Sequential Change–Point Detection in GARCH(p, q) Models *

István Berkes †
A. Rényi Institute of Mathematics, Hungarian Academy of Sciences
P.O. Box 127, H-1364 Budapest, Hungary

Edit Gombay ‡
Department of Mathematical Sciences, University of Alberta
Edmonton, Alberta, Canada T6G 2G1

Lajos Horváth
Department of Mathematics, University of Utah
155 South 1440 East, Salt Lake City, UT 84112-0090, USA

Piotr Kokoszka
Department of Mathematics and Statistics, Utah State University
3900 Old Main Hill, Logan, UT 84322-3900, USA

We suggest a sequential monitoring scheme to detect changes in the parameters of a GARCH(p, q) sequence. The procedure is based on quasi-likelihood scores and does not use model residuals. Unlike for linear regression models, the squared residuals of nonlinear time series models like GARCH do not satisfy a Functional Central Limit Theorem with a Wiener process as a limit, so its boundary crossing probabilities cannot be used. Our procedure nevertheless has an asymptotically controlled size and, moreover, the conditions on the boundary function are very simple; it can be chosen as a constant. We establish the asymptotic properties of our monitoring scheme under both the null of no change in parameters and the alternative of a change in parameters and investigate its finite sample behavior by means of a small simulation study.

AMS 2002 Subject Classification: Primary 62L10 Secondary 62F05
Key words and phrases: GARCH(p, q), change–point, boundary crossing, Wiener process, consistency

*Research partially supported by NSF grant INT-0223262 and NATO grant PST.CLG.977607
†Supported by the Hungarian National Foundation for Scientific Research, Grants T 29621, 37886
‡Supported by NSERC Canada

1 Introduction

The assumption that the parameters remain stable over time plays a crucial role in statistical modeling and inference. If the parameters have changed within the observed sample, then, for example, forecasts loose accuracy and the parameter estimates no longer provide meaningful information. Due to the importance of parameter stability, the detection of possible changes in the data generating process has become an active area of research. There are hundreds of papers studying various aspects of change point detection in econometric settings. We list those most closely related to the present work later in this section. For surveys concerned with the general statistical methodology for change-point detection, we refer to Besseville and Nikifirov (1993), Brodsky and Darkhovsky (1993, 2000) and Csörgő and Horváth (1997).

The present paper is motivated by the work of Chu, Stinchcombe and White (1996) who posed the following problem: "Given a previously estimated model, the arrival of new data invites the question: is yesterday's model capable of explaining today's data ?" These authors developed fluctuation and CUSUM monitoring procedures for linear regression models. Our paper focuses on GARCH models which are important in financial applications, see e.g. Gouriéroux (1997), and, as explained below, differ from linear models in ways which make a direct application of the approach of Chu, Stinchcombe and White (1996) not readily possible.

Before outlining the central ideas of Chu, Stinchcombe and White (1996) and contrasting our approach with them, we note that procedures for detecting parameter changes in a GARCH specification may have a number of useful applications. For example, risk managers use GARCH for calculating portfolio risk measures such as Value at Risk. GARCH parameters are typically estimated over a rolling window of returns and change-points in this window will introduce bias. Similarly, option traders use GARCH to make up for the well-known biases in Black-Scholes option prices. A detection of parameter changes may lead to a more cautious interpretation of the calculated option prices. While our paper is not concerned with such applied issues, we do hope that it will make a contribution to the important problem of monitoring for changes in GARCH models.

As argued convincingly in Chu, Stinchcombe and White (1996), the sequential analysis of economic and financial data is somewhat different from engineering applications. The sampling is costless under the no change null hypothesis and no action is required if the observed processes is "in control" i.e. there is no change in the parameters of the data generating process. Since it is impossible to eliminate false alarm due to chance, the probability of stopping under the no change null hypothesis should be less than a given level $0 < \alpha < 1$. On the other hand, it is desirable to stop with probability one if a change occurred. With these goals in mind, Chu, Stinchcombe and White (1996) proposed to consider decisions functions of the form

$$Q_t^{(m)} = \sum_{m+1 \le i < m+mt} \hat{\omega}_i, \quad t \in [0, \infty), \tag{1.1}$$

where m denotes the number of initial observations used to estimate the model and the $\hat{\omega}_i$ are model residuals which are estimated sequentially. The idea of their procedure is as follows: Denoting by $W(\cdot)$ the standard Wiener process, suppose it can be shown that, as $m \to \infty$,

$$m^{-1/2} Q_t^{(m)} \to W(t) \quad \text{in } D[0,\infty) \tag{1.2}$$

and that a function g can be found such that

$$\lim_{m\to\infty} P\left(|S_n| \geq \sqrt{m} g(n/m), \quad \text{for some } n \geq m\right) = P\left(|W(t)| \geq g(t) \ \text{ for some } t \geq 1\right), \tag{1.3}$$

provided the sequence S_n satisfies a Functional Central Limit Theorem. It can then be concluded from (1.2) and (1.3) that

$$\lim_{m\to\infty} P\left(|Q_t^{(m)}| \geq \sqrt{m} g(n/m), \quad \text{for some } n \geq m\right) = P\left(|W(t)| \geq g(t) \ \text{ for some } t \geq 1\right). \tag{1.4}$$

For several judicious choices of g, the probability on the right hand side of (1.4) can be computed analytically and consequently a monitoring scheme can be developed such that asymptotically, as $m \to \infty$, the decision function $|Q_t^{(m)}|$ crosses the boundary $\sqrt{m} g(n/m)$ at some future date n with a prescribed probability α, provided the parameters have not changed.

The present paper is concerned with the detection of changes in GARCH(p,q) processes which are used to model volatility. Statistics designed for detecting changes in volatility are typically based on the squares of observations or the squares of residuals. The squares of residuals from a GARCH(p,q) process do not satisfy a functional limit theorem with the Wiener process in the limit due to the presence of extra terms in their covariance structure, see e.g. Boldin (1998), Horváth, Kokoszka and Teyssière (2001), Chapter 8 of Koul (2002), Berkes and Horváth (2003b), Kulperger and Yu (2003). Thus the approach of Chu, Stinchcombe and White (1996) cannot be readilly applied. In this paper, we put forward a different approach which does not use model residuals but relies more directly on the quasi-likelihood function. Suppose we have observed $y_1, \ldots, y_m$ for which we have postulated a model with d parameters. Denote by $\mathbf{u}$ a generic element of the parameter space and by $\ell_i(\mathbf{u})$ the conditional quasi-likelihood of y_i given $y_{i-1}, \ldots, y_1$, so that the quasi-likelihood function is $L_m(\mathbf{u}) = \sum_{1\leq i\leq m} \ell_i(\mathbf{u})$. For time series models, the $\ell_i(\mathbf{u})$ typically cannot be computed exactly because of the dependence on the unobserved $y_k, k \leq 0$, and some approximations $\hat{\ell}_i(\mathbf{u})$ must be used instead. Denote by $\hat{\ell}_i'(\mathbf{u})$ the d-dimensional row vector of partial derivatives with respect to the model parameters and consider the matrix

$$\hat{\mathbf{D}}_m = \frac{1}{m} \sum_{1<i\leq m} (\hat{\ell}_i'(\hat{\boldsymbol{\theta}}_m))^T \hat{\ell}_i'(\hat{\boldsymbol{\theta}}_m), \tag{1.5}$$

where $\hat{\boldsymbol{\theta}}_m$ is the quasi maximum likelihood parameter estimate. We can now construct the d-dimensional process

$$G_m(t) = \sum_{m<i\leq mt} \hat{\ell}_i'(\hat{\boldsymbol{\theta}}_m) \hat{\mathbf{D}}_m^{-1/2}, \quad t \in [0,\infty). \tag{1.6}$$

Our approach relies on the realization that the process $G_m(\cdot)$ can be well approximated by the process

$$(1.7) \qquad \left((1+t)W_1\left(\frac{t}{1+t}\right), \ldots, (1+t)W_d\left(\frac{t}{1+t}\right)\right), \quad t \in [0,\infty),$$

where $W_j(\cdot), j = 1, \ldots d$ are independent standard Wiener processes. By taking appropriate functionals, it is then easy to construct a monitoring scheme with a controlled probability of a false rejection of the null hypothesis. The details are presented in Section 3.

While the above idea is applicable to essentially any parametric model for which a reasonable approximate quasi-likelihood function can be found, a rigorous verification for a complex nonlinear model requires some effort. First of all, appropriate approximations $\hat{\ell}_i'(\mathbf{u})$ must be used. To obtain them, we use an expansion developed in Berkes, Horváth and Kokoszka (2003) which is described in Section 2. Secondly, asymptotics for the matrix $\hat{\mathbf{D}}_m$ in (1.5) must be established which allow us to approximate the process $G(\cdot)$ by a bridge type multidimensional Gaussian process (see Lemma 6.6). The corresponding result is stated in Proposition 3.1 which is proved in Section 5. Finally, the partial sum process $\sum_{m<i\le mt} \hat{\ell}_i'(\hat{\boldsymbol{\theta}}_m)$ must be carefully approximated what is accomplished in several stages presented in Section 6.

By considering the quasi-likelihood scores, our approach is related to that of Nyblom (1989), but it is different in that our focus is on controlling the probability of false alarms, whereas Nyblom (1989) concentrated on constructing locally most powerful tests against alternatives which assume that the parameter changes $\boldsymbol{\theta}_k - \boldsymbol{\theta}_{k-1}$ from time $k-1$ to time k form a martingale sequence with known covariance matrices $\mathbf{G}_k = E[(\boldsymbol{\theta}_k - \boldsymbol{\theta}_{k-1})(\boldsymbol{\theta}_k - \boldsymbol{\theta}_{k-1})^T]$. Such assumptions are not appropriate in the context of on-line monitoring for a change in the GARCH(p,q) parameters. Moreover, in order to prove a functional limit theorem similar to our relation (6.21), Nyblom (1989) needed to impose a number of additional technical assumptions on the asymptotic properties of the likelihood scores. Our assumptions pertain only to model parameters and the distribution of the model errors and are very weak, essentially necessary for the consistency and asymptotic normality of the quasi maximum likelihood estimator. Another important contribution to the theory of optimal *a posteriori* (in-sample) change point tests under local alternatives was made by Sowell (1996) who considered a much more general setting. His assumptions however also impose various conditions on score-like objects rather than on the model parameters and errors.

