtaining such thresholds:

**Proposition 4.1.** *Denote* $\stackrel{\mathcal{D}}{=}$ *the equality in distribution. There exists a function* $f : (0,1) \times (1, \infty) \to [0, \infty)$ *defined in (28) such that for any* $T > 1$,

$$U_{d,T} \stackrel{\mathcal{D}}{=} \sup_{0<u<1} f(u,T) \, \|W_d(u)\|$$

*For the open-end procedure, there exists a function* $g : (0,1) \to [0,\infty)$ *defined in (29) such that,*

$$U_{d,\infty} \stackrel{\mathcal{D}}{=} \sup_{0<u<1} g(u) \, \|W_d(u)\|.$$

Table 1 shows the 0.95-quantile of the distribution of $U_{d,T}$ for $d = 1, \cdots, 5$ with $T = 1.5, 10, \infty$.

| | $d=1$ | $d=2$ | $d=3$ | $d=4$ | $d=5$ |
|---|---|---|---|---|---|
| $T=1.5$ | 1.297 | 1.495 | 1.709 | 1.875 | 1.991 |
| $T=10$ | 1.883 | 2.351 | 2.692 | 2.945 | 3.184 |
| $T=\infty$ | 1.954 | 2.432 | 2.760 | 3.073 | 3.334 |

Table 1: Empirical 0.95-quantile of the distribution of $U_{d,T}$.

For any $k > n$, denote

$$\widehat{C}_k = \max_{\ell \in \Pi_{n,k}} \widehat{C}_{k,\ell}.$$

Then, when $\max_{n<k\leq[Tn]+1} \widehat{C}_k$ is larger than the 0.95-quantile of the distribution of $U_{d,T}$, we reject $\mathrm{H}_0$ and decide $\mathrm{H}_1$: a change occurs at a time before $k$.

## 4.2 An illustration

We consider a GARCH(1,1) process : $X_t = \sigma_t \xi_t$ with $\sigma_t^2 = \alpha_0^* + \alpha_1^* X_{t-1}^2 + \beta_1^* \sigma_{t-1}^2$.
Thus, the parameter of the model is $\theta_0^* = (\alpha_0^*, \alpha_1^*, \beta_1^*)$. The historical available data are $X_1, \cdots, X_{1000}$ (therefore $n = 1000$) and the monitoring period is $\{1001, \cdots, 1500\}$. At the nominal level $\alpha = 0.05$, the critical values of the procedure is $C_\alpha = 1.709$ (see Table 1).

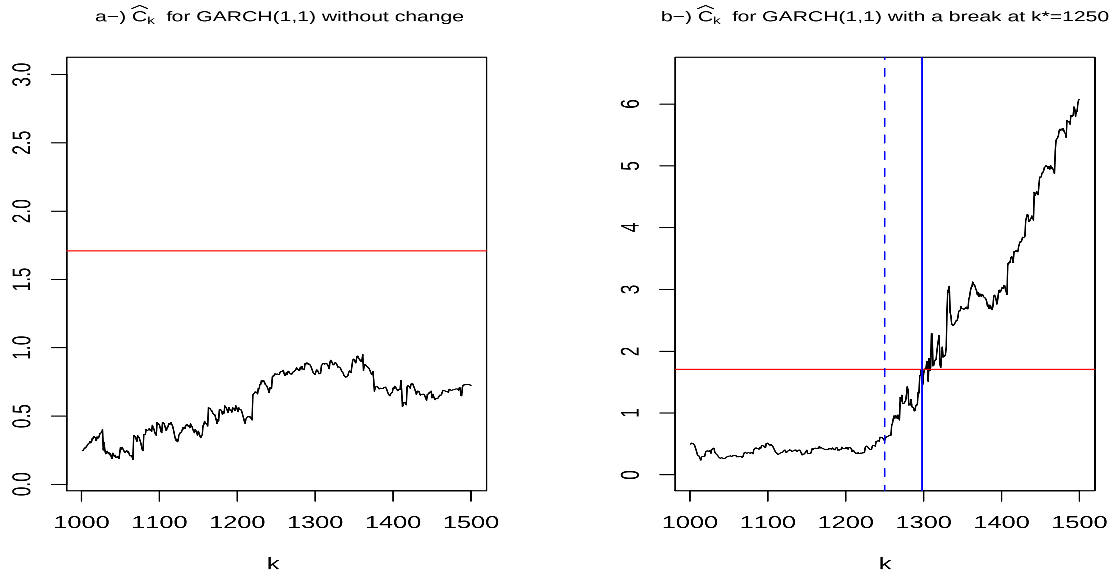


Figure 1: Typical realization of the statistics $\widehat{C}_k$ for GARCH(1,1), $n = 1000$ and $k = 1001, \cdots, 1500$. a-) The parameter $\theta_0^* = (0.01, 0.3, 0.2)$ is constant; b-) the parameter $\theta_0^* = (0.01, 0.3, 0.2)$ changes to $\theta_1^* = (0.03, 0.5, 0.2)$ at $k^* = 1250$. The horizontal solid line represents the limit of the critical region, the vertical dotted line indicates where the change occurs and the vertical solid line indicates the time where the monitoring procedure detecting the change.

Figure 1 exhibits a typical realization of the statistic $(\widehat{C}_k)_{1001 \leq k \leq 1500}$. We consider a scenario without change (see Figure 1 a-)) and a scenario with a change at $k^* = 1250$ (see Figure 1 b-)). Figure 1 a-) shows that the detector $\widehat{C}_k$ is under the horizontal line which represents the limit of the critical region. On the Figure 1 b-), we can see that before change occurs, $\widehat{C}_k$ is under the horizontal line and increases with a high speed after change. Such growth over a long period indicates that something happening in the model.

### 4.3 Monitoring mean shift in times series

Let $(X_1, \cdots, X_n)$ be an (historical) observation of a process $X = (X_t)_{t \in \mathbf{Z}}$. We assume that $X$ satisfy

$$\begin{cases} X_t = \mu_0 + \epsilon_t & \text{for} \quad 1 \leq t \leq k^* \\ X_t = \mu_1 + \epsilon_t & \text{for} \quad t > k^* \end{cases}$$

with $k^* > n$, $\mu_0 \neq \mu_1$ and $(\epsilon_t)$ a zero mean stationary uncorrelated and conditional heteroscedasticity time series belonging to a class $\mathcal{M}_{\mathbf{Z}}(f_\theta, M_\theta)$. We assume that $(\epsilon_t)$ is square integrable with variance $\sigma^2 > 0$. The monitoring procedure starts at $k = n + 1$ and the aim is to test mean shift over the new observations $X_{n+1}, X_{n+2}, \cdots, X_k$ with $k \leq [Tn] + 1$. We compare our procedure to the one based on the detector $\widehat{Q}_k$ (with $\gamma = 0$) defined first in [21] and achieved in Aue *et al.* [2].

