SEQUENTIALLY UPDATED RESIDUALS AND DETECTION OF STATIONARY ERRORS IN POLYNOMIAL REGRESSION MODELS

Ansgar Steland
Institute of Statistics
RWTH Aachen University
Germany

Abstract: The question whether a time series behaves as a random walk or as a stationary process is an important and delicate problem, particularly arising in financial statistics, econometrics, and engineering. This paper studies the problem to detect sequentially that the error terms in a polynomial regression model no longer behave as a random walk but as a stationary process. We provide the asymptotic distribution theory for a monitoring procedure given by a control chart, i.e., a stopping time, which is related to a well known unit root test statistic calculated from sequentially updated residuals. We provide a functional central limit theorem for the corresponding stochastic process which implies a central limit theorem for the control chart. The finite sample properties are investigated by a simulation study.

Keywords: Autoregressive unit root; Change-point; Control chart; Nonparametric smoothing; Sequential analysis; Weighted partial sum process.

Subject Classifications: 62L12; 60G40; 60G50; 62M10; 62E20.

1. INTRODUCTION

Random walks have been proposed as reasonable models for discretely observed data in many disciplines. In engineering, they have been proposed to model production processes with degradation. For instance, the additive damage model assumes that damage cumulates yielding a random walk, and the system fails if the cumulative damage reaches a threshold. We refer to Birnbaum and Saunders (1969), Taguchi (1981, 1985), Taguchi et al. (1989), Adams and Woodall (1989), Doksum and Hóyland (1992), Durham and Padgett (1997), Park and Padgett (2006), Srivastava and Wu (1994, 2003), and the references given therein.

In financial statistics, random walks appear as a model for the (log) prices of an exchange-traded asset. That idea dates back to Bachelier (1900), and nowadays there is an extensive literature on the random walk hypothesis in the empirical finance literature, mainly addressing the question whether the increments are correlated. Random walks have also been proposed as a model for important economic series as the gross domestic product. Therefore, an important problem is to check sequentially whether a time series is compatible with the random walk model or follows an alternative (out-of-control) model under which the series is stationary.

As is well known, a false answer to that question can lead to completely wrong statistical conclusions, since even elementary statistics change their convergence rates and limit distributions. The implications for a rich class of nonparametric kernel control charts covering, e.g., an approximation to the classic EWMA control chart have been discussed in detail in Steland (2004). Another popular approach to monitor both i.i.d. observations and random walks resp. Brownian motions to detect changes in the mean is based on the CUSUM procedure, which is known to be optimal in the sense of Lorden's criterion. We refer to Beibel (1996), Moustakides (1986, 2004, 2007), Ritov (1990), Siegmund (1985), Shiryaev (1996), and to the monograph of Brodsky and Darkhovsky (2000). Having this in mind, it is of particular interest to study sequential monitoring (surveillance) procedures, which are designed to detect departures from the random walk hypothesis as soon as possible.

In this article we investigate a sequential monitoring procedure which is related to a well known unit root test studied in detail by Breitung (2002). To test the unit root null hypothesis against the alternative of stationarity he proposed to use a variance ratio statistic comparing the dispersion of partial sums with the dispersion of the observations. That test statistic is similar to the statistic underlying the so-called variance ratio or KPSS test proposed by Kwiatkowski et al. (1992) to test the inverse testing problem of stationarity against the unit root alternative. Lee and Schmidt (1996) have shown that the KPSS test is also consistent against stationary long-memory alternatives, for a further detailed study we refer to Giraitis et al. (2003). The KPSS test is known to be powerful for many important data generating processes and robust in terms of the type I error rate. For both testing problems (random walk versus stationarity and vice versa) sequential monitoring (surveillance) procedures based on control charts related to the variance ratio statistic have been proposed in Steland (2007a). In that paper the original time series $Y_1, Y_2, \dots$ is monitored. Under mild conditions the asymptotic distributions of the associated stopping times have been established under various in-control and change-point models.

Motivated by promising results from a preliminary study (Steland, 2006), this article considers the more involved and delicate problem to test sequentially whether or not the error terms in a polynomial regression model form a random walk, thus allowing for nonlinear time trends. Assume we observe sequentially a time series $\{Y_t : t \in \mathbb{N}\}$ of real-valued observations satisfying

$$Y_t = m_t + \epsilon_t, \qquad t \in \mathbb{N},$$

with $E(\epsilon_t) = 0$ for all t. In many applications the regression function m is smooth, which motivates to consider polynomials of known degree. Thus, we assume

$$Y_t = \beta_0 + \beta_1 t + \cdots + \beta_p t^p + \epsilon_t, \qquad t \in \mathbb{N}, \tag{1.1}$$

where $\boldsymbol{\beta} = (\beta_0, \dots, \beta_p)' \in \mathbb{R}^{p+1}$ are unknown regression coefficients and $p \in \mathbb{N}_0$. Basically, the aim is to detect a departure from the in-control model that the error terms form a random walk in favor of a stationary process. Note that the model covers the case that before the change a Brownian motion with polynomial drift, $\xi(t) = \mu(t) + \sigma B(t)$, where B denotes standard Brownian motion, $\sigma > 0$ is a constant, and $\mu(t) = \sum_{j=0}^{p} \beta_j t^j$, is discretely sampled at time instances $t = 1, 2, \ldots, q-1$. In this case $\epsilon_t = \sigma B(t) \sim N(0, \sigma^2 t)$, i.e., the variance is a linear function of time. After the change we observe $\xi(t) = \mu(t) + \sigma B(q) + \eta(t)$, $t = q, q+1, \ldots$, where $\eta(t)$ is a stationary process; e.g. given by a continuous-time moving average,

$$\eta(t) = \int_{-\infty}^{t} \varphi(t-s)\, dB(s),$$

for some function φ with $\int \varphi^2(t)\, dt < \infty$. Our results allow for substantially more general error sequences.

Since for many practical applications the most important alternative model is a (polynomial) time trend with stationary errors, we will apply a control chart (stopping time) providing a signal, if there is evidence that the errors are no longer compatible with the random walk hypothesis. We provide the asymptotic distribution theory under the in-control model that the error terms behave as a random walk but allow for an unknown polynomial time trend. Further, we establish results under a change-point model where the errors form a stationary process after an unknown change-point. Since our results provide the asymptotic distribution of the stopping time, one may design a surveillance procedure according to various criteria. Particularly, our results allow to design the procedure to guarantee a specified asymptotic significance level (type I error rate). If we get a signal, the classic polynomial regression model with stationary errors can be regarded as statistically confirmed, which is an attractive property for many applications.

We study the intuitive approach to calculate the least squares residuals and to apply an appropriate monitoring procedure to these residuals. In sequential analysis, recursive residuals are often used, see the classic paper by Brown et al. (1975), and Sen (1982), mainly because they are fast to compute. However, having in mind contemporary computing facilities, we introduce *sequentially updated residuals*, where at each step the full set of residuals is calculated. We consider a monitoring procedure with a time horizon T where monitoring stops, because in many real applications it is unrealistic to assume that monitoring can be conducted forever. However, the modifications of the results to allow for infinite monitoring are straightforward and briefly discussed.

The rest of the paper is organized as follows. In Section 2 we specify and discuss the assumptions on the error terms and introduce the proposed procedure and required regularity conditions. The asymptotic results for the process of sequentially updated residuals, the process associated to the proposed control statistic, and for the resulting stopping time, are discussed in detail in Section 3 under the in-control model that the regression errors behave as a random walk. The results are constructive in the sense that explicit representations of the asymptotic error process can be obtained in terms of the moment functions, $\int_0^s t^k B(t)\,dt$, $s \in [0,1]$, $k \in \mathbb{N}$, associated to a standard Brownian motion B, which makes simulation from the limiting processes feasible. Section 4 gives asymptotic results under a change-point model where the behavior changes after a certain fraction of the data from a random walk behavior to a stationary process. We report in Section 5 about a simulation study which examines some finite sample properties of the method. Proofs of the main results of this paper are postponed to appendices.

2. MODEL, ASSUMPTIONS, AND THE METHOD

2.1. **Model and Assumptions.** It remains to specify model (1.1) in detail. We assume that the error terms, $\epsilon_t = Y_t - m_t$, in model (1.1) form an AR(1) model with possibly correlated but weak dependent innovations (for precise assumptions see below), i.e.,

$$\epsilon_t = \rho_t \epsilon_{t-1} + u_t, \qquad t \in \mathbb{N}, \tag{2.1}$$

where $\rho_t \in (-1, 1]$ are unknown parameters. If $\rho_t = \rho = 1$ for all t, $\{Y_t\}$ is a random walk and *integrated of order* 1, $I(1)$. Here and throughout the paper we simply write $\{Y_n\}$ if the index set is clear. For $|\rho| < 1$ stationary solutions of the above equation exist.

We consider the following change-point testing problem. The null hypothesis,

$$H_0 : \rho_t = 1 \text{ for all } t,$$

states that the error terms of the regression model form a random walk, i.e., are integrated of order 1. The alternative $H_1 = \cup_{q \geq 1} H_1^{(q)}$ with

$$H_1^{(q)} : \rho_t = 1,\ t < q,\ \rho_t = \rho,\ t \geq q,\ |\rho| < 1$$

specifies that there exists a change-point such that the subseries $\{\epsilon_t : t \geq q\}$ satisfies stationary AR(1) equations. It is important to note that the method proposed in this paper does not require any specification of an alternative. In Section 4 we introduce a specific change-point model related to this testing problem.

Let us consider an example.

Example 2.1. *Assume $\epsilon_t = p(L)\xi_t$ with ξ_t i.i.d. $N(0, \sigma_\xi^2)$ for some $\sigma_\xi > 0$, $p(z) = \sum_{j=0}^q \alpha_j z^j$ with coefficients $\alpha_j \in \mathbb{R}$, L the lag operator given by $L\epsilon_t = \epsilon_{t-1}$. Suppose that the characteristic polynomial, $1 - p(z)$, has exactly one unit root of multiplicity 1. Then $p^*(z) = p(z)/(1-z)$ can be inverted, and we obtain the representation $(1-L)\epsilon_t = p^*(L)^{-1}u_t$, i.e.,*

$$\epsilon_t = \epsilon_{t-1} + \sum_{j \geq 0} \beta_j u_{t-j},$$

for certain coefficients β_j, see Brockwell and Davis (1991, Sec. 3.3). Thus, MA(q)-models with an unit root appear as a special case for the error terms in model (1.1) under the null hypothesis.

Concerning the error terms $\{u_t\}$ we shall assume the following mild nonparametric regularity condition making precise our understanding of weak dependence.