There has been a growing literature concerned with the change-point problem specifically in GARCH setting or with a more general problem of detecting changes in volatility. Lamourex and Lastrapes (1990) and Mikosch and Stărică (2002), among others, showed that change points in a GARCH specification may lead to the presence of spurious persistence in volatility. Mikosch and Stărică (2002) also proposed a periodogram based on-line change point detection procedure, but did not develop it theoretically. The following papers deal with change point detection in a historical sample (*a posteriori* methods). Inclan and Tiao (1994), Kokoszka and Leipus (1999, 2000), Kim, Cho and Lee (2000) and

Andreou and Ghysels (2002) studied CUSUM type tests based on squared observations. Tests based on the empirical distribution function of the squared residuals were studied by Horváth, Kokoszka and Teyssière (2001), and Kokoszka and Teyssiere (2002). Inoue (2001) proposed a simulation based method based on the empirical distribution function of the observations and applied it to stochastic volatility models. CUSUM tests based on the partial sums of powers of residuals were developed by Kulperger and Yu (2003). Kokoszka and Teyssière (2002) also studied a generalized likelihood ratio test. Chu (1995) and Lundbergh and Teräsvirta (2002) investigated Lagrange multiplier tets. The above list is not meant to be exhaustive, but is intended to indicate a recent interest in and a variety of approaches to the problem of change point detection in GARCH processes. For a review of sequential testing strategies in case of independent observations, see e.g. Gombay (1996, 2002).

The paper is organized as follows. In Section 2 we present the necessary technical background on GARCH processes which includes very week assumptions on the model parameters and errors. In particular, we do not require that the innovations ϵ_i have a smooth density or that the observations y_i have finite expected value. Section 3 describes the monitoring scheme and contains results establishing its asymptotic behavior under the null and under the alternative. Results of a small simulation study are presented in Section 4.Proofs are collected in Sections 5 - 7.

2 Definitions and assumptions

In this section we present the general framework used troughout the paper. We recall the definition of a GARCH(p,q) process and present recursions needed to define the quasi maximum likelihood estimator. We state the conditions for the existence of a stationary solution to the GARCH(p,q) equations and for the consistency and asymptotic normality of the maximum likelihood estimator. The section is based primarily on the results of Bougerol and Picard (1992a, b) and Berkes, Horváth and Kokoszka (2003).

We assume that

$$y_k = \sigma_k \epsilon_k \tag{2.8}$$

and under the no change null hypothesis

$$\sigma_k^2 = \omega + \sum_{1\le i\le p} \alpha_i y_{k-i}^2 + \sum_{1\le j\le q} \beta_j \sigma_{k-j}^2, \tag{2.9}$$

where $\boldsymbol{\theta} = (\omega, \alpha_1, \ldots, \alpha_p, \beta_1, \ldots, \beta_q)$ is the parameter of the process. Under the alternative

$$\sigma_k^2 = \begin{cases} \omega + \sum_{1\le i\le p} \alpha_i y_{k-i}^2 + \sum_{1\le j\le q} \beta_i \sigma_{k-j}^2, & -\infty < k \le k^* \\ \omega^* + \sum_{1\le i\le p} \alpha_i^* y_{k-i}^2 + \sum_{1\le j\le q} \beta_i^* \sigma_{k-j}^2, & k^* < k < \infty \end{cases} \tag{2.10}$$

i.e. a change in the parameters occurred at time k^* and the new parameter is $\boldsymbol{\theta}^* = (\omega^*, \alpha_1^*, \ldots, \alpha_p^*, \beta_1^*, \ldots, \beta_q^*)$. In the following, we refer to the specification (2.9) as the null hypothesis H_0 and to (2.10) for some integer k^* as the alternative hypothesis H_A.

Throughout this paper we assume that

(2.11) $\epsilon_k, \quad -\infty < k < \infty$ are independent, identically distributed random variables.

Note that as Bougerol and Picard (1992a) we do not assume that the errors ϵ_k have mean zero. Our procedure is not based on the deviations of residuals from zero. Additional assumptions on the distribution of the ϵ_k are stated in conditions (2.17)–(2.19).

Our procedure is based on the quasi maximum likelihood estimator of the parameters of a GARCH process developed by Lee and Hansen (1994), Lumsdaine (1996), Berkes, Horváth and Kokoszka (2003). (For a more general method we refer to Berkes and Horváth (2003)). To define this estimator for general GARCH(p, q) processes, denote by $\mathbf{u} = (x, s_1, \ldots, s_p, t_1, \ldots, t_q)$ the generic element of the parameter space U which, following Berkes, Horváth and Kokoszka (2003), is defined as follows: Let $0 < \underline{u} < \bar{u}$, $0 < \rho_0 < 1$, $q\underline{u} < \rho_0$. Then

$$U = \{\mathbf{u} : t_1 + t_2 + \cdots + t_q \le \rho_0 \text{ and } \underline{u} < \min(x, s_1, s_2, \ldots, s_p, t_1, t_2, \ldots, t_q) \le \max(x, s_1, s_2, \ldots, s_p, t_1, t_2, \ldots, t_q) \le \bar{u}\}.$$

We assume that

$$\boldsymbol{\theta} \in U. \tag{2.12}$$

Define now the log quasi likelihood function as

$$\hat{L}_m(\mathbf{u}) = \frac{1}{m} \sum_{1<k\le m} -\frac{1}{2}\left\{\log \hat{w}_k(\mathbf{u}) + \frac{y_k^2}{\hat{w}_k(\mathbf{u})}\right\},$$

where

$$\hat{w}_k(\mathbf{u}) = c_0(\mathbf{u}) + \sum_{1\le i<k} c_i(\mathbf{u}) y_{k-i}^2.$$

The functions $c_i(\mathbf{u})$, $0 \le i < \infty$ are defined by recursion. If $q \ge p$, then

$$\begin{aligned}
c_0(\mathbf{u}) &= x/(1 - (t_1 + \cdots + t_q)) \\
c_1(\mathbf{u}) &= s_1 \\
c_2(\mathbf{u}) &= s_2 + t_1 c_1(\mathbf{u}) \\
&\vdots \\
c_p(\mathbf{u}) &= s_p + t_1 c_{p-1}(\mathbf{u}) + \cdots + t_{p-1} c_1(\mathbf{u}) \\
c_{p+1}(\mathbf{u}) &= t_1 c_p(\mathbf{u}) + \cdots + t_p c_1(\mathbf{u}) \\
&\vdots \\
c_q(\mathbf{u}) &= t_1 c_{q-1}(\mathbf{u}) + \cdots + t_{q-1} c_1(\mathbf{u})
\end{aligned}$$

and if $q < p$, the equations above are replaced with

$$
\begin{aligned}
c_0(\mathbf{u}) &= x/(1-(t_1+\cdots+t_q)) \\
c_1(\mathbf{u}) &= s_1 \\
c_2(\mathbf{u}) &= s_2 + t_1 c_1(\mathbf{u}) \\
&\vdots \\
c_{q+1}(\mathbf{u}) &= s_{q+1} + t_1 c_q(\mathbf{u}) + \cdots + t_q c_1(\mathbf{u}) \\
&\vdots \\
c_p(\mathbf{u}) &= s_p + t_1 c_{p-1}(\mathbf{u}) + \cdots + t_q c_{p-q}(\mathbf{u}).
\end{aligned}
$$

In general, if $i > R = \max(p, q)$, then

$$c_i(\mathbf{u}) = t_1 c_{i-1}(\mathbf{u}) + t_2 c_{i-2}(\mathbf{u}) + \cdots + t_q c_{i-q}(\mathbf{u}).$$

The above recursions ensure that for the true value $\boldsymbol{\theta}$, $\sigma_k^2 = c_0(\boldsymbol{\theta}) + \sum_{1\le i<\infty} c_i(\boldsymbol{\theta}) y_{k-i}^2$.

In order to formulate a necessary and sufficient condition for the existence of a unique stationary sequence satisfying (2.8) and (2.9) we must introduce further notations. Let

$$
\begin{aligned}
\boldsymbol{\tau}_n &= (\beta_1 + \alpha_1 \epsilon_n^2, \beta_2, \ldots, \beta_{q-1}) \in \mathbb{R}^{q-1} \\
\boldsymbol{\xi}_n &= (\epsilon_n^2, 0, \ldots, 0) \in \mathbb{R}^{q-1}
\end{aligned}
$$

and

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

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

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

where $\mathbf{I}_{q-1}$ and $\mathbf{I}_{p-2}$ are the identity matrices of size $q-1$ and $p-2$, respectively. The norm of any $d \times d$ matrix M is defined by

$$||M|| = \sup\{||M\mathbf{x}||_d / ||\mathbf{x}||_d : \mathbf{x} \in \mathbb{R}^d,\ \mathbf{x} \ne 0\},$$

where $||\cdot||_d$ is the Euclidean norm in $\mathbb{R}^d$. The top Lyapunov exponent γ_L associated with the sequence $\{A_n, -\infty < n < \infty\}$ is

$$\gamma_L = \inf_{1\le n<\infty} \frac{1}{n} E \log ||A_0 A_1 \ldots A_n||,$$

assuming that

(2.13) $$E(\log ||A_0||) < \infty.$$

(We note that $||A_0|| \geq 1$, cf. Berkes, Horváth and Kokoszka (2003).) Bougerol and Picard (1992a, b) showed that if (2.13) holds, then (2.8) and (2.9) have a unique stationary solution if and only if

(2.14) $$\gamma_L < 0.$$

We note that (2.14) implies $\beta_1 + \cdots + \beta_q < 1$ (cf. Bougerol and Picard (1992b)).

The next two conditions are needed to uniquely identify the parameter $\boldsymbol{\theta}$:

(2.15) the polynomials $\alpha_1 x + \alpha_2 x^2 + \cdots + \alpha_p x^p$ and $1 - \beta_1 x - \beta_2 x^2 - \ldots - \beta_q x^q$ are coprimes on the set of polynomials with real coefficients

and

(2.16) ϵ_0^2 is a nondegenerate random variable.

We also assume

(2.17) $$\lim_{t \to 0} t^{-\mu} P\{\epsilon_0^2 \leq t\} = 0 \text{ with some } \mu > 0$$

which will be needed to estimate the moments of $w_0(\mathbf{u})/w_0(\boldsymbol{\theta})$ (cf. Berkes, Horváth and Kokoszka (2003)).

Finally, the last set of conditions concerns the moments of ϵ_0:

(2.18) $$E\epsilon_0^2 = 1$$

and

(2.19) $$E|\epsilon_0^2|^{2+\delta} < \infty \text{ with some } \delta > 0.$$

Berkes, Horváth, and Kokoszka (2003) showed that $m^{1/2}(\hat{\boldsymbol{\theta}}_m - \boldsymbol{\theta})$ is asymptotically normal, if (2.8), (2.9) and (2.13)–(2.19) hold.

