Journal of Multivariate Analysis 132 (2014) 58–81

Contents lists available at ScienceDirect

Journal of Multivariate Analysis

journal homepage: www.elsevier.com/locate/jmva

Robust monitoring of CAPM portfolio betas II

Ondřej Chochola [a], Marie Hušková [a], Zuzana Prášková [a], Josef G. Steinebach [b,*]

[a] Charles University in Prague, Faculty of Mathematics and Physics, Department of Probability and Mathematical Statistics, Sokolovská 83, CZ-186 75 Praha 8, Czech Republic

[b] Universität zu Köln, Mathematisches Institut, Weyertal 86-90, D-50 931 Köln, Germany

ARTICLE INFO

Article history:
Received 13 January 2014
Available online 8 August 2014

AMS 2000 subject classifications:
60F17
60G10
60J65
62F35
62L10
62P05

Keywords:
Robust monitoring
Functional capital asset pricing model
Portfolio beta
M-estimate
Change-point detection

ABSTRACT

In this work, we extend our study in Chochola et al. [7] and propose some robust sequential procedure for the detection of structural breaks in a *Functional* Capital Asset Pricing Model (FCAPM). The procedure is again based on M-estimates and partial weighted sums of M-residuals and "robustifies" the approach of Aue et al. [3], in which ordinary least squares (OLS) estimates have been used. Similar to Aue et al. [3], and in contrast to Chochola et al. [7], high-frequency data can now also be taken into account. The main results prove some null asymptotics for the suggested test as well as its consistency under local alternatives. In addition to the theoretical results, some conclusions from a small simulation study together with an application to a real data set are presented in order to illustrate the finite sample performance of our monitoring procedure.

1. Introduction and statistical framework

Main aim of this work is to continue and extend our study in Chochola et al. [7] concerning the robust monitoring of CAPM portfolio betas. The Capital Asset Pricing Model (CAPM), introduced by Sharpe [18] and subsequently modified by many authors (see, e.g. Lintner [14], Merton [15] and others), is a still very popular and widely used model for evaluating the risk of a portfolio of assets with respect to the market risk. However, it is also well-known that the pricing of assets and predictions of risks may be incorrect and misleading if the model parameters $\boldsymbol{\beta}_i$ are varying over time. As in Aue et al. [3], we adopt here the arguments of Ghysels [9] and study a (piecewise) unconditional CAPM, rather than a conditional version of the latter (cf., e.g., Andersen et al. [1] for a comprehensive review), since in many cases misspecified conditional CAPMs tend to produce larger pricing errors. For a more extensive discussion of this fact, we refer to Aue et al. [3], Sections 1 and 2, and the references mentioned therein.

Indeed, contributing to avoid pricing and prediction errors was the main motivation for Aue et al. [3] in constructing a sequential monitoring procedure for the testing of the stability of portfolio betas. The corresponding stopping rules in [3] are based on comparing the (ordinary) least squares estimate (OLS) of the beta from a historical data set (training period) to that from sequentially incoming new observations, and they were able to take high-frequency data into account which is a typical situation in nowadays' market analyses (see also Chochola et al. [7] and the references mentioned therein).

* Corresponding author.
E-mail addresses: chochola@karlin.mff.cuni.cz (O. Chochola), huskova@karlin.mff.cuni.cz (M. Hušková), praskova@karlin.mff.cuni.cz (Z. Prášková), jost@math.uni-koeln.de (J.G. Steinebach).

http://dx.doi.org/10.1016/j.jmva.2014.07.016
0047-259X/

Since OLS estimates may be sensitive with respect to outliers, we tried to "robustify" the Aue et al. [3] approach in [7] by making use of M-estimates instead of least squares estimates and so are able to deal with heavier tail distributions than the OLS procedure. In a first step, however, we confined ourselves there to a study of the CAPM without high-frequency observations. Aim of our present work now is to extend the latter study to the *Functional* Capital Asset Pricing Model (FCAPM) taking also high-frequency observations into account. It will turn out that, even in this more general situation, some moment conditions may be relaxed (cf., e.g., (B.4) below compared to the corresponding assumption in [7]), but that, on the other hand and similar to Aue et al. [3], certain smoothness conditions have to be added concerning the model's intra-day behavior over time (see, e.g., (A.1)–(A.3), (B.5) and (B.7) below).

Note that, via L_p–m-approximability type conditions (cf. (B.4)–(B.5) below), our model is suitable for covering general types of weak dependencies rather than strong dependencies in the sense of long memory. Monitoring procedures in the latter situation are still open for future work. On the other hand, in contrast to [3], our present approach is now applicable to data sets under heavy-tailed (leptocurtic) and contaminated distributions observed at high frequencies, which is certainly more useful in real data applications. The price to pay, however, is that more involved techniques than those used in Chochola et al. [7] are required now and the computational complexity increases as well. Nevertheless, a similar robust sequential monitoring procedure can be constructed for the FCAPM portfolio betas, now also covering a high-frequency situation as described below.

We would like to mention, however, that our focus here is on the methodological and theoretical side, trying to extend the work of Aue et al. [3] by using a robust approach and that of Chochola et al. [7] by including high-frequency situations. Moreover, for the sake of illustration and comparison, we used the same data set as in [3] for our application and a similar setting in the small simulation study of Section 3.

Our statistical framework in the sequel will be as follows. We consider the model

$$\boldsymbol{r}_i(s) = \boldsymbol{\alpha}_i + \boldsymbol{\beta}_i r_{iM}(s) + \boldsymbol{\varepsilon}_i(s), \quad i \in \mathbb{Z},\ s \in [0,1], \tag{1.1}$$

where $\boldsymbol{r}_i(s) = (r_{i,1}(s), \ldots, r_{i,d}(s))^T$ is a d-dimensional vector of (functional) log-returns at (say) "day" i and "intra-day time" s, $r_{iM}(s)$ is the log-return of the market portfolio at day i and time s, and $\boldsymbol{\varepsilon}_i(s) = (\varepsilon_{i,1}(s), \ldots, \varepsilon_{i,d}(s))^T$ are d-dimensional (functional) error terms. The $\boldsymbol{\alpha}_i$'s and $\boldsymbol{\beta}_i$'s are d-dimensional unknown parameters, and the $\boldsymbol{\beta}_i$'s are the parameters of interest, usually called the "portfolio betas". Note that the sequence $\{(\boldsymbol{r}_i(\cdot), r_{iM}(\cdot))\}$ is a $(d+1)$-dimensional (functional) time series satisfying certain conditions to be specified below.

We assume that a training sample of size m with no instabilities is available, i.e.,

$$\boldsymbol{\alpha}_1 = \ldots = \boldsymbol{\alpha}_m =: \boldsymbol{\alpha}_0 = (\alpha_1^0, \ldots, \alpha_d^0)^T, \qquad \boldsymbol{\beta}_1 = \ldots = \boldsymbol{\beta}_m =: \boldsymbol{\beta}_0 = (\beta_1^0, \ldots, \beta_d^0)^T, \tag{1.2}$$

where $\boldsymbol{\alpha}_0$ and $\boldsymbol{\beta}_0$ are unknown parameters. The problem of the instability of the portfolio betas is formulated as a testing problem, that is, we want to test the null hypothesis

$$H_0:\ \boldsymbol{\beta}_1 = \cdots = \boldsymbol{\beta}_m = \boldsymbol{\beta}_{m+1} = \cdots$$

of "no change" versus the alternative

$$H_A:\ \boldsymbol{\beta}_1 = \cdots = \boldsymbol{\beta}_{m+k^*} \neq \boldsymbol{\beta}_{m+k^*+1} = \cdots$$

of a "structural break" at an unknown change-point $k^* = k_m^*$.

For later convenience we reformulate our model as follows:

$$r_{i,j}(s) = \alpha_j^0 + \beta_j^0 r_{iM}(s) + (\alpha_j^1 + \beta_j^1 r_{iM}(s))\delta_m I\{i > m + k^*\} + \varepsilon_{i,j}(s), \quad j = 1, \ldots, d,\ i = 1, 2, \ldots,\ s \in [0,1], \tag{1.3}$$

where $k^* = k_m^*$ is the change-point and $\alpha_j^0, \beta_j^0, \alpha_j^1, \beta_j^1, \delta_m$ are unknown parameters.

As in [7], our test procedures will be generated by convex loss functions $\varrho_1, \ldots, \varrho_d$ with a.s. derivatives $\varrho_j' = \psi_j$ called score functions having further properties to be specified later. The estimators $\widehat{\alpha}_{jm} = \widehat{\alpha}_{jm}(\psi_j)$, $\widehat{\beta}_{jm} = \widehat{\beta}_{jm}(\psi_j)$ of $\alpha_j^0,\ \beta_j^0$ based on the training sample are defined as minimizers of

$$\sum_{i=1}^{m} \sum_{\nu=1}^{n} \varrho_j(r_{i,j}(s_\nu) - a_j - b_j r_{iM}(s_\nu)) \tag{1.4}$$

w.r.t. a_j, b_j, for $j = 1, \ldots, d$, where $s_\nu = \nu/n$, $\nu = 1, \ldots, n$, are n equidistant intra-day time-points.

The test procedure constructed below will be based on functionals of partial sums of weighted M-residuals, which are defined as follows:

$$\boldsymbol{\psi}(\widehat{\boldsymbol{\varepsilon}}_i(s_\nu)) = (\psi_1(\widehat{\varepsilon}_{i,1}(s_\nu)), \ldots, \psi_d(\widehat{\varepsilon}_{i,d}(s_\nu)))^T \tag{1.5}$$

with

$$\widehat{\boldsymbol{\varepsilon}}_i(s_\nu) = (\widehat{\varepsilon}_{i,1}(s_\nu), \ldots, \widehat{\varepsilon}_{i,d}(s_\nu))^T,$$

$$\widehat{\varepsilon}_{i,j}(s_\nu) = r_{i,j}(s_\nu) - \widehat{\alpha}_{jm} - \widehat{\beta}_{jm} r_{iM}(s_\nu). \tag{1.6}$$

A suitable test statistic based on the first $m + k$ (functional) observations is

$$\widehat{Q}(k, m) = \Big(\frac{1}{\sqrt{m}} \sum_{i=m+1}^{m+k} \frac{1}{n} \sum_{\nu=1}^{n} r_{iM}(s_\nu) \boldsymbol{\psi}(\widehat{\boldsymbol{\varepsilon}}_i(s_\nu))\Big)^T \widehat{\boldsymbol{\Sigma}}_m^{-1} \Big(\frac{1}{\sqrt{m}} \sum_{i=m+1}^{m+k} \frac{1}{n} \sum_{\nu=1}^{n} r_{iM}(s_\nu) \boldsymbol{\psi}(\widehat{\boldsymbol{\varepsilon}}_i(s_\nu))\Big) \tag{1.7}$$

where $n = n(m)$ (see below) and the matrix $\widehat{\boldsymbol{\Sigma}}_m$ is an estimator of the asymptotic variance (matrix)

$$\boldsymbol{\Sigma} = \lim_{m\to\infty} var\Big\{\frac{1}{\sqrt{m}} \sum_{i=1}^{m} \int_0^1 r_{iM}(s) \boldsymbol{\psi}(\boldsymbol{\varepsilon}_i(s)) ds\Big\} \tag{1.8}$$

based on the first m observations. Details will be discussed later.

For notational convenience and later use, we introduce the notations, for $i \in \mathbb{Z}$ and $n \in \mathbb{N}$,

$$\boldsymbol{z}_i = (z_{i,1}, \dots, z_{i,d})^T = \int_0^1 r_{iM}(s) \boldsymbol{\psi}(\boldsymbol{\varepsilon}_i(s)) ds, \tag{1.9}$$

$$\widehat{\boldsymbol{z}}_i = \widehat{\boldsymbol{z}}_{i,n} = (\widehat{z}_{i,1}, \dots, \widehat{z}_{i,d})^T = \frac{1}{n} \sum_{\nu=1}^{n} r_{iM}(s_\nu) \boldsymbol{\psi}(\widehat{\boldsymbol{\varepsilon}}_i(s_\nu)), \tag{1.10}$$

$$\widetilde{\boldsymbol{z}}_i = \widetilde{\boldsymbol{z}}_{i,n} = (\widetilde{z}_{i,1}, \dots, \widetilde{z}_{i,d})^T = \frac{1}{n} \sum_{\nu=1}^{n} r_{iM}(s_\nu) \boldsymbol{\psi}(\boldsymbol{\varepsilon}_i(s_\nu)), \tag{1.11}$$

so that

$$\widehat{Q}(k, m) = \Big(\frac{1}{\sqrt{m}} \sum_{i=m+1}^{m+k} \widehat{\boldsymbol{z}}_i\Big)^T \widehat{\boldsymbol{\Sigma}}_m^{-1} \Big(\frac{1}{\sqrt{m}} \sum_{i=m+1}^{m+k} \widehat{\boldsymbol{z}}_i\Big) \quad \text{and}$$

$$\boldsymbol{\Sigma} = \lim_{m\to\infty} var\Big\{\frac{1}{\sqrt{m}} \sum_{i=1}^{m} \boldsymbol{z}_i\Big\}.$$

Similar to [7], we reject the null hypothesis as soon as the test statistic exceeds a critical level for the first time, i.e., when

$$\widehat{Q}(k, m)/q_\gamma(k/m) \geq c$$

for an appropriately chosen $c = c_\gamma(\alpha)$, where $q_\gamma(t),\ t \in (0, \infty)$, is a suitable boundary (weight) function. In this case we stop the procedure and confirm a structural break, otherwise we continue monitoring. The associated stopping rule is given by

$$\tau_m = \tau_m(\gamma) = \inf\{1 \leq k \leq \lfloor mT \rfloor : \widehat{Q}(k, m)/q_\gamma(k/m) \geq c\}, \tag{1.12}$$

with $\inf \emptyset := \infty$. Here T is a fixed positive number, that is, for practical reasons, we have a so-called *closed-end procedure* again. The following class of weight functions q_γ can be used, e.g.,

$$q_\gamma(t) = (1 + t)^2 \left(\frac{t}{t+1}\right)^{2\gamma}, \quad t \in (0, \infty), \tag{1.13}$$

where γ is a tuning constant taking values in $\big[0, 1/2)$. The critical value c will be chosen such that, under H_0, for $\alpha \in (0, 1)$ (fixed),

$$\lim_{m\to\infty} P\big(\tau_m < \infty\big) = \alpha, \tag{1.14}$$

i.e., the overall asymptotic level (false alarm rate) is α and, under H_A,

$$\lim_{m\to\infty} P\big(\tau_m < \infty\big) = 1, \tag{1.15}$$

i.e., the test is consistent (has asymptotic power 1).

The rest of the paper is organized as follows. The main results including the assumptions and limit properties of the test procedures are presented and discussed in Section 2. Section 3 reports on the results of a small simulation study and an application to a real data set. The proofs of our main results are given in Section 4, whereas Section 5 contains some auxiliary lemmas to be used in the proofs.

2. Assumptions and main results

Compared to [7], the assumptions on the sequence $\{(\varepsilon_{i,1}(\cdot), \dots, \varepsilon_{i,d}(\cdot),\ r_{iM}(\cdot))\}_{i\in\mathbb{Z}}$ and on the loss functions $\varrho_1, \dots, \varrho_d$ (or equivalently on the score functions $\psi_1, \dots, \psi_d$) have to be extended as follows.

We assume on ψ_j, the distributions of $\varepsilon_{0,j}(s)$ and $\lambda_j(x;s) = -E\psi_j(\varepsilon_{0,j}(s) - x), j = 1, \dots, d, s \in [0,1], x \in \mathbb{R}^1$.

(A.1) ψ_j are nondecreasing functions, $\lambda_j(0,s) = 0$, $\lambda_j'(0,\cdot)$ is continuous on $[0,1]$, $\lambda_j'(x,s) := \frac{\partial}{\partial x}\lambda_j(x,s)$ exists in a neighborhood of 0 for all $s \in [0,1]$,

$$|\lambda_j'(x, s+z) - \lambda_j'(0, s+z)| \le D_0|x|, \qquad |x| \le x_0,\ s, s+z \in [0,1],\ |z| \le z_0,$$

and

$$|\lambda_j'(0, x+s) - \lambda_j'(0,s)| \le D_0|x|, \quad |x| \le x_0,\ x+s, s \in [0,1],$$

for some $x_0, z_0, D_0 > 0$;

(A.2) $\int_0^1 \lambda_j'(0,s)\,ds \int_0^1 \lambda_j'(0,s)\,Er_{0M}(s)^2\,ds > \left(\int_0^1 \lambda_j'(0,s)\,Er_{0M}(s)\,ds\right)^2$;
[Note that, via the Cauchy–Schwarz inequality, we have at least "$\ge$" in the latter condition, so we just assume nondegeneracy.]