**Remark 4.1.** *In a more general linear regression, Aue et al. [2] (see also [21]) propose a procedure based on* *the partial sums of residuals; they used the detector define at a monitoring instant $k$ by*

$$\widehat{Q}_k = \frac{1}{c\sqrt{n}\frac{k}{n}(1-\frac{n}{k})^{\gamma}} \Big| \sum_{\ell=n+1}^{k} \widehat{\epsilon}_\ell \Big|$$

*where $c$ is a positive constant, $0 \leq \gamma < 1/2$ and $\widehat{\epsilon}_\ell$ is the residual computed at time $\ell$. The procedure is stopped (and rejects $H_0$) at the first time $k > n$ such that $\widehat{Q}_k > 1$. The stopping time is defined for some $T > 1$ by*

$$\widetilde{\tau}(n) := Inf\big\{n < k < [Tn] + 1 \ / \ \widehat{Q}_k > 1\big\}.$$

*They have shown that, under $H_0$ in the open-end procedure ($T = \infty$),*

$$\lim_{n\to\infty} P\big\{\widetilde{\tau}(n) < \infty\big\} = \lim_{n\to\infty} P\big\{\frac{1}{\sigma}\sup_{k>n} \widehat{Q}_k > 1\big\} = P\Big\{\sup_{0<t<1} \frac{|W_1(t)|}{t^{\gamma}} > c\Big\}$$

*where $W_1$ is the standard Brownian motion. By using the proof of this results and the proof of Theorem 2.1 of [21], one can easily show that under $H_0$ in the closed-end procedure ($T < \infty$),*

$$\lim_{n\to\infty} P\{\widetilde{\tau}(n) < \infty\} = \lim_{n\to\infty} P\big\{\frac{1}{\sigma}\sup_{n<k<[Tn]+1} \widehat{Q}_k > 1\big\} = P\Big\{\sup_{0<t<T/(T+1)} \frac{|W_1(t)|}{t^{\gamma}} > c\Big\}.$$

*We will deal with $\gamma = 0$; one can easily see that*

$$\sup_{0<t<T/(T+1)} |W_1(t)| \stackrel{\mathcal{D}}{=} \sqrt{\frac{T}{T+1}} \sup_{0<t<1} |W_1(t)|.$$

*For $d \geq 1$, the quantiles of $\sqrt{T/(T+1)}\sup_{0<t<1}\|W_d(t)\|$ are known; they can also be computed through Monte-Carlo simulations; see Table 2 where $c_\alpha$ is defined by $P\big\{\sqrt{T/(T+1)}\sup_{0<t<1}\|W_d(t)\| \leq c_\alpha\big\} = 1-\alpha$.*

| | $d=1$ | $d=2$ | $d=3$ | $d=4$ | $d=5$ |
|---|---|---|---|---|---|
| $c_\alpha$ | 1.701 | 1.985 | 2.296 | 2.487 | 2.709 |

Table 2: The $(1-\alpha)$-quantiles $c_\alpha$ of $\sqrt{T/(T+1)}\sup_{0<t<1}\|W_d(t)\|$ with $\alpha = 0.05$ and $T = 1.5$.

We consider the following scenario

$(\epsilon_t)$ is a GARCH(1,1) process; $\epsilon_t = \sigma_t \xi_t$ with $\sigma_t^2 = \alpha_0 + \alpha_1 \epsilon_{t-1}^2 + \beta_1 \sigma_{t-1}^2$ and $(\alpha_0, \alpha_1, \beta_1) = (0.01, 0.3, 0.2)$.

The monitoring horizon is $\{n+1, \cdots, [Tn] = 1.5 \times n\}$ with $n = 500$ and $n = 1000$; under H$_1$, the change occurs at time $k^* = [1.25 \times n]$. Tables 3 indicates the empirical levels and the empirical powers based on 200 replications. The elementary statistics of the empirical detection delay are reported in Table 4.

| | | Detector | $n=500$ | $n=1000$ |
|---|---|---|---|---|
| Empirical levels : | $\mu_0 = \mu_1 = 0$ | $\widehat{C}_k$ | 0.075 | 0.065 |
| | | $\widehat{Q}_k$ | 0.005 | 0.010 |
| Empirical powers : | $\mu_0 = 0;\ \mu_1 = 0.07$ | $\widehat{C}_k$ | 0.945 | 0.995 |
| | | $\widehat{Q}_k$ | 0.635 | 0.940 |

Table 3: Empirical levels and powers for monitoring means shift in GARCH(1,1) with $(\alpha_0^*, \alpha_1^*, \beta_1^*) = (0.01, 0.3, 0.2)$. The empirical levels are computed when $\mu_0 = 0$ and the empirical powers are computed when the mean $\mu_0 = 0$ changes to $\mu_1 = 0.07$.

| $\widehat{d}_n$ | | Mean | SD | Min | $Q_1$ | Med | $Q_3$ | Max |
|---|---|---|---|---|---|---|---|---|
| $n = 500$ ; $k^* = 650$ | $\widehat{C}_k$ | 63.21 | 21.67 | 11 | 47 | 63 | 81 | 121 |
| | $\widehat{Q}_k$ | 90.29 | 25.21 | 14 | 71 | 91 | 109 | 125 |
| $n = 1000$ ; $k^* = 1300$ | $\widehat{C}_k$ | 86.18 | 28.88 | 15 | 67 | 85 | 106 | 178 |
| | $\widehat{Q}_k$ | 144.18 | 42.59 | 47 | 113 | 146 | 176 | 245 |

Table 4: Elementary statistics of the empirical detection delay for monitoring mean shift in GARCH$(1, 1)$ processes.

The results of Table 3 show that the procedure based on the detector $\widehat{Q}_k$ is too conservative. The same observations have been made by Aue *et al.* [2] when they considered monitoring change in linear regression model (in their simulation study, they also considered a scenario with null slope which is similar to the setting considered here). The empirical levels of the procedure based on the detector $\widehat{C}_k$ slightly decreases as $n$ increases for approaching the nominal level.

The empirical powers of the procedure based on the detector $\widehat{C}_k$ is better than those based on the detector $\widehat{Q}_k$. Moreover, as we mentioned above, the challenge of this problem is to minimize the detection delay. It can be seen in Table 4 that the detection delay provided by $\widehat{C}_k$ satisfied the relation $\widehat{d}_{1000} \leq \sqrt{1000/500}\ \widehat{d}_{500}$ (from Theorem 3.2, we deduced that $\widehat{d}_n = \mathcal{O}_P\big(n^{1/2}\log n\big)$ when $n$ is large enough). Moreover, our test procedure outperforms the Horváth *et al.'s* test in terms of mean and quantiles of the detection delay.

## 4.4 Monitoring parameter changes in AR$(1)$ and GARCH$(1, 1)$ processes

In this subsection, we present some simulations results for monitoring parameter changes in AR(1) and GARCH$(1, 1)$ models and compare our procedure to the one proposed by Na *et al.* [31] based the detector defined at (13). The comparisons between their procedure based on $\widehat{D}_k$ and ours based on $\widehat{C}_k$ are made in the following situations:

1. For **AR**$(1)$ **model** : $X_t = \phi_1^* X_{t-1} + \xi_t$ . Under $H_0$, $\theta_0 = \phi_1^* = 0.2$. Under $H_1$, $\theta_0$ changes to $\theta_1 = 0.6$ at $k^* = [1.25 \times n]$.
2. For **GARCH**$(1, 1)$ **model** : $X_t = \sigma_t \xi_t$ with $\sigma_t^2 = \alpha_0^* + \alpha_1^* X_{t-1}^2 + \beta_1^* \sigma_{t-1}^2$. Under $H_0$, $\theta_0 = (\alpha_0^*, \alpha_1^*, \beta_1^*) = (0.01, 0.3, 0.2)$, while under $H_1$, $\theta_0 = (0.01, 0.3, 0.2)$ changes to $\theta_1 = (0.03, 0.5, 0.2)$ at $k^* = [1.25 \times n]$.

Tables 5 indicates the empirical levels and the empirical powers based on 200 replications. The elementary statistics of the empirical detection delays are reported in Tables 6.

We considered AR and GARCH processes with zero mean. Contrary to the mean shift studied above, this mean is not estimated. For both AR and GARCH models, it appears in Table 5 that the procedure based on detector $\widehat{D}_k$ is conservative.