(E) $\{u_t : t \in \mathbb{N}\}$ is strictly stationary with mean zero and $E|u_1|^2 < \infty$ such that

$$\sum_{t=1}^{\infty} |\mathrm{Cov}\,(u_1, u_{1+t})| < \infty,$$

and satisfies a functional central limit theorem (FCLT), i.e.,

$$T^{-1/2} \sum_{i \leq \lfloor Ts \rfloor} u_i \xrightarrow{w} \eta B(s), \qquad T \to \infty, \tag{2.2}$$

for some constant $0 < \eta < \infty$. Here B denotes a Brownian motion with $B(0) = 0$, and $\xrightarrow{w}$ stands for weak convergence in the Skorohod space $D[0,1]$. Skorohod spaces are briefly discussed at the end of this section.

Remark 2.1. (i) *By the Skorohod-Wichura-Dudley representation theorem (Pollard (1984), Ch. IV.3, Theorem 13), a condition as (2.2) is equivalent to the condition:*

There are Brownian motions B_T, $T \geq 1$, such that

$$\sup_{s\in[0,1]} \big| T^{-1/2} \sum_{i\leq \lfloor Ts \rfloor} u_i - \eta B_T(s) \big| = o_P(1), \qquad T \to \infty.$$

(ii) *Combining model (2.1) with $\rho_t = 1$ for all t under the assumption (E) yields a nonparametric approach to define the $I(1)$-property of a time series.*

As an example satisfying the assumption (E) let us discuss briefly $ARCH(\infty)$ models, a popular parametric class of time series models.

Example 2.2. *Recall that $\{X_t\}$ satisfies ARCH(∞) equations, if there exists a sequence of i.i.d. non-negative random variables $\{\xi_j\}$, such that*

$$X_j = \eta_t \xi_t, \qquad \eta_t = a + \sum_{j=1}^{\infty} b_j X_{t-j},$$

where $a \geq 0, b_j \geq 0$ for $j \in \mathbb{N}$. Suppose now that

$$u_t = \sigma_t e_t$$

where $\{e_t\}$ are i.i.d. with $E(e_t) = 0$ and $E(e_t^2) = 1$. Put $\sigma_t = \eta_t$ and $\xi_t = e_t^2$ to embed the classic ARCH model into the above $ARCH(\infty)$ framework. Giraitis, Kokoszka and Leipus (2001, Example 2.2 and Theorem 2.1) have shown that $\{u_t\}$ satisfies (E) provided $E|e_1|^4 < \infty$ and

$$(E\xi_1^4)^{1/4} \sum_{j=1}^{\infty} b_j < 1.$$

2.2. **Monitoring Procedure.** Our stopping time defining the detector essentially relies on a weighted version of the KPSS test statistic, see Kwiatkowski et al. (1992), Breitung (2002), and Steland (2007a). At each time point $t \leq T$ when a new observation is available, we calculate the full set of residuals $\widehat{\epsilon}_1(t), \dots, \widehat{\epsilon}_t(t)$ using all available observations $Y_1, \dots, Y_t$. Using these sequentially updated residuals, we calculate an appropriately weighted version of the unit root test statistic. Define

$$U_t = \frac{t^{-4} \sum_{i=1}^t \big(\sum_{j=1}^i \widehat{\epsilon}_j(t) \big)^2 K((i-t)/h)}{t^{-2} \sum_{j=1}^t \widehat{\epsilon}_j^2(t)}, \qquad t \geq p+1.$$

In these formulas K, called *kernel*, is a nonnegative function with $\int K(z)\,dz < \infty$. Kernels such that $K(z)$ is decreasing for increasing $|z|$ as the Gaussian kernel or the Epanechnikov kernel given by $z \mapsto (3/4)\mathbf{1}_{[-1,1]}(1-z^2)$, $z \in \mathbb{R}$, have the intuitive appeal that recent summands get higher weights than past ones. However, our main results work under the following weak conditions:

(K1) $\|K\|_\infty < \infty$, $\int K(z)\,dz = 1$, and $\int zK(z)\,dz = 0$.
(K2) K is Lipschitz continuous.

The parameter $h = h_T$ is used as a scaling constant in the kernel and defines the *memory* of the procedure. For instance, if $K(z) > 0$ for $z \in [-1,1]$, and $K(z) = 0$ otherwise, U_t looks back h observations. We will assume that

$$\lim_{T\to\infty} T/h_T = \zeta \in [1,\infty).$$

That condition ensures that the number of observations used by the procedure gets larger as T increases.

The KPSS or variance ratio control chart is defined as

$$R_T = \inf\{k \le t \le T : U_t \le c_R\}, \qquad T \ge k,$$

with the convention $\inf\emptyset = \infty$. T is the time horizon where monitoring stops. For our asymptotic results we assume $T \to \infty$, since for applications approximations of the distribution of R_T for moderate and large time horizons T are of interest. c_R is a control limit (critical value) chosen by the statistician.

It remains to discuss how to choose the control limit c_R. Since monitoring stops latest at time T, we may interpret the stopping time as a hypothesis test with early stopping in favor of the alternative. Thus, one may choose c_R to control asymptotically the type I error rate of a false decision in favor of stationarity, i.e.,

$$\lim_{T\to\infty} P_0(R_T \le T) = \alpha, \tag{2.3}$$

for some given $\alpha \in (0,1)$. Here P_0 indicates that the probability is calculated under the null hypothesis. Alternatively, one may control a conditional version of the in-control average run length (CARL). Note that the stopping time R_T takes values in the set $\{k,\dots,T\} \cup$

$\{\infty\}$, where ∞ represents no signal, which is the preferred event under the in-control model. Now we may choose c_R such that $\mathrm{CARL}_0 = E_0(R_T | R_T < \infty)$ is greater or equal to some given value $\xi \in (k, T)$. Since our results provide the asymptotic distribution of the stopping time R_T, one may also choose the control limit to control other characteristics, e.g., the (conditional) median average run length. For simplicity of exposition we shall assume in the sequel that c_R is chosen such that (2.3) holds.

We will assume that monitoring starts after a certain fraction of the data, i.e.,

$$k = \lfloor T\kappa \rfloor, \qquad \text{for some } \kappa \in (0,1),$$

to avoid that inference is based on too few observations at the beginning. The event $R_T \leq T$ is interpreted as evidence for stationary innovations, and we get that information after R_T observations instead of waiting until time T. If $R_T = \infty$, the random walk hypothesis for the error terms is regarded as compatible with the observed data.

2.3. **Extension to Infinite Time Horizon.** Suppose we observe sequentially an infinite sequence $Y_1, Y_2, \ldots$ and want to monitor this series with the detection rule

$$\inf\{k \leq t < \infty : U_t \leq c_R\}.$$

In this context T is simply used to define an appropriate time scale to determine the bandwidth sequence h_T with $T/h_T \to \zeta$. The FCLT (2.2) is replaced by

$$\{T^{-1/2} \sum_{i \leq \lfloor Ts \rfloor} u_i : s \in [0,\infty)\} \stackrel{w}{\to} \{\eta B(s) : s \in [0,\infty)\},$$

as $T \to \infty$, where convergence takes place in the space $D[0,\infty)$ instead of $D[0,1]$. All limit theorems in this paper are formulated for the time interval $[\kappa, 1]$, i.e., in the space $D[\kappa, 1]$, but are valid for $D[\kappa, z]$ for any fixed $1 < z < \infty$. $X_n \stackrel{w}{\to} X$ in $D[0,\infty)$ is equivalent to

$$g_m(t) X_n(t)|_{[0,m]} \stackrel{w}{\to} g_m(t) X(t)|_{[0,m]}$$

in $D[0,m]$ for each integer m, where $g_m(t) = \mathbf{1}_{[0,m-1]}(t) + (m-t)\mathbf{1}_{[m-1,m]}(t)$, $t \in [0,\infty)$, see Billingsley (1999, Sec. 16) or Pollard (1984, Ch. VI.) Thus the results carry over to $D[\kappa,\infty)$. Hence, there is no loss in generality to consider the time interval $[0,\kappa]$.

2.4. **Skorohod Spaces.** In this paper we will use the notion of weak convergence in the Skorohod space $D([0,1]^2;\mathbb{R}^k)$. Denote the Skorohod space of cadlag functions $[0,1] \to \mathbb{R}$ by $D[0,1] = D([0,1];\mathbb{R})$. Compared to $D[0,1]$ the space $D([0,1]^2;\mathbb{R}^k)$ has been only rarely used in the literature. Therefore, we close this section with a brief exposition of the most important definitions and facts.

Recall that a sequence $\{X, X_n\}$ of random elements with values in a metric space converges weakly, denoted by $X_n \xrightarrow{w} X$, as $n \to \infty$, if $Eh(X_n) \to Eh(X)$, $n \to \infty$, for all measureable real functions which are bounded and continuous w.r.t. the metric. For a detailed classic treatment of these issues we refer to Billingsley (1999).

Equip $D[0,1]$ with the Skorohod metric d yielding a complete and separable metric space. For $p \in \mathbb{N}$ let $D_{\mathbb{R}^p}[\kappa,1] = D([\kappa,1];\mathbb{R}^p)$ denote the space of all cadlag functions $[\kappa,1] \to \mathbb{R}^p$ which we equip with the metric $d_p(f,g) = \sum_{i=1}^p d(f_i,g_i)$, $f = (f_1,\dots,f_p)'$, $g = (g_1,\dots,g_p)'$, $f_i, g_i : [0,1] \to \mathbb{R}$, $i = 1,\dots,p$. The treatment of its generalization to the index set $[0,1]^2$, i.e., $D([0,1]^2;\mathbb{R})$, is more subtle. Let us briefly recall some facts about this function space and weak convergence of sequences of $D([0,1]^2;\mathbb{R})$-valued random elements, as studied by Straf (1970), Bickel and Wichura (1971), and Neuhaus (1971). The space $D([0,1]^2;\mathbb{R})$ can be defined as the uniform closure of the vector subspace of all simple functions, i.e., linear combinations of functions of the form $t \mapsto 1_{E_1 \times E_2}(t)$ where each E_i is either a left-closed, right-open subinterval of $[0,1]$, or the singleton $\{1\}$. Here the closure is taken in the space of all bounded functions $[0,1]^2 \to \mathbb{R}$. For functions $f,g \in D([0,1]^2;\mathbb{R})$ an appropriate metric, $d_2(f,g)$, is defined as the smallest $\varepsilon > 0$ such that there exist continuous bijections $\lambda_1,\lambda_2 : [0,1] \to [0,1]$ with $\|\lambda - \mathrm{id}\,\|_\infty \le \varepsilon$ and $\|f - g \circ \lambda\|_\infty \le \varepsilon$. Here $\lambda = (\lambda_1,\lambda_2)$ and $g \circ \lambda(r,s) = g(\lambda_1(r),\lambda_2(s))$ for $(r,s) \in [0,1]^2$. A sequence $\{f_n\} \subset D([0,1]^2;\mathbb{R})$ converges to some $f \in D([0,1]^2;\mathbb{R})$ iff there exists some sequence $\{\lambda_n\}$ of pairs of continuous bijections $[0,1] \to [0,1]$ such that $\|f_n \circ \lambda_n - f\|_\infty \to 0$ and $\|\lambda - \mathrm{id}\,\|_\infty \to 0$, as $n \to \infty$. Further, if $f \in C([0,1]^2;\mathbb{R})$ convergence in the Skorohod metric implies uniform convergence, since in this case f is uniformly continuous. It turns out that $(D([0,1]^2;\mathbb{R}), d_2)$ is a separable metric space, a common framework to define weak convergence of probability measures and random elements.