3 Main results

Suppose we have observed $y_1, \ldots, y_m$ which represent available historical data. The estimator for the unknown parameter $\boldsymbol{\theta}$ based on these data is defined by

(3.20) $$\hat{\boldsymbol{\theta}}_m = \arg\max\left\{\hat{L}_m(\mathbf{u}) : \quad \mathbf{u} \in U\right\},$$

where U is a suitably chosen compact set defined in Section 2. Using the notation introduced in Section 2, we also define the conditional likelihoods

$$\hat{\ell}_i(\mathbf{u}) = -\frac{1}{2}\left\{\log \hat{w}_i(\mathbf{u}) + \frac{y_i^2}{\hat{w}_i(\mathbf{u})}\right\}$$

and the matrix

$$\hat{\mathbf{D}}_m = \frac{1}{m}\sum_{1<i\leq m}(\hat{\ell}_i'(\hat{\boldsymbol{\theta}}_m))^T \hat{\ell}_i'(\hat{\boldsymbol{\theta}}_m).$$

(^{T} denotes the transpose of vectors and matrices.)

Let $|\cdot|$ denote the maximum norm of vectors and matrices. We can now define the stopping time k_m as

$$k_m = \min\left\{k : \left|\sum_{m<i\le m+k} \hat{\ell}_i'(\hat{\boldsymbol{\theta}}_m)\hat{\mathbf{D}}_m^{-1/2}\right| > m^{1/2}\left(1+\frac{k}{m}\right) b\left(\frac{k}{m}\right)\right\}.$$

If $k_m < \infty$, we say that a change occurred. We choose the boundary function $b(t)$ so that

$$\lim_{m\to\infty} P_{H_0}\{k_m < \infty\} = \alpha, \tag{3.21}$$

where $0 < \alpha < 1$ is a prescribed number and

$$\lim_{m\to\infty} P_{H_A}\{k_m < \infty\} = 1. \tag{3.22}$$

Recall that H_0 is defined by (2.8) and (2.9) and H_A by (2.8) and (2.10). Conditions on the boundary function $b(\cdot)$ are specified in the following two subsections which study the asymptotic behavior of the monitoring scheme, respectively, under H_0 and H_A. Unlike for the scheme of Chu, Stinchcombe and White (1996), which requires more complex boundary functions, in our setting, the simplest choice satisfying all the assumptions is to take a constant $b(t) = b > 0$.

3.1 Monitoring scheme under H_0

Throughout this section we assume that (2.8) and (2.9) hold.

It follows from the definition that the matrix $\hat{\mathbf{D}}_m$ is non–negative definite. We show in Proposition 3.1 that $\hat{\mathbf{D}}_m$ is nonsingular with probability tending to one as $m \to \infty$. Hence $\hat{\mathbf{D}}_m^{-1/2}$ exists with probability tending to one as $m \to \infty$.

To formulate Proposition 3.1, we define

$$w_k(\mathbf{u}) = c_0(\mathbf{u}) + \sum_{1\le i<\infty} c_i(\mathbf{u}) y_{k-i}^2$$

and

$$\ell_k(\mathbf{u}) = -\frac{1}{2}\left(\log w_k(\mathbf{u}) + \frac{y_k^2}{w_k(\mathbf{u})}\right).$$

Finally, define

$$\mathbf{D}(\mathbf{u}) = E\left[(\ell_0'(\mathbf{u}))^T \ell_0'(\mathbf{u})\right]$$

and

$$\mathbf{D} = \mathbf{D}(\boldsymbol{\theta}).$$

Proposition 3.1 *If (2.8), (2.9) and (2.13)–(2.19) are satisfied, then*

$$\hat{\mathbf{D}}_m \to \mathbf{D} \quad a.s.$$

Also, $\mathbf{D}$ *is a positive definite, non–singular matrix.*

The proof of Proposition 3.1 is presented in Section 5.

We impose the following conditions on the boundary function $b(t)$:

(3.23) $$b(t) \text{ is continuous on } (0, \infty)$$

and

(3.24) $$\inf_{0<t<\infty} b(t) > 0.$$

Theorem 3.1 *If (2.8), (2.9),(2.13) – (2.19) and conditions (3.23), (3.24) are satisfied, then*

$$\lim_{m\to\infty} P_{H_0}\{k_m < \infty\} = 1 - \left(P\left\{\sup_{0\le t<1} \frac{|W(t)|}{b(t/(1-t))} \le 1\right\}\right)^{p+q+1},$$

where $\{W(t), 0 \le t \le 1\}$ *denotes a Wiener process.*

Proof of Theorem 3.1 is presented in Section 6.

Choosing $b(t) = b$, a constant function, and using the well–known formula for the distribution function of $\sup_{0\le t\le 1} |W(t)|$ (cf. Csörgő and Révész (1981, p. 43)) we obtain the following corollary.

Corollary 3.1 *If (2.8), (2.9),(2.13) – (2.19) and conditions (3.23), (3.24) are satisfied, then*

$$\begin{aligned}\lim_{m\to\infty} P_{H_0}\{k_m < \infty\} &= 1 - \left(P\left\{\sup_{0\le t\le 1} |W(t)| \le b\right\}\right)^{p+q+1} \\ &= 1 - \left\{\frac{4}{\pi} \sum_{0\le k<\infty} \frac{(-1)^k}{2k+1} \exp\left(-\frac{\pi^2(2k+1)^2}{8b^2}\right)\right\}^{p+q+1}.\end{aligned}$$

Corollary 3.1 allows to specify the critical level b for any significance level α in (3.21).

3.2 Monitoring Scheme under H_A

Under the alternative (2.10) the parameter changes from $\boldsymbol{\theta}$ to $\boldsymbol{\theta}^*$ at time $k^* > m$. We define the sequence $\{z_k, \ -\infty < k < \infty\}$, which represents the model after a change in parameters, by the equations

(3.25) $$z_k = \gamma_k \epsilon_k$$

and

(3.26) $$\gamma_k^2 = \omega^* + \sum_{1\le i\le p} \alpha_i^* z_{k-i}^2 + \sum_{1\le j\le q} \beta_j^* \gamma_{k-j}^2.$$

The error sequence ϵ_k has not changed.

Our first condition is that the parameter $\boldsymbol{\theta}^*$ defines a GARCH(p,q) process. The vectors $\boldsymbol{\tau}_n^* = (\beta_1^* + \alpha_1^*\epsilon_n^2, \beta_2^*, \ldots, \beta_{q-1}^*) \in \mathbb{R}^{q-1}$ and $\boldsymbol{\alpha}^* = (\alpha_2^*, \ldots, \alpha_{p-1}^*) \in \mathbb{R}^{p-2}$ are defined as $\boldsymbol{\tau}_n$ and $\boldsymbol{\alpha}$ in Section 2 but α_i and β_i are replaced by α_i^* and β_j^*. Similarly to A_n we define

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

The top Lyapunov exponent of the sequence A_n^*, $-\infty < n < \infty$ is denoted by γ_L^*. As in Section 2, the equations (3.25) and (3.26) have a unique stationary solution if and only if

$$\gamma_L^* < 0 \tag{3.27}$$

assuming that

$$E(\log ||A_0^*||) < \infty. \tag{3.28}$$

Similarly to $\ell_i(\mathbf{u})$ we define

$$g_i(\mathbf{u}) = -\frac{1}{2}\left\{\log v_i(\mathbf{u}) + \frac{z_i^2}{v_i(\mathbf{u})}\right\},$$

where

$$v_i(\mathbf{u}) = c_0(\mathbf{u}) + \sum_{1\le j<\infty} c_j(\mathbf{u}) z_{j-i}^2.$$

Hence

$$g_i'(\mathbf{u}) = \frac{1}{2}\left(\epsilon_i^2 \frac{\gamma_i^2}{v_i(\mathbf{u})}\frac{v_i'(\mathbf{u})}{v_i(\mathbf{u})} - \frac{v_i'(\mathbf{u})}{v_i(\mathbf{u})}\right).$$

Let

$$\bar{g}(\mathbf{u}) = Eg_0(\mathbf{u}).$$

We assume that

$$\bar{g}'(\boldsymbol{\theta}) \neq \mathbf{0} \tag{3.29}$$

and

$$\text{there is } U^*, \text{ a neighborhood of } \boldsymbol{\theta}, \text{ such that } \bar{g}(\mathbf{u}),\ \bar{g}'(\mathbf{u}),\ \bar{g}''(\mathbf{u}) \text{ exist and are continuous for each } \mathbf{u} \in U^*. \tag{3.30}$$

We note that under some regularity conditions $\bar{g}'(\mathbf{u}) = \mathbf{0}$ if and only if $\mathbf{u} = \boldsymbol{\theta}^*$ (cf. Lemma 5.5 in Berkes, Horváth and Kokoszka (2003)). We also assume that

$$k^* = k^*(m) \text{ and } \limsup_{m\to\infty}(k^*/m) < \infty \tag{3.31}$$

and

$$\sup_{0<t<\infty} b(t) < \infty. \tag{3.32}$$

Theorem 3.2 *If (2.8), (2.10),(2.13) – (2.19), (3.23), (3.24) and (3.25)–(3.32) hold, then*

$$\lim_{m\to\infty} P_{H_A}\{k_m < \infty\} = 1.$$

Proof of Theorem 3.2 is presented in Section 7.

4 Simulations

In this section we report the results of a simulation study intended to assess the behavior of the procedure in finite samples. Even before we commenced the numerical experiments, it was clear to us that much larger sample sizes than those considered by Chu, Stinchcombe and White (1996) (CSW in the following) will be needed for the asymptotic behavior to manifest itself. There are two main reasons for it. First, the required model estimation in CSW consisted essentially in estimating the mean of normal variables and it is well–known that already for samples of size 50 such estimates are very accurate. In our context, we have to estimate the parameters of a GARCH process via nonlinear optimization. It is well–known that although these estimates are optimal when the innovations are normal (the case we considered in the simulation study), they may have large biases and standard errors even for samples as large as one thousand. For a sample size of one thousand, these estimators are acurate for a wide range of parameters, but for some choices of parameters they exhibit large biases. Second, our procedure requires the estimation of the covariance matrix $\hat{\mathbf{D}}_m$. Since there is no close formula for $\hat{\mathbf{D}}_m$, it would have to be estimated even if the parameters were known. We conducted a number of experiments, not reported here in a systematic way, in which we used the exact values of the GARCH parameters rather than their estimates. Even in this situation, samples of size about one thousand are required to obtain relatively stable estimates of $\hat{\mathbf{D}}_m$.