Unlike Na *et al.* [31], we consider a scenario of GARCH model with moderate change in parameter. One can see in Table 5 and 6 that the procedure based on the detector $\widehat{C}_k$ is uniformly better (in term of power and detection delay) than the one based on $\widehat{D}_k$.

| | | | Detector | $n=500$ | $n=1000$ |
|---|---|---|---|---|---|
| Empirical levels | AR(1) | $\theta_0=0.2$ | $\widehat{C}_k$ | 0.065 | 0.055 |
| | | | $\widehat{Q}_k$ | 0.015 | 0.010 |
| | GARCH(1,1) | $\theta_0=(0.01,0.3,0.2)$ | $\widehat{C}_k$ | 0.110 | 0.075 |
| | | | $\widehat{Q}_k$ | 0.040 | 0.015 |
| Empirical powers | AR(1) | $\theta_0=0.2;\ \theta_1=0.6$ | $\widehat{C}_k$ | 0.895 | 0.995 |
| | | | $\widehat{Q}_k$ | 0.665 | 0.940 |
| | GARCH(1,1) | $\theta_0=(0.01,0.3,0.2);\ \theta_1=(0.03,0.5,0.2)$ | $\widehat{C}_k$ | 0.790 | 0.965 |
| | | | $\widehat{Q}_k$ | 0.645 | 0.825 |

Table 5: Empirical levels and powers for monitoring parameter change in AR(1) and GARCH(1,1) processes. The empirical levels are computed when $\theta_0=\phi_1^*=0.2$ and $\theta_0=(0.01,0.3,0.2)$ for AR(1) and GARCH(1,1) respectively. Under the alternative, change occurs at $k^*=[1.25\times n]$ and the empirical powers are computed when $\theta_0$ change to $\theta_1=0.6$ for AR(1) process and $\theta_1=(0.03,0.5,0.2)$ for GARCH(1,1) process.

| $\widehat{d}_n$ | | | Mean | SD | Min | $Q_1$ | Med | $Q_3$ | Max |
|---|---|---|---|---|---|---|---|---|---|
| AR(1) | $n=500$ ; $k^*=625$ | $\widehat{C}_k$ | 68.85 | 28.40 | 4 | 49 | 69 | 91 | 123 |
| | | $\widehat{D}_k$ | 78.20 | 28.98 | 6 | 56 | 84 | 99 | 124 |
| | $n=1000$ ; $k^*=1250$ | $\widehat{C}_k$ | 96.81 | 42.41 | 12 | 68 | 95 | 128 | 212 |
| | | $\widehat{D}_k$ | 131.15 | 54.32 | 25 | 89 | 127 | 164 | 248 |
| GARCH(1,1) | $n=500$ ; $k^*=625$ | $\widehat{C}_k$ | 54.66 | 28.07 | 8 | 34 | 53 | 87 | 122 |
| | | $\widehat{D}_k$ | 70.75 | 27.73 | 12 | 47 | 70 | 96 | 125 |
| | $n=1000$ ; $k^*=1250$ | $\widehat{C}_k$ | 83.24 | 39.66 | 11 | 54 | 88 | 108 | 243 |
| | | $\widehat{D}_k$ | 97.70 | 44.56 | 15 | 65 | 91 | 122 | 236 |

Table 6: Elementary statistics of the empirical detection delay for monitoring parameter change in AR(1) and GARCH(1,1).

# 5 Real-Data Applications

We consider the returns of the daily closing values of the FTSE 100 stock index (the share index of the 100 most highly capitalized UK companies listed on the London Stock Exchange) from 2 January 2004 to 30 April 2013. These data are available on Yahoo! Finance at `http://finance.yahoo.com/`. They are represented on Figure 2 . The returns of this series are known to represent ARCH effect and GARCH$(1, 1)$ can be used to capture it, see the book of Francq and Zakoïan [18] for more details.

Recently, the (off-line) multiple change detection has retained attention of many researchers (see for instance [24], [16], [5], [19]) and the changes of FTSE 100 time series have been studied for instance in [5] and [19]. However, an off-line procedure is extremely numerically time consuming when the number of changes is unknown. We propose the monitoring approach as an alternative procedure of the off-line multiple change procedures. Assume that the all available data are $X_1, \cdots, X_N$. The procedure works as follows:

1. Choice the historical observations at the first part of the data. This can be done by applying a retrospective test (see for instance [25]) on all the observations; if change is detected, then divided the observations in two parts and applied the retrospective test on the first part. Repeat this procedure until obtained the historical data. Let $X_1, \cdots, X_n$ be the obtained historical data.

2. Start monitoring from $n + 1$; when monitoring stops (for example at $\widetilde{t}_1$), applied the retrospective test on $X_1, \cdots, X_{\widetilde{t}_1}$. This will detected and estimated the first break in the observations (for example $\widehat{t}_1$).

3. Update the historical data (by taking the observation $X_{\widehat{t}_1}, \cdots, X_{\widetilde{t}_1}$) and go back to the previous step. Repeat the procedure until the monitoring arrives to the last observation $X_N$.

A future work will be devoted to the efficiency of such procedure and a comparison to the existing ones.

Consider the observations from 2 January 2004 to 30 June2006 as historical data. This period is known to be stable in the financial community; a retrospective test (proposed in [25]) has not detected any change on these data. Hence, the monitoring started at 3 July 2006; Figure 2 gives the realizations of the statistics $\widehat{C}_k$.
The retrospective test is applied on the data going from 2 January 2004 to 5 September 2007 and the first break is detected at 4 June 2007. We have continued the procedure and the following results are obtained

- $\widehat{t}_1 \simeq$ 4 June 2007 corresponding to the period just before the beginning of the Subprime Crisis in US;

- $\widehat{t}_2 \simeq$ 6 September 2008 corresponding to the Lehman Brothers Bankruptcy;

- $\widehat{t}_3 \simeq$ 2 January 2009 corresponding to the worldwide governments intervention to solve the financial crisis;

- $\widehat{t}_4 \simeq$ 30 June 2009, $\widehat{t}_5 \simeq$ 15 July 2011, $\widehat{t}_6 \simeq$ 3 January 2012. These breaks indicates the turmoils periods in the $2010 - 2012$ Greece and European debt crisis.

See also Figure 2. The breaks point $\widehat{t}_1$, $\widehat{t}_2$ and $\widehat{t}_3$ are closed to those obtained by Bardet *et al.* [5] (by using penalized quasi-likelihood)and Fryzlewicz and Subba Rao [19] (by using BASTA algorithm).

# 6 Proofs of main results

Let us prove first some useful lemmas. 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.

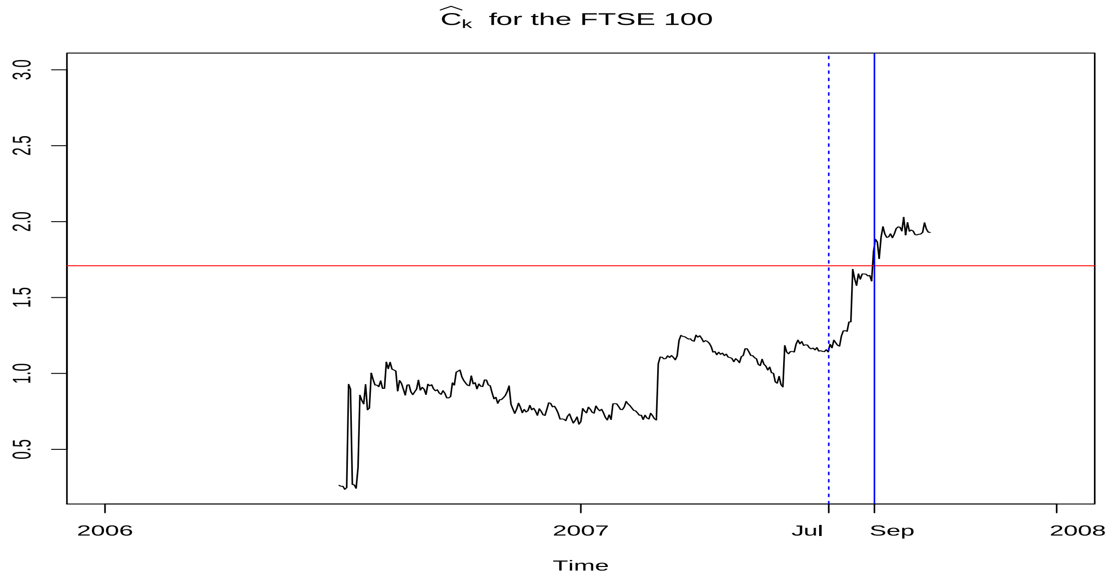


Figure 2: Realizations of the statistics $\widehat{C}_k$ with $k$ going from 3 July 2006 to 3 October 2007 ; the historical data considered are the time series going from 2 January 2004 to 30 June 2006. The horizontal solid line represents the limit of the critical region, the vertical dotted line indicates the date of the beginning of the Subprime Crisis in US and the vertical solid line indicates the time where the monitoring procedure stopped.