(A.3) $\sup_{s\in[0,1]} E|\psi_j(\varepsilon_{0,j}(s))|^3 < \infty$ and

$$E\left|\psi_j(\varepsilon_{0,j}(s) + t_2) - \psi_j(\varepsilon_{0,j}(s) + t_1)\right|^2 \le C_1|t_2 - t_1|, \quad |t_1|, |t_2| \le c_0,\ s \in [0,1],$$

for some $c_0, C_1 > 0$.

For later applications, let us briefly recall some of the most often considered ψ_j-functions. The classical choice $\psi_j(x) = x$, $x \in \mathbb{R}^1$, leads to the ordinary least squares (OLS) and L_2-residuals. A choice of $\psi_j(x) = \operatorname{sign} x,\ x \in \mathbb{R}^1$, leads to L_1-estimators and L_1-residuals. Huber [12] introduced $\psi_j(x) = x\,I\{|x| \le K\} + K \operatorname{sign} x\, I\{|x| > K\},\ x \in \mathbb{R}^1$, for some $K > 0$, which is one of the most often used score functions, usually known as the Huber function.

For a vector-valued random variable $\boldsymbol{X}$ define

$$\|\boldsymbol{X}\|_p = \left(E|\boldsymbol{X}|^p\right)^{1/p}, \quad p \ge 1,$$

the L_p-norm of $\boldsymbol{X}$, where $|\boldsymbol{X}|$ denotes the Euclidean norm of $\boldsymbol{X}$.

Concerning the assumptions on $\{r_{iM}(\cdot)\}$ and $\{\boldsymbol{\varepsilon}_i(\cdot)\}$ we follow the setup in Aue et al. [3], but instead of fourth moment assumptions used there it typically suffices here to have second or $(2+\Delta)$-moment conditions:

(B.1) For any $i \in \mathbb{Z}$, $r_{iM}(\cdot) = h(\boldsymbol{\xi}_i(\cdot), \boldsymbol{\xi}_{i-1}(\cdot), \dots)$, where $h(\cdot)$ is a measurable function, $\{\boldsymbol{\xi}_i(\cdot)\}$ is a sequence of i.i.d. random functions, and $\sup_{s\in[0,1]} E|r_{0M}(s)|^3 < \infty$.
[Note that $\{r_{iM}(\cdot) : i \in \mathbb{Z}\}$ is a stationary and ergodic sequence.]

Remark 2.1. For the sake of simplicity, we assume a third moment condition in Assumptions (A.3) and (B.1). With some more technical effort, the latter can be replaced by a $(2+\Delta)$-moment condition with some $\Delta > 0$ (cf. Lemma 5.1(i)–(ii)).

(B.2) For any $i \in \mathbb{Z}$, $\boldsymbol{\varepsilon}_i(\cdot) = \boldsymbol{g}(\boldsymbol{\zeta}_i(\cdot), \boldsymbol{\zeta}_{i-1}(\cdot), \dots)$, where $\boldsymbol{g}(\cdot)$ is a measurable function, $\{\boldsymbol{\zeta}_i(\cdot)\}$ is a sequence of i.i.d. random functions having some further properties to be specified later.
[Note that $\{\boldsymbol{\varepsilon}_i(\cdot) : i \in \mathbb{Z}\}$ is also a stationary and ergodic sequence.]

(B.3) The sequences $\{\boldsymbol{\xi}_i(\cdot)\}$ and $\{\boldsymbol{\zeta}_i(\cdot)\}$ are independent.

(B.4) For all $i \in \mathbb{Z}$,

$$\sup_{s\in[0,1]} \sum_{L=1}^{\infty} \|r_{iM}(s) - r_{iM}^{(L)}(s)\|_2 < \infty,$$

where

$$r_{iM}^{(L)}(\cdot) = h(\boldsymbol{\xi}_i(\cdot), \boldsymbol{\xi}_{i-1}(\cdot), \dots, \boldsymbol{\xi}_{i-L+1}(\cdot), \boldsymbol{\xi}_{i-L}^{(L)}(\cdot), \boldsymbol{\xi}_{i-L-1}^{(L)}(\cdot), \dots),$$

with $\boldsymbol{\xi}_{i-L}^{(L)}(\cdot), \boldsymbol{\xi}_{i-L-1}^{(L)}(\cdot), \dots$ being i.i.d. with the same distribution as $\boldsymbol{\xi}_0(\cdot)$ and independent of $\{\boldsymbol{\xi}_i(\cdot)\}$.
[Note that $r_{iM}^{(L)}(\cdot) \overset{\mathcal{D}}{=} r_{iM}(\cdot) \overset{\mathcal{D}}{=} r_{0M}(\cdot)$ for all $i \in \mathbb{Z}$ and $L \ge 1$.]

(B.5) With $\boldsymbol{\psi}(\boldsymbol{\varepsilon}_i(\cdot)) = (\psi_1(\varepsilon_{i,1}(\cdot)), \dots, \psi_d(\varepsilon_{i,d}(\cdot)))^T$, for all $i \in \mathbb{Z}$, it holds that

$$\sup_{s\in[0,1]} \sup_{|\boldsymbol{a}|\le a_0} \sum_{L=1}^{\infty} \|\boldsymbol{\psi}(\boldsymbol{\varepsilon}_i(s) - \boldsymbol{a}) - \boldsymbol{\psi}(\boldsymbol{\varepsilon}_i^{(L)}(s) - \boldsymbol{a})\|_2 < \infty$$

for some $a_0 > 0$, where

$$\boldsymbol{\varepsilon}_i^{(L)}(\cdot) = \boldsymbol{g}(\boldsymbol{\zeta}_i(\cdot), \boldsymbol{\zeta}_{i-1}(\cdot), \dots, \boldsymbol{\zeta}_{i-L+1}(\cdot), \boldsymbol{\zeta}_{i-L}^{(L)}(\cdot), \boldsymbol{\zeta}_{i-L-1}^{(L)}(\cdot), \dots),$$

with $\boldsymbol{\zeta}_{i-L}^{(L)}(\cdot), \boldsymbol{\zeta}_{i-L-1}^{(L)}(\cdot), \dots$ being i.i.d. with the same distribution as $\boldsymbol{\zeta}_0(\cdot)$ and independent of $\{\boldsymbol{\zeta}_i(\cdot)\}$.

Remark 2.2. Assumption (B.5) could be weakened as follows, but then the proofs would require somewhat more technicalities:

(B.5′) $\limsup_{n\to\infty} \frac{1}{n}\sum_{\nu=1}^{n} \sup_{|\boldsymbol{a}|\le a_0} \sum_{L=1}^{\infty} \|\boldsymbol{\psi}(\boldsymbol{\varepsilon}_i(s_\nu) - \boldsymbol{a}) - \boldsymbol{\psi}(\boldsymbol{\varepsilon}_i^{(L)}(s_\nu) - \boldsymbol{a})\|_2 < \infty$

for some $a_0 > 0$, with $s_\nu = \nu/n$, $\nu = 1, \dots, n$, and $\{\boldsymbol{\varepsilon}_i^{(L)}(\cdot)\}$ as in (B.5).

As in Aue et al. [3] and Chochola et al. [7], the above assumptions are motivated by the work of Hörmann and Kokoszka [10] on the concept of L_p–m-approximability, but could be relaxed here to a certain extent.

The following conditions, assuming that the processes under consideration are smooth functions of the intra-day parameter $s \in [0, 1]$, are weakened versions of the corresponding conditions in Aue et al. [3].

First we also make the following "high-frequency" assumption:

(B.6) We let $n = n(m) \to \infty$ as $m \to \infty$.

Secondly, we assume smoothness of the $r_{iM}(\cdot)$'s and $\psi_j(\varepsilon_{i,j}(\cdot))$'s:

(B.7) For all $i \in \mathbb{Z}, j = 1, \dots, d$, with $s_\nu = 1/n$ as above and $n = n(m) \to \infty$,
(a) $\lim_{m\to\infty} (\log m) \frac{1}{n} \sum_{\nu=1}^{n} \sup_{h\in[0,1/n]} \|r_{iM}(s_\nu) - r_{iM}(s_\nu - h)\|_2 = 0$
and
(b) $\lim_{m\to\infty} (\log m) \frac{1}{n} \sum_{\nu=1}^{n} \sup_{h\in[0,1/n]} \|\psi_j(\varepsilon_{i,j}(s_\nu)) - \psi_j(\varepsilon_{i,j}(s_\nu - h))\|_2 = 0.$

Remark 2.3. It will be obvious from the proofs below that, if the L_2-approximability conditions in Assumptions (B.4) and (B.5) are replaced by corresponding $L_{2+\Delta}$-approximability (with some $\Delta > 0$), then the convergence rate condition in (B.7) can be avoided, i.e., (B.7) can be replaced by

(B.7′) For all $i \in \mathbb{Z}, j = 1, \dots, d$, with $s_\nu = 1/n$ as above and $n = n(m) \to \infty$,
(a) $\lim_{m\to\infty} \frac{1}{n} \sum_{\nu=1}^{n} \sup_{h\in[0,1/n]} \|r_{iM}(s_\nu) - r_{iM}(s_\nu - h)\|_{2+\Delta} = 0$
and
(b) $\lim_{m\to\infty} \frac{1}{n} \sum_{\nu=1}^{n} \sup_{h\in[0,1/n]} \|\psi_j(\varepsilon_{i,j}(s_\nu)) - \psi_j(\varepsilon_{i,j}(s_\nu - h))\|_{2+\Delta} = 0.$

Remark 2.4. The theoretical results below as well as the applications to the real data set work with equidistant grid points being the same for all components. Nevertheless, going through the proofs this assumption can be relaxed, e.g., working with more general $s_{\nu,j}$'s, $j = 1, \dots, d$, under accordingly modified assumptions. Moreover, having a closer look at the test statistic defined through (1.7), (1.10) and (2.4), we realize that the test procedures depend on the observations through

$$\widehat{\boldsymbol{z}}_i = \frac{1}{n} \sum_{\nu=1}^{n} r_{iM}(s_\nu) \boldsymbol{\psi}(\widehat{\boldsymbol{\varepsilon}}_i(s_\nu)),$$

that are averages over time grids s_ν, i.e., averages over the intra-day behavior, which also work for asynchronous data.

Next we present our results on the limit behavior of the test procedures, both under the null hypothesis H_0 as well as under the alternative H_A.

2.1. Asymptotic results

Theorem 2.1. *Let Assumptions* (A.1)–(A.2), (B.1)–(B.7) *and* (1.13) *with* $\gamma \in [0, 1/2)$ *be satisfied and*

$$\widehat{\boldsymbol{\Sigma}}_m - \boldsymbol{\Sigma} = o_P(1) \quad (m \to \infty), \tag{2.1}$$

where, with the $\boldsymbol{z}_i$*'s from* (1.9),

$$\boldsymbol{\Sigma} = \lim_{m\to\infty} var\left\{ \frac{1}{\sqrt{m}} \sum_{i=1}^{m} \boldsymbol{z}_i \right\} = E[\boldsymbol{z}_0 \boldsymbol{z}_0^T] + \sum_{i=1}^{\infty} E[\boldsymbol{z}_0 \boldsymbol{z}_i^T + \boldsymbol{z}_i \boldsymbol{z}_0^T], \tag{2.2}$$

and $\boldsymbol{\Sigma}$ *is a positive definite matrix. Then, under the null hypothesis* H_0,

$$\max_{1\le k\le \lfloor mT\rfloor} \left(\frac{\widehat{Q}(k, m)}{q_\gamma(k/m)} \right) \xrightarrow{\mathcal{D}} \sup_{0<t<T/(T+1)} \left(\frac{\sum_{j=1}^{d} W_j^2(t)}{t^{2\gamma}} \right) \quad (m \to \infty),$$

where $\{W_j(t),\ t \in [0, 1]\},\ j = 1, \dots, d$, *are independent (standard) Brownian motions (Wiener processes).*

The proof of Theorem 2.1 is postponed to Section 4.

It follows from Assumptions (A.1)–(A.2) and (B.1)–(B.5) that $\{r_{iM}(\cdot)\}$ and $\{\boldsymbol{\psi}(\boldsymbol{\varepsilon}_i(\cdot))\}$ are independent sequences. Then Lemma 2.1 and Theorem 4.2 in Hörmann and Kokoszka [10] imply that the series in (2.2) converges (component-wise) absolutely.

Now we turn to the model under local alternatives, i.e.

$$r_{i,j}(s) = \alpha_j^0 + \beta_j^0 r_{iM}(s) + (\alpha_j^1 + \beta_j^1 r_{iM}(s))\delta_m I\{i > m + k^*\} + \varepsilon_{i,j}(s), \quad j = 1, \dots, d,\ i = 1, 2, \dots,\ s \in [0, 1], \tag{2.3}$$

with $\delta_m \to 0$ and $k^* < \lfloor mT \rfloor$.

Theorem 2.2. *Let Assumptions* (A.1)–(A.2), (B.1)–(B.7) *and* (1.13) *with* $\gamma \in [0, 1/2)$ *be satisfied and*

$$\widehat{\mathbf{\Sigma}}_m - \mathbf{\Sigma} = o_P(1) \quad (m \to \infty),$$

where $\mathbf{\Sigma}$ *is as in* Theorem 2.1. *Then, under* (2.3), *with* $\delta_m \to 0,\ |\delta_m|m^{1/2} \to \infty$, $\liminf_{m\to\infty}(\lfloor mT \rfloor - k^*)/m > 0$, *and* $\beta_j^1 \neq 0$ *for at least one j,*

$$\max_{1\le k\le \lfloor mT\rfloor}\left(\frac{\widehat{Q}(k,m)}{q_\gamma(k/m)}\right) \xrightarrow{P} \infty \quad (m \to \infty).$$

The proof of Theorem 2.2 is also postponed to Section 4.

Remark 2.5. (a) By Theorem 2.1, the assertion (1.14) holds true if $c_\gamma(\alpha)$ satisfies

$$P\left(\sup_{0<t<T/(T+1)}\left(\frac{\sum_{j=1}^{d} W_j^2(t)}{t^{2\gamma}}\right) \ge c_\gamma(\alpha)\right) = \alpha,$$

where $c_\gamma(\alpha)$ can either be obtained by simulation of the limit distribution or by an application of a suitable form of bootstrap based on the training sample.

(b) Theorem 2.2 implies the consistency of the test, i.e., the validity of (1.15) (asymptotic power 1).

2.2. Estimation of the variance matrix

In this section we deal with an estimator of the asymptotic variance (matrix) $\mathbf{\Sigma}$ as given in (2.2). Notice that $\mathbf{\Sigma} = \sum_{k=-\infty}^{\infty}\mathbf{\Gamma}_k$, where $\mathbf{\Gamma}_k = E[\boldsymbol{z}_0\boldsymbol{z}_k^T]$ for $k \ge 0$ and $\mathbf{\Gamma}_{-k} = \mathbf{\Gamma}_k^T$.

We consider an estimator of $\mathbf{\Sigma}$ based on the first m (functional) observations defined as

$$\widehat{\mathbf{\Sigma}}_m = \sum_{|k|<q}\omega_q(k)\widehat{\mathbf{\Gamma}}_k \tag{2.4}$$

where $q = q(m)$, $\omega_q(k) = \omega(k/q)$ and ω is a kernel specified below, and $\widehat{\mathbf{\Gamma}}_k$ is the kth lag sample covariance corresponding to $\mathbf{\Gamma}_k$, i.e.,

$$\widehat{\mathbf{\Gamma}}_k = \begin{cases} \dfrac{1}{m}\displaystyle\sum_{i=1}^{m-k}\widehat{\boldsymbol{z}}_i\widehat{\boldsymbol{z}}_{i+k}^T, & k \ge 0,\\ \widehat{\mathbf{\Gamma}}_{-k}^T, & k < 0,\end{cases} \tag{2.5}$$

with the $\widehat{\boldsymbol{z}}_i$'s as defined in (1.10), based on the $r_{iM}(\cdot)$'s from (1.1) and $\boldsymbol{\psi}(\widehat{\varepsilon}_i)$'s according to the M-residuals as given in (1.5) and (1.6).

Theorem 2.3. *Let Assumptions* (A.1), (A.2), *and* (B.1)–(B.7) *be satisfied. Let* $\widehat{\mathbf{\Sigma}}_m$ *be the estimator of* $\mathbf{\Sigma}$ *given in* (2.4) *with a kernel* $\omega_q(k) = \omega(k/q)$ *satisfying the following conditions:*

(i) $\omega(0) = 1$;
(ii) ω *is a symmetric and Lipschitz-continuous function;*
(iii) ω *has bounded support;*
(iv) *the Fourier transform of* ω *is also Lipschitz-continuous and integrable;*
(v) $q(m) = O(\log m)$ $(m \to \infty)$.