We now proceed with a detailed description of our simulation study and the conclusions it leads to. We focused on the popular GARCH(1,1) models and considered a wide range of the parameters ω, α_1 and β_1. The GARCH models were simulated and estimated using the S+ module GARCH, whereas for the estimation of $\hat{\mathbf{D}}_m$ it was necessary to write a much faster C++ code and interface it with S+.

We report the results for three GARCH(1,1) models:

Model I: $\omega = 0.05,\ \alpha_1 = 0.4,\ \beta_1 = 0.3;$

Model II: $\omega = 0.05,\ \alpha_1 = 0.5,\ \beta_1 = 0.0;$

Model III: $\omega = 1.0,\ \alpha_1 = 0.3,\ \beta_1 = 0.2;$

The results for these three models are fairly representative of the overall conclusions.

To facilitate the graphical presentation of the results, we work with the normalized decision function

$$C(k) = \left(1+\frac{k}{m}\right)^{-1} \frac{1}{\sqrt{m}} \left| \left[\sum_{m<i\le m+k} l_i'(\hat{\theta}_m)\hat{\mathbf{D}}_m^{-1/2} \right] \right|, \quad 1 \le k < \infty.$$

A change in parameters is signaled if $C(k) > c(\alpha)$, where $c(\alpha)$ satisfies

$$1 - P\left(\sup_{0<t<1} |W(t)| \le c(\alpha)\right)^3 = \alpha. \tag{4.1}$$

The following table gives the critical values for the conventional significance levels.

α	0.10	0.05	0.01
$c(\alpha)$	2.382	2.632	3.150

Table 1: Critical values calculated according to relation (4.1).

In Table 2 we report the empirical rejection probabilities of the null hypothesis of no change in the model parameters assuming this hypothesis is true. It is seen that for models with pronounced GARCH characteristics, i.e. parameters α_1 and β_1 much larger than ω, the procedure has correct size for monitoring horizons of about 500 time units with the 10 percent bound being somewhat more reliable. By a monitoring horizon we understand here the length of time we are willing to use the procedure without updating the parameters. We note that the theory developed in this paper shows that the empirical size tends to the nominal size as $m \to \infty$, so for any finite m size distortions will be present. This is particularly visible if the GARCH parameters are difficult to estimate, as in Model III (the process looks more like a white noise); the procedure has a high probability of type I error. We conjecture, that in such situations, m much larger than one thousand would be required to obtain empirical size close to the nominal size. Using the true values of α_1, β_1 and ω, leads to entries about half the size of those reported in Table 2. With $m = 1000$, the method is not acurate for monitoring horizons longer than 500 and cannot be used in an automatic way. As we mention in the discussion toward the end of this section, a visual real-time inspection of the graph of $C(k)$ following an alarm (critical level exceeded) might indicate that there is no reason to suspect a change in model parameters, see Figure 3. Alternatively, finite sample corrections could be obtained by simulation for specific values of m and monitoring horizons of interest. Such simulations would be specific to a problem at hand and have not been conducted.

The power of the procedure for three change–point scenarios is reported in Table 3. As can be expected, large changes in parameters are detected more reliably.

From practical point of view, it is more useful to study the distribution of the detection delay time or, equivalently, the distribution of the random time when the decision function $C(k)$ first exceeds a critical level. In Table 4, we report selected descriptive statistics for such distributions. The estimated densities are depicted in Figure 1.

Focusing first on the first three change–point models reported in Table 4, we note that the distribution of the delay time is fairly symmetric but its spread increases as the change–point moves further away from the point where the monitorig was initiated. Similar findings were reported in CSW. However, unlike for the FL monitoring scheme investigated in CSW, the average delay time does not appear to increase with the distance

k	100	200	300	400	500	600	700	800	900	1000
Model I: $\omega = 0.05,\ \alpha_1 = 0.4,\ \beta_1 = 0.3$										
α=.10	.001	.008	.021	.045	.072	.099	.132	.146	.165	.183
α=.05	.001	.003	.009	.028	.052	.070	.095	.112	.127	.139
Model II: $\omega = 0.05,\ \alpha_1 = 0.5,\ \beta_1 = 0.0$										
α=.10	.013	.027	.041	.054	.068	.080	.099	.116	.129	.137
α=.05	.013	.026	.038	.051	.066	.076	.090	.109	.120	.128
Model III: $\omega = 1.0,\ \alpha_1 = 0.3,\ \beta_1 = 0.2$										
α=.10	.047	.146	.237	.329	.407	.465	.519	.561	.593	.616
α=.05	.029	.101	.190	.256	.334	.390	.440	.483	.516	.539

Table 2: Empirical sizes for monitoring horizons k. Simulations were done with $m = 1000$ and are based on 2000 replications.

of the change–point from the initiation point and it is about 20 for a change from Model I to Model III. For relatively less significant changes in parametes, like the change from Model I to Model II, the delay time is much longer. Even in such situations, however, a visual real-time inspection of the graph of $C(k)$ may suggest that something is happening to the parameters of the model. In the panel in the right-bottom corner of Figure 2, five randomly selected trajectories of $C(k)$ for the change from Model I to Model II are shown. A picture of this type may be fairly typical in real-data applications as the parameters need not switch immediately into a new regime, but may evolve gradually through a number of smaller changes. In contrast, as shown in Figure 3, if there is no change, the trajectory of $C(k)$ may occasionaly exceed the critical value, but it will not show a pronounced upward trend manifest in Figure 2.

5 Proof of Proposition 3.1

Let

$$\hat{\mathbf{D}}_m(\mathbf{u}) = \frac{1}{m} \sum_{1<i\leq m} (\hat{\ell}_i'(\mathbf{u}))^T \hat{\ell}_i'(\mathbf{u})$$

and

$$\mathbf{D}_m(\mathbf{u}) = \frac{1}{m} \sum_{1\leq i\leq m} (\ell_i'(\mathbf{u}))^T \ell_i'(\mathbf{u}).$$

In the proof of their Lemma 5.8, Berkes, Horváth and Kokoszka (2003) showed that there is a constant $0 < \varrho < 1$ and a positive random variable ξ such that

$$\sup_{\mathbf{u}\in U} \left| \hat{\ell}_k'(\mathbf{u}) - \ell_k'(\mathbf{u}) \right| \leq \xi \varrho^k. \tag{5.2}$$

k	100	200	300	400	500	600	700	800	900	1000
Model I changes to Model III at $k^* = 50$										
α=.10	.975	1	1	1	1	1	1	1	1	1
α=.05	.962	1	1	1	1	1	1	1	1	1
Model I changes to Model III at $k^* = 250$										
α=.10	.002	.004	.947	1	1	1	1	1	1	1
α=.05	.000	.003	.932	1	1	1	1	1	1	1
Model I changes to Model III at $k^* = 500$										
α=.10	.000	.009	.020	.037	.055	.999	1	1	1	1
α=.05	.000	.005	.012	.024	.034	.998	1	1	1	1
Model I changes to Model II at $k^* = 250$										
α=.10	.002	.004	.017	.071	.252	.523	.735	.859	.927	.966
α=.05	.000	.003	.008	.045	.177	.407	.646	.796	.892	.946

Table 3: Empirical power the test. For each k, the fraction of trajectories crossing the critical levels From Table 1 is reported. Simulations were done with $m = 1000$ and are based on 1000 replications.

Berkes, Horváth and Kokoszka (2003) also showed that $\sup_{\mathbf{u}\in U} |\ell_i(\mathbf{u})|$ is a stationary sequence and

$$E \sup_{\mathbf{u}\in U} |\ell_0(\mathbf{u})| < \infty.$$

Hence

$$\begin{aligned} \sup\nolimits_{\mathbf{u}\in U} \left|\hat{\mathbf{D}}_m(\mathbf{u}) - \mathbf{D}_m(\mathbf{u})\right| &\le \frac{1}{m}\left(2\xi \sum_{1<i\le m} \rho^i \sup_{\mathbf{u}\in U} |\ell_i(\mathbf{u})| + \xi^2 \sum_{1<i\le m} \rho^{2i}\right) \\ &\overset{\text{a.s.}}{=} O\left(\frac{1}{m}\right) \end{aligned} \tag{5.3}$$

by Lemma 2.2 of Berkes, Horváth and Kokoszka (2003).

Next we show that

$$E \sup_{\mathbf{u}\in U} \left|(\ell_0'(\mathbf{u}))^T \ell_0'(\mathbf{u})\right| < \infty. \tag{5.4}$$

By (6.17) we have that

$$\begin{aligned} \sup_{\mathbf{u}\in U} \left|(\ell_0'(\mathbf{u}))^T \ell_0(\mathbf{u})\right| &\le \sup_{\mathbf{u}\in U} |\ell'(\mathbf{u})|^2 \\ &\le \epsilon_0^4 \left(\sup_{\mathbf{u}\in U} \frac{\sigma_0^2}{w_0(\mathbf{u})}\right)^2 \left(\sup_{\mathbf{u}\in U} \frac{|w_0'(\mathbf{u})|}{w_0(\mathbf{u})}\right)^2 + \left(\sup_{\mathbf{u}\in U} \frac{|w_0'(\mathbf{u})|}{w_0(\mathbf{u})}\right)^2. \end{aligned}$$

Mean	SE	Min	Q1	Med	Q3	Max
Model I changes to Model III at $k^* = 50$						
71.4	12.4	50	63	71	79	135
Model I changes to Model III at $k^* = 250$						
272.3	18.1	89	262	271	282	338
Model I changes to Model III at $k^* = 500$						
516.4	54.7	121	511	523	538	618
Model I changes to Model II at $k^* = 250$						
612.9	166.5	89	498	589	710	1000

Table 4: Elementary statistics for the distribution of the first exceedance of the 10 percent critical level. The estimated densities are graphed in Figure 1. Simulations were done with $m = 1000$ and are based on 1000 replications.