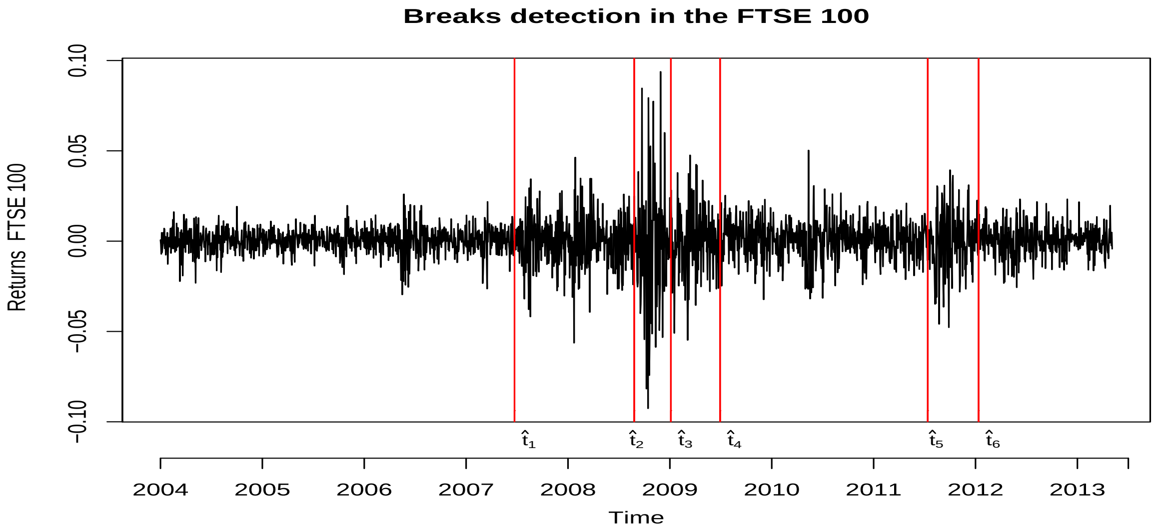


Figure 3: Break detection in the returns of FTSE 100 data using monitoring procedure based on $\widehat{C}_k$. The verticals lines indicate the dates when breaks have been detected.

Recall that $(X_1, \cdots, X_n)$ is an observed trajectory of a process $\mathcal{M}_{\mathbf{Z}}(M_{\theta_0^*}, f_{\theta_0^*})$.
Let $k > n$ and $T_{1,n} = \{1, \cdots, n\}$, $T_{\ell,k} = \{\ell, \ell+1, \cdots, k\}$ with $\ell \in \Pi_{n,k} = \{v_n, v_n+1, \cdots, k - v_n\}$, and define

$$C_{k,\ell} := \sqrt{n}\, \frac{k-\ell}{k} \big\| G^{-1/2} F \cdot \big(\widehat{\theta}(T_{\ell,k}) - \widehat{\theta}(T_{1,n})\big)\big\|,$$

with $\widehat{\theta}$ defined in (7).

**Lemma 6.1.** *Under the assumptions of Theorem 3.1,*

$$\sup_{k>n} \max_{\ell \in \Pi_{n,k}} \frac{1}{b((k-\ell)/n)} \big| \widehat{C}_{k,\ell} - C_{k,\ell} \big| = o_P(1) \quad \textit{as } n \to \infty.$$

**Proof.** For any $n \geq 1$, we have

$$\sup_{k>n} \max_{\ell \in \Pi_{n,k}} \frac{1}{b((k-\ell)/n)} \big| \widehat{C}_{k,\ell} - C_{k,\ell} \big| = \frac{1}{\inf_{s>0} b(s)} \sup_{k>n} \max_{\ell \in \Pi_{n,k}} \big| \widehat{C}_{k,\ell} - C_{k,\ell} \big|.$$

Now, proceed similarly as in the proof of Lemma 3 of [25]. ∎

**Lemma 6.2.** *Under the assumptions of Theorem 3.1*

$$\sup_{k>n} \max_{\ell \in \Pi_{n,k}} \frac{1}{b((k-\ell)/n)} \frac{\sqrt{n}}{k} \Big\| (k-\ell) F \left(\widehat{\theta}(T_{\ell,k}) - \widehat{\theta}(T_{1,n})\right) - 2\Big(\frac{\partial}{\partial\theta} L(T_{\ell,k}, \theta_0^*) - \frac{k-\ell}{n} \frac{\partial}{\partial\theta} L(T_{1,n}, \theta_0^*)\Big) \Big\| = o_P(1) \quad \textit{as } n \to \infty.$$

**Proof.** Let $k \geq n$ and $\mathcal{T} \subset \{1, \cdots, k\}$. By applying the Taylor expansion to the coordinates of $\partial \widehat{L}(\mathcal{T}, \cdot)/\partial\theta$, and using the fact that $\partial \widehat{L}(\mathcal{T}, \widehat{\theta}(\mathcal{T}))/\partial\theta = 0$ we have

$$\frac{2}{\mathrm{Card}(\mathcal{T})} \frac{\partial}{\partial\theta} \widehat{L}(\mathcal{T}, \theta_0^*) = \widetilde{F}(T) \cdot \big(\widehat{\theta}(\mathcal{T}) - \theta_0^*\big) \quad \text{where} \quad \widetilde{F}(\mathcal{T}) = -2\Big(\frac{1}{\mathrm{Card}(\mathcal{T})} \frac{\partial^2 \widehat{L}(\mathcal{T}, \widetilde{\theta}_i(\mathcal{T}))}{\partial\theta\partial\theta_i}\Big)_{1 \leq i \leq d}$$

for some $\widetilde{\theta}_i(\mathcal{T})$ between $\widehat{\theta}(\mathcal{T})$ and $\theta_0^*$.
Hence for any $\ell \in \Pi_{n,k}$

$$F\left(\widehat{\theta}(T_{\ell,k}) - \theta_0^*\right) = \frac{2}{k-\ell} \frac{\partial}{\partial\theta} L(T_{\ell,k}, \theta_0^*) + \big(F - \widetilde{F}(T_{\ell,k})\big)\big(\widehat{\theta}(T_{\ell,k}) - \theta_0^*\big) + \frac{2}{k-\ell}\Big(\frac{\partial}{\partial\theta} \widehat{L}(T_{\ell,k}, \theta_0^*) - \frac{\partial}{\partial\theta} L(T_{\ell,k}, \theta_0^*)\Big).$$

and

$$F\left(\widehat{\theta}(T_{1,n}) - \theta_0^*\right) = \frac{2}{n} \frac{\partial}{\partial\theta} L(T_{1,n}, \theta_0^*) + \big(F - \widetilde{F}(T_{1,n})\big)\big(\widehat{\theta}(T_{1,n}) - \theta_0^*\big) + \frac{2}{n}\Big(\frac{\partial}{\partial\theta} \widehat{L}(T_{1,n}, \theta_0^*) - \frac{\partial}{\partial\theta} L(T_{1,n}, \theta_0^*)\Big).$$