3. ASYMPTOTIC RESULTS FOR INTEGRATED PROCESSES

This section is devoted to a detailed study of the proposed procedure unter the null hypothesis that the error terms of the regression model behave as a random walk. Our approach is to represent the KPSS control chart as an inf-functional of the stochastic process associated to the sequence $\{U_t\}$. That process turns out to be a functional of the stochastic process associated to the residuals up to negligible terms. We provide functional central limit theorems for these processes and a central limit theorem for the stopping time R_T.

We need some notations. Let $\mathbf{X}_n$ denote the design matrix for a polynomial regression of order p with intercept based on n observations, i.e.,

$$\mathbf{X}_n = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ 1 & 2 & \cdots & 2^p \\ \vdots & \vdots & & \vdots \\ 1 & n & \cdots & n^p \end{bmatrix} = [\mathbf{x}_1, \ldots, \mathbf{x}_n]',$$

where

$$\mathbf{x}_t = (1, t, \ldots, t^p)'.$$

Define for $p+1 \leq t \leq T$ the random vectors

$$\begin{aligned} \boldsymbol{\epsilon}_t &= (\epsilon_1, \ldots, \epsilon_t)' \\ \widehat{\boldsymbol{\epsilon}}_t &= (\mathbf{I}_t - \mathbf{X}_t(\mathbf{X}_t'\mathbf{X}_t)^{-1}\mathbf{X}_t')\mathbf{Y}_t, \end{aligned}$$

with $\mathbf{Y}_t = (Y_1, \ldots, Y_t)'$. $\mathbf{I}_t$ denotes the t-dimensional identity matrix.

3.1. **Residual Process without Updating.** Let us first consider the natural process associated to the sequence $\widehat{\epsilon}_1, \ldots, \widehat{\epsilon}_T$ of residuals, where at time t the current residual $\widehat{\epsilon}_t$ is simply added to the residuals $\widehat{\epsilon}_i$, $\lfloor T\kappa \rfloor \leq i < t$. Here the former residuals are *not* updated. In the sequel $\lfloor Ts \rfloor$ stands for the current time point. The stochastic process associated to $\widehat{\epsilon}_1, \ldots, \widehat{\epsilon}_T$ is given by

$$\widehat{E}_T(s) = T^{-1/2}\widehat{\epsilon}_{\lfloor Ts \rfloor}, \qquad s \in [\kappa, 1],$$

where $\widehat{\epsilon}_t = 0$ for $0 \le t < p+1$, and

$$\widehat{\epsilon}_{\lfloor Ts \rfloor} = Y_{\lfloor Ts \rfloor} - \mathbf{x}'_{\lfloor Ts \rfloor}(\mathbf{X}'_{\lfloor Ts \rfloor}\mathbf{X}_{\lfloor Ts \rfloor})^{-1}\mathbf{X}'_{\lfloor Ts \rfloor}\mathbf{Y}_{\lfloor Ts \rfloor}$$

is the last coordinate of the vector $\widehat{\boldsymbol{\epsilon}}_{\lfloor Ts \rfloor} = (\widehat{\epsilon}_1, \ldots, \widehat{\epsilon}_{\lfloor Ts \rfloor})'$.

We have to introduce the weighting matrix

$$\mathbf{W}_t = \operatorname{diag}(1, t^{-1}, \ldots, t^{-p}), \qquad t \in \mathbb{N},$$

to take into account the order of the polynomial regressors.

Lemma 3.1. *Fix $\kappa \in (0,1)$. Assume (E). Then*

$$T^{-3/2}\mathbf{W}_{\lfloor Ts \rfloor}\mathbf{X}'_{\lfloor Ts \rfloor}\boldsymbol{\epsilon}_{\lfloor Ts \rfloor} \xrightarrow{w} \eta \int_0^s (1, r/s, \ldots, (r/s)^p)' B(r)\, dr, \qquad \text{in } D_{\mathbb{R}^p}[\kappa, 1],$$

as $T \to \infty$, where the limit is almost surely (a.s.) continuous.

Lemma 3.1 plays a crucial role in the proofs of the main results, but it is also interesting in its own right. Notice that $\mathbf{X}'_{\lfloor Ts \rfloor}\boldsymbol{\epsilon}_{\lfloor Ts \rfloor}$ is the natural sufficient statistic when the errors are i.i.d. normal. The lemma states that for random walk error terms with weak dependent increments the correct scaling operator for the natural sufficient statistic is given by $T^{-3/2}\mathbf{W}_{\lfloor T \cdot \rfloor}$ to obtain a non-degenerate distributional limit. The limit process is given by the vector of weighted integrals of Brownian motion, $\eta \int_0^s (r/s)^k B(r)\, dr$, $k = 0, \ldots, p$, where the integral is a Riemann integral. The factor η summarizes the impact of the correlation of the increments.

Let us introduce the Hilbert matrix of dimension $p+1$ given by

$$\mathbf{H} = (1/(i+j-1))_{i,j \in \{1,\ldots,p+1\}}.$$

It is known that its inverse, $\mathbf{H}^{-1}$, has entries

$$(\mathbf{H}^{-1})_{i,j} = (-1)^{i+j}(i+j-1)\binom{p+i}{p+1-j}\binom{p+j}{p+1-i}\binom{i+j-2}{i-1}^2,$$

see Choi (1983).

We need the following simple result about sufficient conditions for uniform convergence of the inverse of a sequence of invertible matrix-valued functions $\mathbf{A}_n(x)$, $\mathbf{A}_n : \mathbb{R} \to \mathbb{R}^{l \times l}$, to the inverse of its limit $\mathbf{A}(x)$. Let $\|\cdot\|$ denote the Euclidean vector and matrix norm, respectively.

Lemma 3.2. *Suppose* $\{\mathbf{A}(x), \mathbf{A}_n(x) : n \geq 1\}$, *is a sequence of k-dimensional matrix-valued functions such that*

$$\sup_x \|\mathbf{A}_n(x) - \mathbf{A}(x)\|_2 = o(1).$$

Suppose

$$0 < \inf_x \sigma_1(x) \quad \text{and} \quad \sup_x \sigma_k(x) < \infty, \tag{3.1}$$

where $\sigma_1(x)$ ($\sigma_k(x)$) denotes the smallest (largest) eigenvalue of $\mathbf{A}(x)^*\mathbf{A}(x)$. Then

$$\sup_x \|\mathbf{A}_n^{-1}(x) - \mathbf{A}^{-1}(x)\|_2 = o(1).$$

Theorem 3.1. *Fix $\kappa \in (0,1)$. Assume (E). Then, under the null hypothesis H_0,*

$$\widehat{E}_T \xrightarrow{w} \mathcal{E}, \qquad \text{in } D[\kappa, 1],$$

as $T \to \infty$, where the a.s. continuous process $\mathcal{E}$ is given by

$$\mathcal{E}(s) = \eta \left\{ B(s) - s^{-1}\mathbf{1}'\mathbf{H}^{-1} \int_0^s (1, r/s, \dots, (r/s)^p)' B(r)\, dr \right\},$$

for $s \in [\kappa, 1]$.

This theorem provides an explicit formula for the limit process of $\widehat{E}_T$. The limit process is a linear function of Brownian motion $B(s)$ and the limit process appearing in Lemma 3.1.

3.2. **Sequentially updated Residual Process.** Again, $\lfloor Ts \rfloor$ denotes the current time point and $\lfloor Tr \rfloor$ stands for another time point, in most cases a previous one. Let us now consider the two-parameter stochastic process

$$\widehat{E}_{\lfloor Tr \rfloor}(\lfloor Ts \rfloor) = T^{-1/2}\widehat{\epsilon}_{\lfloor Tr \rfloor}(\lfloor Ts \rfloor), \qquad r \in [\kappa, s],\ s \in [\kappa, 1],$$

where for $p+1 \leq k \leq t \leq T$ we denote by $\widehat{\epsilon}_k(t)$ the k-th residual associated to the observation Y_k, calculated using the data $Y_1, \dots, Y_t$. This means, having observed the nth observation, *all* residuals are updated. We call $\widehat{E}_{\lfloor Tr \rfloor}(\lfloor Ts \rfloor)$ the *sequentially updated residual* *process.* Extend the definition by putting $\widehat{E}_{\lfloor Tr \rfloor}(\lfloor Ts \rfloor) = 0$ if $r > s$ or $r, s \in [0, \kappa)$ to obtain a $D([0,1]^2)$-valued process. Note that

$$\widehat{\epsilon}_{\lfloor Tr \rfloor}(\lfloor Ts \rfloor) = Y_{\lfloor Tr \rfloor} - \mathbf{x}'_{\lfloor Tr \rfloor} \widehat{\boldsymbol{\beta}}_{\lfloor Ts \rfloor},$$

with

$$\begin{aligned} \widehat{\boldsymbol{\beta}}_{\lfloor Ts \rfloor} &= (\mathbf{X}'_{\lfloor Ts \rfloor} \mathbf{X}_{\lfloor Ts \rfloor})^{-1} \mathbf{X}'_{\lfloor Ts \rfloor} \boldsymbol{\epsilon}_{\lfloor Ts \rfloor} \\ \boldsymbol{\epsilon}_{\lfloor Ts \rfloor} &= (\epsilon_1, \ldots, \epsilon_{\lfloor Ts \rfloor})'. \end{aligned} \tag{3.2}$$

Theorem 3.2. *Fix $\kappa \in (0,1)$ and assume (E). We have under H_0*

$$\widehat{E}_{\lfloor Tr \rfloor}(\lfloor Ts \rfloor) \xrightarrow{w} \mathcal{E}(r, s),$$

in $D([\kappa, 1]^2; \mathbb{R})$, as $T \to \infty$, where the process $\mathcal{E}$ is given by

$$\mathcal{E}(r, s) = \eta \left\{ B(r) - \mathbf{v}(r, s) s^{-1} \mathbf{H}^{-1} \int_0^s \mathbf{v}(u, s) B(u) \, du \right\} \tag{3.3}$$

with

$$\mathbf{v}(r, s) = (1, r/s, \ldots, (r/s)^p)', \tag{3.4}$$

for $\kappa \le r \le s \le 1$.

Notice that the limit process for sequentially updated residuals has a similar structure as for residuals without updating, but the vector functions appearing in the definition of $\mathcal{E}(r, s)$ now depend on both r and s. Again, the influence of the dependence structure of the error terms is summarized by the factor η.