Using the independence of ϵ_0 and $(w_0(\mathbf{u}), \sigma_0^2)$, $\mathbf{u} \in U$ we get that

$$\begin{aligned}
&E\epsilon_0^4 \left(\sup_{\mathbf{u} \in U} \frac{\sigma_0^2}{w_0(\mathbf{u})} \right)^2 \left(\sup_{\mathbf{u} \in U} \frac{|w_0'(\mathbf{u})|}{w_0(\mathbf{u})} \right)^2 \\
&= E\epsilon_0^4 E \left(\sup_{\mathbf{u} \in U} \frac{\sigma_0^2}{w_0(\mathbf{u})} \right)^2 \left(\sup_{\mathbf{u} \in U} \frac{|w_0'(\mathbf{u})|}{w_0(\mathbf{u})} \right)^2 \\
&\le E\epsilon_0^4 E \left(E \left(\sup_{\mathbf{u} \in U} \frac{\sigma_0^2}{w_0(\mathbf{u})} \right)^{2+\delta/2} \right)^{2/(2+\delta/2)} \left(E \left(\sup_{\mathbf{u} \in U} \frac{|w_0'(\mathbf{u})|}{w_0(\mathbf{u})} \right)^{2(4+\delta)/\delta} \right)^{\delta/(4+\delta)}.
\end{aligned}$$

Lemma 5.1 of Berkes, Horváth and Kokoszka (2003) yields that

$$E \left(\sup_{\mathbf{u} \in U} \frac{\sigma_0^2}{w_0(\mathbf{u})} \right)^{2+\delta/2} < \infty, \tag{5.5}$$

and Lemma 3.6 of Berkes and Horváth (2003) gives

$$E \left(\sup_{\mathbf{u} \in U} \frac{|w_0'(\mathbf{u})|}{w_0(\mathbf{u})} \right)^{2(4+\delta)/\delta} < \infty.$$

Hence the proof of (5.4) is complete. For each $\mathbf{u} \in U$, $\ell_i'(\mathbf{u})$ is a stationary and ergodic sequence. So by (5.4) we can use the ergodic theorem resulting in

$$\mathbf{D}_m(\mathbf{u}) \to \mathbf{D}(\mathbf{u}) \quad \text{a.s. for any } \mathbf{u} \in U. \tag{5.6}$$

Next we show that there is a constant C_1 and $U^* \subseteq U$, a neighborhood of $\boldsymbol{\theta}$, such that

$$\left| E(\ell_0'(\mathbf{u}))^T \ell_0'(\mathbf{u}) - E(\ell_0'(\mathbf{v}))^T \ell_0'(\mathbf{v}) \right| \le C_1 |\mathbf{u} - \mathbf{v}|, \quad \text{if } \mathbf{u}, \mathbf{v} \in U^*. \tag{5.7}$$

Using (6.17) we can write

$$E\sup\left\{\frac{1}{|\mathbf{u}-\mathbf{v}|}\left|(\ell_0'(\mathbf{u}))^T(\ell_0'(\mathbf{u})-\ell_0'(\mathbf{v}))\right|:\ \mathbf{u},\mathbf{v}\in U^*\right\}\le A_1+\cdots+A_4, \tag{5.8}$$

where

$$A_1=E\epsilon_0^4E\sup\left\{\frac{1}{|\mathbf{u}-\mathbf{v}|}\frac{\sigma_0^2}{w_0(\mathbf{u})}\left|\left(\frac{w_0'(\mathbf{u})}{w_0(\mathbf{u})}\right)^T\sigma_0^2\left(\frac{w_0'(\mathbf{u})}{w_0^2(\mathbf{u})}-\frac{w_0'(\mathbf{v})}{w_0^2(\mathbf{v})}\right)\right|:\mathbf{u},\mathbf{v}\in U^*\right\},$$

$$A_2=E\epsilon_0^2E\sup\left\{\frac{1}{|\mathbf{u}-\mathbf{v}|}\left|\frac{\sigma_0^2}{w_0(\mathbf{u})}\left(\frac{w_0'(\mathbf{u})}{w_0(\mathbf{u})}\right)^T\left(\frac{w_0'(\mathbf{u})}{w_0(\mathbf{u})}-\frac{w_0'(\mathbf{v})}{w_0(\mathbf{v})}\right)\right|:\mathbf{u},\mathbf{v}\in U^*\right\},$$

$$A_3=E\epsilon_0^2E\sup\left\{\frac{1}{|\mathbf{u}-\mathbf{v}|}\left|\left(\frac{w_0'(\mathbf{u})}{w_0(\mathbf{u})}\right)^T\sigma_0^2\left(\frac{w_0'(\mathbf{u})}{w_0^2(\mathbf{u})}-\frac{w_0'(\mathbf{v})}{w_0^2(\mathbf{v})}\right)\right|:\mathbf{u},\mathbf{v}\in U^*\right\},$$

and

$$A_4=E\sup\left\{\frac{1}{|\mathbf{u}-\mathbf{v}|}\left|\left(\frac{w_0'(\mathbf{u})}{w_0(\mathbf{u})}\right)^T\left(\frac{w_0'(\mathbf{u})}{w_0(\mathbf{u})}-\frac{w_0'(\mathbf{v})}{w_0(\mathbf{v})}\right)\right|:\mathbf{u},\mathbf{v}\in U^*\right\}.$$

Using the mean value theorem coordinate–wise we get that

$$A_1\le C_2E\left[\left(\sup\left\{\frac{\sigma_0^2}{w_0(\mathbf{u})},\mathbf{u}\in U\right\}\right)^2\sup\left\{\frac{|w_0'(\mathbf{u})|}{w_0(\mathbf{u})}:\mathbf{u}\in U\right\}\sup\left\{\frac{w_0(\mathbf{v})}{w_0(\mathbf{u})};\mathbf{u},\mathbf{v}\in U^*\right\}\right.$$
$$\left.+\left(\sup\left\{\frac{|w_0'(\mathbf{u})|}{w_0(\mathbf{u})}:\mathbf{u}\in U\right\}\right)^2\left(\sup\left\{\frac{w_0(\mathbf{v})}{w_0(\mathbf{u})}:\mathbf{u},\mathbf{v}\in U^*\right\}\right)^2\right].$$

By the Hölder inequality we have

$$A_1\le C_2\left\{E\left(\sup\left\{\frac{\sigma_0^2}{w_0(\mathbf{u})}\mathbf{u}\in U\right\}\right)^{2+\delta/2}\right\}^{2/(4+\delta)}$$
$$\times\left(E\left\{\left(\sup\left\{\frac{|w_0'(\mathbf{u})|}{w_0(\mathbf{u})}:\mathbf{u}\in U\right\}\right)^2\left(\sup\frac{w_0(\mathbf{v})}{w_0(\mathbf{u})}:\mathbf{u},\mathbf{v}\in U^*\right)^2+1\right\}^{(4+\delta)/\delta}\right)^{\delta/(4+\delta)}. \tag{5.9}$$

By (5.5), the first expected value is finite in (5.9). Using the Cauchy inequality we get that

$$E\left[\left\{\sup\left(\frac{|w_0'(\mathbf{u})|}{w_0(\mathbf{u})}:\mathbf{u}\in U\right)\right\}^{(8+2\delta)/\delta}\left\{\sup\left(\frac{w_0(\mathbf{v})}{w_0(\mathbf{u})}:\mathbf{u},\mathbf{v}\in U^*\right)\right\}^{(8+2\delta)/\delta}\right]$$
$$\le\left\{E\left(\sup\frac{|w_0'(\mathbf{u})|}{w_0(\mathbf{u})}:\mathbf{u}\in U\right)^{(16+4\delta)/\delta}E\left(\sup\left(\frac{w_0(\mathbf{v})}{w_0(\mathbf{u})}:\mathbf{u},\mathbf{v}\in U^*\right)\right)^{(16+4\delta)/\delta}\right\}^{1/2}. \tag{5.10}$$

The first expected value on the right hand side of (5.10) is finite according to Lemma 3.6 of Berkes and Horváth (2003). the second expected value on the right hand side of (5.10) is finite by Lemma 3.7 of Berkes and Horváth (2003) assuming that U^* is a small enough neighborhood of $\boldsymbol{\theta}$. Hence

$$A_1 < \infty. \tag{5.11}$$

Similar arguments show that

$$A_i < \infty, \qquad i = 2, 3, 4, \tag{5.12}$$

implying that

$$E \sup \left\{ \frac{1}{|\mathbf{u} - \mathbf{v}|} \left| (\ell_0'(\mathbf{u}))^T (\ell_0'(\mathbf{u}) - \ell_0'(\mathbf{v})) \right| : \mathbf{u}, \mathbf{v} \in U^* \right\} < \infty \tag{5.13}$$

assuming that U^* is a small enough neighborhood of $\boldsymbol{\theta}$. By symmetry, we have that

$$E \sup \left\{ \frac{1}{|\mathbf{u} - \mathbf{v}|} \left| (\ell_0'(\mathbf{u}) - \ell_0'(\mathbf{v}))^T \ell_0'(\mathbf{v}) \right| : \mathbf{u}, \mathbf{v} \in U^* \right\} < \infty. \tag{5.14}$$

Hence the proof of (5.7) is complete.

We note that

$$\begin{aligned} &\sup_{\mathbf{u}, \mathbf{v} \in U^*} \frac{1}{|\mathbf{u} - \mathbf{v}|} |\mathbf{D}_m(\mathbf{u}) - \mathbf{D}_m(\mathbf{v})| \\ &\qquad \leq \frac{2}{m} \sum_{1 < i \leq m} \sup_{\mathbf{u}, \mathbf{v} \in U^*} \frac{1}{|\mathbf{u} - \mathbf{v}|} \left| (\ell_i'(\mathbf{u}))^T \ell_i'(\mathbf{u}) - (\ell_i'(\mathbf{v}))^T \ell_i'(\mathbf{v}) \right|. \end{aligned}$$

Since $\sup\{|(\ell_i'(\mathbf{u}))^T \ell_i'(\mathbf{u}) - (\ell_i'(\mathbf{v}))^T \ell_i'(\mathbf{v})|/|\mathbf{u} - \mathbf{v}| : \mathbf{u}, \mathbf{v} \in U^*\}$ is a stationary and ergodic sequence with finite mean by (5.7), the ergodic theorem implies that

$$\lim_{m \to \infty} \frac{1}{|\mathbf{u} - \mathbf{v}|} |\mathbf{D}_m(\mathbf{u}) - \mathbf{D}_m(\mathbf{v})| = c^*$$

with some constant c^*. Hence by (5.6) we have that

$$\sup_{\mathbf{u} \in U^*} |\mathbf{D}_m(\mathbf{u}) - \mathbf{D}(\mathbf{u})| \to 0 \quad \text{a.s.} \tag{5.15}$$

Berkes, Horváth and Kokoszka (2003) showed that

$$\hat{\boldsymbol{\theta}}_m \to \boldsymbol{\theta} \quad \text{a.s.} \tag{5.16}$$

and therefore the first part of Proposition 3.1 follows from (5.15).

The non—singularity of $\mathbf{D} = \mathbf{D}(\boldsymbol{\theta})$ was proven by Berkes Horváth and Kokoszka (2003) and the positive definiteness of $\mathbf{D}$ is obvious.

6 Proof of Theorem 3.1

The proof of Theorem 3.1 is based on several lemmas which we present after introducing some additional notation.