Then

$$\widehat{\mathbf{\Sigma}}_m = \mathbf{\Sigma} + o_P(1) \quad (m \to \infty).$$

We can work, e.g., either with the Bartlett kernel

$$\omega(x) = (1 - |x|)I\{|x| \le 1\} \tag{2.6}$$

or with the flat-top kernel

$$\omega(x) = \begin{cases} 1, & |x| \le \dfrac{1}{2},\\ 2(1-|x|), & \dfrac{1}{2} < |x| < 1,\\ 0, & |x| \ge 1.\end{cases} \tag{2.7}$$

3. Applications and simulations

In this section we present some results from a small simulation study as well as an application to a real data set in order to illustrate the finite sample performance of our monitoring procedure based on the test statistic (1.7) with boundary function (1.13).

First we discuss some aspects which are common to both the simulation study and the application. Since the asymptotic distribution of the test statistic given in Theorem 2.1 coincides with the one derived in Chochola et al. [7] (cf. also Remark 2.3), we can use the critical values given in Table 1 of [7].

The question of the choice of the tuning constant γ has also been discussed in [7] and the recommendation given there remains valid, i.e., if a change is to be expected "early" after the training period, then γ near to 0.5 is advisable, whereas for "late change scenarios", small γ's are recommended. A choice of $\gamma = 0.25$ provides a reasonably good balance between these two scenarios and is thus used here.

We consider the L_2, Huber and L_1 ψ-functions and always apply the same function to all coordinates.

It remains to choose the kernel function and especially its bandwidth q in the estimator of the variance matrix suggested in (2.4). In this aspect, we use the results of Chochola [6] which show that it can be difficult to set a proper q a priori for the Bartlett or the flat-top kernel, because it depends on the degree of dependency of the data. Thus better results can be obtained using a data-driven adaptive choice of the bandwidth based on the work of Andrews [2] and implemented in the statistical software R as described in Zeileis [19]. Differences between possible kernel choices are not too big, so that we always use the Bartlett kernel here.

As an illustration of a possible application of our robust monitoring, we investigate the data set used in Aue et al. [3] in more detail. Recalling this data set, it consists of five stocks from different sectors of S&P 100, namely Boeing (BA), Bank of America (BAC), Microsoft (MSFT), AT&T (T), and Exxon Mobile (XOM). As the market portfolio, the S&P 100 index itself is used.

The intra-day behavior of the process $\{\boldsymbol{r}_i(s) : s \in [0; 1]; i \in \mathbb{Z}\}$, which is defined at time s as the difference between the log-prices of the stocks at time s and $s + 15$ min, is thus sampled every 15 min during any trading day i. The process $r_{iM}(\cdot)$ is defined analogously.

The historical training period starts on January 29, 2001 and consists of 120 trading days for which the values of the portfolio betas under consideration appear reasonably stable. The choice of the beginning of the period is motivated by the fact that, prior to January 29, 2001, the tick size (i.e. the smallest value the price can change) was different. The monitoring horizon for the closed-end procedure was selected as 360 days, corresponding to $T = 3$ for our stopping rule in (1.12). This covers the 9/11 event, the influence of which we want to study.

The stability of the historical portfolio betas was checked via moving windows estimates presented in Fig. 1. The figure shows Huber estimates of portfolio betas based on moving windows of 10 trading days for each company throughout the historical and monitoring periods, but the figures look similar for L_2 estimates. The solid black vertical line marks the end of the historical period (120 days), whereas the dashed black line marks the last day, when the estimate is not influenced by the observations from the monitoring period. The gray lines refer in the same way to the 9/11 event. Since "no change" during the historical period is assumed, we tested for a change in this period via L_2 and Huber retrospective procedures and this assumption could be confirmed.

The BAC and T estimates seem to be stable throughout the whole period, whereas there is a small temporary influence of the 9/11 event on MSFT and a very big one on BA. Finally there seems to be a shift in the portfolio beta of XOM right after the end of the training period. We come back to these observations later on.

Next we discuss the robust monitoring itself. Fig. 2 shows values of the normalized test statistic, i.e., $\widehat{Q}(k, m)/(c_{0.25}(0.05)\, q_{\gamma}(k/m))$, for the L_2 (dashed line), Huber (solid line) and L_1 (dotted line) monitoring procedure and for various combinations of stocks, which are given in the heading of each chart. On the x-axis the number of trading days is shown starting from the beginning of the monitoring. A vertical gray dashed line marks the September 11, 2001, terrorist attack, the horizontal one (at value 1) indicates the critical line, due to the normalization of the statistic.

When all companies are considered together, we get the same results as in Aue et al. [3] for the L_2 procedure. The critical value is extremely exceeded. For the Huber and L_1 procedures the crossing still occurs, but in a much more moderate way.

It is possible to get further insight by looking at the stocks individually. In view of the conclusions from Fig. 1, we examined Boeing (BA) and Exxon (XOM). Portfolio betas of the three remaining companies (BAC, MSFT and T) do not show any sign of a change as can be seen from the last chart. For Boeing (BA) we can see the extreme influence of the 9/11 event on the L_2 monitoring procedure. In case of robust procedures this has a much smaller impact, the critical value, however, is still crossed right after the event. For Exxon (XOM) and robust monitoring, the critical line is crossed already before the 9/11 attack—there is a steady increase in the test statistic from the beginning of the monitoring on, which is in line with the conclusions from Fig. 1. By applying the retrospective procedure to the XOM data for the first 120 and 240 trading days, respectively, it turned out that no change could be detected based on the period of length 120, but using 240 days a change was indicated close to trading day 110. This explains why the critical line is already crossed before the 9/11 event.

It is of further interest, whether the change in Boeing's (BA) portfolio betas after the 9/11 was only temporary or persistent. In order to find out, we use the same monitoring procedure, but exclude 5 or 10 trading days after the 9/11 from the monitoring. This can be seen in Fig. 3. We can see that, if 5 days are excluded, then the crossings are much smaller and, if 10 days are excluded, then the terrorist attack has no impact at all and the change is not indicated until mid of March 2002.

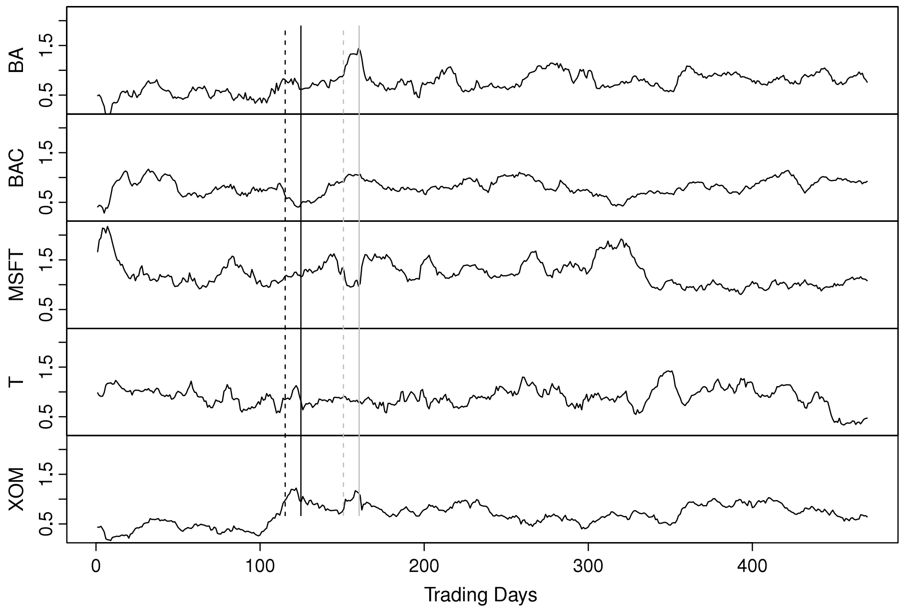

Fig. 1. Huber estimates of portfolio beta based on moving windows of 10 trading days. Black solid vertical line marks end of training period, gray one marks the 9/11 event. Dashed lines indicate the beginning of moving windows that are already influenced by these events.

Table 1
Empirical sizes at nominal level $\alpha = 5\%$ under H_0.

r_{iM}	$\boldsymbol{\varepsilon}_i$	m	L_2	Huber	L_1
B_i	B_i	100	8.4	7.0	5.9
		200	5.3	4.8	4.1
B_i	Mix	100	25.7	7.4	5.8
		200	14.5	4.6	3.9
AR(1; 0.1)	B_i	100	8.4	7.0	5.0
		200	6.4	5.4	4.4
AR(1; 0.4)	B_i	100	9.3	7.6	6.2
		200	6.7	6.1	5.6

In order to further quantify the finite sample properties of the monitoring procedure, a small simulation study has been conducted. We simulated data according to the model (1.1), with $d = 2, \boldsymbol{\alpha}_0 = (1/2, 1/2)^T, \boldsymbol{\beta}_0 = (1, 1)^T$ for simplicity. Various settings have been used for the market portfolio log-returns $r_{iM}(\cdot)$ and the error terms $\boldsymbol{\varepsilon}_i(\cdot)$. The $r_{iM}(\cdot)$'s were either independent standard Brownian motions (denoted B_i) or, similarly as in Aue et al. [3], chosen as a functional AR(1) process, i.e.,

$$r_{iM}(s) = \rho \int_0^1 K(s, t) r_{i-1,M}(t) dt + \eta_i(s), \quad s \in [0, 1],$$

where $\{\eta_i(\cdot) : i \in \mathbb{Z}\}$ denotes a sequence of independent standard Brownian motions and $K(s; t) = c \exp(-|t - s|)$, with c such that the norm of K equals one. We chose $\rho = 0.1$ and $\rho = 0.4$ as the dependency coefficient and denote the models as AR(1;0.1) or AR(1;0.4). The random errors, in both coordinates, are either standard Brownian motions or, to illustrate the robustness of the monitoring procedures, we use a 5% contamination with Brownian motion having larger variance, i.e. $10B_i$ (denoted Mix).

Finally $m = 100$ or $m = 200$ and $T = 5$ were chosen, with a tuning constant $\gamma = 0.25$ in the boundary function, nominal level $\alpha = 5\%$, and the Bartlett kernel is used with an adaptive choice of the bandwidth q, as discussed at the beginning of this section. All results are based on 2000 repetitions.

First we have a look at the empirical levels presented in Table 1. We can see that the levels are approximately kept for the Huber and L_1 procedures in all scenarios considered. This, however, is no longer true for the L_2 procedure, especially in the case of the contaminated model. So, in order to compare the different procedures one would have to adjust them to possess the same empirical size.

In order to illustrate the properties of the test under the alternative hypothesis, we chose $k^* = 10$ and a unit change in both parameters $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ and in both coordinates. Fig. 4 shows the densities of the detection delays $\tau_m - k^*$ for various choices of distributions of $r_{i,M}$ and $\boldsymbol{\varepsilon}_i$. As long as both are standard Brownian motions (B_i, B_i), the L_2 procedure performs better than

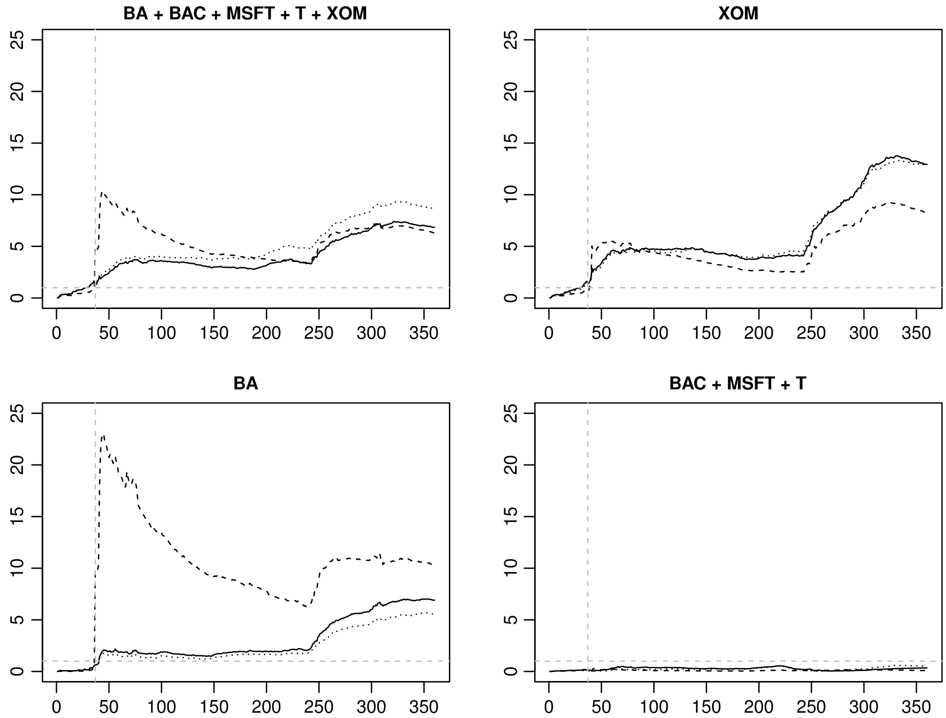

Fig. 2. Normalized test statistics for the L_2 (dashed line), Huber (solid line) and L_1 (dotted line) monitoring procedures, various combinations of stocks—given in the heading of each chart. x-axis shows number of trading days from the beginning of the monitoring on.

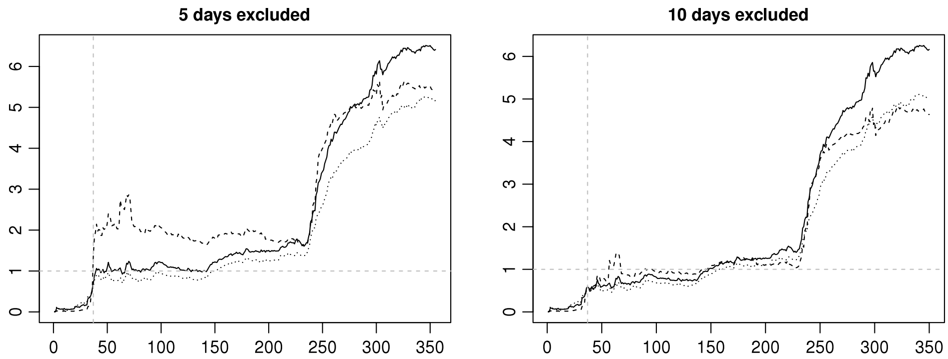

Fig. 3. Boeing stock, normalized test statistics for L_2 (dashed line), Huber (solid line) and L_1 (dotted line) monitoring procedures. 5 or 10 days excluded from the monitoring after the 9/11.

Huber and L_1, while in case of (B_i, Mix) the L_2 procedure is outperformed by the robust ones, in particular by the Huber procedure. The latter effect is even more visible if all procedures are adjusted to the same empirical size (see also Table 1).

In conclusion, in certain situations the robust monitoring procedures suggested in this work show definite advantages over the much more sensitive L_2 approach. They usually avoid overrejection of the test and are able to keep the approximate size. A choice of Huber's ψ-function seems to provide a good balance between robust and sensitive monitoring. If no prior knowledge is available on where to expect a possible change, a choice of the tuning constant $\gamma = 0.25$ in (1.13) appears to be appropriate.

4. Proofs

Proof of Theorem 2.1. Similar to Chochola et al. [7], the proof can be given in three steps. Let us recall that we work with the model

$$r_{i,j}(s) = \alpha_j^0 + \beta_j^0 r_{iM}(s) + (\alpha_j^1 + \beta_j^1 r_{iM}(s))\delta_m I\{i > m + k^*\} + \varepsilon_{i,j}(s), \quad j = 1, \ldots, d, \; i = 1, 2, \ldots, \; s \in [0, 1], \qquad (4.1)$$

as defined in (1.3).

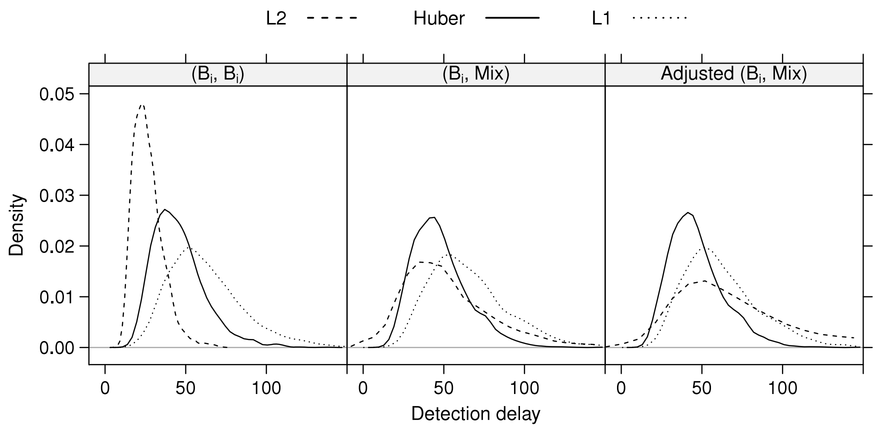

Fig. 4. Densities of the detection delays for L_2 (dashed line), Huber (solid line) and L_1 (dotted line) monitoring procedures.