Therefore, for any $\ell \in \Pi_{n,k}$

$$\begin{aligned} &\frac{\sqrt{n}}{k}\Big((k-\ell) F\left(\widehat{\theta}(T_{\ell,k}) - \widehat{\theta}(T_{1,n})\right) - 2\Big(\frac{\partial}{\partial\theta} L(T_{\ell,k}, \theta_0^*) - \frac{k-\ell}{n} \frac{\partial}{\partial\theta} L(T_{1,n}, \theta_0^*)\Big)\Big) \\ &\quad = \sqrt{n}\frac{k-\ell}{k}\big(F - \widetilde{F}(T_{\ell,k})\big)\big(\widehat{\theta}(T_{\ell,k}) - \theta_0^*\big) + 2\frac{\sqrt{n}}{k}\Big(\frac{\partial}{\partial\theta} \widehat{L}(T_{\ell,k}, \theta_0^*) - \frac{\partial}{\partial\theta} L(T_{\ell,k}, \theta_0^*)\Big) \\ &\quad\quad - \sqrt{n}\frac{k-\ell}{k}\big(F - \widetilde{F}(T_{1,n})\big)\big(\widehat{\theta}(T_{1,n}) - \theta_0^*\big) - 2\frac{k-\ell}{k} \frac{1}{\sqrt{n}}\Big(\frac{\partial}{\partial\theta} \widehat{L}(T_{1,n}, \theta_0^*) - \frac{\partial}{\partial\theta} L(T_{1,n}, \theta_0^*)\Big). \end{aligned} \tag{15}$$

For $k > n$ and with some $\ell_k \in \Pi_{n,k}$, we have

$$\max_{\ell \in \Pi_{n,k}} \frac{1}{b((k-\ell)/n)} \sqrt{n}\frac{k-\ell}{k} \|F - \widetilde{F}(T_{\ell,k})\| \, \|\widehat{\theta}(T_{\ell,k}) - \theta_0^*\| \leq \frac{1}{\underset{s>0}{\mathrm{Inf}} b(s)} \sqrt{k-\ell_k} \|F - \widetilde{F}(T_{\ell_k,k})\| \, \|\widehat{\theta}(T_{\ell_k,k}) - \theta_0^*\|.$$

According to [4] and [5], $\|F - \widetilde{F}(T_{\ell_k,k})\| = o_P(1)$ and $\|\widehat{\theta}(T_{\ell_k,k}) - \theta_0^*\| = O_P(1/\sqrt{k-\ell_k})$ as $k - \ell_k \to \infty$. Hence

$$\sup_{k>n} \max_{\ell\in\Pi_{n,k}} \frac{1}{b((k-\ell)/n)} \sqrt{n}\frac{k-\ell}{k} \|F - \widetilde{F}(T_{\ell,k})\| \, \|\widehat{\theta}(T_{\ell,k}) - \theta_0^*\| = o_P(1) \quad \text{as} \quad n \to \infty. \tag{16}$$

Similar arguments imply that

$$\sup_{k>n} \max_{\ell\in\Pi_{n,k}} \frac{1}{b((k-\ell)/n)} \sqrt{n}\frac{k-\ell}{k} \|F - \widetilde{F}(T_{1,n})\| \, \|\widehat{\theta}(T_{1,n}) - \theta_0^*\| = o_P(1) \quad \text{as} \quad n \to \infty. \tag{17}$$

For $k > n$ and for some $\ell_k \in \Pi_{n,k}$, we have

$$\max_{\ell\in\Pi_{n,k}} \frac{1}{b((k-\ell)/n)} \frac{\sqrt{n}}{k} \|\frac{\partial}{\partial\theta}\widehat{L}(T_{\ell,k},\theta_0^*) - \frac{\partial}{\partial\theta}L(T_{\ell,k},\theta_0^*)\|$$
$$\leq \frac{1}{\operatorname{Inf}_{s>0} b(s)} \frac{1}{\sqrt{k-\ell_k}} \|\frac{\partial}{\partial\theta}\widehat{L}(T_{\ell_k,k},\theta_0^*) - \frac{\partial}{\partial\theta}L(T_{\ell_k,k},\theta_0^*)\|.$$

According to [4], $\frac{1}{\sqrt{k-\ell_k}} \|\frac{\partial}{\partial\theta}\widehat{L}(T_{\ell_k,k},\cdot) - \frac{\partial}{\partial\theta}L(T_{\ell_k,k},\cdot)\|_\Theta = o_P(1)$ as $k - \ell_k \to \infty$. Hence

$$\sup_{k>n} \max_{\ell\in\Pi_{n,k}} \frac{1}{b((k-\ell)/n)} \frac{\sqrt{n}}{k} \|\frac{\partial}{\partial\theta}\widehat{L}(T_{\ell,k},\theta_0^*) - \frac{\partial}{\partial\theta}L(T_{\ell,k},\theta_0^*)\| = o_P(1) \quad \text{as} \quad n \to \infty. \tag{18}$$

Similar arguments show that

$$\sup_{k>n} \max_{\ell\in\Pi_{n,k}} \frac{1}{b((k-\ell)/n)} \frac{k-\ell}{k} \frac{1}{\sqrt{n}} \|\frac{\partial}{\partial\theta}\widehat{L}(T_{1,n},\theta_0^*) - \frac{\partial}{\partial\theta}L(T_{1,n},\theta_0^*)\| = o_P(1) \quad \text{as} \quad n \to \infty. \tag{19}$$

Thus, Lemma 6.2 follows from (15), (16), (17), (18) and (19). ■

**Lemma 6.3.** *Under the assumptions of Theorem 3.1, with $W_G$ a d-dimensional Gaussian centered process such as $E\left(W_G(s)W_G(\tau)'\right) = \min(s,\tau)G$,*

- *(i) For any $T > 1$ ,*

$$\sup_{n<k[Tn]+1} \max_{\ell\in\Pi_{n,k}} \frac{1}{b((k-\ell)/n)} \sqrt{n}\frac{k-\ell}{k} \|F\left(\widehat{\theta}(T_{\ell,k}) - \widehat{\theta}(T_{1,n})\right)\| \xrightarrow[n\to\infty]{\mathcal{D}} \sup_{1<t<T} \sup_{0<s<t} \frac{\|W_G(s) - sW_G(1)\|}{t\ b(s)} \ ;$$

- *(ii) When $T = \infty$,*

$$\sup_{k>n} \max_{\ell\in\Pi_{n,k}} \frac{1}{b((k-\ell)/n)} \sqrt{n}\frac{k-\ell}{k} \|F\left(\widehat{\theta}(T_{\ell,k}) - \widehat{\theta}(T_{1,n})\right)\| \xrightarrow[n\to\infty]{\mathcal{D}} \sup_{t>1} \sup_{0<s<t} \frac{\|W_G(s) - sW_G(1)\|}{t\ b(s)}.$$

**Proof.**

We are going to apply Lemma 6.2 for specifying the asymptotic behaviour of $\widehat{\theta}(T_{\ell,k}) - \widehat{\theta}(T_{1,n})$.

For $k > n$ and $\ell \in \Pi_{n,k}$, we have

$$2\frac{\sqrt{n}}{k}\left(\frac{\partial}{\partial\theta}L(T_{\ell,k},\theta_0^*) - \frac{k-\ell}{n}\frac{\partial}{\partial\theta}L(T_{1,n},\theta_0^*)\right) = -\frac{n}{k}\frac{1}{\sqrt{n}}\left(\sum_{i=\ell+1}^{k} \frac{\partial q_i(\theta_0^*)}{\partial\theta} - \frac{k-\ell}{n}\sum_{i=1}^{n} \frac{\partial q_i(\theta_0^*)}{\partial\theta}\right).$$

(i) Let $T > 1$. We have

$$\max_{n<k<nT} \max_{\ell\in\Pi_{n,k}} \frac{1}{b((k-\ell)/n)} \frac{2\sqrt{n}}{k} \|\frac{\partial}{\partial\theta}L(T_{\ell,k},\theta_0^*) - \frac{k-\ell}{n}\frac{\partial}{\partial\theta}L(T_{1,n},\theta_0^*)\|$$
$$= \max_{n<k<nT} \max_{\ell\in\Pi_{n,k}} \frac{1}{b((k-\ell)/n)} \frac{\sqrt{n}}{k} \|\sum_{i=\ell}^{k} \frac{\partial q_i(\theta_0^*)}{\partial\theta} - \frac{k-\ell}{n}\sum_{i=1}^{n} \frac{\partial q_i(\theta_0^*)}{\partial\theta}\|$$
$$= \max_{t\in\{1,1+\frac{1}{n},\cdots,T\}} \max_{s\in\{1-\frac{v_n}{n},2-\frac{v_n}{n},\cdots,t-\frac{v_n}{n}\}} \frac{1}{b(([nt]-[ns])/n)} \frac{n}{[nt]} \|\frac{1}{\sqrt{n}}\left(\sum_{i=[ns]+1}^{[nt]} \frac{\partial q_i(\theta_0^*)}{\partial\theta} - \frac{[nt]-[ns]}{n}\sum_{i=1}^{n} \frac{\partial q_i(\theta_0^*)}{\partial\theta}\right)\|.$$