Example 3.1. *Explicit representations of the limit processes are now easy to obtain. Let us consider dimensions $p = 1$ and $p = 2$, which are of special importance for applications.*

(i) *For $p = 1$ we have $\mathbf{H}^{-1} = \begin{bmatrix} 4 & -6 \\ -6 & 12 \end{bmatrix}$ and*

$$s^{-1} \mathbf{H}^{-1} \left[\int_0^s B(r) \, dr, \int_0^s r B(r) \, dr \right]' = \begin{pmatrix} \frac{4}{s} \int_0^s B(r) \, dr - \frac{6}{s^2} \int_0^s r B(r) \, dr \\ -\frac{6}{s} \int_0^s B(r) \, dr + \frac{12}{s^2} \int_0^s r B(r) \, dr \end{pmatrix}.$$

Thus,

$$\mathcal{E}(r,s) = \eta \left\{ B(r) + \left(\frac{6r}{s^2} - \frac{4}{s} \right) \int_0^s B(u)\,du + \left(\frac{6}{s^2} - \frac{12r}{s^3} \right) \int_0^s uB(u)\,du \right\},$$

for $r,s \in [\kappa,1]$, $r \leq s$.

(ii) *If* $p=2$, $\mathbf{H}^{-1} \int_0^s [1, r/s, r^2/s^2]' B(r)\,dr$ *is given by*

$$\begin{pmatrix} 9\int_0^s B(r)\,dr - \frac{36}{s}\int_0^s rB(r)\,dr + \frac{30}{s^2}\int_0^s r^2 B(r)\,dr \\ -36\int_0^s B(r)\,dr + \frac{192}{s}\int_0^s rB(r)\,dr - \frac{180}{s^2}\int_0^s r^2 B(r)\,dr \\ 30\int_0^s B(r)\,dr - \frac{180}{s}\int_0^s rB(r)\,dr + \frac{180}{s^2}\int_0^s r^2 B(r)\,dr \end{pmatrix}$$

We obtain

$$\begin{aligned} \mathcal{E}(r,s) &= \eta \Bigg\{ B(s) - \left(\frac{9}{s} - \frac{36r}{s^2} + \frac{30r^2}{s^3} \right) \int_0^s B(r)\,dr \\ &\quad - \left(-\frac{36}{s^2} + \frac{192r}{s^3} - \frac{180r^2}{s^4} \right) \int_0^s rB(r)\,dr \\ &\quad - \left(\frac{30}{s^3} - \frac{180r}{s^4} + \frac{180r^2}{s^5} \right) \int_0^s r^2 B(r)\,dr \Bigg\}, \end{aligned}$$

for $r,s \in [\kappa,1]$, $r \leq s$.

Remark 3.1. *Based on these explicit formulas, simulating trajectories of the process* $\mathcal{E}(r,s)$ *becomes a feasible task. Using Donsker's theorem one may simulate trajectories of* $B(r)$ *and employ numerical integration to simulate the moment functions* $\int_0^s r^k B(r)\,dr$, $s \in [\kappa,1]$, $k \in \mathbb{N}$, *appearing in the formulae.*

3.3. **Weighted Variance Ratio Process.** We are now in a position to examine the process associated to the sequence of control statistics $\{U_t\}$. For brevity of exposition we present the results for the sequentially updated residuals. The required modifications when using the sequential residuals without updating are straightforward.

Define the *kernel-weighted variance ratio process*

$$V_T(s) = \frac{\lfloor Ts \rfloor^{-4} \sum_{i=g_T}^{\lfloor Ts \rfloor} \left(\sum_{j=1}^{i} \widehat{\epsilon}_j(\lfloor Ts \rfloor) \right)^2 K((i - \lfloor Ts \rfloor)/h)}{\lfloor Ts \rfloor^{-2} \sum_{j=g_T}^{\lfloor Ts \rfloor} \widehat{\epsilon}_j^2(\lfloor Ts \rfloor)}, \qquad s \in [0,1].$$

Here and in the sequel we agree to put $0/0 = 0$. g_T denotes the time point where calculations start. To ensure both that the residuals can be calculated and the sums appearing in the definition of $V_T(s)$ have a reasonable number of summands for all $s \in [\kappa, 1]$, we assume $p+1 \le g_T < \lfloor T\kappa \rfloor$. A plausible choice is

$$g_T = \lfloor T\gamma \rfloor, \qquad \text{for some } \gamma \in (0, \kappa).$$

Then $g_T/T \to \gamma > 0$. More generally, let

$$\gamma = \lim_{T\to\infty} g_T/T \in [0, \kappa]. \tag{3.5}$$

The stopping time R_T can now be represented as

$$R_T = T \inf\{s \in [\kappa, 1] : V_T(s) \le c\}.$$

We are now in a position to formulate the main result.

Theorem 3.3. *Fix $\kappa \in (0, 1)$ and assume (E). Under H_0 we have*

$$V_T(s) \xrightarrow{w} \mathcal{V}(s) = \frac{s^{-2} \int_\gamma^s \left(\int_0^r \mathcal{E}(s,t)\, dt \right)^2 K(\zeta(r-s)) dr}{\int_\gamma^s \mathcal{E}^2(s, r)\, dr}, \qquad T \to \infty,$$

in the space $D[\kappa, 1]$. The limit process is continuous w.p. 1 and depends only on K, ζ, and Brownian motion B, but not on the quantity η.

We discuss this theorem at the end of this section in greater detail.

3.4. **KPSS (Variance Ratio) Residual Control Chart.** The central limit theorem for the stopping time R_T of the KPSS residual control chart appears now as a corollary to Theorem 3.3.

Corollary 3.1. *For the stopping time R_T we have under the conditions of Theorem 3.3*

$$R_T/T \xrightarrow{d} \mathcal{R} = \inf\{s \in [\kappa, 1] : \mathcal{V}(s) \le c_R\},$$

as $T \to \infty$.

As a consequence, the KPSS residual control chart can be designed to achieve a given nominal significance level $\alpha \in (0, 1)$. Indeed, Corollary 3.1 implies that $P_0(R_T \le T) \to P_0(\mathcal{R} \le 1)$, as $T \to \infty$. Since

$$\mathcal{R} \le 1 \Leftrightarrow \inf_{s\in[\kappa,1]} \mathcal{V}(s) \le c_R,$$

we select the control limit as

$$c_R = F^{-1}(1-\alpha),$$

where F denotes the distribution function of $\inf_{s\in[\kappa,1]} \mathcal{V}(s)$.

Remark 3.2. *Having in mind practical applications it is worth discussing the following issues.*

(i) *V_T converges weakly to a stochastic process which does not depend on any nuisance parameter. When a kernel K and the parameter ζ are selected, the process $\mathcal{V}$ is known. This means, the asymptotic law of V_T is distribution-free. As a consequence, the asymptotic distribution of R_T is also asymptotically distribution-free.*

(ii) *In practice, one can simulate trajectories from the limit process and calculate for each trajectory the time point where the control limit c_R is reached. In this way one can simulate the asymptotic distribution of R_T to determine a control limit c_R such that the resulting asymptotic type I error rate is α.*

4. ASYMPTOTIC RESULTS FOR A CHANGE-POINT MODEL

The results of the previous section allow to design monitoring procedures and to study the behavior of the resulting procedure under the null hypothesis (in-control model) that the underlying time series of observations follows a polynomial regression model with random walk error terms under the stated regularity assumptions.

In this section we discuss the asymptotic behavior of the KPSS residual monitoring approach under a change-point model, where the first part of the time series behaves as a random walk and the second part is stationary. We assume

$$\epsilon_t = \begin{cases} \sum_{j=0}^{t} u_j, & t = 0,\ldots,\lfloor T\vartheta\rfloor - 1, \\ \xi_T u_t, & t = \lfloor T\vartheta\rfloor,\ldots,T. \end{cases} \tag{4.1}$$

After the change-point $q = \lfloor T\vartheta\rfloor$, which is given by the fixed but unknown parameter $\vartheta \in (0,1)$, the error terms change and are no longer a random walk. $\{\xi_T\}$ is a sequence of scale constants satisfying the condition

$$\xi_T \sim T^{\beta}, \qquad \text{for some } \beta \geq 0. \tag{4.2}$$

We shall need a further constraint on β which will be discussed below. If $\beta = 0$, the error process after the change, i.e., $\{\epsilon_t : q \le t \le T\}$ with $q = \lfloor T\vartheta \rfloor$, is stationary. However, we allow for positive values of β. In this case the error terms form a row-wise stationary array. For simplicity of exposition we omit the dependence of ϵ_t on T in our notation.

Our asymptotic results require the following additional assumptions.

(C1) $\{u_t\}$ is a strictly stationary process with

$$\lim_{x \to \infty} \frac{P(|u_1| > x)}{x^{-\gamma}} < \infty,$$

for some $\gamma > 2$ and satisfies the FCLT

$$T^{-1/2} \sum_{i \le \lfloor Ts \rfloor} u_i \xrightarrow{w} \eta B(s), \qquad T \to \infty, \tag{4.3}$$

for some constant $0 < \eta < \infty$, where again B denotes standard Brownian motion starting at 0.

(C2) The parameters α and β satisfy the relations $0 \le \beta < 1/2$ and $\gamma > \frac{1}{1/2-\beta}$.

Note that the condition on the tail probabilities ensures that the $E|u_t|^2 < \infty$.

In the sequel we use the same notation for the quantities defined for the polynomial regression model with error terms $\{\epsilon_t\}$ satisfying the change-point model above.

Let us again start with the residual process. We only discuss the FCLT for the process of sequentially updated residuals, $\widehat{E}_{\lfloor Tr \rfloor}(\lfloor Ts \rfloor)$, $\kappa \le r \le s \le 1$, which is defined as before.

Theorem 4.1. *Suppose the change-point model (4.1) holds. Additionally, assume that (E), (C1), and (C2) are satisfied. Then, for any fixed $\kappa \in (0,1)$, the following assertions hold true.*

(i) *We have in the space $D([\kappa, 1]; \mathbb{R}^p)$,*

$$T^{-3/2} \mathbf{W}_{\lfloor Ts \rfloor} \mathbf{X}'_{\lfloor Ts \rfloor} \boldsymbol{\epsilon}_{\lfloor Ts \rfloor} \xrightarrow{w} \eta \int_0^s (1, r/s, \dots, (r/s)^p)' B(r)\, dr \mathbf{1}_{[\kappa,\vartheta)}(s),$$

as $T \to \infty$.

(ii) *The sequentially updated LS residual process converges weakly,*

$$\widehat{E}_T \xrightarrow{w} \mathcal{E}_\vartheta, \qquad \text{in } D([\kappa, 1]^2; \mathbb{R}),$$

as $T \to \infty$, where the cadlag process $\mathcal{E}$ is given by

$$\mathcal{E}_\vartheta(r,s) = \eta \left\{ B(s) - s^{-1}\mathbf{v}(r,s)\mathbf{H}^{-1} \int_0^s \mathbf{v}(u,s)B(u)\,du \right\} \mathbf{1}_{[\kappa,\vartheta)}(s),$$

for $\kappa \le r \le s \le 1$.