1. In a first step we make use of asymptotic representations of the estimators $\widehat{\alpha}_{jm},\ \widehat{\beta}_{jm}$ of $\alpha_j^0,\ \beta_j^0,\ j = 1, \dots, d$, from (1.4). These estimators are based on the training sample only, so that we are in a non-sequential setup and can proceed in the same way as in treating the behavior of multivariate M-estimators. However, we need to take care of the dependency structure of the random error functions.

In the following it is convenient to introduce auxiliary estimators $\widehat{\alpha}^*_{jm}$ and $\widehat{\beta}^*_{jm}$ as minimizers of

$$\sum_{i=1}^{m}\sum_{\nu=1}^{n} \varrho_j(\varepsilon_{i,j}(s_\nu) - a_j^*/\sqrt{m} - b_j^* r_{iM}(s_\nu)/\sqrt{m}) \tag{4.2}$$

w.r.t. a_j^* and b_j^*, for $j = 1, \dots, d$, where $s_\nu = \nu/n$, $\nu = 1, \dots, n$. Clearly,

$$\widehat{\alpha}^*_{jm} = \sqrt{m}(\widehat{\alpha}_{jm} - \alpha_j^0), \qquad \widehat{\beta}^*_{jm} = \sqrt{m}(\widehat{\beta}_{jm} - \beta_j^0). \tag{4.3}$$

Usually, the estimators $\widehat{\alpha}^*_{jm}$ and $\widehat{\beta}^*_{jm}$ can be obtained as solutions of the equations

$$\sum_{i=1}^{m}\sum_{\nu=1}^{n} \psi_j(\varepsilon_{i,j}(s_\nu) - (a_j^* + b_j^* r_{iM}(s_\nu))/\sqrt{m}) = 0, \tag{4.4}$$

$$\sum_{i=1}^{m}\sum_{\nu=1}^{n} \psi_j(\varepsilon_{i,j} - (a_j^* + b_j^* r_{iM}(s_\nu))/\sqrt{m})\widetilde{r}_{iM} = 0, \tag{4.5}$$

w.r.t. a_j^*, b_j^*, for $j = 1, \dots, d$.

Lemmas 5.2 and 5.3 below ensure that $\widehat{\alpha}^*_{jm} = O_P(1)$ and $\widehat{\beta}^*_{jm} = O_P(1)$ and, moreover, we get the asymptotic representations, as $m \to \infty$,

$$\widehat{\alpha}^*_m = \frac{1}{\int_0^1 \lambda'(0,z)dz}\frac{1}{\sqrt{m}}\sum_{i=1}^{m}\int_0^1 \psi(\varepsilon_i(s))ds - \widehat{\beta}^*_m \frac{\int_0^1 \lambda'(0,z)Er_{0M}^2(z)dz}{\int_0^1 \lambda'(0,z)dz} + O_P(m^{-\eta}), \tag{4.6}$$

$$\widehat{\beta}^*_m = \frac{\frac{1}{\sqrt{m}}\sum_{i=1}^{m}\int_0^1 \psi(\varepsilon_i(s))\left(r_{iM}(s) - \frac{\int_0^1 \lambda'(0,z)Er_{0M}(z)dz}{\int_0^1 \lambda'(0,z)dz}\right)ds}{\int_0^1 \lambda'(0,z)Er_{0M}^2(z)dz - \frac{\left(\int_0^1 \lambda'(0,z)Er_{0M}(z)dz\right)^2}{\int_0^1 \lambda'(0,z)dz}} + O_P(m^{-\eta}), \tag{4.7}$$

with some $\eta > 0$ (cf. Remark 5.1).

2. Next, as a consequence of Lemmas 5.2–5.4 in combination with Remarks 5.1–5.2, we observe that the limit behavior of the weighted partial sums

$$\widehat{\boldsymbol{H}}(m,k) = (\widehat{H}_1(m,k), \dots, \widehat{H}_d(m,k))^T = \frac{1}{\sqrt{m}}\sum_{i=m+1}^{m+k}\frac{1}{n}\sum_{\nu=1}^{n} r_{iM}(s_\nu)\boldsymbol{\psi}(\widehat{\boldsymbol{\varepsilon}}_i(s_\nu)), \quad k = 1, \dots, \lfloor mT \rfloor,$$

is the same as that of

$$\widetilde{\boldsymbol{H}}(m,k)=\frac{1}{\sqrt{m}}\Big(\sum_{i=m+1}^{m+k}\frac{1}{n}\sum_{\nu=1}^{n}r_{iM}(s_\nu)\boldsymbol{\psi}(\boldsymbol{\varepsilon}_i(s_\nu))-\frac{k}{m}\sum_{i=1}^{m}\frac{1}{n}\sum_{\nu=1}^{n}r_{iM}(s_\nu)\boldsymbol{\psi}(\boldsymbol{\varepsilon}_i(s_\nu))\Big),\quad k=1,\dots,\lfloor mT\rfloor.$$

In view of Lemma 5.5(ii) together with Assumption (2.1), this further implies that the limit behavior of

$$\max_{1\le k\le\lfloor mT\rfloor}\widehat{Q}(k,m)/q_\gamma(k/m)$$

is the same as that of

$$\max_{1\le k\le\lfloor mT\rfloor}Q(k,m)/q_\gamma(k/m),$$

where

$$Q(k,m)=\boldsymbol{H}(m,k)^T\boldsymbol{\Sigma}^{-1}\boldsymbol{H}(m,k),\tag{4.8}$$

with

$$\boldsymbol{H}(m,k)=\frac{1}{\sqrt{m}}\Big(\sum_{i=m+1}^{m+k}\int_0^1 r_{iM}(s)\boldsymbol{\psi}(\boldsymbol{\varepsilon}_i(s))\,ds-\frac{k}{m}\sum_{i=1}^{m}\int_0^1 r_{iM}(s)\boldsymbol{\psi}(\boldsymbol{\varepsilon}_i(s))\,ds\Big),\quad k=1,\dots,\lfloor mT\rfloor.$$

3. In order to obtain the limit behavior of

$$\max_{1\le k\le\lfloor mT\rfloor}Q(k,m)/q_\gamma(k/m),$$

with $Q(k,m)$ from (4.8), we follow the lines of proof of Theorem 2.1 in Chochola et al. [7]. We just have to replace the random sequences and processes $\{\boldsymbol{Z}_i\}$, $\{\boldsymbol{Z}_i^{(L)}\}$ and $\{\boldsymbol{Z}_m(t)\}$ introduced there by

$$\boldsymbol{Z}_i=(Z_{i,1},\dots,Z_{i,d})^T=\int_0^1 r_{iM}(s)\boldsymbol{\psi}(\boldsymbol{\varepsilon}_i(s))\,ds,\quad i=1,2,\dots,$$

$$\boldsymbol{Z}_i^{(L)}=(Z_{i,1}^{(L)},\dots,Z_{i,d}^{(L)})^T=\int_0^1 r_{iM}^{(L)}(s)\boldsymbol{\psi}(\boldsymbol{\varepsilon}_i^{(L)}(s)),\quad i=1,2,\dots,\quad\text{and}$$

$$\boldsymbol{Z}_m(t)=\frac{1}{\sqrt{m}}\sum_{i=1}^{\lfloor mt\rfloor}\boldsymbol{Z}_i,\quad 0\le t\le T+1,$$

where $\int_0^1$ is to be taken componentwise.

The main step, that is, the weak convergence in the Skorokhod space $\mathcal{D}^d[0,T+1]$

$$\boldsymbol{Z}_m(\cdot)\xrightarrow{\mathcal{D}^d[0,T+1]}\boldsymbol{W}_{\boldsymbol{\Sigma}}(\cdot),$$

where $\{\boldsymbol{W}_{\boldsymbol{\Sigma}}(t):t\in[0,T+1]\}$ is a centered Gaussian process with covariance function $E[\boldsymbol{W}_{\boldsymbol{\Sigma}}(t)\boldsymbol{W}_{\boldsymbol{\Sigma}}^T(s)]=\min(t,s)\boldsymbol{\Sigma}$, is again a consequence of Billingsley [5], Theorem 21.1. An application of the continuous mapping theorem then completes the proof. For details we refer to Chochola et al. [7], pp. 383–385. □

Proof of Theorem 2.2. It suffices to show that

$$\frac{\widehat{Q}(\tilde{k},m)}{q_\gamma(\tilde{k}/m)}\xrightarrow{P}\infty$$

for suitably chosen $\tilde{k}$. We take $\tilde{k}=k^*+(mT-k^*)/2$. In view of our assumptions on $\widehat{\boldsymbol{\Sigma}}_m$ and the choice of $\tilde{k}$ it suffices to treat

$$\begin{aligned}&\frac{1}{\sqrt{m}}\sum_{i=k^*+1}^{\tilde{k}}\frac{1}{n}\sum_{\nu=1}^{n}r_{iM}(s_\nu)\psi_j(\widehat{\varepsilon}_{i,j}(s_\nu))\\&\quad=\frac{1}{\sqrt{m}}\sum_{i=k^*+1}^{\tilde{k}}\frac{1}{n}\sum_{\nu=1}^{n}r_{iM}(s_\nu)\psi_j\big(\varepsilon_{i,j}(s_\nu)-(\widehat{\alpha}_{mj}^*+\widehat{\beta}_{mj}^*r_{iM}(s_\nu))/\sqrt{m}+(\alpha_j^1+\beta_j^1r_{iM}(s_\nu))\delta_m\big),\end{aligned}$$

where $\widehat{\alpha}_{mj}^*=O_P(1)$ and $\widehat{\beta}_{mj}^*=O_P(1)$. Therefore it is enough to study

$$\frac{1}{\sqrt{m}}\sum_{i=k^*+1}^{\tilde{k}}\frac{1}{n}\sum_{\nu=1}^{n}r_{iM}(s_\nu)\psi_j\big(\varepsilon_{i,j}(s_\nu)-(a+br_{iM}(s_\nu))/\sqrt{m}+(\alpha_j^1+\beta_j^1r_{iM}(s_\nu))\delta_m\big)$$

for $|a| + |b| \le C, C > 0$. Proceeding analogously to the proof of Lemma 5.4 and recalling that $\delta_m \to 0$, but $|\delta_m|\sqrt{m} \to \infty$, we get

$$\left| E^* \frac{1}{\sqrt{m}} \sum_{i=k^*+1}^{\tilde{k}} \frac{1}{n} \sum_{\nu=1}^{n} r_{iM}(s_\nu)\psi_j\big(\varepsilon_{i,j}(s_\nu) - (a + br_{iM}(s_\nu))/\sqrt{m} + (\alpha_j^1 + \beta_j^1 r_{iM}(s_\nu)\delta_m)\big)\right| \xrightarrow{P} \infty$$

and

$$var^*\left\{ \frac{1}{\sqrt{m}} \sum_{i=k^*+1}^{\tilde{k}} \frac{1}{n} \sum_{\nu=1}^{n} r_{iM}(s_\nu)\psi_j\big(\varepsilon_{i,j}(s_\nu) - (a + br_{iM}(s_\nu))/\sqrt{m} + (\alpha_j^1 + \beta_j^1 r_{iM}(s_\nu))\delta_m\big)\right\} = O_P(1),$$

uniformly in $|a| + |b| \le C, C > 0$.

From here, after some standard steps, we receive the desired assertion. □

Proof of Theorem 2.3. Let $\boldsymbol{z}_i, \widehat{\boldsymbol{z}}_i$ and $\widetilde{\boldsymbol{z}}_i$ be as given in (1.9)–(1.11), respectively. Recall $\widehat{\boldsymbol{\Gamma}}_k$ from (2.5) and further define, for $k \ge 0$,

$$\boldsymbol{\Gamma}_k = \frac{1}{m} \sum_{i=1}^{m-k} \boldsymbol{z}_i \boldsymbol{z}_{i+k}^T,$$

$$\widetilde{\boldsymbol{\Gamma}}_k = \frac{1}{m} \sum_{i=1}^{m-k} \widetilde{\boldsymbol{z}}_i \widetilde{\boldsymbol{z}}_{i+k}^T,$$

and, for $k < 0$, put $\boldsymbol{\Gamma}_k = \boldsymbol{\Gamma}_{-k}^T$ and $\widetilde{\boldsymbol{\Gamma}}_k = \widetilde{\boldsymbol{\Gamma}}_{-k}^T$, respectively.

Let $\widehat{\boldsymbol{\Sigma}}_m$ be as given in (2.4) and put

$$\boldsymbol{\Sigma}_m = \sum_{|k|<q} \omega_q(k)\boldsymbol{\Gamma}_k$$

and

$$\widetilde{\boldsymbol{\Sigma}}_m = \sum_{|k|<q} \omega_q(k)\widetilde{\boldsymbol{\Gamma}}_k.$$

Then we have

$$\widehat{\boldsymbol{\Sigma}}_m = \boldsymbol{\Sigma}_m + (\widehat{\boldsymbol{\Sigma}}_m - \widetilde{\boldsymbol{\Sigma}}_m) + (\widetilde{\boldsymbol{\Sigma}}_m - \boldsymbol{\Sigma}_m).$$

First, let us consider $\boldsymbol{\Sigma}_m$. Note that $\{\boldsymbol{z}_i : i \in \mathbb{Z}\}$ is a stationary, L_2-approximable, centered sequence with $E\|\boldsymbol{z}_0\|^2 < \infty$, which follows from Assumptions (B.1)–(B.5) together with Lemma 2.1 in Hörmann and Kokoszka [10]. With a kernel ω_q satisfying conditions (i)–(v), all assumptions of Theorem 16.6 in Horváth and Kokoszka [11] are fulfilled. According to the latter theorem, we get

$$\boldsymbol{\Sigma}_m \xrightarrow{P} \boldsymbol{\Sigma} \quad \text{as } m \to \infty. \tag{4.9}$$

In the next step we will show that

$$\widehat{\boldsymbol{\Sigma}}_m - \widetilde{\boldsymbol{\Sigma}}_m = O_p(q(m)m^{-1/4}). \tag{4.10}$$

Here we can proceed quite analogously to the corresponding part of the proof of Theorem 2.3 in Chochola et al. [7]. Obviously,

$$\widehat{\boldsymbol{\Sigma}}_m - \widetilde{\boldsymbol{\Sigma}}_m = \sum_{|k|<q} \omega_q(k)(\widehat{\boldsymbol{\Gamma}}_k - \widetilde{\boldsymbol{\Gamma}}_k)$$

and, since

$$\widehat{\boldsymbol{z}}_i\widehat{\boldsymbol{z}}_{i+k}^T - \widetilde{\boldsymbol{z}}_i\widetilde{\boldsymbol{z}}_{i+k}^T = (\widehat{\boldsymbol{z}}_i - \widetilde{\boldsymbol{z}}_i)(\widehat{\boldsymbol{z}}_{i+k} - \widetilde{\boldsymbol{z}}_{i+k})^T + (\widehat{\boldsymbol{z}}_i - \widetilde{\boldsymbol{z}}_i)\widetilde{\boldsymbol{z}}_{i+k}^T + \widetilde{\boldsymbol{z}}_i(\widehat{\boldsymbol{z}}_{i+k} - \widetilde{\boldsymbol{z}}_{i+k})^T,$$

we have

$$\sum_{0\le k<q} \omega_q(k)(\widehat{\boldsymbol{\Gamma}}_k - \widetilde{\boldsymbol{\Gamma}}_k) = \boldsymbol{S}_1 + \boldsymbol{S}_2 + \boldsymbol{S}_3,$$

where

$$\boldsymbol{S}_1 = \sum_{0\le k<q} \omega_q(k)\frac{1}{m}\sum_{i=1}^{m-k}\frac{1}{n^2}\sum_{\mu=1}^{n}\sum_{\nu=1}^{n} r_{iM}(s_\mu)r_{i+k,M}(s_\nu)[\boldsymbol{\psi}(\widehat{\boldsymbol{\varepsilon}}_i(s_\mu)) - \boldsymbol{\psi}(\boldsymbol{\varepsilon}_i(s_\mu))][\boldsymbol{\psi}(\widehat{\boldsymbol{\varepsilon}}_{i+k}(s_\nu)) - \boldsymbol{\psi}(\boldsymbol{\varepsilon}_{i+k}(s_\nu))]^T,$$

$$\boldsymbol{S}_2 = \sum_{0\le k<q} \omega_q(k) \frac{1}{m} \sum_{i=1}^{m-k} \frac{1}{n^2} \sum_{\mu=1}^{n} \sum_{\nu=1}^{n} r_{iM}(s_\mu) r_{i+k,M}(s_\nu) [\boldsymbol{\psi}(\widehat{\boldsymbol{\varepsilon}}_i(s_\mu)) - \boldsymbol{\psi}(\boldsymbol{\varepsilon}_i(s_\mu))] \boldsymbol{\psi}(\boldsymbol{\varepsilon}_{i+k}(s_\nu))^T,$$

$$\boldsymbol{S}_3 = \sum_{0\le k<q} \omega_q(k) \frac{1}{m} \sum_{i=1}^{m-k} \frac{1}{n^2} \sum_{\mu=1}^{n} \sum_{\nu=1}^{n} r_{iM}(s_\mu) r_{i+k,M}(s_\nu) \boldsymbol{\psi}(\boldsymbol{\varepsilon}_i(s_\mu)) [\boldsymbol{\psi}(\widehat{\boldsymbol{\varepsilon}}_{i+k}(s_\nu)) - \boldsymbol{\psi}(\boldsymbol{\varepsilon}_{i+k}(s_\nu))]^T.$$