Let

$$\ell_i(\mathbf{u}) = -\frac{1}{2}\left\{\log w_i(\mathbf{u}) + \frac{y_i^2}{w_i(\mathbf{u})}\right\},$$

where

$$w_i(\mathbf{u}) = c_0(\mathbf{u}) + \sum_{1\le j<\infty} c_j(\mathbf{u}) y_{j-i}^2.$$

We note that

$$\ell_i'(\mathbf{u}) = \frac{1}{2}\left(\epsilon_i^2 \frac{\sigma_i^2}{w_i(\mathbf{u})}\frac{w_i'(\mathbf{u})}{w_i(\mathbf{u})} - \frac{w_i'(\mathbf{u})}{w_i(\mathbf{u})}\right) \tag{6.17}$$

and we define

$$T_m(k,\mathbf{u}) = \sum_{m<i\le m+k} \ell_i'(\mathbf{u}), \qquad 1\le k<\infty.$$

Lemma 6.1 *If the conditions of Theorem 3.1 are satisfied, then*

$$\sup_{1\le k<\infty}\left|\sum_{m<i\le m+k} \hat{\ell}_i'(\hat{\boldsymbol\theta}_m) - T_m(k,\hat{\boldsymbol\theta}_m)\right| \Big/ \left(m^{1/2}\left(1+\frac{k}{m}\right) b\left(\frac{k}{m}\right)\right) \overset{a.s.}{=} O(m^{-1/2}),$$

as $m\to\infty$.

PROOF: By Lemmas 5.8 and 5.9 of Berkes, Horváth and Kokoszka (2003) we have that

$$\sup_{\mathbf{u}\in U}\left|\sum_{1\le i\le n}\left(\hat{\ell}_i'(\mathbf{u}) - \ell_i'(\mathbf{u})\right)\right| \overset{\text{a.s.}}{=} O(1),$$

implying Lemma 6.1.

Let

$$\mathbf{Z}(\mathbf{u}) = E\ell_0''(\mathbf{u}).$$

Lemma 6.2 *If the conditions of Theorem 3.1 are satisfied, then there is* U^**, a neighborhood of* $\boldsymbol\theta$*, such that*

$$\sup_{\mathbf{u}\in U^*}\left|\frac{1}{n}\sum_{1\le i\le n}\ell_i''(\mathbf{u}) - \mathbf{Z}(\mathbf{u})\right| \to 0 \quad a.s.$$

PROOF: This is Lemma 5.6 in Berkes, Horváth and Kokoszka (2003).

Lemma 6.3 *If the conditions of Theorem 3.1 are satisfied, then*

$$\sup_{1\le k<\infty} \frac{\left|\sum_{m<i\le m+k} \ell_i'(\hat{\boldsymbol{\theta}}_m) - \left(\sum_{m<i\le m+k} \ell_i'(\boldsymbol{\theta}) + (\hat{\boldsymbol{\theta}}_m - \boldsymbol{\theta})k\mathbf{Z}(\boldsymbol{\theta})\right)\right|}{m^{1/2}\left(1+\frac{k}{m}\right) b\left(\frac{k}{m}\right)} = o_P(1),$$

as $m \to \infty$.

Proof: First we show that there is a neighborhood of $\boldsymbol{\theta}$, say U^*, such that

$$\sup_{1\le k<\infty}\ \sup_{\mathbf{u}\in U^*} \frac{\left|\sum_{m<i\le m+k} \ell_i''(\mathbf{u}) - k\mathbf{Z}(\mathbf{u})\right|}{m\left(1+\frac{k}{m}\right) b\left(\frac{k}{m}\right)} = o_P(1), \tag{6.18}$$

as $m\to\infty$. Since $\ell_i''(\mathbf{u})$, $\mathbf{u}\in U^*$ is a stationary sequence, by (3.24) it is enough to prove that

$$\sup_{1\le k<\infty}\ \sup_{\mathbf{u}\in U^*} \frac{\left|\sum_{k<i\le m+k} \ell_i''(\mathbf{u}) - k\mathbf{Z}(\mathbf{u})\right|}{m+k} = o_P(1), \quad \text{as } m\to\infty. \tag{6.19}$$

However, (6.19) is an immediate consequence of Lemma 6.2. Theorem 4.4 of Berkes, Horváth and Kokoszka (2003) implies that

$$\left|\hat{\boldsymbol{\theta}}_m - \boldsymbol{\theta}\right| = O_P(m^{-1/2}). \tag{6.20}$$

Using the mean value theorem coordinate–wise for $\sum \ell_i'(\hat{\boldsymbol{\theta}}_m)$ and then (6.18), (6.20) for the coordinates of $\sum \ell_i''(\mathbf{u})$ we get Lemma 6.3.

Lemma 6.4 *If (2.8), (2.9) and (2.13) – (2.19) are satisfied, then*

$$\hat{\boldsymbol{\theta}}_m - \boldsymbol{\theta} = -\frac{1}{m}\sum_{1\le i\le m} \ell_i'(\boldsymbol{\theta})\mathbf{Z}^{-1}(\boldsymbol{\theta})(1+o(1)) \quad a.s.$$

as $m\to\infty$.

Proof: Lemma 6.4 follows from Theorem 4.4 of Berkes, Horváth and Kokoszka (2003).

Lemma 6.5 *If the conditions of Theorem 3.1 are satisfied, then*

$$\sup_{1\le k<\infty} \frac{\left| \sum_{m<i\le m+k} \ell_i'(\hat{\boldsymbol{\theta}}_m) - \left(\sum_{m<i\le m+k} \ell_i'(\boldsymbol{\theta}) - \frac{k}{m} \sum_{1\le i\le m} \ell_i'(\boldsymbol{\theta}) \right) \right|}{m^{1/2}\left(1+\frac{k}{m}\right) b\left(\frac{k}{m}\right)} = o_P(1),$$

as $m\to\infty$.

PROOF: We note that

$$\max_{1<k\le m} \frac{\left| \frac{k}{m} \sum_{1\le i\le m} \ell_i'(\boldsymbol{\theta}) \right|}{m^{1/2}\left(1+\frac{k}{m}\right) b\left(\frac{k}{m}\right)} \le \sup_{1\le k<\infty} \frac{\frac{k}{m}}{1+\frac{k}{m}} \sup_{1\le j<\infty} \frac{1}{b\left(\frac{j}{k}\right)} m^{-1/2} \left| \sum_{1\le i\le m} \ell_i'(\boldsymbol{\theta}) \right| = O_P(1),$$

by conditions (3.23), (3.24) and the asymptotic normality of $m^{-1/2}\sum_{1\le i\le m}\ell_i'(\boldsymbol{\theta})$ (cf. Berkes, Kokoszka and Horváth (2003)). Hence putting together Lemmas 6.3 and 6.4 we get the result in Lemma 6.5.

Lemma 6.6 *If the conditions of Theorem 3.1 are satisfied, then*

$$\sup_{1\le k<\infty} \frac{\left| \sum_{m<i\le m+k} \ell_i'(\boldsymbol{\theta}) - \frac{k}{m} \sum_{1\le i\le m} \ell_i'(\boldsymbol{\theta}) \right|}{m^{1/2}\left(1+\frac{k}{m}\right) b\left(\frac{k}{m}\right)} \xrightarrow{\mathcal{D}} \sup_{0<t<\infty} \frac{|W_{\mathbf{D}}(1+t)-(1+t)W_{\mathbf{D}}(1)|}{(1+t)b(t)},$$

as $m\to\infty$, where $W_{\mathbf{D}}(s)$ is a Gaussian process with $EW_{\mathbf{D}}(s)=0$ and $EW_{\mathbf{D}}^T(s)W_{\mathbf{D}}(s')=\min(s,s')\mathbf{D}$.

PROOF: As it is shown in Berkes, Horváth and Kokoszka (2003), $\ell_i'(\boldsymbol{\theta})$ is a stationary ergodic martingale difference sequence; clearly $\mathrm{cov}(\ell_i'(\boldsymbol{\theta}))=\mathbf{D}$. Hence the Cramér–Wold device (cf. Billingsley (1968, p. 206)) yields that for any $T>0$

$$m^{-1/2}\sum_{1\le i\le mt}\ell_i'(\boldsymbol{\theta}) \xrightarrow{\mathcal{D}[0,T]} W_{\mathbf{D}}(t) \text{ as } m\to\infty. \tag{6.21}$$

Hence

$$m^{-1/2}\left\{ \sum_{m<i\le m(1+t)} \ell_i'(\boldsymbol{\theta}) - t\sum_{1\le i\le m}\ell_i'(\boldsymbol{\theta}) \right\} \xrightarrow{\mathcal{D}[0,T]} W_{\mathbf{D}}(1+t)-(1+t)W_{\mathbf{D}}(1) \tag{6.22}$$

for any $T > 0$ as $m \to \infty$. By the Hájek–Rényi–Chow inequality (cf. Chow (1960)) we have

$$\lim_{T\to\infty} \limsup_{m\to\infty} P\left\{ \sup_{mT\le k<\infty} \left| \sum_{1\le i\le m+k} \ell_i'(\boldsymbol{\theta}) \right| \Big/ \left(m^{1/2}\left(1+\frac{k}{m}\right) b\left(\frac{k}{m}\right)\right) \ge x \right\} = 0 \tag{6.23}$$

for any $x > 0$. The coordinates of $W_{\mathbf{D}}(t)$ are Brownian motions, so by the law of the iterated logarithm and (3.24) we have

$$\sup_{T\le t<\infty} \frac{|W_{\mathbf{D}}(1+t)|}{(1+t)b(t)} \to 0 \quad \text{a.s.} \;\; (T\to\infty). \tag{6.24}$$

Lemma 6.6 now follows from (6.21)–(6.24).