The next result shows that under the aforementioned conditions the asymptotic distribution of the kernel-weighted variance ratio process is obtained by replacing formally $\mathcal{E}$ by $\mathcal{E}_\vartheta$ in the limit process.

Theorem 4.2. *Suppose the change-point model (4.1), assumption (E), (C1) and (C2) are satisfied. Then, for any fixed $\kappa \in (0,1)$,*

$$V_T(s) \xrightarrow{w} \mathcal{V}_\vartheta(s) = \frac{s^{-2}\int_\gamma^s \left(\int_0^r \mathcal{E}_\vartheta(s,t)\,dt\right)^2 K(\zeta(r-s))dr}{\int_\gamma^s \mathcal{E}_\vartheta^2(s,r)\,dr}, \qquad T \to \infty,$$

in the space $D[\kappa, 1]$. The limit process depends only on K, ζ, and Brownian motion B, and the change-point parameter ϑ.

Again, the central limit theorem for the KPSS residual control chart under the change-point model appears as a corollary.

Corollary 4.1. *Under the assumptions of Theorem 4.2 the stopping time R_T satisfies*

$$R_T/T \xrightarrow{d} \mathcal{R}_\vartheta = \inf\{s \in [\kappa, 1] : \mathcal{V}_\vartheta(s) > c_R\},$$

as $T \to \infty$.

5. SIMULATIONS

We conducted a Monte Carlo study to investigate the properties of the KPSS monitoring procedure when applied to residuals. Time series of length $T = 500$ according to model

$$Y_t = \beta_0 + \beta_1 \cdot t + (\beta_1 + \Delta)t\mathbf{1}_{\{q,q+1,\dots\}}(t) + \epsilon_t,$$

where

$$\epsilon_t = \begin{cases} \sum_{i=1}^t \eta_i, & \text{for } t < q, \\ \sum_{i=1}^{q-1} \eta_i + \eta_t, & \text{for } q \le t \le T, \end{cases}$$

with

$$\eta_t = \rho\eta_{t-1} + \xi_t - \beta\xi_{t-1}, \qquad \xi_t \overset{i.i.d.}{\sim} N(0,1),$$

were simulated. Let us first discuss the construction of the innovation terms η_t. The AR parameter was chosen as $\rho = 0.3$ and the MA parameter β from the set $\{-0.8, 0, 0.8\}$. Thus, $\{\eta_t\}$ is a correlated but weakly dependent sequence with mean 0. For time points $t < q$ the obervations Y_t are given by a random walk with correlated increments η_i. At the change-point q the process changes its behavior. The random walk stops and correlated error terms η_t determine the behavior of ϵ_t.

Concerning the design of the monitoring procedures we used the Gaussian kernel, $K(z) = (2\pi)^{-1}\exp(-z^2/2)$, $z \in \mathbb{R}$, and the bandwidths $h \in \{25, 50\}$, yielding $\zeta \in \{20, 10\}$. The deterministic component of the model is given by a linear trend whose slope, depending on Δ, may change at the change-point, too. If $\Delta \neq 0$, there is both a change in the error terms and a change in the slope. That should make the detection of the change to stationarity of the errors more difficult, since the residuals are estimated assuming a constant slope.

In a first step we examined for the setting $\Delta = 0$, $h = 25$, and $T = 500$, the relationship between the control limit c and, firstly, the probability that the method gives a signal (Figure 1) and, secondly, the conditional average run length (CARL) given that the method gives a signal at all (Figure 2). Since monitoring stops latest at the 500th observation, trajectories crossing the control limit later are not taken into account. The CARL is the average run length corresponding to all trajectories yielding a signal until time 500. The curves, which can also be used to choose the control limit, are quite similar for $\beta \in \{-0.8, 0\}$, but there is an effect for positive values of β. For the considered setting it also becomes apparent that common type I error rates correspond to rather large CARL values. On the other side, if the procedure is designed to yield CARL values of, say, 300, the chart works on a type I error rate which is usually regarded as unacceptable from a hypothesis testing viewpoint. However, note that this is partly due to the fact that we studied monitoring with a time horizon. Without a time horizon the average run lengths would be substantially higher yielding smaller control limits and, as a consequence, smaller associated type I error rates.

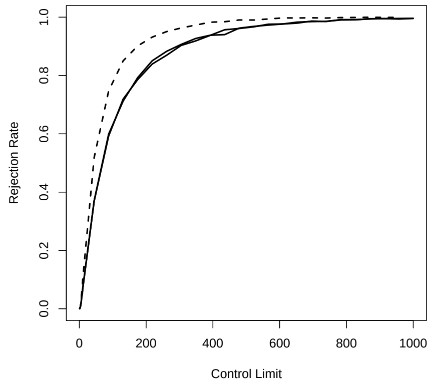

Figure 1. Empirical rejection rates as a function of 10^6 times the control limit c for $\beta = -0.8$, $\beta = 0$, and $\beta = 0.8$ (dashed).

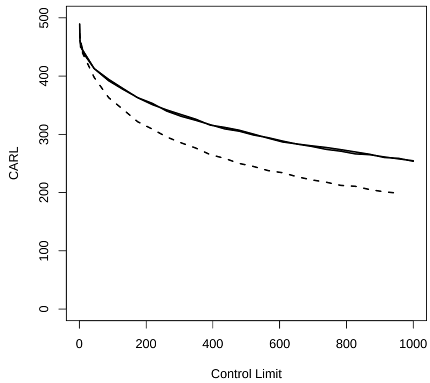

Figure 2. Conditional average run length (CARL) as a function of 10^6 times the control limit c for $\beta = -0.8$, $\beta = 0$, and $\beta = 0.8$ (dashed).

We also simulated the power of the KPSS variance ratio residual control chart when designed to achieve a type I error rate of $\alpha = 0.05$. The corresponding control limit was obtained by simulating from the limit distributions. We examine the cases $\Delta = 0$ and $\Delta = 0.5$, where the latter case corresponds to a change to stationary errors term with an additional change of the slope.

Table 1 provides the simulated rejection rates. It can be seen that the KPSS control chart is quite robust with respect to the parameter β determining the degree of correlation for the increments of the random walk. That behavior is consistent with the theoretical and empirical findings in Steland (2007a), where related monitoring procedures for time series without trends have been studied in detail. The results show that an early change can be detected quite well, but late changes are very hard to detect. However, this is, of course, a problem for all methods, and for the statistical problem at hand, $N = 500$ is not a large maximal number of observations. The results also indicate that the power is quite robust with respect to additional changes in slope.

We may summarize that the KPSS control chart for residuals provides a quite reliable tool to detect stationary errors in polynomial regression models.

Table 1. Empirical rejection rates of the KPSS control chart. The right columns ($\Delta = 0.25$) correspond to a change of the slope

	$\Delta = 0$			$\Delta = 0.25$		
		β			β	
change-point	-0.8	0	0.8	-0.8	0	0.8
Results for $h = 25$						
25	0.44	0.60	0.94	0.45	0.57	0.94
75	0.16	0.18	0.41	0.18	0.19	0.38
100	0.17	0.16	0.29	0.16	0.16	0.29
no-change	0.06	0.06	0.07	0.06	0.06	0.10
Results for $h = 50$						
25	0.52	0.61	0.97	0.53	0.60	0.97
75	0.18	0.19	0.44	0.17	0.18	0.44
100	0.15	0.16	0.30	0.15	0.16	0.33
no-change	0.03	0.03	0.06	0.03	0.03	0.07

ACKNOWLEDGMENTS

The author gratefully acknowledges a grant (STE 1034/6-1) from Deutsche Forschungsgemeinschaft (DFG). He also thanks S. Teller for her careful proofreading, H. Satvat and H. Schottmüller for building up a Linux cluster for parallel computing which was used for simulations, and B. Giese for preparing the final Latex layout. This is a preprint of an article forthcoming in *Sequential Analysis*.

APPENDIX A: PROOFS OF RESULTS FROM SECTION 3

A.1. Proof of Lemma 3.1

Note that for each $s \in [\kappa, 1]$ we have

$$\mathbf{X}'_{\lfloor Ts \rfloor}\boldsymbol{\epsilon}_{\lfloor Ts \rfloor} = \left(\sum_{t=1}^{\lfloor Ts \rfloor} \epsilon_t, \sum_{t=1}^{\lfloor Ts \rfloor} t\epsilon_t, \dots, \sum_{t=1}^{\lfloor Ts \rfloor} t^p \epsilon_t \right)'$$

yielding

$$\begin{aligned} T^{-3/2}\mathbf{W}_{\lfloor Ts \rfloor}\mathbf{X}'_{\lfloor Ts \rfloor}\boldsymbol{\epsilon}_{\lfloor Ts \rfloor} &= T^{-3/2} \left(\sum_{t=1}^{\lfloor Ts \rfloor} (t/\lfloor Ts \rfloor)^{i-1}\epsilon_t \right)_{i=1,\dots,p+1} \\ &= T^{-1/2} \left(\int_0^s (\lfloor Tr \rfloor/\lfloor Ts \rfloor)^{i-1}\epsilon_{\lfloor Tr \rfloor}\,dr \right)_{i=1,\dots,p+1} \qquad \text{(A.1)} \\ &= \left(\int_0^s (\lfloor Tr \rfloor/\lfloor Ts \rfloor)^{i-1} T^{-1/2}\epsilon_{\lfloor Tr \rfloor}\,dr \right)_{i=1,\dots,p+1}. \end{aligned}$$

It is straightforward to check that

$$\sup_{\kappa \le r \le s \le 1} \max_{1 \le i \le p} |(\lfloor Tr \rfloor/\lfloor Ts \rfloor)^i - (r/s)^i| = O(1/T). \tag{A.2}$$

Hence,

$$\sup_{0 \le r \le s \le 1} \max_{1 \le i \le p} \Big| \int_0^s (\lfloor Tr \rfloor/\lfloor Ts \rfloor)^{i-1} z(r)\,dr - \int_0^s (r/s)^{i-1} z(r)\,dr \Big| = O(1/T).$$