For $s \in [0, 1]$, set $\boldsymbol{d}_i(s) = \boldsymbol{a} + \boldsymbol{b} r_{iM}(s)$, where $\boldsymbol{a} = (a_1, \ldots, a_d)^T$, $\boldsymbol{b} = (b_1, \ldots, b_d)^T$, $\boldsymbol{d}_i(s) = (d_{i,1}(s), \ldots, d_{i,d}(s))^T$, and introduce

$$\begin{aligned}\boldsymbol{S}_1^0 = &\sum_{0\le k<q} \omega_q(k) \frac{1}{m} \sum_{i=1}^{m-k} \frac{1}{n^2} \sum_{\mu=1}^{n} \sum_{\nu=1}^{n} r_{iM}(s_\mu) r_{i+k,M}(s_\nu) [\boldsymbol{\psi}(\boldsymbol{\varepsilon}_i(s_\mu) - \boldsymbol{d}_i(s_\mu)/\sqrt{m}) - \boldsymbol{\psi}(\boldsymbol{\varepsilon}_i(s_\mu))] \\ &\times [\boldsymbol{\psi}(\boldsymbol{\varepsilon}_{i+k}(s_\nu) - \boldsymbol{d}_{i+k}(s_\nu)/\sqrt{m}) - \boldsymbol{\psi}(\boldsymbol{\varepsilon}_{i+k}(s_\nu))]^T,\end{aligned}$$

$$\boldsymbol{S}_2^0 = \sum_{0\le k<q} \omega_q(k) \frac{1}{m} \sum_{i=1}^{m-k} \frac{1}{n^2} \sum_{\mu=1}^{n} \sum_{\nu=1}^{n} r_{iM}(s_\mu) r_{i+k,M}(s_\nu) [\boldsymbol{\psi}(\boldsymbol{\varepsilon}_i(s_\mu) - \boldsymbol{d}_i(s_\mu)/\sqrt{m}) - \boldsymbol{\psi}(\boldsymbol{\varepsilon}_i(s_\mu))] \times \boldsymbol{\psi}(\boldsymbol{\varepsilon}_{i+k}(s_\nu))^T,$$

$$\begin{aligned}\boldsymbol{S}_3^0 = &\sum_{0\le k<q} \omega_q(k) \frac{1}{m} \sum_{i=1}^{m-k} \frac{1}{n^2} \sum_{\mu=1}^{n} \sum_{\nu=1}^{n} r_{iM}(s_\mu) r_{i+k,M}(s_\nu) \boldsymbol{\psi}(\boldsymbol{\varepsilon}_i(s_\mu)) \\ &\times [\boldsymbol{\psi}(\boldsymbol{\varepsilon}_{i+k}(s_\nu) - \boldsymbol{d}_{i+k}(s_\nu)/\sqrt{m}) - \boldsymbol{\psi}(\boldsymbol{\varepsilon}_{i+k}(s_\nu))]^T.\end{aligned}$$

Now, for any $1 \le j, \ell \le d$,

$$\begin{aligned}&E|r_{iM}(s_\mu) r_{i+k,M}(s_\nu) [\psi_j(\varepsilon_{i,j}(s_\mu) - d_{i,j}(s_\mu)/\sqrt{m}) - \psi_j(\varepsilon_{i,j}(s_\mu))] \\ &\quad \times [\psi_\ell(\varepsilon_{i+k,\ell}(s_\nu) - d_{i+k,\ell}(s_\nu)/\sqrt{m}) - \psi_\ell(\varepsilon_{i+k,\ell}(s_\nu))]| \\ &\le E|r_{iM}(s_\mu) r_{i+k,M}(s_\nu)| (E^*|\psi_j(\varepsilon_{i,j}(s_\mu) - d_{i,j}(s_\mu)/\sqrt{m}) - \psi_j(\varepsilon_{i,j}(s_\mu))|^2)^{1/2} \\ &\quad \times (E^*|\psi_\ell(\varepsilon_{i+k,\ell}(s_\nu) - d_{i+k,\ell}(s_\nu)/\sqrt{m}) - \psi_\ell(\varepsilon_{i+k,\ell}(s_\nu))|^2)^{1/2} \le Cm^{-1/2},\end{aligned}$$

uniformly in $\boldsymbol{a}, \boldsymbol{b}$ such that $\max_{1\le j\le d}(|a_j| + |b_j|) < C$ for some constant $C > 0$, where we have used the rule of iterated expectations (with E^* being the conditional expectation given r_{iM}, $i = 1, \ldots, m$), independence of $\{r_{iM}(\cdot)\}$ and $\{\boldsymbol{\psi}(\boldsymbol{\varepsilon}_i(\cdot))\}$ (cf. Assumption (B.3)), the Cauchy–Schwarz inequality, Assumptions (B.1) and (A.3) and the boundedness of ω_q. From here we can conclude that, as $m \to \infty$, each (j, ℓ)th element of $\boldsymbol{S}_1^0$ is $O_p(q(m)m^{-1/2})$, and so is $\boldsymbol{S}_1^0$, uniformly in $\boldsymbol{a}, \boldsymbol{b}$ such that $\max_{1\le j\le d}(|a_j| + |b_j|) < C$, with some $C > 0$.

Proceeding in the same way, we obtain $\boldsymbol{S}_2^0 = O_p(q(m)m^{-1/4})$ and $\boldsymbol{S}_3^0 = O_p(q(m)m^{-1/4})$, as $m \to \infty$, uniformly in $\boldsymbol{a}, \boldsymbol{b}$ such that $\max_{1\le j\le d}(|a_j| + |b_j|) < C$ for some constant $C > 0$.

Since $\widehat{\varepsilon}_{i,j}(s) = \varepsilon_{i,j}(s) - \widehat{\alpha}^*_{jm}/\sqrt{m} - \widehat{\beta}^*_{jm} r_{iM}(s)/\sqrt{m}$ and $\widehat{\alpha}^*_{jm} = O_P(1)$, $\widehat{\beta}^*_{jm} = O_P(1)$, for all $j = 1, \ldots, d$ (see (4.6) and (4.7), respectively), we obtain, due to the monotonicity of the ψ_j's, that $\boldsymbol{S}_1 + \boldsymbol{S}_2 + \boldsymbol{S}_3 = O_p(q(m)m^{-1/4})$. Combining this with the corresponding estimates for $-q < k < 0$, we get

$$\widehat{\boldsymbol{\Sigma}}_m - \widetilde{\boldsymbol{\Sigma}}_m = O_P\big(q(m)m^{-1/4}\big) \quad (m \to \infty),$$

i.e. (4.10).

It remains to estimate $\widetilde{\boldsymbol{\Sigma}}_m - \boldsymbol{\Sigma}_m$.

First, notice that, for $k \ge 0$,

$$\widetilde{\boldsymbol{\Gamma}}_k - \boldsymbol{\Gamma}_k = \frac{1}{m} \sum_{i=1}^{m-k} [(\widetilde{\boldsymbol{z}}_i - \boldsymbol{z}_i)(\widetilde{\boldsymbol{z}}_{i+k} - \boldsymbol{z}_{i+k})^T + (\widetilde{\boldsymbol{z}}_i - \boldsymbol{z}_i)\boldsymbol{z}_{i+k}^T + \boldsymbol{z}_i(\widetilde{\boldsymbol{z}}_{i+k} - \boldsymbol{z}_{i+k})^T].$$

Further, for $i \in \mathbb{Z}$, $n \in \mathbb{N}$,

$$\begin{aligned}\widetilde{\boldsymbol{z}}_i - \boldsymbol{z}_i &= \frac{1}{n} \sum_{\nu=1}^{n} r_{iM}(s_\nu) \boldsymbol{\psi}(\boldsymbol{\varepsilon}_i(s_\nu)) - \sum_{\nu=1}^{n} \int_{s_{\nu-1}}^{s_\nu} r_{iM}(s) \boldsymbol{\psi}(\boldsymbol{\varepsilon}_i(s)) ds \\ &= \sum_{\nu=1}^{n} \int_{s_{\nu-1}}^{s_\nu} [r_{iM}(s_\nu) \boldsymbol{\psi}(\boldsymbol{\varepsilon}_i(s_\nu)) - r_{iM}(s) \boldsymbol{\psi}(\boldsymbol{\varepsilon}_i(s))] ds \\ &= \sum_{\nu=1}^{n} \int_{s_{\nu-1}}^{s_\nu} [r_{iM}(s_\nu) - r_{iM}(s)] \boldsymbol{\psi}(\boldsymbol{\varepsilon}_i(s_\nu)) ds + \sum_{\nu=1}^{n} \int_{s_{\nu-1}}^{s_\nu} r_{iM}(s) [\boldsymbol{\psi}(\boldsymbol{\varepsilon}_i(s_\nu)) - r_{iM}(s) \boldsymbol{\psi}(\boldsymbol{\varepsilon}_i(s))] ds \\ &= \boldsymbol{u}_i + \boldsymbol{v}_i.\end{aligned}$$

Thus,

$$\sum_{0\le k<q} \omega_q(k)(\widetilde{\boldsymbol{\Gamma}}_k - \boldsymbol{\Gamma}_k) = \boldsymbol{A}_1 + \boldsymbol{A}_2 + \boldsymbol{A}_3,$$

where

$$\boldsymbol{A}_1 = \sum_{0\le k<q} \omega_q(k)\frac{1}{m}\sum_{i=1}^{m-k}[\boldsymbol{u}_i\boldsymbol{u}_{i+k}^T + \boldsymbol{u}_i\boldsymbol{v}_{i+k}^T + \boldsymbol{v}_i\boldsymbol{u}_{i+k}^T + \boldsymbol{v}_i\boldsymbol{v}_{i+k}^T], \tag{4.11}$$

$$\boldsymbol{A}_2 = \sum_{0\le k<q} \omega_q(k)\frac{1}{m}\sum_{i=1}^{m-k}(\boldsymbol{u}_i + \boldsymbol{v}_i)\widetilde{\boldsymbol{z}}_{i+k}^T, \tag{4.12}$$

$$\boldsymbol{A}_3 = \sum_{0\le k<q} \omega_q(k)\frac{1}{m}\sum_{i=1}^{m-k}\boldsymbol{z}_i(\boldsymbol{u}_{i+k} + \boldsymbol{v}_{i+k})^T. \tag{4.13}$$

Since all the matrices appearing on the right-hand side of (4.11) are of the same type, we shall only treat one of them. Consider, for example, the (j, ℓ)th element of the matrix $\boldsymbol{u}_i\boldsymbol{v}_{i+k}^T$. We have

$$u_{i,j}v_{i+k,\ell} = \sum_{\mu=1}^{n}\sum_{\nu=1}^{n}\int_{s_{\mu-1}}^{s_\mu}\int_{s_{\nu-1}}^{s_\nu}\Big([r_{iM}(s_\mu) - r_{iM}(s)]\psi_j(\varepsilon_{i,j}(s_\mu))r_{i+k,M}(t)[\psi_\ell(\varepsilon_{i+k,\ell}(s_\nu)) - \psi_\ell(\varepsilon_{i+k,\ell}(t))]\Big)dsdt,$$

and from here, using the independence of $\{r_{iM}(\cdot)\}$ and $\{\boldsymbol{\psi}(\boldsymbol{\varepsilon}_i(\cdot))\}$, the Cauchy–Schwarz inequality and stationarity,

$$\begin{aligned}
E\left|u_{i,j}v_{i+k,\ell}\right| &\le \sum_{\mu=1}^{n}\sum_{\nu=1}^{n}\int_{s_{\mu-1}}^{s_\mu}\int_{s_{\nu-1}}^{s_\nu}\Big(\|r_{iM}(s_\mu) - r_{iM}(s)\|_2 \cdot \|r_{i+k,M}(t)\|_2 \\
&\quad \times \|\psi_j(\varepsilon_{i,j}(s_\mu))\|_2 \cdot \|\psi_\ell(\varepsilon_{i+k,\ell}(s_\nu)) - \psi_\ell(\varepsilon_{i+k,\ell}(t))\|_2\Big)dsdt \\
&\le \sum_{\mu=1}^{n}\sum_{\nu=1}^{n}\int_{s_{\mu-1}}^{s_\mu}\int_{s_{\nu-1}}^{s_\nu}\Big(\sup_{s\in[0,1]}\|\psi_j(\varepsilon_{i,j}(s))\|_2 \cdot \sup_{h\in[0,1/n]}\|r_{iM}(s_\mu) - r_{iM}(s_\mu - h)\|_2 \\
&\quad \times \sup_{t\in[0,1]}\|r_{i+k,M}(t)\|_2 \cdot \sup_{h\in[0,1/n]}\|\psi_\ell(\varepsilon_{i+k,\ell}(s_\nu)) - \psi_\ell(\varepsilon_{i+k,\ell}(s_\nu - h))\|_2\Big)dsdt \\
&= \sup_{s\in[0,1]}\|\psi_j(\varepsilon_{0,j}(s))\|_2 \cdot \sup_{t\in[0,1]}\|r_{0M}(t)\|_2 \cdot \frac{1}{n}\sum_{\mu=1}^{n}\sup_{h\in[0,1/n]}\|r_{0M}(s_\mu) - r_{0M}(s_\mu - h)\|_2 \\
&\quad \times \frac{1}{n}\sum_{\nu=1}^{n}\sup_{h\in[0,1/n]}\|\psi_\ell(\varepsilon_{0,\ell}(s_\nu)) - \psi_\ell(\varepsilon_{0,\ell}(s_\nu - h))\|_2.
\end{aligned}$$

Now, using Assumptions (A.3), (B.1) and (B.7a)–(B.7b), together with the fact that ω_q is bounded and $q(m) = O(\log m)$, we can easily deduce that

$$\sum_{0\le k<q}\omega_q(k)\frac{1}{m}\sum_{i=1}^{m-k}u_{i,j}v_{i+k,\ell} = o_p(1) \quad (m\to\infty).$$

The same result holds for all elements of the matrix $\boldsymbol{A}_1$. Concerning the matrices $\boldsymbol{A}_2$ and $\boldsymbol{A}_3$, we can proceed in the same way. It suffices to write $\boldsymbol{z}_i = \sum_{\nu=1}^{n}\int_{s_{\nu-1}}^{s_\nu} r_{iM}(s)\boldsymbol{\psi}(\varepsilon_i(s))ds$ and make use of Assumptions (A.3) and (B.1) again together with either (B.7a) or (B.7b). Combining all the asymptotics above with the corresponding estimates for $-q < k < 0$, we get

$$\widetilde{\boldsymbol{\Sigma}}_m - \boldsymbol{\Sigma}_m = o_p(1) \quad (m\to\infty), \tag{4.14}$$

which together with (4.9) and (4.10) concludes the proof. □

5. Some auxiliary results

In the sequel, C and D denote generic positive constants, which may vary from case to case.

For the sake of brevity, we let $\{x_i(\cdot)\}$ denote any of the sequences $\{r_{iM}(\cdot) - Er_{iM}(\cdot)\}$, $\{\psi_j(\varepsilon_{i,j}(\cdot))\}$ or $\{r_{iM}(\cdot)\psi_j(\varepsilon_{i,j}(\cdot))\}$ and write $\{x_i^{(L)}(\cdot)\}$ for the corresponding counterparts of $\{r_{iM}^{(L)}(\cdot) - Er_{iM}^{(L)}(\cdot)\}$, $\{\psi_j(\varepsilon_{i,j}^{(L)}(\cdot))\}$ or $\{r_{iM}^{(L)}(\cdot)\psi_j(\varepsilon_{i,j}^{(L)}(\cdot))\}$, respectively.