Define the set $S := \{(t,s) \in [1,T] \times [1,T]/\ s < t\}$. According to [4], $\big(\dfrac{\partial q_i(\theta_0^*)}{\partial\theta}\big)_{t\in\mathbf{Z}}$ is a stationary ergodic martingale difference sequence with covariance matrix $G$. By Cramér-Wold device (see [11] p. 206), it holds that

$$\frac{1}{\sqrt{n}} \sum_{i=[ns]+1}^{[nt]} \frac{\partial q_i(\theta_0^*)}{\partial\theta} \ \underset{n\to\infty}{\overset{\mathcal{D}(S)}{\longrightarrow}} \ W_G(t-s).$$

with $\underset{n\to\infty}{\overset{\mathcal{D}(S)}{\longrightarrow}}$ means the weak convergence on the Skorohod space $\mathcal{D}(S)$. Hence

$$\frac{1}{\sqrt{n}}\Big( \sum_{i=[ns]+1}^{[nt]} \frac{\partial q_i(\theta_0^*)}{\partial\theta} - \frac{[nt]-[ns]}{n} \sum_{i=1}^{n} \frac{\partial q_i(\theta_0^*)}{\partial\theta}\Big) \underset{n\to\infty}{\overset{\mathcal{D}(S)}{\longrightarrow}} \ W_G(t-s) - (t-s)W_G(1).$$

Therefore

$$\begin{aligned} \max_{n<k<nT} \max_{\ell\in\Pi_{n,k}} \frac{1}{b((k-\ell)/n)} \frac{2\sqrt{n}}{k} \Big\| \frac{\partial}{\partial\theta} L(T_{\ell,k},\theta_0^*) - \frac{k-\ell}{n}\frac{\partial}{\partial\theta} L(T_{1,n},\theta_0^*)\Big\| & \\ \underset{n\to\infty}{\overset{\mathcal{D}}{\longrightarrow}} \ \sup_{1<t<T}\ \sup_{1<s<t} \frac{\|W_G(t-s)-(t-s)W_G(1)\|}{t\ b(t-s)} & \\ \underset{n\to\infty}{\overset{\mathcal{D}}{\longrightarrow}} \ \sup_{1<t<T}\ \sup_{1<s<t} \frac{\|W_G(s)-s\,W_G(1)\|}{t\ b(s)}. & \qquad (20) \end{aligned}$$

(ii) According to (i), it suffices to show that the limit distribution (as $n,T\to\infty$) of

$$\sup_{k>nT} \max_{\ell\in\Pi_{n,k}} \frac{1}{b((k-\ell)/n)} \frac{2\sqrt{n}}{k} \Big\| \frac{\partial}{\partial\theta} L(T_{\ell,k},\theta_0^*) - \frac{k-\ell}{n}\frac{\partial}{\partial\theta} L(T_{1,n},\theta_0^*)\Big\|$$

exists and is equal to the limit distribution (as $T\to\infty$) of

$$\sup_{t>T}\ \sup_{1<s<t} \frac{\|W_G(s)-s\,W_G(1)\|}{t\ b(s)}.$$

Let $k > nT$. We have

$$\max_{\ell\in\Pi_{n,k}} \frac{1}{b((k-\ell)/n)} \frac{\sqrt{n}}{k} \Big\| \frac{\partial}{\partial\theta} L(T_{\ell,k},\theta_0^*)\Big\| \leq \frac{1}{\operatorname*{Inf}_{s>0} b(s)} \frac{\sqrt{n}}{k} \Big\| \sum_{i=\ell_k+1}^{k} \frac{\partial q_i(\theta_0^*)}{\partial\theta}\Big\| \ \text{ for some } \ \ell_k \in \Pi_{n,k}.$$

It comes from the Hájek-Rényi-Chow inequality (see [14]) that, for any $\varepsilon > 0$

$$\lim_{T\to\infty} \lim_{n\to\infty} P\Big( \sup_{k>nT} \frac{\sqrt{n}}{k} \Big\| \sum_{i=\ell_k+1}^{k} \frac{\partial q_i(\theta_0^*)}{\partial\theta}\Big\| > \varepsilon\Big) = 0.$$

Hence

$$\sup_{k>nT} \max_{\ell\in\Pi_{n,k}} \frac{1}{b((k-\ell)/n)} \frac{2\sqrt{n}}{k} \Big\| \frac{\partial}{\partial\theta} L(T_{\ell,k},\theta_0^*)\Big\| = o_P(1) \ \text{ as } \ T,n\to\infty. \qquad (21)$$

Moreover, since the function $b(\cdot)$ is non-increasing, for any $n,T>1$, we have:

$$\begin{aligned} \sup_{k>nT} \max_{\ell\in\Pi_{n,k}} \frac{1}{b((k-\ell)/n)} \frac{2\sqrt{n}}{k} \Big\| \frac{k-\ell}{n} \frac{\partial}{\partial\theta} L(T_{1,n},\theta_0^*)\Big\| &= \Big\|\frac{1}{\sqrt{n}} \sum_{i=1}^{n} \frac{\partial q_i(\theta_0^*)}{\partial\theta}\Big\| \times \sup_{k>nT} \max_{\ell\in\Pi_{n,k}} \frac{1}{b((k-\ell)/n)} \frac{k-\ell}{k} \\ &= \Big\|\frac{1}{\sqrt{n}} \sum_{i=1}^{n} \frac{\partial q_i(\theta_0^*)}{\partial\theta}\Big\| \times \sup_{k>nT} \frac{1}{b((k-v_n)/n)} \frac{k-v_n}{k} \\ &= \frac{1}{\operatorname*{Inf}_{s>0} b(s)} \Big\|\frac{1}{\sqrt{n}} \sum_{i=1}^{n} \frac{\partial q_i(\theta_0^*)}{\partial\theta}\Big\| \\ &\underset{n\to\infty}{\overset{\mathcal{D}}{\longrightarrow}} \ \frac{1}{\operatorname*{Inf}_{s>0} b(s)} \|W_G(1)\|, \qquad (22) \end{aligned}$$

using again the Cramèr-Wold device. It comes from (21) and (22) that

$$\sup_{k>nT}\max_{\ell\in\Pi_{n,k}}\frac{1}{b((k-\ell)/n)}\frac{2\sqrt{n}}{k}\|\frac{\partial}{\partial\theta}L(T_{\ell,k},\theta_0^*)-\frac{k-\ell}{n}\frac{\partial}{\partial\theta}L(T_{1,n},\theta_0^*)\| \xrightarrow[T,n\to\infty]{\mathcal{D}} \frac{1}{\operatorname{Inf}_{s>0} b(s)}\|W_G(1)\|. \tag{23}$$

Furthermore, since the coordinates of $W_G$ are Brownian motions, by the law of the iterated logarithm there exists $t_0>\exp(1)$ such as

$$s>t_0\Rightarrow \|W_G(s)\|\le\sqrt{s}\log(s)\ \text{ almost surely.}$$

Thus, for any $t>t_0$, we obtain almost surely

$$\sup_{1<s<t}\|W_G(s)\|\le\sup_{1<s<t_0}\|W_G(s)\|+\sqrt{t}\log(t).$$

Therefore, for $T$ large enough, we have

$$\sup_{t>T}\sup_{1<s<t}\frac{\|W_G(s)\|}{t\,b(s)}\le\frac{1}{\inf_{s>0}b(s)}\big(\frac{1}{T}\sup_{1<s<t_0}\|W_G(s)\|+\sup_{t>T}\frac{\log(t)}{\sqrt{t}}\big)\xrightarrow[T\to\infty]{\text{a.s.}}0. \tag{24}$$