If we define the functional $\tau : (D[\kappa, 1], d) \to (D_{\mathbb{R}^p}[\kappa, 1], d_p)$ by

$$\tau(z)(s) = \left(\int_0^s (r/s)^{i-1} z(r)\,dr \right)_{i=1,\dots,p+1}, \qquad s \in [\kappa, 1],$$

for any $z \in D[\kappa, 1]$, we obtain

$$T^{-3/2}\mathbf{W}_{\lfloor Ts \rfloor}\mathbf{X}'_{\lfloor Ts \rfloor}\boldsymbol{\epsilon}_{\lfloor Ts \rfloor} = \tau(T^{-1/2}\epsilon_{\lfloor T\cdot \rfloor})(s) + o_P(1), \tag{A.3}$$

the $o_P(1)$ being uniform in $s \in [\kappa, 1]$. It is easy to see that for any sequence $\{z, z_n\} \subset D_{\mathbb{R}^p}[\kappa, 1]$ such that $\lim_{n\to\infty} d_p(z_n, z) = 0$, as $n \to \infty$, with $z \in C[\kappa, 1]$, we have

$$\lim_{n\to\infty} d_p(\tau(z_n), \tau(z)) = 0.$$

Thus, the continuous mapping theorem in general separable metric spaces (Shorack and Wellner (1986), Th. 4, p. 47, and Remark 2, p. 49) and (E) yield the result. □

A.2. Proof of Lemma 3.2

Let $\mathrm{cond}_2(\mathbf{A}(x)) = \sigma_k(x)/\sigma_1(x)$ denote the condition of $\mathbf{A}(x)$ w.r.t. the spectral vector norm $\|\cdot\|_2$. Let $\varepsilon > 0$. If $\|\mathbf{A}_n(x) - \mathbf{A}(x)\|_2 < \varepsilon$, the a-priori error estimate for linear equations with disturbed coefficient matrices yields

$$\|a_{nj}^{-1}(x) - a_j^{-1}\|_2 \leq \frac{\mathrm{cond}_2(\mathbf{A}(x))}{\|\mathbf{A}(x)\|_2 - \varepsilon \mathrm{cond}_2(\mathbf{A}(x))} \varepsilon \|\mathbf{a}_j(x)^{-1}\|,$$

where $a_{nj}^{-1}(x)$ ($a_j^{-1}(x)$) denotes the jth column of $\mathbf{A}_n^{-1}(x)$ ($\mathbf{A}^{-1}(x)$). □

A.3. Proof of Theorem 3.1

Recall the representations

$$\widehat{\boldsymbol{\beta}}_{\lfloor Ts \rfloor} - \boldsymbol{\beta} = (\mathbf{X}'_{\lfloor Ts \rfloor}\mathbf{X}_{\lfloor Ts \rfloor})^{-1}\mathbf{X}'_{\lfloor Ts \rfloor}\boldsymbol{\epsilon}_{\lfloor Ts \rfloor} \qquad \text{and} \qquad \widehat{\epsilon}_{\lfloor Ts \rfloor} = \epsilon_{\lfloor Ts \rfloor} - \mathbf{x}'_{\lfloor Ts \rfloor}(\widehat{\boldsymbol{\beta}}_{\lfloor Ts \rfloor} - \boldsymbol{\beta}),$$

where $\widehat{\boldsymbol{\beta}}_{\lfloor Ts \rfloor}$ is defined in (3.2) yielding

$$\begin{aligned}\widehat{\epsilon}_{\lfloor Ts \rfloor} &= \epsilon_{\lfloor Ts \rfloor} - \mathbf{x}'_{\lfloor Ts \rfloor}(\mathbf{X}'_{\lfloor Ts \rfloor}\mathbf{X}_{\lfloor Ts \rfloor})^{-1}\mathbf{X}'_{\lfloor Ts \rfloor}\boldsymbol{\epsilon}_{\lfloor Ts \rfloor} \\ &= \epsilon_{\lfloor Ts \rfloor} - T^{1/2}\mathbf{x}'_{\lfloor Ts \rfloor}\mathbf{W}_{\lfloor Ts \rfloor}(\mathbf{W}_{\lfloor Ts \rfloor}T^{-1}\mathbf{X}'_{\lfloor Ts \rfloor}\mathbf{X}_{\lfloor Ts \rfloor}\mathbf{W}_{\lfloor Ts \rfloor})^{-1}T^{-3/2}\mathbf{W}_{\lfloor Ts \rfloor}\mathbf{X}'_{\lfloor Ts \rfloor}\boldsymbol{\epsilon}_{\lfloor Ts \rfloor}.\end{aligned}$$

Since $\mathbf{x}'_{\lfloor Ts \rfloor}\mathbf{W}_{\lfloor Ts \rfloor} = (1, \lfloor Ts \rfloor, \ldots, \lfloor Ts \rfloor^p)\mathbf{W}_{\lfloor Ts \rfloor} = \mathbf{1}'$ where $\mathbf{1} = (1, \ldots, 1)' \in \mathbb{R}^{p+1}$, we have

$$T^{-1/2}\widehat{\epsilon}_{\lfloor Ts \rfloor} = T^{-1/2}\epsilon_{\lfloor Ts \rfloor} - \mathbf{1}'\widetilde{\mathbf{H}}_{\lfloor Ts \rfloor}^{-1}T^{-3/2}\mathbf{W}_{\lfloor Ts \rfloor}\mathbf{X}'_{\lfloor Ts \rfloor}\boldsymbol{\epsilon}_{\lfloor Ts \rfloor}$$

where

$$\widetilde{\mathbf{H}}_{\lfloor Ts \rfloor} = \mathbf{W}_{\lfloor Ts \rfloor}T^{-1}\mathbf{X}'_{\lfloor Ts \rfloor}\mathbf{X}_{\lfloor Ts \rfloor}\mathbf{W}_{\lfloor Ts \rfloor}. \tag{A.4}$$

Obviously,

$$\widetilde{\mathbf{H}}_{\lfloor Ts \rfloor} = \left(\frac{\lfloor Ts \rfloor}{T} \lfloor Ts \rfloor^{-(i+j-2)} \sum_{t=1}^{\lfloor Ts \rfloor} t^{i+j-2} \right)_{i,j}.$$

We will show that this matrix converges to $s\mathbf{H}$ uniformly in $s \in [\kappa, 1]$. Recall that

$$\sum_{t=1}^{\lfloor Ts \rfloor} t^{i+j-2} = \frac{(\lfloor Ts \rfloor + 1)^{i+j-2}}{i+j-1} + O((\lfloor Ts \rfloor + 1)^{i+j-3}).$$

Hence

$$\lfloor Ts \rfloor^{i+j-2} \sum_{t=1}^{\lfloor Ts \rfloor} t^{i+j-2} = \frac{1}{i+j-1} + O(1/(\lfloor T\kappa \rfloor + 1)),$$

yielding

$$\sup_{s \in [\kappa,1]} |(\widetilde{\mathbf{H}}_{\lfloor Ts \rfloor})_{ij} - s/(i+j-1)| = O(\lfloor T\kappa \rfloor^{-1}) \tag{A.5}$$

for $i, j \in \{1, \dots, p+1\}$. Recall the representation (A.1) and (A.3). Since $\sup_{s \in [\kappa,1]} \tau(T^{-1/2}\boldsymbol{\epsilon}_{\lfloor T\cdot \rfloor})$ converges weakly to the random variable

$$\sup_{s \in [\kappa,1]} \tau(\eta B)(s) = \sup_{s \in [\kappa,1]} \left(\eta \int_0^s (r/s)^{i-1} B(r)\, dr \right)_{i=1,\dots,p+1},$$

we may conclude that

$$\sup_{s \in [\kappa,1]} T^{-3/2} \mathbf{W}'_{\lfloor Ts \rfloor} \mathbf{X}'_{\lfloor Ts \rfloor} \boldsymbol{\epsilon}_{\lfloor Ts \rfloor} = O_P(1). \tag{A.6}$$

(A.5), (A.6), and Lemma 3.2 imply

$$\widehat{E}_T(s) = T^{-1/2} \epsilon_{\lfloor Ts \rfloor} - \mathbf{1}' s^{-1} \mathbf{H}^{-1} T^{-3/2} \mathbf{W}_{\lfloor Ts \rfloor} \mathbf{X}'_{\lfloor Ts \rfloor} \boldsymbol{\epsilon}_{\lfloor Ts \rfloor} + O_P(1/\lfloor Ts \rfloor)$$

Using the result (A.3) we obtain

$$\widehat{E}_T(s) = T^{-1/2} \epsilon_{\lfloor Ts \rfloor} - \mathbf{1}' s^{-1} H^{-1} \tau(T^{-1/2} \boldsymbol{\epsilon}_{\lfloor T\cdot \rfloor}) + o_P(1)$$

which shows that up to terms of order $o_P(1)$ the process $\widehat{E}_T$ is a continuous functional of $\{T^{-1/2}\epsilon_{\lfloor Ts \rfloor} : s \in [\kappa, 1]\}$. Consequently,

$$\widehat{E}_T(s) \xrightarrow{w} \eta \left\{ B(s) - s^{-1} \mathbf{1}' H^{-1} \int_0^s (1, r/s, \dots, (r/s)^p)' B(r)\, dr \right\},$$

in $D[\kappa, 1]$, as $T \to \infty$. □

A.4. Proof of Theorem 3.2

The proof is similar as the proof of Theorem 3.1. We have

$$\widehat{\epsilon}_{\lfloor Tr \rfloor}(\lfloor Ts \rfloor) = \epsilon_{\lfloor Tr \rfloor} - \mathbf{x}'_{\lfloor Tr \rfloor} \mathbf{W}_{\lfloor Ts \rfloor} (\mathbf{W}_{\lfloor Ts \rfloor} \mathbf{X}'_{\lfloor Ts \rfloor} \mathbf{X}_{\lfloor Ts \rfloor} \mathbf{W}_{\lfloor Ts \rfloor})^{-1} \mathbf{W}_{\lfloor Ts \rfloor} \mathbf{X}'_{\lfloor Ts \rfloor} \boldsymbol{\epsilon}_{\lfloor Ts \rfloor}.$$

Note that

$$\mathbf{x}'_{\lfloor Tr\rfloor}\mathbf{W}_{\lfloor Ts\rfloor} = (1, \lfloor Tr\rfloor/\lfloor Ts\rfloor, \dots, (\lfloor Tr\rfloor/\lfloor Ts\rfloor)^p)'$$

and let $\mathbf{v}(r,s) = (1, r/s, \dots, (r/s)^p)'$. Due to (A.2) we have

$$\sup_{0\le r\le s\le 1} \|\mathbf{x}'_{\lfloor Tr\rfloor}\mathbf{W}_{\lfloor Ts\rfloor} - \mathbf{v}(r,s)\| = O(1/T).$$