Lemma 5.1. *Under the assumptions of Theorem 2.1, possibly extended to an $L_{2+\Delta}$-approximability condition in* (B.4) *and* (B.5) *(cf. Remark 2.5),*

(i) *there is a constant $C > 0$ such that, for every $\ell \in \mathbb{Z}$, $K \in \mathbb{N}$, and $s \in [0, 1]$,*

$$E\left|\sum_{i=\ell+1}^{\ell+K} x_i(s)\right|^p \le C \sup_{s\in[0,1]} \|x_0(s)\|_p^p K^{p/2}, \quad 2 \le p \le 2+\Delta,$$

and, for $b_1 \ge b_2 \ge \cdots \ge b_K > 0$,

$$E \max_{1\le k\le K} \left|b_k \sum_{i=\ell+1}^{\ell+k} x_i(s)\right|^2 \le C \sup_{s\in[0,1]} \|x_0(s)\|_2^2 (\log K)^2 \sum_{k=1}^{K} b_k^2, \tag{5.1}$$

$$E \max_{1\le k\le K} \left|b_k \sum_{i=\ell+1}^{\ell+k} x_i(s)\right|^p \le C \sup_{s\in[0,1]} \|x_0(s)\|_p^p \sum_{k=1}^{K} b_k^p k^{p/2-1}, \quad 2 < p \le 2+\Delta; \tag{5.2}$$

(ii) *for some $D > 0$ and all $m \in \mathbb{N}$, $s \in [0, 1]$,*

$$E\left(\max_{1\le k\le \lfloor mT\rfloor} \frac{1}{\sqrt{m}\,(k/m)^\gamma}\left|\sum_{i=m+1}^{m+k} x_i(s)\right|\right)^2 \le D \sup_{s\in[0,1]} \|x_0(s)\|_2^2 (\log m)^2, \tag{5.3}$$

$$E\left(\max_{1\le k\le \lfloor mT\rfloor} \frac{1}{\sqrt{m}\,(k/m)^\gamma}\left|\sum_{i=m+1}^{m+k} x_i(s)\right|\right)^p \le D \sup_{s\in[0,1]} \|x_0(s)\|_p^p, \quad 2 < p \le 2+\Delta; \tag{5.4}$$

For the proof of (5.2) and (5.4), however, it is necessary to replace the L_2-approximability conditions in Assumptions (B.4) *and* (B.5) *by a corresponding $L_{2+\Delta}$-approximability assumption, with some $\Delta > 0$.*

(iii) *uniformly in $s \in [0, 1]$ and for any $q_m \to \infty$,*

$$\max_{1\le i\le \lfloor m(T+1)\rfloor} |r_{iM}(s)| = O_P(m^{1/3}), \tag{5.5}$$

$$\sup_{s\in[0,1]} P\Big(\max_{1\le i\le \lfloor m(T+1)\rfloor} |r_{iM}(s)| \ge q_m m^{1/3}\Big) \to 0. \tag{5.6}$$

Proof. (i) Making use of the $L_{2+\Delta}$-approximability from Assumption B.4 (with $\Delta \ge 0$, cf. Remarks 2.1 and 2.3), the first bound has been obtained in Berkes et al. [4], Proposition 4. Observe that, in our case,

$$\|x_i(s)\|_p = \|x_i^{(L)}(s)\|_p \le \sup_{s\in[0,1]} \|x_0(s)\|_p, \quad \text{for } 2 \le p \le 2+\Delta,\ s \in [0, 1].$$

Similarly, for the two other bounds confer, e.g., Kirch [13], Theorems B.1 and B.3, which are based on earlier results of Móricz [16] and Móricz et al. [17] in combination with Fazekas and Klesov [8].

Note that the sequence $\{r_{iM}(s)\psi_j(\varepsilon_{i,j}(s))\}$ also satisfies the $L_{2+\Delta}$-approximability condition, uniformly in $s \in [0, 1]$, since

$$\begin{aligned}
&\|r_{iM}(s)\psi_j(\varepsilon_{i,j}(s)) - r_{iM}^{(L)}(s)\psi_j(\varepsilon_{i,j}^{(L)}(s))\|_{2+\Delta} \\
&\le \|(r_{iM}(s) - r_{iM}^{(L)}(s))\psi_j(\varepsilon_{i,j}(s))\|_{2+\Delta} + \|r_{iM}^{(L)}(s)(\psi_j(\varepsilon_{i,j}(s)) - \psi_j(\varepsilon_{i,j}^{(L)}(s)))\|_{2+\Delta} \\
&\le \sup_{s\in[0,1]} \|r_{iM}(s) - r_{iM}^{(L)}(s)\|_{2+\Delta} \sup_{s\in[0,1]} \|\psi_j(\varepsilon_{i,j}(s))\|_{2+\Delta} + \sup_{s\in[0,1]} \|r_{iM}^{(L)}(s)\|_{2+\Delta} \sup_{s\in[0,1]} \|\psi_j(\varepsilon_{i,j}(s)) - \psi_j(\varepsilon_{i,j}^{(L)}(s))\|_{2+\Delta},
\end{aligned}$$

where, for the second inequality, we have used the independence of the sequences $\{r_{iM}(s)\}$ and $\{\psi_j(\varepsilon_{i,j}(s))\}$.

(ii) It follows immediately from the fact that the sequence $\{x_i(s)\}$, $s \in [0, 1]$ fixed, satisfies Assumptions (B.1) and (B.4) together with the estimates in (5.1) and (5.2).

(iii) By (i),

$$E\left|\sum_{i=\ell+1}^{\ell+K} (r_{iM}(s) - Er_{iM}(s))\right|^3 \le C \sup_{s\in[0,1]} \|r_{0M}(s) - Er_{0M}(s)\|_3^3 K^{3/2}, \quad s \in [0, 1].$$

We also have, for $s \in [0, 1]$,

$$\max_{1\le i\le \lfloor m(T+1)\rfloor} |r_{iM}(s)| \le \max_{1\le i\le \lfloor m(T+1)\rfloor} |r_{iM}(s) - Er_{iM}(s)| + \max_{1\le i\le \lfloor m(T+1)\rfloor} |Er_{iM}(s)|, \quad \text{and}$$

$$\max_{1\le i\le \lfloor m(T+1)\rfloor} |r_{iM}(s) - Er_{iM}(s)| \le D\Big(\frac{1}{\lfloor m(T+1)\rfloor} \sum_{i=1}^{\lfloor m(T+1)\rfloor} |r_{iM}(s) - Er_{iM}(s)|^3\Big)^{1/3} \lfloor m(T+1)\rfloor^{1/3}.$$

Since, by our assumptions, for fixed $s \in [0, 1]$, $\{r_{iM}(s) - Er_{iM}(s)\}$ is a stationary and ergodic sequence and $\sup_{s\in[0,1]} E|r_{iM}(s) - Er_{iM}(s)|^3 < \infty$, the ergodic theorem implies, as $m \to \infty$,

$$\frac{1}{\lfloor m(T+1)\rfloor}\sum_{i=1}^{\lfloor m(T+1)\rfloor} |(r_{iM}(s) - Er_{iM}(s))|^3 \to E|(r_{0M}(s) - Er_{0M}(s))|^3 \le \sup_{s\in[0,1]} E|r_{0M}(s) - Er_{0M}(s)|^3 < \infty.$$

Combining all these we get (5.5), which immediately implies (5.6). □

In the following E^* and var^* denote the conditional expectation and conditional variance given $r_{iM}(\cdot), i = 1, \ldots, m; m+1, \ldots, \lfloor mT\rfloor$. We omit the index j, i.e., we write $\varepsilon_i(s), \psi, \ldots$ instead of $\varepsilon_{i,j}(s), \psi_j(s), \ldots$.

Lemma 5.2. *Let the assumptions of Theorem 2.1 be satisfied. Then, as $m \to \infty$,*

$$\sup_{|a|+|b|\le C} |Z_m(a, b) - E^*Z_m(a, b)| = O_P\big(m^{-\eta}\big),$$

$$E^*Z_m(a, b) = \frac{1}{n}\sum_{\nu=1}^{n}\frac{\lambda'(0, s_\nu)}{2}\frac{1}{m}\sum_{i=1}^{m}\big(a + br_{iM}(s_\nu)\big)^2 + O_P\big(m^{-\eta}(|a|^3 + |b|^3)\big),$$

and

$$\sup_{|a|+|b|\le C}\left|Z_m(a, b) - \frac{1}{n}\sum_{\nu=1}^{n}\frac{\lambda'(0, s_\nu)}{2}\frac{1}{m}\sum_{i=1}^{m}\big(a + br_{iM}(s_\nu)\big)^2\right| = O_P(m^{-\eta}),$$

for some $\eta > 0$, where

$$Z_m(a, b) = \frac{1}{n}\sum_{\nu=1}^{n}\sum_{i=1}^{m}\big(\rho(\varepsilon_i(s_\nu) - a/\sqrt{m} - br_{iM}(s_\nu)/\sqrt{m}) - \rho(\varepsilon_i(s_\nu)) + (a/\sqrt{m} + br_{iM}(s_\nu)/\sqrt{m})\psi(\varepsilon_i(s_\nu))\big).$$

Proof. The lines of the proof are quite standard. We just need to derive a proper approximation for the conditional expectation and variance of $Z_m(a, b)$.

Whenever convenient we use the short-hand notations

$$d_i(s_\nu) = a + br_{iM}(s_\nu) \quad \text{and}$$
$$g(\varepsilon_i(s_\nu), x, d_i(s_\nu)) = \operatorname{sign} d_i(s_\nu)\big(-\psi(\varepsilon_i(s_\nu) - x \operatorname{sign} d_i(s_\nu)) + \psi(\varepsilon_i(s_\nu))\big), \quad i \in \mathbb{Z}.$$

Note that, for any d,

$$\rho(\varepsilon_i - d) - \rho(\varepsilon_i) + d\psi(\varepsilon_i) = \operatorname{sign} d\int_0^{|d|}\big(-\psi(\varepsilon_i - x \operatorname{sign} d) + \psi(\varepsilon_i)\big)dx \ge 0, \quad i \in \mathbb{Z}.$$

Direct calculations in combination with Lemma 5.1 result in

$$\begin{aligned}
E^*Z_m(a, b) &= \frac{1}{n}\sum_{\nu=1}^{n} E^*\sum_{i=1}^{m}\int_0^{|d_i(s_\nu)|/\sqrt{m}} g(\varepsilon_i(s_\nu), x, d_i(s_\nu))dx \\
&= \frac{1}{n}\sum_{\nu=1}^{n}\sum_{i=1}^{m}\lambda'(0, s_\nu)d_i^2(s_\nu)\frac{1}{2m} + O_P\Big(\frac{1}{n}\sum_{\nu=1}^{n}\sum_{i=1}^{m}|d_i(s_\nu)|^3\frac{1}{m^{3/2}}\Big) \\
&= \sum_{\nu=1}^{n}\frac{1}{2}\lambda'(0, s_\nu)\left(a + 2ab\frac{1}{m}\sum_{i=1}^{m} r_{iM}(s_\nu) + b^2\frac{1}{m}\sum_{i=1}^{m} r_{iM}^2(s_\nu)\right) + O_P\big(m^{-\eta}(|a|^3 + |b|^3)\big),
\end{aligned}$$

for some $\eta > 0$ and uniformly in $|a| + |b| \le C$.

For the conditional variance we obtain

$$\begin{aligned}
var^*\big\{Z_m(a, b)\big\} &= E^*\Big(\frac{1}{n}\sum_{\nu=1}^{n}\sum_{i=1}^{m}\int_0^{|d_i(s_\nu)|/\sqrt{m}}\big(g(\varepsilon_i(s_\nu), x, d_i(s_\nu)) - E^*g(\varepsilon_i(s_\nu), x, d_i(s_\nu))\big)dx\Big)^2 \\
&= \sum_{i_1=1}^{m} E^*\Big(\frac{1}{n}\sum_{\nu=1}^{n}\int_0^{|d_{i_1}(s_\nu)|/\sqrt{m}}\big(g(\varepsilon_{i_1}(s_\nu), x, d_{i_1}(s_\nu)) - E^*g(\varepsilon_{i_1}(s_\nu), x, d_{i_1}(s_\nu))\big)dx\Big)^2 \\
&\quad + 2E^*\sum_{1\le i_1<i_2\le m}\frac{1}{n}\sum_{\nu_1=1}^{n}\Big\{\Big(\int_0^{|d_{i_1}(s_{\nu_1}|/\sqrt{m})}\big(g(\varepsilon_{i_1}(s_{\nu_1}), x, d_{i_1}(s_{\nu_1})) - E^*g(\varepsilon_{i_1}(s_{\nu_1}), x, d_{i_1}(s_{\nu_1}))\big)dx\Big)
\end{aligned}$$

$$\times \Big(\frac{1}{n}\sum_{\nu_2=1}^{n}\int_0^{|d_{i_2}(s_{\nu_2})|/\sqrt{m}}\big(g(\varepsilon_{i_2}(s_{\nu_2}),y,d_{i_2}(s_{\nu_2}))-E^*g(\varepsilon_{i_2}(s_{\nu_2}),y,d_{i_2}(s_{\nu_2}))\big)dy\Big)\Big\}$$
$$= I_1 + I_2 \quad \text{(say)}.$$

Using Assumption (A.3) together with the Cauchy–Schwarz inequality, we get

$$\begin{aligned}
I_1 &= \sum_{i_1=1}^{m} E^*\Big(\frac{1}{n}\sum_{\nu=1}^{n}\int_0^{|d_{i_1}(s_\nu)|/\sqrt{m}}\big(g(\varepsilon_{i_1}(s_\nu),x,d_{i_1}(s_\nu))-E^*g(\varepsilon_{i_1}(s_\nu),x,d_{i_1}(s_\nu))\big)dx\Big)^2\\
&\le D\sum_{i_1=1}^{m} E^*\Big(\frac{1}{n}\sum_{\nu=1}^{n}\int_0^{|d_{i_1}(s_\nu)|/\sqrt{m}} g(\varepsilon_{i_1}(s_\nu),x,d_{i_1}(s_\nu))dx\Big)^2\\
&\le D\sum_{i_1=1}^{m}\frac{1}{n}\sum_{\nu_1=1}^{n}\frac{1}{n}\sum_{\nu_2=1}^{n}E^*\Big[\Big(\Big|\int_0^{|d_{i_1}(s_{\nu_1})|/\sqrt{m}} g(\varepsilon_{i_1}(s_{\nu_1}),x,d_{i_1}(s_{\nu_1})dx)\Big|\Big)\\
&\quad\times\Big(\Big|\int_0^{|d_{i_1}(s_{\nu_2})|/\sqrt{m}} g(\varepsilon_{i_1}(s_{\nu_2}),z,d_{i_1}(s_{\nu_2}))dz\Big|\Big)\Big]\\
&\le D\sum_{i_1=1}^{m}\frac{1}{n}\sum_{\nu_1=1}^{n}\frac{1}{n}\sum_{\nu_2=1}^{n}\Big[\Big(|d_{i_1}(s_{\nu_1})|/\sqrt{m}\Big)^3\Big(|d_{i_1}(s_{\nu_2})|/\sqrt{m}\Big)^3\Big]^{1/2}\\
&= D\Big(\frac{1}{m}\Big)^{3/2}\Big(|a|^3 m + |b|^3\sum_{i_1=1}^{m}\Big(\frac{1}{n}\sum_{\nu_1=1}^{n}|r_{i_1M}(s_{\nu_1})|^{3/2}\Big)^2 O_P((|a|^3+|b|^3)m^{-\eta})\Big),
\end{aligned}$$

uniformly in $|a|+|b|\le C$.