Proof of Theorem 3.1: Putting together Lemmas 6.1–6.6 we get that

$$\sup_{1\le k<\infty} \frac{\left| \sum_{m<i\le m+k} \ell_i'\left(\hat{\boldsymbol{\theta}}_m\right) \hat{\mathbf{D}}_m^{-1/2} \right|}{m^{1/2}\left(1+\dfrac{k}{m}\right) b\left(\dfrac{k}{m}\right)} \xrightarrow{\mathcal{D}} \sup_{0<t<\infty} \frac{\left| \left(W_{\mathbf{D}}(1+t) - (1+t)\mathbf{W}_{\mathbf{D}}(1)\right)\mathbf{D}^{-1/2} \right|}{(1+t)b(t)}.$$

Elementary arguments show that

$$E\left\{ \left(W_{\mathbf{D}}(1+t) - (1+t)\mathbf{W}_{\mathrm{D}}(1)\right)\mathbf{D}^{-1/2} \right\}^T \left\{ \left(W_{\mathbf{D}}(1+s) - (1+s)W_{\mathbf{D}}(1)\right) D^{-1/2} \right\}$$

$$= t(1+s)\mathbf{I}_{p+q+1}, \qquad t\le s,$$

where $\mathbf{I}_{p+q+1}$ is the identity matrix in $\mathbb{R}^{p+q+1}$. Computing the covariances one can verify that

$$\left\{ \left(W_{\mathbf{D}}(1+t) - (1+t)\mathbf{W}_{\mathbf{D}}(1)\right)\mathbf{D}^{-1/2}, t\ge 0 \right\}$$

$$\stackrel{\mathcal{D}}{=} \left\{ (1+t)W_1\left(\frac{t}{1+t}\right), \ldots, (1+t)W_{p+q+1}\left(\frac{t}{1+t}\right), t\ge 0 \right\},$$

where $W_1, W_2, \ldots, W_{p+q+1}$ are independent Wiener processes. Hence

$$\sup_{0<t<\infty} \frac{\left| \left(W_{\mathbf{D}}(1+t) - (1+t)W_{\mathbf{D}}(1)\right)\mathbf{D}^{-1/2} \right|}{(1+t)b(t)} \stackrel{\mathcal{D}}{=} \max_{1\le i\le p+q+1} \sup_{0<s<1} \frac{|W_i(s)|}{b\left(\dfrac{s}{1-s}\right)},$$

completing the proof of the Theorem 3.1.

7 Proof of Theorem 3.2

By Proposition 3.1 it is enough to show that

$$\sup_{1\le k<\infty}\left|\sum_{m<i\le m+k}\hat{\ell}_i'(\hat{\boldsymbol{\theta}}_m)\mathbf{D}^{-1/2}\right|\Big/\left(m^{1/2}\left(1+\frac{k}{m}\right)\right)\xrightarrow{P}\infty, \tag{7.25}$$

as $m\to\infty$. Theorem 3.1 yields that

$$\sup_{1\le k\le k^*-m}\left|\sum_{m<i\le m+k}\hat{\ell}_i'(\hat{\boldsymbol{\theta}}_m)\mathbf{D}^{-1/2}\right|\Big/\left(m^{1/2}\left(1+\frac{k}{m}\right)\right)=O_P(1),$$

as $m\to\infty$. Let $\mathbf{X}_n=(\sigma^2_{n+1},\dots,\sigma^2_{n-p+2},y^2_n,\dots,y^2_{n-q+2})^T\in\mathbb{R}^{p+q-1}$ and $\mathbf{d}=(\omega^*,0,\dots,0)^T$. Using (2.9) and (2.10) we get that

$$\mathbf{X}_{n+1}=A^*_{n+1}\mathbf{X}_n+\mathbf{d}\;\text{ if }\;n\ge k^*$$

and induction yields

$$\mathbf{X}_{k^*+k}=A^*_{k^*+k}\cdots A^*_{k^*+1}\mathbf{X}_{k^*}+\left(\sum_{0\le i\le k-1}A^*_{k^*+k}A^*_{k^*+k-1}\cdots A^*_{k^*+k-i}+\mathbf{I}_{p+q-1}\right)\mathbf{d}.$$

Condition (3.25) and the independence of the matrices A^*_j yield that there is a constant $0<\varrho_*<1$ such that

$$|A^*_{k^*+k}\cdots A^*_{k^*+1}\mathbf{X}_{k^*}|=O_P(\varrho_*^k).$$

Thus

$$|y^2_{k^*+k}-z^2_{k^*+k}|=O_P(\varrho_*^k) \tag{7.26}$$

and

$$|\sigma^2_{k^*+k}-\gamma^2_{k^*+k}|=O_P(\varrho_*^k). \tag{7.27}$$

Hence following the proof of Lemma 6.1, one can easily derive from (7.26) and (7.27) that

$$\sup_{k^*-m<k<\infty}\left|\sum_{m<i\le m+k}\left(\hat{\ell}_i'(\hat{\boldsymbol{\theta}}_m)-g_i'(\hat{\theta}_m)\right)\right|\Big/\left(m^{1/2}\left(1+\frac{k}{m}\right)\right)=o_P(1).$$

Using the mean value coordinate–wise and the ergodic theorem we get that

$$\begin{aligned}\sum_{m<i\le m+k}g_i'(\hat{\boldsymbol{\theta}}_m)&=\sum_{m<i\le m+k}g_i'(\boldsymbol{\theta})+(\hat{\boldsymbol{\theta}}_m-\boldsymbol{\theta})k(\bar{g}''(\boldsymbol{\theta})^T+o_P(1))\\&=k(\bar{g}'(\boldsymbol{\theta})+o_P(1)).\end{aligned} \tag{7.28}$$

Choosing any sequence $\tilde{k} = \tilde{k}(m)$ satisfying $\tilde{k}/m^{1/2} \to \infty$ and $\tilde{k} = O(m)$ we get from (7.28)

$$\left| \sum_{m<i\leq m+\tilde{k}} g_i'(\hat{\boldsymbol{\theta}}_m)\mathbf{D}^{-1/2} \right| \Big/ \left(m^{1/2}\left(1+\frac{\tilde{k}}{m}\right)\right)$$

$$= \left| \bar{g}'(\boldsymbol{\theta})\mathbf{D}^{-1/2} + o_P(1) \right| \frac{\tilde{k}}{m^{1/2}\left(1+\dfrac{\tilde{k}}{m}\right)} \xrightarrow{P} \infty,$$

since $|\bar{g}'(\boldsymbol{\theta})\mathbf{D}^{-1/2}| \neq 0$ by (3.29).

8 References

Andreou, E. and Ghysels, E. (2002) Detecting multiple breaks in financial market volatility dynamics. *Journal of Applied Econometrics* **17** 579-600

Berkes, I. and Horváth, L. (2003). The efficiency of the estimators of the parameters in GARCH processes. *Annals of Statistics*, forthcoming

Berkes, I. and Horváth, L. (2003b). Limit results for the empirical process of squared residuals in GARCH models. *Stochastic Processes and their Applications* **105**, 279–298

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

Besseville, M. and Nikifirov, I. V. (1993) *Detection of Abrupt Changes: Theory and Applications.* Prentice Hall, Upper Saddle River.

Billingsley, P. (1968) *Convergence of Probability Measures.* Wiley, New York.

Brodsky, B. E. and Darkhovsky, B. S. (1993). *Nonparametric Methods in Change–Point Problems.* Kluwer, Dordrecht.

Brodsky, B. E. and Darkhovsky, B. S. (2000). *Non–Parametric Statistical Diagnosis.* Kluwer, Dordrecht.

Boldin, M. V. (1998). On residual empirical distribution functions in ARCH models with applications to testing and estimation. *Mitt. Math. Sem. Giessen.* **235** 49–66

Bougerol, P. and Picard, N. (1992a). Strict stationary of generalized autoregressive processes. *Ann. Probab.* **20** 1714-1730.

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

Chow, Y. S. (1960). A martingale inequality and the law of large numbers. *Proc. Amer. Math. Soc.* **11** 107–111.

Chu, C.-S. J (1995). Detecting parameter shift in GARCH models. *Econometric Reviews* **14** 241-266

Chu, C.-S. J., Stinchcombe, M. and White, H. (1996). Monitoring structural change. *Econometrica* **64** 1045–1065.

Csörgő, M. and Horváth, L. (1997). *Limit Theorems in Change–Point Analysis.* Wiley, Chichester.

Csörgő, M. and Révész, P. (1981). *Strong Approximations in Probability and Statistics.* Academic Press, New York.

Gombay, E. (1996). The weighted sequential likelihood ratio. *Canad. J. Statist.* **24** 229–239.

Gombay, E. (2002). Sequential change–point detection and estimation. Preprint.

Gouriéroux, C. (1997) *ARCH Models and Financial Applications.* Springer-Verlag, New York

Horváth, L. Kokoszka, P. and Teyssiére, G. (2001). Empirical process of the squared residuals of an ARCH sequence. *The Annals of Statistics* **29** 445-469

Inclan, C. and Tiao, G. (1994). Use of cumulative sums of squares for retrospective detection of change in variance. *Journal of American Statistical Association* **89** 913-929

Inoue, A. (2001). Testing for distributional change in time series. *Econometric Theory* **17** 156-187

Kim, S. Cho, S. and Lee, S. (2000). On the cusum test for parameter changes in GARCH(1,1) model. *Communications in Statistics - Theory and Methodology* **29** 445-462

Kokoszka, P. and Leipus, R. (1999). Testing for parameter changes in ARCH models *Lithuanian Mathematical Journal* **39** 231-247

Kokoszka, P. and Leipus, R. (2000). Change-point estimation in ARCH models. *Bernoulli* **6** 513-539

Kokoszka, P. and Teyssière, G. (2002). Change point detection in GARCH models: asymptotic and bootstrap tests. Preprint available at `http://stat.usu.edu/~piotr/research.html`

Koul, H. (2002). *Weighted Empirical Processes in Dynamic Nonlinear Models.* Springer

Kulperger, R. and Yu, H. (2003). High moment partial sum processes of residuals in GARCH models. Preprint

Lamoureux, C. and Lastrapes, W. D. (1990). Persistence in variance, structural change, and the GARCH model. *Journal of Business and Economic Statistics.* **8** 225–234.

Lee, S.-W. and Hansen, B. E. (1994). Asymptotic theory for the GARCH$(1,1)$ quasi–maximum likelihood estimator. *Econometric Theory* **10** 29–52.

Lumsdaine, R. L. (1996). Consistency and asymptotic normality of the quasi–maximum likelihood estimator in IGARCH $(1,1)$ and covariance stationary GARCH$(1,1)$ models. *Econometrica* **6** 575–596.