Combining this fact with (A.3) yields

$$\begin{aligned} T^{-1/2}\widehat{\epsilon}_{\lfloor Tr\rfloor}(\lfloor Ts\rfloor) &= \frac{\epsilon_{\lfloor Tr\rfloor}}{T^{1/2}} - \mathbf{x}'_{\lfloor Tr\rfloor}\mathbf{W}_{\lfloor Ts\rfloor}\widetilde{\mathbf{H}}^{-1}_{\lfloor Ts\rfloor}\boldsymbol{\epsilon}_{\lfloor Ts\rfloor} \\ &= \frac{\epsilon_{\lfloor Tr\rfloor}(\lfloor Ts\rfloor)}{T^{1/2}} - \left\{\mathbf{v}(r,s)s^{-1}\mathbf{H}^{-1} + o_P(1)\right\}\left\{\tau(T^{-1/2}\epsilon_{\lfloor T\circ\rfloor})(s) + o_P(1)\right\}, \end{aligned}$$

where the $o_P(1)$ terms are uniform in $r, s \in [\kappa, 1]$. Hence, uniformly in $r, s \in [\kappa, 1]$,

$$T^{-1/2}\widehat{\epsilon}_{\lfloor Tr\rfloor}(\lfloor Ts\rfloor) = \varphi(T^{-1/2}\epsilon_{\lfloor T\cdot\rfloor})(r,s) + o_P(1),$$

where the functional $\varphi : D[\kappa, 1] \to D([\kappa, 1]^2; R)$ is given by

$$\text{(A.7)} \quad \varphi(z)(r,s) = z(r) - s^{-1}\mathbf{v}(r,s)\mathbf{H}^{-1}\int_0^s (1, u/s, \dots, (u/s)^p)' z(u)\, du, \qquad r, s \in [\kappa, 1],$$

for $z \in D[\kappa, 1]$. It is easy to see that for any sequence $\{z, z_n\} \subset D[\kappa, 1]$ with $d(z_n, z) \to 0$, as $n \to \infty$, and $z \in C[\kappa, 1]$, we have $\|\varphi(z_n) - \varphi(z)\|_\infty \to 0$, as $n \to \infty$. Hence, an application of the continuous mapping theorem yields

$$T^{-1/2}\widehat{\epsilon}_{\lfloor Tr\rfloor}(\lfloor Ts\rfloor) \xrightarrow{w} \varphi(\sigma B) = \eta\left\{B(r) - s^{-1}\mathbf{v}(r,s)\mathbf{H}^{-1}\int_0^s (1, u/s, \dots, (u/s)^p)' B(u)\, du\right\},$$

as $T \to \infty$. □

A.5. Proof of Theorem 3.3

We formulate the proof such that the corresponding result for the change-point model of Section 4 can be obtained by straightforward modifications. To simplify exposition we assume $\gamma = 0$. Note that for any $\lambda_1, \lambda_2 \in \mathbb{R}$ the process

$$W_{\lambda_1,\lambda_2}(s) = \frac{\lambda_1}{\lfloor Ts\rfloor^4}\sum_{i=1}^{\lfloor Ts\rfloor}\Big(\sum_{j=1}^{i}\widehat{\epsilon}_j(\lfloor Ts\rfloor)\Big)^2 K((i - \lfloor Ts\rfloor)/h) + \frac{\lambda_2}{\lfloor Ts\rfloor^2}\sum_{i=1}^{\lfloor Ts\rfloor}\widehat{\epsilon}_i^2(\lfloor Ts\rfloor)$$

can be written as

$$\tau_{\lambda_1,\lambda_2}(\widehat{E}) = \lambda_1\left(\frac{T}{\lfloor Ts\rfloor}\right)^4 \int_0^s \left(\int_0^r \widehat{E}_{\lfloor Tz\rfloor}(\lfloor Ts\rfloor)\,dz\right)^2 K((\lfloor Tr\rfloor - \lfloor Ts\rfloor)/h)\,dr$$
$$+ \lambda_2\left(\frac{T}{\lfloor Ts\rfloor}\right)^2 \int_0^s \widehat{E}^2_{\lfloor Tr\rfloor}(\lfloor Ts\rfloor)\,dr,$$

where $\tau_{\lambda_1,\lambda_2}$ maps elements of $D([0,1]^2)$ to elements of $D[0,1]$. Let us show continuity of $\tau_{\lambda_1,\lambda_2}$ w.r.t. the supnorm. W.l.o.g. we may assume $\|K\|_\infty = 1$. Using the inequality $|a^2 - b^2| \le (|a| + |b|)|a - b|$ for real numbers a, b we can bound $|\tau_{\lambda_1,\lambda_2}(z_1) - \tau_{\lambda_1,\lambda_2}(z_2)|$ by

$$[\lambda_1(T/\lfloor Ts\rfloor)^4\|z_1\|_\infty + \lambda_2(T/\lfloor Ts\rfloor)^2(\|z_1\|_\infty + \|z_2\|_\infty)]\|z_1 - z_2\|_\infty = O(\|z_1 - z_2\|_\infty).$$

Hence, for $0 \le s_1 \le \cdots \le s_L \le 1$, $L \in \mathbb{N}$, any associated linear combination $\sum_{k=1}^L \rho_k W_{\lambda_1,\lambda_2}(s_k)$, $\rho_1, \dots, \rho_L \in \mathbb{R}$, of the coordinates of the random vectors $(W_{\lambda_1,\lambda_2}(s_1), \dots, W_{\lambda_1,\lambda_2}(s_L))$, converges in distribution to $\sum_{k=1}^L \rho_k \tau_{\lambda_1,\lambda_2}(\mathcal{E})(s_k)$, since $\widehat{E}_T \xrightarrow{w} \mathcal{E}$, $T \to \infty$, by Theorem 3.2. This verifies convergence of the finite-dimensional distributions of the $(D[\kappa,1])^2$-valued stochastic process (Z_{T1}, Z_{T2}), where

$$Z_{T1}(s) = \lfloor Ts\rfloor^{-4}\sum_{i=1}^{\lfloor Ts\rfloor}\left(\sum_{j=1}^{i}\widehat{\epsilon}_j(\lfloor Ts\rfloor)\right)^2 K((i - \lfloor Ts\rfloor)/h), \quad Z_{T2}(s) = \lfloor Ts\rfloor^{-2}\sum_{i=1}^{\lfloor Ts\rfloor}\widehat{\epsilon}_i(\lfloor Ts\rfloor)$$

for $s \in [\kappa, 1]$. Tightness w.r.t. the product topology is a consequence of Prohorov's theorem, since both coordinate processes converge weakly. Thus, $(Z_{T1}, Z_{T2}) \xrightarrow{w} (Z_1, Z_2)$, $T \to \infty$, in $(D[\kappa, 1])^2$, where

$$Z_1(s) = s^{-4}\int_0^s \Big(\int_0^r \mathcal{E}(z, s)\,dz\Big)^2 K(\zeta(s - r))\,dr, \quad Z_2(s) = s^{-2}\int_0^s \mathcal{E}^2(r, s)\,dr,$$

for $s \in [\kappa, 1]$. Now a straightforward argument implies that the ratio, V_T, converges weakly, as $T \to \infty$. Finally, by Lipschitz continuity of the kernel K the limit process $\mathcal{V}$ is continuous w.p. 1. □

A.6. Proof of Corollary 3.1

The result is shown using similar arguments as in Steland (2007b, Corollary 2.1.) □

APPENDIX B: PROOFS OF RESULTS FROM SECTION 4

By virtue of the method of proof used in the previous section, we are in a position to extend the results for the kernel weighted variance ratio process and its associated stopping time to the change-point model of Section 3, if we have a FCLT for the process of sequentially updated residuals. Thus we provide a detailed proof of Theorem 4.1 and indicate the required modifications to prove Theorem 4.2.

B.1. Proof of Theorem 4.1

Under the change-point model we have

$$\{T^{-1/2}\epsilon_{\lfloor Ts \rfloor} : \kappa \le s < \vartheta\} \stackrel{w}{\to} \{\eta B(s) : \kappa \le s < \vartheta\},$$

as $T \to \infty$. Consider the process $T^{-1/2}\epsilon_{\lfloor Ts \rfloor}$ for $\vartheta \le s \le 1$. First note that

$$T^{-1/2}\epsilon_{\lfloor Ts \rfloor} \le \sup_{z \in [\kappa,1]} |T^{1/2}\epsilon_{\lfloor Tz \rfloor}|.$$

Let $\delta > 0$. By assumptions (C1) and (C2)

$$\begin{aligned}
P\left(\sup_{z \in [\vartheta,1]} |T^{-1/2}\epsilon_{\lfloor Tz \rfloor}| > \delta \right) &= P\left(\max_{t = \lfloor T\vartheta \rfloor, \dots, T} |u_t| > T^{1/2}\delta/\xi_T \right) \\
&\le (T - \lfloor T\vartheta \rfloor + 1) P(|u_1| > T^{1/2}\delta/\xi_T) \\
&= O(T^{1-\gamma(1/2-\beta)}) \\
&= o_P(1).
\end{aligned}$$

if $\beta < 1/2$ and $\gamma > (1/2 - \beta)^{-1}$. Again using the Skorohod-Dudley-Wichura representation theorem we may assume that

$$\sup_{s \in [\kappa,\vartheta]} |T^{-1/2}\epsilon_{\lfloor Ts \rfloor} - \eta B(s)| \stackrel{a.s.}{\to} 0,$$

and

$$\sup_{s \in [\kappa,\vartheta]} |T^{-1/2}\epsilon_{\lfloor Ts \rfloor}| \stackrel{a.s.}{\to} 0,$$

as $T \to \infty$. This implies a.s. convergence in the Skorohod metric to the cadlag process $B\mathbf{1}_{[\kappa,\vartheta)}$, i.e.,

$$d(T^{-1/2}\epsilon_{\lfloor T\cdot \rfloor}, \eta B\mathbf{1}_{[\kappa,\vartheta)}) \stackrel{a.s.}{\to} 0,$$

as $T \to \infty$, which in turn yields weak convergence,

$$T^{-1/2}\epsilon_{\lfloor Ts \rfloor} \xrightarrow{w} \eta B(s)\mathbf{1}_{[\kappa,\vartheta)}(s),$$

in $D[\kappa, 1]$, as $T \to \infty$. Combining this fact with (A.3), the continuity of the functional τ (Jacod and Shiryaev (2003), VI, Proposition 1.22, p. 329) yields

$$T^{-3/2}\mathbf{W}_{\lfloor Ts \rfloor}\mathbf{X}'_{\lfloor Ts \rfloor}\boldsymbol{\epsilon}_{\lfloor Ts \rfloor} \xrightarrow{w} \eta \int_0^s (1, u/s, \dots, (u/s)^p)' B(r)\, dr \mathbf{1}_{[\kappa,\vartheta)}(s),$$

as $T \to \infty$. The same arguments as in the proof of Theorem 3.2 show that

$$\widehat{E}_{\lfloor Tr \rfloor}(\lfloor Ts \rfloor) = T^{-1/2}\widehat{\epsilon}_{\lfloor Tr \rfloor}(\lfloor Ts \rfloor) = \varphi(T^{-1/2}\epsilon_{\lfloor T\cdot \rfloor})(r, s) + o_P(1),$$

as $T \to \infty$, where the functional φ is defined in (A.7). We have by linearity

$$\begin{aligned}
\varphi(T^{-1/2}\epsilon_{\lfloor T\cdot \rfloor})(r, s) &= \left[T^{-1/2}\epsilon_{\lfloor Ts \rfloor} - s^{-1}\mathbf{v}(r, s)\mathbf{H}^{-1} \int_0^s \mathbf{v}(u, s) T^{-1/2}\epsilon_{\lfloor Tu \rfloor}\, du \right] \\
&= \left[\eta B(s) - s^{-1}\mathbf{v}(r, s)\mathbf{H}^{-1} \int_0^s \mathbf{v}(u, s)\eta B(u)\, du \right] \mathbf{1}_{[\kappa,\vartheta)}(s) \\
&\quad + R_{T1}(s) - R_{T2}(r, s) \\
&= \varphi(\eta B \mathbf{1}_{[\kappa,\vartheta)})(r, s) + R_1(s) + R_2(r, s),
\end{aligned}$$

where the remainder terms are given by

$$\begin{aligned}
R_{T1}(s) &= T^{-1/2}\epsilon_{\lfloor Ts \rfloor} - \eta B(s)\mathbf{1}_{[\kappa,\vartheta)}(s), \\
R_{T2}(r, s) &= s^{-1}\mathbf{v}(r, s)\mathbf{H}^{-1} \int_0^s \mathbf{v}(u, s)[T^{-1/2}\epsilon_{\lfloor Tu \rfloor} - \eta B(u)]\mathbf{1}_{[\kappa,\vartheta)}(s)\, du.
\end{aligned}$$