Finally, since $b(\cdot)$ is non-increasing, for any $T>1$, we have

$$\sup_{t>T}\sup_{1<s<t}\frac{\|sW_G(1)\|}{t\,b(s)}=\|W_G(1)\|\sup_{t>T}\frac{1}{t}\sup_{1<s<t}\frac{s}{b(t)}=\|W_G(1)\|\sup_{t>T}\frac{1}{b(t)}=\frac{1}{\operatorname{Inf}_{s>0} b(s)}\|W_G(1)\|. \tag{25}$$

It comes from (24) and (25) that the limit of (20) satisfies when $T\to\infty$,

$$\sup_{t>T}\sup_{1<s<t}\frac{\|W_G(s)-sW_G(1)\|}{t\,b(s)}\xrightarrow[T\to\infty]{\mathcal{D}}\frac{1}{\operatorname{Inf}_{s>0} b(s)}\|W_G(1)\|. \tag{26}$$

From the relations (20), (23) and (26), it comes that

$$\sup_{k>nT}\max_{\ell\in\Pi_{n,k}}\frac{1}{b((k-\ell)/n)}\frac{2\sqrt{n}}{k}\|\frac{\partial}{\partial\theta}L(T_{\ell,k},\theta_0^*)-\frac{k-\ell}{n}\frac{\partial}{\partial\theta}L(T_{1,n},\theta_0^*)\|\xrightarrow[n\to\infty]{\mathcal{D}}\sup_{t>T}\sup_{1<s<t}\frac{\|W_G(s)-sW_G(1)\|}{t\,b(s)}.$$

Hence, (ii) follows according to (i) and Lemma 6.2. ■

**Proof of Theorem 3.1**

(i) Assume that $T=\infty$. We know that

$$\begin{aligned}P\{\tau(n)<\infty\}&=P\Big\{\sup_{k>n}\max_{\ell\in\Pi_{n,k}}\frac{\widehat{C}_{k,\ell}}{b((k-\ell)/n)}>1\Big\}\\&=P\Big\{\sup_{k>n}\max_{\ell\in\Pi_{n,k}}\frac{1}{b((k-\ell)/n)}\sqrt{n}\frac{k-\ell}{k}\|\widehat{G}(T_{1,n})^{-1/2}\widehat{F}(T_{1,n})\cdot\big(\widehat{\theta}(T_{\ell,k})-\widehat{\theta}(T_{1,n})\big)\|>1\Big\}.\end{aligned}$$

Since $\widehat{G}(T_{1,n})\xrightarrow[n\to\infty]{a.s.}G$ and $\widehat{F}(T_{1,n})\xrightarrow[n\to\infty]{a.s.}F$, it comes from Lemma 6.1 and 6.3 that

$$\begin{aligned}&\sup_{k>n}\max_{\ell\in\Pi_{n,k}}\frac{1}{b((k-\ell)/n)}\sqrt{n}\frac{k-\ell}{k}\|\widehat{G}(T_{1,n})^{-1/2}\widehat{F}(T_{1,n})\cdot\big(\widehat{\theta}(T_{\ell,k})-\widehat{\theta}(T_{1,n})\big)\|\\&\qquad\xrightarrow[n\to\infty]{\mathcal{D}}\sup_{t>1}\sup_{1<s<t}\frac{\|G^{-1/2}(W_G(s)-sW_G(1))\|}{t\,b(s)}.\end{aligned}$$

Since the covariance matrix of $\{W_G(s)\ ;\ s\ge0\}$, is $\min(s,\tau)G$, the covariance matrix of $\{G^{-1/2}W_G(s)\ ;\ s\ge0\}$ is $\min(s,\tau)I_d$ (where $I_d$ is the $d$-dimensional identity matrix). Hence (i) follows.

(ii) The results of closed-end procedure ($T<\infty$) is obtained by going the same lines as in (i) and by applying Lemma 6.1 and 6.3. ■

**Proof of Theorem 3.2**

Denote $k_n = k^* + n^\delta$ for $\delta \in (1/2, 1)$. For $n$ large enough, we have $v_n < n^\delta$ and thus $k_n - v_n = k^* + n^\delta - v_n \geq k^*$. Moreover since $k^*(n) = [T^* n]$ for some $T^* < T$, for $n$ large enough and for both open-end and closed-end procedure $k_n = k^* + n^\delta < T^* n + n^\delta < Tn$. Hence, $k_n$ belongs between $n$ and $[Tn] + 1$ and $k^* \in \Pi_{n,k_n}$ for $n$ large enough. Therefore, according to assumption **B**, there exists a constant $c > 0$ such that

$$
\begin{aligned}
\max_{\ell \in \Pi_{n,k_n}} \frac{\widehat{C}_{k_n,\ell}}{b((k_n-\ell)/n)} &= \max_{\ell \in \Pi_{n,k_n}} \frac{1}{b((k_n-\ell)/n)} \sqrt{n} \frac{k_n - \ell}{k_n} \|\widehat{G}(T_{1,n})^{-1/2} \widehat{F}(T_{1,n}) \cdot \big(\widehat{\theta}_{k_n}(T_{\ell,k_n}) - \widehat{\theta}(T_{1,n})\big) \\
&\geq \frac{1}{b((k_n-k^*)/n)} \sqrt{n} \frac{k_n - k^*}{k_n} \|\widehat{G}(T_{1,n})^{-1/2} \widehat{F}(T_{1,n}) \cdot \big(\widehat{\theta}_{k_n}(T_{k^*,k_n}) - \widehat{\theta}(T_{1,n})\big) \\
&\geq c\, \sqrt{n}\, \frac{n^\delta}{k^* + n^\delta} \big\|\widehat{G}(T_{1,n})^{-1/2} \widehat{F}(T_{1,n}) \big(\widehat{\theta}_{k_n}(T_{k^*,k_n}) - \widehat{\theta}(T_{1,n})\big)\big\| \\
&\geq c\, \frac{n^{1/2+\delta}}{c_0 n + n^\delta} \big\|\widehat{G}(T_{1,n})^{-1/2} \widehat{F}(T_{1,n}) \big(\widehat{\theta}_{k_n}(T_{k^*,k_n}) - \widehat{\theta}(T_{1,n})\big)\big\| \\
&\geq c\, \frac{n^{\delta-1/2}}{(c_0+1)} \big\|\widehat{G}(T_{1,n})^{-1/2} \widehat{F}(T_{1,n}) \big(\widehat{\theta}_{k_n}(T_{k^*,k_n}) - \widehat{\theta}(T_{1,n})\big)\big\|. \qquad (27)
\end{aligned}
$$

According to [4] and [25], $\widehat{G}(T_{1,n}) \xrightarrow[n\to\infty]{a.s.} G$, $\widehat{F}(T_{1,n}) \xrightarrow[n\to\infty]{a.s.} F$, $\widehat{\theta}(T_{1,n}) \xrightarrow[n\to\infty]{a.s.} \theta_0^*$ and $\widehat{\theta}_{k_n}(T_{k^*,k_n})) \xrightarrow[n\to\infty]{a.s.} \theta_1^*$. Since $G$ is symmetric positive definite, $F$ is invertible, $\theta_0^* \neq \theta_1^*$ and $\delta > 1/2$ , then (27) implies that

$$
\max_{\ell \in \Pi_{n,k_n}} \frac{\widehat{C}_{k_n,\ell}}{b((k_n-\ell)/n)} \xrightarrow[n\to\infty]{a.s.} \infty.
$$

■

**Proof of Proposition 4.1**

- Let $T > 1$. For any $t > 1$, by computing the covariance, one can easily verify that

$$
\sup_{1<s<t} \frac{1}{t} \|W_d(s) - sW_d(1))\| \stackrel{\mathcal{D}}{=} \sup_{1<s<t} \frac{s}{t} \|W_d\big(\frac{s-1}{s}\big)\| = \sup_{0<u<1-1/t} \frac{1}{t(1-u)} \|W_d(u)\|.
$$