Mikosch, T. and Stărică, C. (2002). Long-range dependence effects and ARCH modeling. *Theory and Applications of Long-range Dependence.*, 439–459, P. Doukhan, G. Oppenheim and M. S. Taqqu, eds. Birkhäuser, Boston

Lundbergh, S. Teräsvirta, T. (2002). Evaluating GARCH models. *Journal of Econometrics* **110** 417-435

Nyblom, Y. (1989). Testing for the constancy of parameters over time. *Journal of the American Statistical Association* **6** 223–230

Sowell, F. (1996). Optimal tests for parameter instability in the generalized method of moments framework. *Econometrica*, **64** 1085-1107

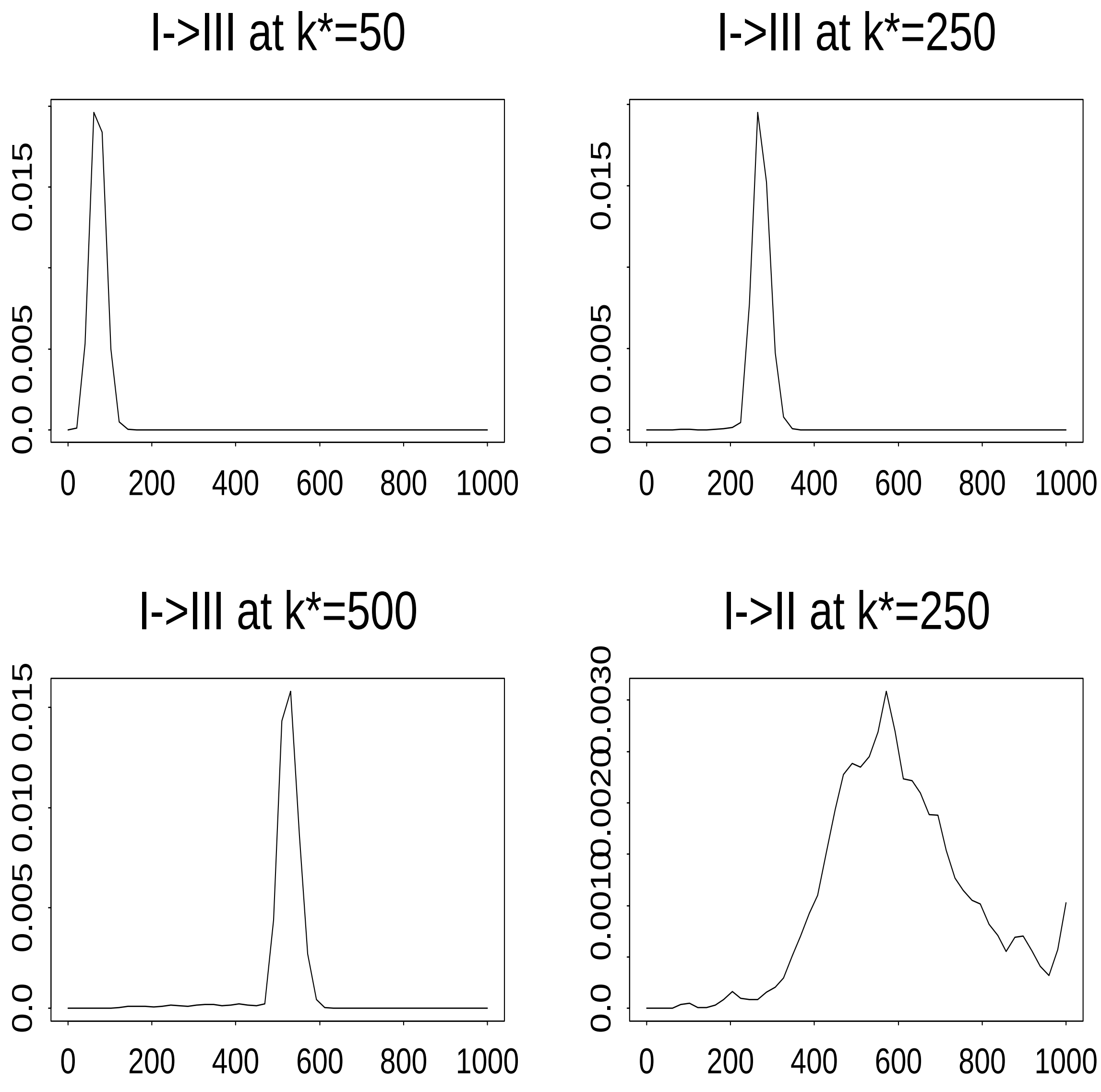

Figure 1: Estimated densities of the first exceedance of the 10 percent critical level. Estimates were obtained using cosine kernel with support of lenght 70. Simulations were done with $m = 1000$ and are based on 1000 replications.

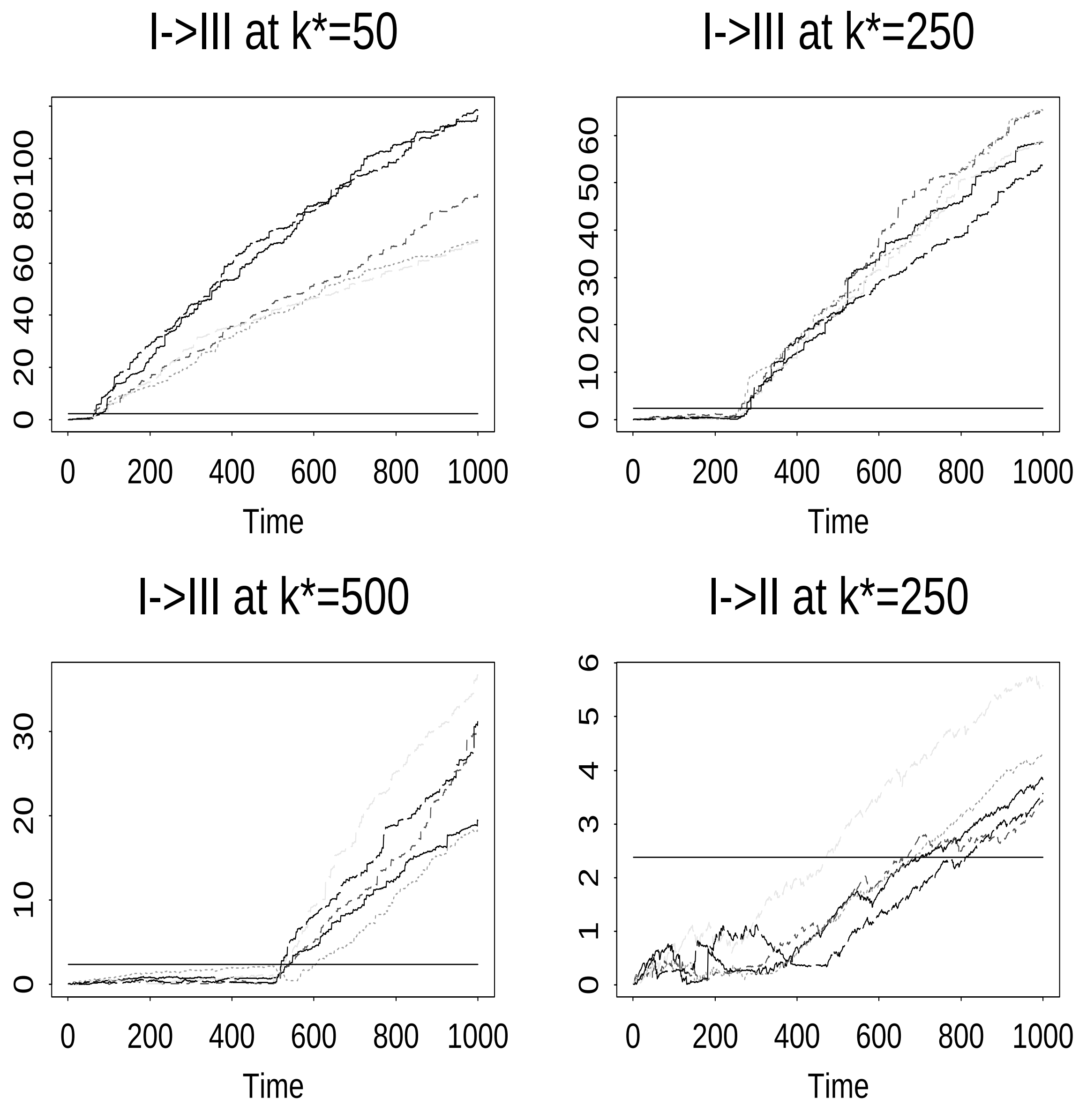

Figure 2: Five randomly selected realizations of the sequence $C(k)$ for the data summarized in Table 4. The vertical lines correspond to 10 percent critical values. The corresponding estimated densities of the first hitting time are depicted in Figure 1.

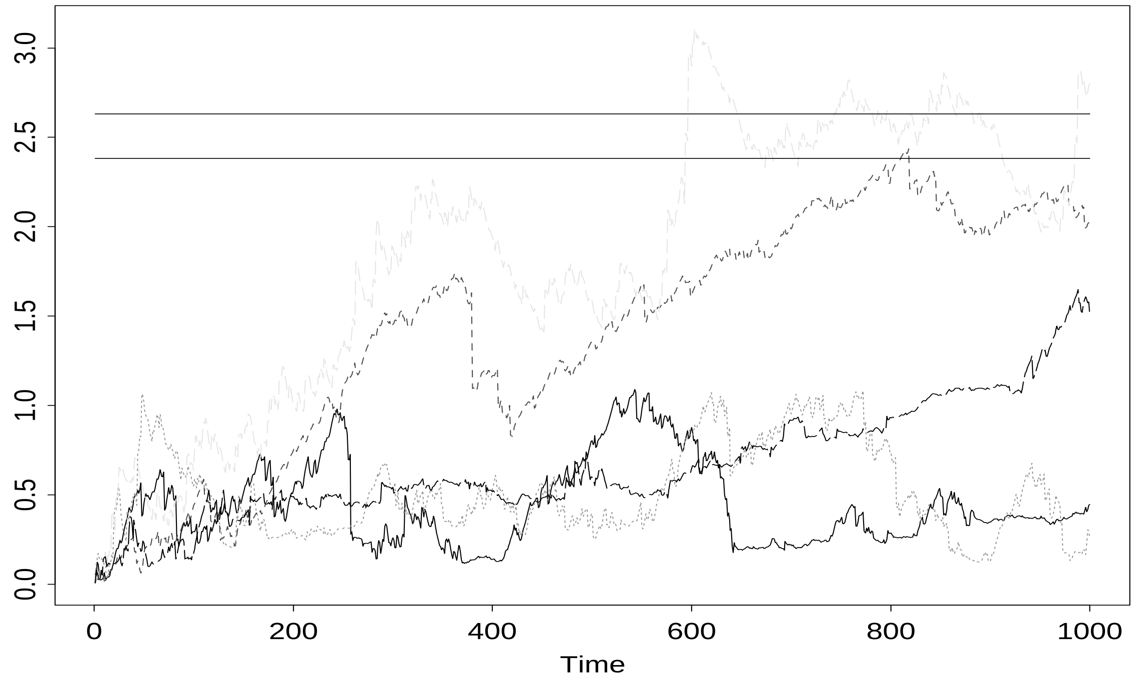

Figure 3: Five randomly selected realizations of the sequence $C(k)$ for Model I and $m = 1000$. The two horizontal lines correspond to five and ten percent critical values from Table 1. The fractions of trajectories crossing these lines before time k are reported in Table 2.