Concerning I_2 we have, due to the independence of $\{r_{iM}(s)\}$ and $\{\boldsymbol{\varepsilon}_i(s)\}$,

$$\begin{aligned}
I_2 &\le 2\sum_{i_1=1}^{m-1}\sum_{i_2=1}^{m-i_1}\frac{1}{n}\sum_{r_2=1}^{n}\frac{1}{n}\sum_{r_1=1}^{n}\int_0^{|d_{i_1}(s_{r_1})|/\sqrt{m}}\int_0^{|d_{i_1+i_2}(s_{r_2})|/\sqrt{m}}\\
&\quad\Big(E^*(g(\varepsilon_{i_1}(s_{r_1}),x,d_{i_1}(s_{r_1}))))^2\big(E^*(-\psi(\varepsilon_{i_1+i_2}(s_{r_2})-y)+\psi(\varepsilon^{(i_2)}_{i_1+i_2}(s_{r_2})-y))^2\\
&\quad+E^*(-\psi(\varepsilon_{i_1+i_2}(s_{r_2}))+\psi(\varepsilon^{(i_2)}_{i_1+i_2}(s_{r_2})))^2\big)\Big)^{1/2}dxdy\\
&\le D\sum_{i_1=1}^{m-1}\frac{1}{n}\sum_{r_2=1}^{n}\frac{1}{n}\sum_{r_1=1}^{n}|d_{i_1}(s_{r_1})/\sqrt{m}|^{1/2+1}\\
&\quad\times\sum_{i_2=1}^{m-i_1}|d_{i_1+i_2}(s_{r_2})/\sqrt{m}|\sup_{|a|\le a_0}\Big(E^*\big(\psi(\varepsilon_{i_1+i_2}(s_{r_2})-a)-\psi(\varepsilon^{(i_2)}_{i_1+i_2}(s_{r_2})-a)\big)^2\Big)^{1/2}\\
&\le D\frac{1}{n}\sum_{r_2=1}^{n}\frac{1}{n}\sum_{r_1=1}^{n}\frac{1}{m^{3/2}}\sum_{i_1=1}^{m-1}|d_{i_1}(s_{r_1})|^{3/2}\sup_{|a|\le a_0}\sum_{i_2=1}^{m-i_1}|d_{i_1+i_2}(s_{r_2})|\Big(E\big(\psi(\varepsilon_0(s_{r_2})-a)-\psi(\varepsilon_0(s_{r_2})^{(i_2)}-a)\big)^2\Big)^{1/2}\\
&= O_P(m^{-\eta}),
\end{aligned}$$

where we used the fact

$$E|d_{i_1+i_2}(s_{r_1})|^{3/2}\cdot|d_{i_1}(s_{r_2})|\le\Big(E|d_1(s_{r_1})|^3 Ed_1^2(s_{r_2})\Big)^{1/2}.$$

On combining the above estimates for $E^*Z_m(a,b), I_1, I_2$, we conclude that Lemma 5.2 holds true. □

Lemma 5.3. *Let the assumptions of Theorem 2.1 be satisfied. Then, as $m\to\infty$,*

$$\sup_{|a|+|b|\le C}|\boldsymbol{M}_m(a,b)-E^*\boldsymbol{M}_m(a,b)| = O_P(m^{-\eta}),$$

$$E^*\boldsymbol{M}_m(a,b) = -\frac{1}{n}\sum_{\nu=1}^{n}\lambda'(0,s_\nu)\Big(a+b\frac{1}{m}\sum_{i=1}^{m}r_{iM}(s_\nu),\ a\frac{1}{m}\sum_{i=1}^{m}r_{iM}(s_\nu)+b\frac{1}{m}\sum_{i=1}^{m}r^2_{iM}(s_\nu)\Big)^T + O_P\big(m^{-\eta}\big),$$

and

$$\sup_{|a|+|b|\le C}\Big|\boldsymbol{M}_m(a,b)+\frac{1}{n}\sum_{\nu=1}^{n}\frac{1}{m}\lambda'(0,s_\nu)\Big(a\,m+b\sum_{i=1}^{m}r_{iM}(s_\nu),\ a\sum_{i=1}^{m}r_{iM}(s_\nu)+b\sum_{i=1}^{m}r^2_{iM}(s_\nu)\Big)^T\Big| = O_P\big(m^{-\eta}\big),$$

with some $\eta > 0$, *where*

$$\boldsymbol{M}_m(a,b) = \frac{1}{n}\sum_{\nu=1}^{n}\frac{1}{\sqrt{m}}\sum_{i=1}^{m}\big(1, r_{iM}(s_\nu)\big)^T\big(\psi(\varepsilon_i(s_\nu) - (a + br_{iM}(s_\nu))/\sqrt{m}) - \psi(\varepsilon_i(s_\nu))\big).$$

Proof. Again one has to get suitable approximations for the conditional expectation $\boldsymbol{M}_m(a,b)$ and the conditional (2×2)-variance matrix

$$var^*\{\boldsymbol{M}_n(a,b)\} = E^*\big(\boldsymbol{M}_n(a,b) - E^*\boldsymbol{M}_n(a,b)\big)\big(\boldsymbol{M}_n(a,b) - E^*\boldsymbol{M}_n(a,b)\big)^T.$$

We start with the conditional expectation

$$\begin{aligned}
E^*\boldsymbol{M}_m^T(a,b) &= \frac{1}{n}\sum_{\nu=1}^{n}\frac{1}{\sqrt{m}}\sum_{i=1}^{m}\big(1, r_{iM}(s_\nu)\big)\big(-\lambda(d_i(s_\nu)/\sqrt{m}, s_\nu)\big)\\
&= -\frac{1}{n}\sum_{\nu=1}^{n}\frac{1}{m}\lambda'(0,s_\nu)\sum_{i=1}^{m}\big(1, r_{iM}(s_\nu)\big)d_i(s_\nu) + O_P\Big(\frac{1}{n}\sum_{\nu=1}^{n}\frac{1}{m^{3/2}}\sum_{i=1}^{m}\big(1, |r_{iM}(s_\nu)|\big)|d_i^2(s_\nu)|\Big)\\
&= -\frac{1}{n}\sum_{\nu=1}^{n}\frac{1}{m}\lambda'(0,s_\nu)\sum_{i=1}^{m}\big(a + br_{iM}(s_\nu), ar_{iM}(s_\nu) + br_{iM}^2(s_\nu)\big)\\
&\quad + O_P\Big(a^2m^{-1/2} + b^2\frac{1}{n}\sum_{\nu=1}^{n}\frac{1}{m^{3/2}}\sum_{i=1}^{m}|r_{iM}(s_\nu)|^3\Big) = O_P\big((a^2+b^2)m^{-\eta}\big),
\end{aligned}$$

uniformly in $|a| + |b| \le C$, where the rates above are to be understood componentwise.

For the conditional variance matrix we only calculate one term. The calculation of the others is similar and will therefore be omitted. We have

$$\begin{aligned}
& var^*\Big\{\frac{1}{n}\sum_{\nu=1}^{n}\frac{1}{\sqrt{m}}\sum_{i=1}^{m}r_{iM}(s_\nu)\big(\psi_j(\varepsilon_{i,j}(s_\nu) - d_i(s_\nu)/\sqrt{m}) - \psi_j(\varepsilon_{i,j}(s_\nu))\big)\Big\}\\
&= \frac{1}{m}\sum_{i=1}^{m}E^*\Big(\frac{1}{n}\sum_{\nu=1}^{n}r_{iM}(s_\nu)\big(\psi(\varepsilon_i(s_\nu) - d_i(s_\nu)/\sqrt{m}) - \psi(\varepsilon_i(s_\nu)) + \lambda_j(d_i(s_\nu)/\sqrt{m}, s_\nu)\big)\Big)^2\\
&\quad + 2\frac{1}{n}\sum_{r_1=1}^{n}\frac{1}{n}\sum_{r_2=1}^{n}\frac{1}{m}\sum_{i=1}^{m}\sum_{j=1}^{m-i}r_{iM}(s_{r_1})r_{i+j,M}(s_{r_2})\\
&\quad \times E^*\big(\psi_j(\varepsilon_{i,j}(s_{r_1}) - d_i(s_{r_1})/\sqrt{m}) - \psi_j(\varepsilon_{i,j}(s_{r_1})) + \lambda_j(d_i(s_{r_1})/\sqrt{m}, s_{r_1})\big)\\
&\quad \times \big(\psi(\varepsilon_{i+j}(s_{r_2}) - d_{i+j}(s_{r_2})/\sqrt{m}) - \psi_j(\varepsilon_{i+j}(s_{r_2})) + \lambda_j(d_{i+j}(s_{r_2})/\sqrt{m}, s_{r_2})\big)\\
&= J_1 + 2J_2 \quad \text{(say)}.
\end{aligned}$$

In view of Assumption (A.2), a similar estimate as that for I_1 in the proof of Lemma 5.2 gives

$$\begin{aligned}
J_1 &= \frac{1}{m}\sum_{i=1}^{m}\frac{1}{n}\sum_{\nu=1}^{n}\frac{1}{n}\sum_{r=1}^{n}r_{iM}(s_\nu)r_{iM}(s_r)\\
&\quad \times E^*\big(\psi(\varepsilon_i(s_\nu) - d_i(s_\nu)/\sqrt{m}) - \psi(\varepsilon_i(s_\nu)) + \lambda(d_i(s_\nu)/\sqrt{m}, s_\nu)\big)\\
&\quad \times \big(\psi(\varepsilon_i(s_r) - d_i(s_r)/\sqrt{m}) - \psi(\varepsilon_i(s_r)) + \lambda(d_i(s_r)/\sqrt{m}, s_r)\big)\\
&\le D\frac{1}{m}\sum_{i=1}^{m}\frac{1}{n}\sum_{\nu=1}^{n}\frac{1}{n}\sum_{r=1}^{n}|r_{iM}(s_\nu)||r_{iM}(s_r)|\\
&\quad \times E^*\big(\psi(\varepsilon_i(s_\nu) - d_i(s_\nu)/\sqrt{m}) - \psi(\varepsilon_i(s_\nu))^2E^*\big(\psi(\varepsilon_i(s_r) - d_i(s_r)/\sqrt{m}) - \psi(\varepsilon_i(s_r))^2\Big)^{1/2}\\
&\le D\frac{1}{m}\sum_{i=1}^{m}\Big(\frac{1}{n}\sum_{\nu=1}^{n}|r_{iM}(s_\nu)||d_i(s_\nu)/\sqrt{m}|^{1/2}\Big)^2\\
&\le D\frac{1}{n}\sum_{\nu=1}^{n}\frac{1}{m}\sum_{i=1}^{m}|r_{iM}(s_\nu)|^2|d_i(s_\nu)/\sqrt{m})| = O_P(m^{-\eta}).
\end{aligned}$$

Concerning J_2 we obtain

$$\begin{aligned}
J_2 =\ & \frac{1}{n}\sum_{r_1=1}^{n}\frac{1}{n}\sum_{r_2=1}^{n}\frac{1}{m}\sum_{i=1}^{m}\sum_{j=1}^{m-i} r_{iM}(s_{r_1})r_{i+j,M}(s_{r_2})\\
&\times E^*\big(\psi(\varepsilon_i(s_{r_1}) - d_i(s_{r_1})/\sqrt{m}) - \psi(\varepsilon_i(s_{r_1})) + \lambda(d_i(s_{r_1})/\sqrt{m}, s_{r_1})\big)\\
&\times \big(\psi(\varepsilon_{i+j}(s_{r_2}) - d_{i+j}(s_{r_2})/\sqrt{m}) - \psi(\varepsilon_{i+j}(s_{r_2})) - (\psi(\varepsilon^{(j)}_{i+j}(s_{r_2}) - d_{i+j}(s_{r_2})/\sqrt{m}) - \psi(\varepsilon^{(j)}_{i+j}(s_{r_2})))\big),
\end{aligned}$$

and, uniformly in $|a| + |b| \le C$,

$$\begin{aligned}
|J_2| \le\ & D\frac{1}{n}\sum_{r_1=1}^{n}\frac{1}{n}\sum_{r_2=1}^{n}\frac{1}{m}\sum_{i=1}^{m}\sum_{j=1}^{m-i} |r_{iM}(s_{r_1})r_{i+j,M}(s_{r_2})|(|d_i(s_{r_1})|/\sqrt{m})^{1/2}\\
&\times \Big\{E^*\big(\psi(\varepsilon_{i+j}(s_{r_2}) - d_{i+j}(s_{r_2})/\sqrt{m}) - \psi(\varepsilon_{i+j}(s_{r_2})) - (\psi(\varepsilon^{(j)}_{i+j}(s_{r_2}) - d_{i+j}(s_{r_2})/\sqrt{m}) - \psi(\varepsilon^{(j)}_{i+j}(s_{r_2})))\big)^2\Big\}^{1/2}\\
\le\ & D\frac{1}{n}\sum_{r_1=1}^{n}\frac{1}{n}\sum_{r_2=1}^{n}\frac{1}{m^{3/2}}\sum_{i=1}^{m}\sum_{j=1}^{m-i}\big(a^{1/2} + b^{1/2}|r_{iM}(s_{r_1}|)\big)|r_{iM}(s_{r_1})r_{i+j,M}(s_{r_2})|\\
&\times \sup_{|a|\le a_0}\big(E(\psi(\varepsilon_0(s_{r_2}) - a) - \psi(\varepsilon_0^{(j)}(s_{r_2}) - a))^2\big)^{1/2}\\
=\ & O_P\big((a^{1/2} + b^{1/2})m^{-\eta}\big).
\end{aligned}$$

Now, a similar estimate as that for I_2 in the proof of Lemma 5.2 gives

$$\sup_{|a|+|b|\le C} |J_2| = O_P\big(m^{-\eta}\big),$$

with some $\eta > 0$, so that altogether we have

$$\sup_{|a|+|b|\le C} var^*\Big\{\frac{1}{n}\sum_{i=1}^{n}\frac{1}{\sqrt{m}}\sum_{i=1}^{m} r_{iM}(s_r)\big(\psi(\varepsilon_i(s_r) - d_i(s_r)/\sqrt{m}) - \psi(\varepsilon_i(s_r))\big)\Big\} = O_P\big(m^{-\eta}\big),$$

for some $\eta > 0$. □

Remark 5.1. Inserting $\widehat{\alpha}^*_{j,m}$ and $\widehat{\beta}^*_{j,m}$ (as defined in (4.3)) for a, b into the assertion of Lemma 5.3 and omitting the index j for the sake of simplicity, we receive

$$\begin{aligned}
&\frac{1}{n}\sum_{\nu=1}^{n}\frac{1}{\sqrt{m}}\sum_{i=1}^{m}\big(1, r_{iM}(s_\nu)\big)^T\big(\psi(\varepsilon_i(s_\nu) - (\widehat{\alpha}^*_m + \widehat{\beta}^*_m r_{iM}(s_\nu))/\sqrt{m}) - \psi(\varepsilon_i(s_\nu))\big)\\
&+ \frac{1}{n}\sum_{\nu=1}^{n}\lambda'(0, s_\nu)\Big(\widehat{\alpha}_m + \widehat{\beta}^*_m\frac{1}{m}\sum_{i=1}^{m} r_{iM}(s_\nu), \widehat{\alpha}^*_m\frac{1}{m}\sum_{i=1}^{m} r_{iM}(s_\nu) + \widehat{\beta}_m\frac{1}{m}\sum_{i=1}^{m} r^2_{iM}(s_\nu)\Big)^T = O_P\big(m^{-\eta}\big).
\end{aligned}$$

Due to the definition of $\widehat{\alpha}^*_m$ and $\widehat{\beta}^*_m$ and by our assumptions, we have the following asymptotic representation:

$$\widehat{\alpha}^*_m = \frac{1}{\int_0^1\lambda'(0, z)dz}\frac{1}{\sqrt{m}}\sum_{i=1}^{m}\int_0^1\psi(\varepsilon_i(s))ds - \widehat{\beta}^*_m\frac{\int_0^1\lambda'(0, z)Er^2_{0M}(z)dz}{\int_0^1\lambda'(0, z)dz} + o_P(m^{-\eta}),$$

$$\widehat{\beta}^*_m = \frac{\frac{1}{\sqrt{m}}\sum_{i=1}^{m}\int_0^1\psi(\varepsilon_i(s))\Big(r_{iM}(s) - \frac{\int_0^1\lambda'(0,z)Er_{0M}(z)dz}{\int_0^1\lambda'(0,z)dz}\Big)ds}{\int_0^1\lambda'(0, z)Er^2_{0M}(z)dz - \frac{\left(\int_0^1\lambda'(0,z)Er_{0M}(z)dz\right)^2}{\int_0^1\lambda'(0,z)dz}} + o_P(m^{-\eta}).$$

The last two relations are important for getting the limit distribution of our test procedure.

The next lemma follows along the arguments of Lemma 5.4 in Chochola et al. [7], modified along the lines of the proofs of the previous lemmas. So the proof will only be sketched and not be given in detail.

Lemma 5.4. *Let the assumptions of Theorem 2.1 be satisfied. Then, for any $T > 0$, as $m \to \infty$,*

$$\max_{1\le k\le\lfloor mT\rfloor}\left(\frac{|\{N_{k,m}(a, b) - E^*N_{km}(a, b)\}_{a=\widehat{\alpha}^*_m, b=\widehat{\beta}^*_m}|}{(k/m)^\gamma}\right) = O_P(m^{-\eta}),$$

for some $\eta > 0$, *where* $\widehat{\alpha}_m^*$, $\widehat{\beta}_m^*$ *are as in* (4.3), *and*

$$N_{k,m}(a,b) = \frac{1}{n}\sum_{\nu=1}^{n}\frac{1}{\sqrt{m}}\sum_{i=m+1}^{m+k} r_{iM}(s_\nu)\big(\psi(\varepsilon_i(s_\nu) - a/\sqrt{m} - br_{iM}(s_\nu)/\sqrt{m}) - \psi(\varepsilon_i(s_\nu))\big).$$

Proof. Lemma 5.4 is related to Lemma 5.3, but it is somewhat more complicated.

Direct calculations give

$$\begin{aligned} E^*N_{k,m}(a,b) &= -\frac{1}{n}\sum_{\nu=1}^{n}\frac{1}{\sqrt{m}}\sum_{i=m+1}^{m+k} r_{iM}(s_\nu)\lambda\big((a+br_{iM}(s_\nu), s_\nu)/\sqrt{m}\big) \\ &= -\frac{1}{n}\sum_{\nu=1}^{n}\lambda'(0,s_\nu)\frac{1}{m}\Big(a\sum_{i=m+1}^{m+k} r_{iM} + b\sum_{i=m+1}^{m+k} r_{iM}^2(s_\nu)\Big) \\ &\quad + O_P\Big(\frac{1}{n}\sum_{\nu=1}^{n}\frac{1}{\sqrt{m}}\sum_{i=m+1}^{m+k}|r_{iM}(s_\nu)||(a+br_{iM}(s_\nu))/\sqrt{m}|^2\Big), \end{aligned}$$

uniformly for $|a| + |b| \le C$, with some $\eta > 0$. In fact we need to study more carefully the properly standardized remainder

$$\max_{1\le k\le\lfloor mT\rfloor}\frac{1}{(k/m)^\gamma}\Big(\frac{1}{n}\sum_{\nu=1}^{n}\frac{1}{m^{3/2}}\sum_{i=m+1}^{m+k}|r_{iM}(s_\nu)|\Big) + \max_{1\le k\le\lfloor mT\rfloor}\frac{1}{(k/m)^\gamma}\Big(\frac{1}{n}\sum_{\nu=1}^{n}\frac{1}{m^{3/2}}\sum_{i=m+1}^{m+k}|r_{iM}(s_\nu)|^3\Big).$$

Both terms above are $O_P(m^{-\eta})$ for some $\eta > 0$.