Clearly, $\sup_{s\in[\kappa,\vartheta)} |R_{T1}(s)| \to 0$, as $T \to \infty$, a.s. To estimate R_{T2}, denote the maximum vector norm and the induced matrix norm by $\| \circ \|_\infty$ and observe that

$$\begin{aligned}
&\left\| \int_0^s \mathbf{v}(u, s)[T^{-1/2}\epsilon_{\lfloor Tu \rfloor} - \eta B(u)]\mathbf{1}_{[\kappa,\vartheta)}(s)\, du \right\|_\infty \\
&\qquad \le \int_0^s \|\mathbf{v}\|_\infty \sup_{z\in[\kappa,\vartheta)} |T^{-1/2}\epsilon_{\lfloor Tz \rfloor} - \eta B(z)| \mathbf{1}_{[\kappa,\vartheta)}(s)\, du \\
&\qquad \le \|\mathbf{v}\|_\infty \vartheta \sup_{z\in[\kappa,\vartheta)} |T^{-1/2}\epsilon_{\lfloor Tz \rfloor} - \eta B(z)| \xrightarrow{a.s.} 0,
\end{aligned}$$

where $\|\mathbf{v}\|_\infty = \sup_{r,s\in[\kappa,\vartheta)} \|\mathbf{v}(r, s)\|_\infty < \infty$. Hence,

$$\sup_{s\in[\kappa,\vartheta)} |R_{T2}(r, s)| \le \kappa^{-1}\|\mathbf{v}\|_\infty^2 \|\mathbf{H}^{-1}\|_\infty \vartheta \sup_{z\in[\kappa,\vartheta]} |T^{-1/2}\epsilon_{\lfloor Tz \rfloor} - \eta B(z)|\, du \xrightarrow{a.s.} 0,$$

as $T \to \infty$. Consequently,

$$\varphi(T^{-1/2}\epsilon_{\lfloor T\cdot \rfloor})(r,s) \overset{a.s.}{\to} \varphi(\eta B \mathbf{1}_{[\kappa,\vartheta)}),$$

as $T \to \infty$, which implies via

$$d(\varphi(\epsilon_{\lfloor T\cdot \rfloor}), \varphi(\eta B \mathbf{1}_{[\kappa,\vartheta)})) \overset{a.s.}{\to} 0,$$

as $T \to \infty$, weak convergence which completes the proof. $\square$

B.2. Proof of Theorem 4.2

To proof goes along the lines of the proof of Theorem 3.3. Notice that now the linear combinations $\sum_{k=1}^{L} \rho_k W_{\lambda_1,\lambda_2}(s_k)$ converge weakly in distribution to $\sum_{k=1}^{L} \rho_k \tau_{\lambda_1,\lambda_2}(\mathcal{E}_\vartheta)(s_k)$. $\square$

References

[1] Adam, B. M. and Woodall, W. H. (1989). An Analysis of Taguchi's On-Line Process-Control Procedure Under a Random-Walk Model, *Technometrics* 31: 401-413.

[2] Bachelier, L. (1900). Théorie de la Spéculation, *Annales Scientifiques de l'École Normale Supérieure* 17: 21-86.

[3] Beibel, M. (1996). A Note on Ritov's Bayes Approach to the Minimax Property of the CUSUM Procedure, *Annals of Statistics* 24: 1804-1812.

[4] Bickel, P. J. and Wichura, M. J. (1971). Convergence Criteria for Multiparameter Stochastic Processes and Some Applications, *Annals of Statistics* 42: 1656-1670.

[5] Birnbaum, Z. W. and Saunders, S. C. (1969). A New Family of Life Distributions. *Journal of Applied Probability* 6: 319-327.

[6] Billingsley, P. (1999). *Convergence of Probability Measures*, New York: Wiley.

[7] Breitung, J. (2002). Nonparametric Tests for Unit Roots and Cointegration, *Journal of Econometrics* 108: 343-363.

[8] Brockwell, P. J. and Davis, R. A. (1991). *Time Series: Theory and Methods*, 2nd edition, New York: Springer.

[9] Brodsky, B. E. and Darkhovsky, B. S. (2000). *Non-Parametric Statistical Diagnosis: Problems and Methods*, Dordrecht: Kluwer.

[10] Brown, R. L., Durbin, J., and Evans, J. M. (1975). Techniques for Testing Constancy of Regression Relationships over Time, *Journal of Royal Statistical Society, Series B* 37: 149-163.

[11] Choi, M. D. (1983). Tricks and Treats with the Hilbert Matrix, *American Mathematician Monthly* 90: 301-312.

[12] Durham, S. D. and Padgett, W. J. (1997). A Cumulative Damage Model for System Failure with Application to Carbon Fibers and Composites, *Technometrics* 39: 34-44.

[13] Doksum, K. A. and Hóyland, A. (1992). Models for Variable-Stress Accelerated Life Testing Experiments Based on Wiener Processes and the Inverse Gaussian Distribution. *Technometrics* 34: 74-82.

[14] Giraitis, L., Kokoszka, P., and Leipus, R. (2001). Testing for Long Memory in the Presence of a General Trend, *Journal of Applied Probabilty* 4: 1033-1054.

[15] Giraitis, L., Kokoszka, P., Leipus, R., and Teyssiere, G. (2003). Rescaled Variance and Related Tests for Long Memory in Volatility and Levels, *Journal of Econometrics* 112: 265-294.

[16] Jacod, J. and Shiryaev, A. N. (2003). *Limit Theorems for Stochastic Processes*, Berlin: Springer.

[17] Kwiatkowski, D., Phillips, P. C. B., Schmidt, P., and Shin, Y. (1992). Testing the Null Hypothesis of Stationary Against the Alternative of a Unit Root: How Sure Are We That Economic Time Series Have a Unit Root? *Journal of Econometrics* 54: 159-178.

[18] Lee, D. and Schmidt, P. (1996). On the Power of the KPSS Test of Stationarity against Fractionally-Integrated Alternatives, *Journal of Econometrics* 73: 285-302.

[19] Moustakides, G. V. (1986). Optimal Stopping Times for Detecting Changes in Distributions, *Annals of Statistics* 14: 1379-1387.

[20] Moustakides, G. V. (2004). Optimality of the CUSUM Procedure in Continuous Time, *Annals of Statistics* 32: 302-316.

[21] Moustakides, G. V. (2007). Sequential Change Detection Revisited, *Annals of Statistics*, in press.

[22] Neuhaus, G. (1971). On Weak Convergence of Stochastic Processes with Multidimensional Time Parameter, *Annals of Statistics* 42: 1285-1295.

[23] Ritov, Y. (1990). Decision Theoretic Optimality of the CUSUM Procedure, *Annals of Statistics* 18: 1464-1469.

[24] Park, C. and Padgett, W. J. (2006). Stochastic Degradation Models with Several Accelerating Variables, *IEEE Transactions on Reliability* 55: 379-390.

[25] Pollard, D. (1984). *Optimal Stopping Rules*, New York: Springer.

[26] Sen, P. K. (1982). Invariance Principles for Recursive Residuals, *Annals of Statistics* 10: 307-312.

[27] Shiryayev, A. N. (1978). *Optimal Stopping Rules*, New York: Springer.

[28] Shorack, G. R. and Wellner, J. A. (1986). *Empirical Processes with Applications to Statistics*, New York: Wiley.

[29] Siegmund, D. (1985). *Sequential Analysis*, New York: Springer.

[30] Srivastava, M. S. and Wu, Y. (1994). On-Line Procedures under the Random Walk Model with Measurement Error and Attribute Observationns, *Canadian Journal of Statistics* 22: 377-386.

[31] Srivastava, M. S. and Wu, Y. (2003). Taguchi's Approach to On-Line Quality Procedure, in *Handbook of Statistics*, vol. 22, R. Khattree and C. R. Rao, eds., pp. 657-694, Amsterdam: Elsevier.

[32] Steland, A. (2004). Random Walks with Drift - A Sequential View, *Journal of Time Series Analysis* 26: 917-942.

[33] Steland, A. (2006). On Monitoring a Sequential Linear LS Residual Process for Integrated Errors, *Proceedings of the 7th Prague Symposium on Asymptotic Statistics and 15th Prague Conference on Information Theory, Statistical Decision Functions and Random Processes*, M. Hušková and M. Janžura, eds., pp. 655-665, Prague: MATFYZPRESS.

[34] Steland, A. (2007a). Monitoring Procedures to Detect Unit Roots and Stationarity, *Econometric Theory* 23: 1108-1135.

[35] Steland, A. (2007b). Weighted Dickey-Fuller Processes for Detecting Stationarity, *Journal of Statistical Planning and Inference* 137: 4011-4030.

[36] Straf, M. L. (1970). Weak Convergence of Stochastic Processes with Several Parameters, *Proceedings of the Sixth Berkley Symposium on Mathematical Statistics and Probability*, L. Le Cam, J. Neyman, and E. L. Scott, eds., pp. 187-221, Berkeley: University of California Press.

[37] Taguchi, G. (1981). *On-Line Quality Control During Production*, Tokyo: Japanese Standard Association.

[38] Taguchi, G. (1985). Quality Engineering in Japan, *Communications in Statistics - Theory & Methods* 14: 2785-2801.

[39] Taguchi, G., Elsayed, E. A., and Hsian, T. (1989). *Quality Engineering in Production Systems*, New York: McGraw-Hill.