  Thus

$$
\begin{aligned}
\sup_{1<t<T} \sup_{1<s<t} \frac{1}{t} \|W_d(s) - sW_d(1))\| &\stackrel{\mathcal{D}}{=} \sup_{1<t<T} \sup_{0<u<1-1/t} \frac{1}{t(1-u)} \|W_d(u)\| \\
&= \sup_{1/T<\tau<1} \sup_{0<u<1-\tau} \frac{\tau}{1-u} \|W_d(u)\| \\
&= \sup_{0<v<1-1/T} \sup_{0<u<v} \frac{1-v}{1-u} \|W_d(u)\|.
\end{aligned}
$$

  But, $\|W_d(u)\| \stackrel{\mathcal{D}}{=} v^{1/2} \big\|W_d\big(\frac{u}{v}\big)\big\|$. Therefore with $u = u'v$, we have

$$
\begin{aligned}
\sup_{0<v<1-1/T} \sup_{0<u<v} \frac{1-v}{1-u} \|W_d(u)\| &\stackrel{\mathcal{D}}{=} \sup_{0<v<1-1/T} \sup_{0<u'<1} \frac{(1-v)v^{1/2}}{1-u'v} \|W_d(u')\| \\
&\stackrel{\mathcal{D}}{=} \sup_{0<u'<1} \Big( \sup_{0<v<1-1/T} \frac{(1-v)v^{1/2}}{1-u'v} \Big) \|W_d(u')\|.
\end{aligned}
$$

  Hence, take

$$
f(u,T) = \sup_{0<v<1-1/T} \frac{(1-v)v^{1/2}}{1-uv} \qquad (28)
$$

  to complete the first part of the proposition.

- If $T=\infty$, proceeding as above, we obtain

$$\sup_{t>1}\sup_{1<s<t}\frac{1}{t}\|W_d(s)-sW_d(1))\| \stackrel{\mathcal{D}}{=} \sup_{0<u'<1}\Big(\sup_{0<v<1}\frac{(1-v)v^{1/2}}{1-u'v}\Big)\|W_d(u')\|.$$

Thus, take

$$g(u)=\sup_{0<v<1}\frac{(1-v)v^{1/2}}{1-uv}. \tag{29}$$

In this case, $g$ can be easily and explicitly computed. The supremum $\sup_{0<v<1}\frac{(1-v)v^{1/2}}{1-uv}$ is obtained at $v=2\left(3-u+\sqrt{(9-u)(1-u)}\right)^{-1}$. Therefore,

$$g(u)=\frac{\sqrt{9-u}+\sqrt{1-u}}{\sqrt{9-u}+3\sqrt{1-u}}\Big(\frac{2}{3-u+\sqrt{(9-u)(1-u)}}\Big)^{1/2}.$$

This completes the proof of the Proposition 4.1. ■

**Acknowledgements.** The authors are grateful to the referees for many relevant suggestions and comments which helped to improve the contents of the paper.

# 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] Aue, A., Horváth, L., Husková, M., and Kokoszka, P. Change point monitoring in linear models. *Econometrics Journal 9*, (2006), 373-403.

[3] Aue, A., Horváth, L. and Reimherr, M. Delay times of sequential procedures for multiple time series regression models. *Journal of Econometrics 149*, (2009), 174-190.

[4] Bardet, J.-M. and Wintenberger, O. Asymptotic normality of the quasi-maximum likelihood estimator for multidimensional causal processes. *Ann. Statist. 37*, (2009), 2730-2759.

[5] Bardet, J.-M. , Kengne, W. and Wintenberger, O. Detecting multiple change-points in general causal time series using penalized quasi-likelihood. *Electronic Journal of Statistics 6*, (2012), 435-477.

[6] Bardet, J.-M. and Kengne, W. Monitoring procedure for parameter change in causal time series. *Preprint, arXiv:1209.4746.*

[7] Basseville, M. and Nikiforov, I. Detection of Abrupt Changes: Theory and Application. Prentice Hall, Englewood Cliffs, NJ, 1993.

[8] Berkes, I., Gombay, E., Horváth, L., and Kokoszka, P. Sequential change-point detection in GARCH(p,q) models. *Econometric Theory 20*, (2004), 1140-1167.

[9] Berkes, I., Horváth, L., and Kokoszka, P. GARCH processes: structure and estimation. *Bernoulli 9* (2003), 201–227.

[10] Berkes, I., Horváth, L., Kokoszka, P. and Shao, Q.-M. Almost sure convergence of the Bartlett estimator. *Period Math Hung 51* (2005), 11–25.

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

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

[13] Černíková, A., Hušková, M., Prášková, Z. and Steinebach, Josef G Delay time in monitoring jump changes in linear models. *Statistics, 47*, (2013) 1-25.

[14] Chow, Y.S. A martingale inequality and the law of large numbers. *Proceedings of the American Mathematics Society , 11*, (1960), 107-111.

[15] Chu, C-SJ., Stinchcombe, M., and White, W. Monitoring structural change. *Econometrica, 64*, (1996), 1045-1065.

[16] Davis, R. A., Lee, T. C. M. and Rodriguez-Yam, G. A. Break detection for a class of nonlinear time series models. *Journal of Time Series Analysis 29*, (2008), 834–867.

[17] Doukhan, P. and Wintenberger, O. Weakly dependent chains with infinite memory. *Stochastic Process. Appl. 118*, (2008) 1997-2013.

[18] Francq, C. and Zakoïan, J.-M. GARCH Models: Structure, Statistical Inference and Financial Applications. *Wiley, Chichester, UK* (2010).

[19] Fryzlewicz, P. and Subba Rao, S. Multiple-change-point detection for auto-regressive conditional heteroscedastic processes. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* , (2013), DOI: 10.1111/rssb.12054.

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

[21] Horváth, L., Hušková, M., Kokoszka, P. and Steinebach, J. Monitoring changes in linear models. *Journal of Statistical Planning and Inference 126*, (2004), 225-251.

[22] Hušková, M. and Kirch, C. Bootstrapping sequential change-point tests for linear regression. *Metrika 75*, (2012), 673-708.

[23] Hušková, M., Praskova, Z. and Steinebach, J. Delay time in monitoring jump changes in linear models. *ISWM-2009 Invited paper 16*, (2009).

[24] Inclan, C. and 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.

[25] Kengne W. Testing for parameter constancy in general causal time-series models. *J. Time Ser. Anal. 33*, (2012), 503-518.

[26] Kirch, C. Bootstrapping sequential change-point tests. *Sequential Analysis 27*, (2008), 330-349.

[27] Kokoszka, P. and Leipus, R. Testing for parameter changes in ARCH models. *Lithuanian Mathematical Journal 39*, (1999), 182-195.

[28] Lai, T.L. and Shan, Z. Efficient Recursive Algorithms for Detection of Abrupt Changes in Signals and Control Systems. *IEEE Transactions on Automatic Control, 44*, (1999), 952-966.

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

[30] Leisch, F., Hornik, K., and Kuan, C-H. Monitoring structural changes with the generalized fluctuation test. *Econometric Theory 16*, (2000), 835-854.

[31] Na, O., Lee, Y. and Lee, S. Monitoring parameter change in time series models. *Stat Methods Appl 20*, (2011), 171-199.

[32] Pollak, M. and Siegmund, D. Sequential detection of a change in a normal mean when the initial value is unknown. *Annals of Statistics, 19* , (1991), 394-416.

[33] Rabemananjara, R. and Zakoïan, J.M. Threshold ARCH models and asymmetries in volatility. *Journal of Applied Econometrics 8*, (1993), 31-49.

[34] Robinson, P. M. Testing for strong serial correlation and dynamic conditional heteroskedasticity in multiple regression. *J. Econometrics 47*, (1991), 67-84.

[35] Zeileis, A., Leisch F., Kleiber, C., and Hornik, K. Monitoring structural change in dynamic econometric models. *J. Appl. Econometrics 20*, (2005), 99-121.