Next, we try to get an upper bound for $var^*\{N_{k,m}(a,b)\}$. We have

$$\begin{aligned} var^*\{N_{k,m}(a,b)\} &= \frac{1}{m}\sum_{i=m+1}^{m+k} E^*\Big(\frac{1}{n}\sum_{\nu=1}^{n} r_{iM}(s_\nu)\times(\psi(\varepsilon_i(s_\nu) - d_i(s_\nu)/\sqrt{m}) \\ &\qquad - \psi(\varepsilon_i(s_\nu)) - E^*\psi(\varepsilon_i(s_\nu) - d_i(s_\nu)/\sqrt{m}))\Big)^2 \\ &\qquad + 2\frac{1}{m}\sum_{i=m+1}^{m+k}\frac{1}{n}\sum_{r_1=1}^{n} r_{iM}(s_{r_1})\,E^*\Big(\psi(\varepsilon_i(s_{r_1}) - d_i(s_{r_1})/\sqrt{m}) \\ &\qquad - \psi(\varepsilon_i(s_{r_1})) - E^*\psi(\varepsilon_i(s_{r_1}) - d_i(s_{r_1})/\sqrt{m})\Big)\times\Big(\sum_{\nu=1}^{m+k-i}\frac{1}{n}\sum_{r_2=1}^{n} r_{i+\nu,M}(s_{r_2}) \\ &\qquad \times\big(\psi(\varepsilon_{i+\nu}^{(\nu)}(s_{r_2}) - d_{i+\nu}(s_{r_2})/\sqrt{m}) - \psi(\varepsilon_{i+\nu}^{(\nu)}(s_{r_2})) - E^*\psi(\varepsilon_{i+\nu}^{(\nu)}(s_{r_2}) - d_{i+\nu}(s_{r_2})/\sqrt{m})\big)\Big) \\ &= L_{1,k} + 2L_{2,k} \quad \text{(say)}, \end{aligned}$$

and, along the lines of the proof of Lemma 5.3 (see the estimation of the terms J_1, J_2 there), we get

$$L_{1,k} = \frac{k}{m}m^{-1/2}\big(|a| + |b|\big)O_P(1),$$

$$|L_{2,k}| = \frac{1}{m^{1+1/2}}(|a|^{1/2}k + |b|^{1/2}k)O_P(1),$$

uniformly in $|a| + |b| \le C$ and in $1 \le k \le \lfloor mT\rfloor$. So, altogether we have

$$var^*\{N_{k,m}(a,b)\} = \frac{k}{m}m^{-\eta}(|a| + |b|)O_P(1),$$

uniformly in $|a| + |b| \le C$ and in $1 \le k \le \lfloor mT\rfloor$, with some $\eta > 0$.

Quite similarly we get, for $\ell = 1, 2, \ldots,$

$$var^*\{N_{k+\ell,m}(a,b) - N_{k,m}(a,b)\} = \frac{\ell}{m}m^{-\eta}(|a| + |b|)O_P(1).$$

Then, on applying Theorem B.4 of Kirch [13],

$$\begin{aligned} & m^{-1+2\gamma}E^*\max_{1\le k\le\lfloor mT\rfloor}\Big(\frac{1}{k^\gamma}|N_{m,k}(a,b) - E^*N_{m,k}(a,b)|\Big)^2 \\ &= m^{-1+2\gamma}(\log m)^2\sum_{k=1}^{\lfloor mT\rfloor}\frac{1}{k^{2\gamma}}m^{-\eta}(|a| + |b|)O_P(1) = (\log m)^2 m^{-\eta}(|a| + |b|)O_P(1). \end{aligned}$$

We need to replace a, b by the estimators $\widehat{\alpha}_m^*$, $\widehat{\beta}_m^*$. However our $N_{k,m}(a,b)$ depends on $\varepsilon_1(\cdot),\dots,\varepsilon_m(\cdot)$. Therefore we try to replace $N_{k,m}(a,b)$ by something that is asymptotically equivalent, but does not depend on $\varepsilon_1(\cdot),\dots,\varepsilon_m(\cdot)$.

Toward this note that

$$N_{k,m}^{(m)}(a,b) = \frac{1}{\sqrt{m}}\sum_{i=1}^{k}\frac{1}{n}\sum_{\nu=1}^{n} r_{i+m,M}(s_\nu)\big(\psi(\varepsilon_{i+m}^{(i)}(s_\nu) - d_i(s_\nu)/\sqrt{m}) - \psi(\varepsilon_{m+i}^{(i)}(s_\nu))\big)$$

has all the properties of $N_{k,m}(a,b)$ above, but it is independent of $\varepsilon_1(\cdot),\dots,\varepsilon_m(\cdot)$. This together with the consistency of $\widehat{\alpha}_m^*$ and $\widehat{\beta}_m^*$ implies

$$\begin{aligned}
&m^{-1+2\gamma}\max_{1\le k\le\lfloor mT\rfloor}\left(\frac{1}{k^\gamma}\left|\{N_{m,k}^{(m)}(a,b) - E^*N_{m,k}^{(m)}(a,b)\}_{a=\widehat{\alpha}_m^*,b=\widehat{\beta}_m^*}\right|\right)^2\\
&\quad = O_P\big((\log m)^2 m^{-\eta}\max(|\widehat{\alpha}_m^*| + |\widehat{\beta}_m^*|, |\widehat{\alpha}_m^*|^{1/2} + |\widehat{\beta}_m^*|^{1/2})\big) = O_P\big((\log m)^2 m^{-\eta}\big).
\end{aligned}$$

It is still necessary to show the closeness of $N_{k,m}(a,b)$ and $N_{k,m}^{(m)}(a,b)$. Clearly, $N_{k,m}^{(m)}(a,b)$ is independent of $\varepsilon_1(\cdot),\dots,\varepsilon_m(\cdot)$ and

$$\begin{aligned}
&E^*\big(N_{k,m}(a,b) - N_{k,m}^{(m)}(a,b)\big) = 0,\\
&N_{k,m}(a,b) - N_{k,m}^{(m)}(a,b) = \frac{1}{\sqrt{m}}\frac{1}{n}\sum_{\nu=1}^{n}\sum_{i=1}^{k}(r_{i+m,M}(s_\nu))\\
&\quad\times\left(\left(\psi\left(\varepsilon_{i+m}(s_\nu) - \frac{d_{i+m}(s_\nu)}{\sqrt{m}}\right) - \psi\left(\varepsilon_{i+m}^{(i)}(s_\nu) - \frac{d_{i+m}(s_\nu)}{\sqrt{m}}\right)\right) - \big(\psi(\varepsilon_{i+m}(s_\nu)) - \psi(\varepsilon_{i+m}^{(i)}(s_\nu))\big)\right),\\
&E^*|N_{k,m}(a,b) - N_{k,m}^{(m)}(a,b)| \le \frac{D}{\sqrt{m}}\frac{1}{n}\sum_{\nu=1}^{n}\sum_{i=1}^{k}|r_{i+m,M}(s_\nu)|\\
&\qquad\qquad\times\sup_{|a|\le a_0}E\big(|\psi(\varepsilon_0(s_\nu)-a) - \psi(\varepsilon_0^{(i)}(s_\nu)-a)| + |\psi(\varepsilon_0)(s_\nu) - \psi(\varepsilon_0^{(i)}(s_\nu))|\big)\\
&\qquad\qquad\le \frac{D}{\sqrt{m}}\frac{1}{n}\sum_{\nu=1}^{n}\sum_{i=1}^{\lfloor mT\rfloor}|r_{i+m,M}(s_\nu)|\sup_{|a|\le a_0}E|\psi(\varepsilon_0(s_\nu)-a) - \psi(\varepsilon_0^{(i)}(s_\nu)-a)|,
\end{aligned}$$

which holds for any $1\le k\le\lfloor mT\rfloor$. So, in view of our assumptions,

$$\sup_{|a|+|b|\le C}E^*|N_{k,m}(a,b) - N_{k,m}^{(m)}(a,b,j)| = O_P(m^{-1/2}),$$

whence

$$\sup_{1\le k\le\lfloor mT\rfloor}\left(\frac{\sup_{|a|+|b|\le C}E^*|N_{k,m}(a,b) - N_{k,m}^{(m)}(a,b)|}{(k/m)^\gamma}\right) = O_P(m^{-\eta}),$$

for some $\eta > 0$. A combination of the above estimates completes the proof of Lemma 5.4. □

Remark 5.2. From Lemma 5.4 we get the following approximations:

$$\begin{aligned}
&\frac{1}{\sqrt{m}}\sum_{i=m+1}^{m+k}\sum_{\nu=1}^{n}\frac{1}{n}r_{iM}(s_\nu)\psi(\varepsilon_i(s_\nu) - \widehat{\alpha}_m^*/\sqrt{m} - \widehat{\beta}_m^* r_{iM}(s_\nu)/\sqrt{m})\\
&\quad = \frac{1}{\sqrt{m}}\sum_{i=m+1}^{m+k}\frac{1}{n}\sum_{\nu=1}^{n}r_{iM}(s_\nu)\psi(\varepsilon_i(s_\nu)) - \Big(\widehat{\alpha}_m\frac{1}{m}\sum_{i=m+1}^{m+k}\sum_{\nu=1}^{n}\lambda'(0,s_\nu)r_{iM}(s_\nu)\\
&\qquad + \widehat{\beta}_m^*\frac{1}{m}\sum_{i=m+1}^{m+k}\frac{1}{n}\sum_{\nu=1}^{n}\lambda'(0,s_\nu)r_{iM}^2(s_\nu)\Big) + O_P(m^{-\eta}).
\end{aligned}$$

In view of Remark 5.1, we similarly get

$$\begin{aligned}
&\frac{1}{\sqrt{m}}\sum_{i=1}^{m}\frac{1}{n}\sum_{\nu=1}^{n}r_{iM}(s_\nu)\psi(\varepsilon_i(s_\nu))\\
&\quad = \widehat{\alpha}_m^*\frac{1}{m}\sum_{i=1}^{m}\frac{1}{n}\sum_{\nu=1}^{n}\lambda'(0,s_\nu)r_{iM}(s_\nu) + \widehat{\beta}_m^*\frac{1}{m}\sum_{i=1}^{m}\frac{1}{n}\sum_{\nu=1}^{n}\lambda'(0,s_\nu)r_{iM}^2(s_\nu) + O_P\big(m^{-\eta}\big).
\end{aligned}$$

After some standard steps this results in

$$\frac{1}{\sqrt{m}}\sum_{i=m+1}^{m+k}\frac{1}{n}\sum_{\nu=1}^{n} r_{iM}(s_\nu)\psi(\varepsilon_i(s_\nu)-\widehat{\alpha}_m^*/\sqrt{m}-\widehat{\beta}_m^* r_{iM}(s_\nu)/\sqrt{m})$$
$$=\frac{1}{\sqrt{m}}\Big(\sum_{i=m+1}^{m+k}\frac{1}{n}\sum_{\nu=1}^{n} r_{iM}(s_\nu)\psi(\varepsilon_i(s_\nu))-\frac{k}{m}\sum_{i=1}^{m}\frac{1}{n}\sum_{\nu=1}^{n} r_{iM}(s_\nu)\psi(\varepsilon_i(s_\nu))+O_P\big(m^{-\eta}\big)\Big).$$

Here we also used that

$$\max_{1\le k\le\lfloor mT\rfloor}\frac{1}{m(k/m)^\gamma}\Big|\frac{1}{n}\sum_{\nu=1}^{n}\lambda'(0,s_\nu)\Big(\frac{k}{m}\sum_{i=1}^{m} r_{iM}(s_\nu)-\sum_{i=m+1}^{m+k} r_{iM}(s_\nu)\Big)$$
$$+\frac{1}{n}\sum_{\nu=1}^{n}\lambda'(0,s_\nu)\Big(\frac{k}{m}\sum_{i=1}^{m} r_{iM}^2(s_\nu)-\sum_{i=m+1}^{m+k} r_{iM}^2(s_\nu)\Big)\Big|=o_P(1)$$

as $m\to\infty$ (cf. Lemma 5.5 (iii)).

Lemma 5.5. *Let Assumptions* (B.1), (B.4), *and* (B.6)–(B.7) *be satisfied. Then,*

(i) *there is a constant $C>0$ such that*

$$E\max_{1\le k\le\lfloor mT\rfloor}\frac{1}{\sqrt{m}\,(k/m)^\gamma}\Big|\sum_{i=m+1}^{m+k}\sum_{\nu=1}^{n}\int_{s_{\nu-1}}^{s_\nu}\big(r_{iM}(s_\nu)-r_{iM}(s)\big)\psi_j(\varepsilon_{i,j}(s))ds\Big|$$
$$\le C\,(\log m)\,\frac{1}{n}\sum_{\nu=1}^{n}\sup_{h\in[0,1/n]}\|r_{0M}(s_\nu)-r_{0M}(s_\nu-h)\|_2; \tag{5.7}$$

(ii) *for $j=1,\dots,d$, as $m\to\infty$,*

$$\max_{1\le k\le\lfloor mT\rfloor}\frac{1}{\sqrt{m}\,(k/m)^\gamma}\Big|\sum_{i=m+1}^{m+k}\big(\widetilde{z}_{i,j}-z_{i,j}\big)\Big|$$
$$=\max_{1\le k\le\lfloor mT\rfloor}\frac{1}{\sqrt{m}\,(k/m)^\gamma}\Big|\sum_{i=m+1}^{m+k}\Big(\frac{1}{n}\sum_{\nu=1}^{n} r_{iM}(s_\nu)\psi_j(\varepsilon_{i,j}(s_\nu))-\int_0^1 r_{iM}(s)\psi_j(\varepsilon_{i,j}(s))ds\Big)\Big|=o_P(1). \tag{5.8}$$

Moreover, due to strict stationarity, the above relations also hold with $\max_{1\le k\le\lfloor mT\rfloor}\frac{1}{\sqrt{m}\,(k/m)^\gamma}\sum_{i=m+1}^{m+k}$ *being replaced by* $\frac{1}{\sqrt{m}}\sum_{i=1}^{m}$.

(iii) *for $j=1,\dots,d$, as $m\to\infty$,*

$$\max_{1\le k\le\lfloor mT\rfloor}\frac{1}{m\,(k/m)^\gamma}\Big|\sum_{i=m+1}^{m+k}\frac{1}{n}\sum_{\nu=1}^{n}\lambda'(0,s_\nu)\big(r_{iM}(s_\nu)-Er_{0M}(s_\nu)\big)\Big|=o_P(1); \tag{5.9}$$
$$\max_{1\le k\le\lfloor mT\rfloor}\frac{1}{m\,(k/m)^\gamma}\Big|\sum_{i=m+1}^{m+k}\frac{1}{n}\sum_{\nu=1}^{n}\lambda'(0,s_\nu)\big(r_{iM}^2(s_\nu)-Er_{0M}^2(s_\nu)\big)\Big|=o_P(1). \tag{5.10}$$

Moreover, the above relations also hold with $\max_{1\le k\le\lfloor mT\rfloor}\frac{1}{m\,(k/m)^\gamma}\sum_{i=m+1}^{m+k}$ *being replaced by* $\frac{1}{m}\sum_{i=1}^{m}$.

Proof. Again, for the sake of simplicity, we omit the index j in the sequel.

(i) On interchanging summation, expectation and integration, a similar argument as in the proof of (5.3) gives

$$E\max_{1\le k\le\lfloor mT\rfloor}\frac{1}{\sqrt{m}\,(k/m)^\gamma}\Big|\sum_{i=m+1}^{m+k}\sum_{\nu=1}^{n}\int_{s_{\nu-1}}^{s_\nu}\big(r_{iM}(s_\nu)-r_{iM}(s)\big)\psi(\varepsilon_i(s))ds\Big|$$
$$\le\sum_{\nu=1}^{n}\int_{s_{\nu-1}}^{s_\nu}\Big\{E\Big(\max_{1\le k\le\lfloor mT\rfloor}\frac{1}{\sqrt{m}\,(k/m)^\gamma}\Big|\sum_{i=m+1}^{m+k}\big(r_{iM}(s_\nu)-r_{iM}(s)\big)\psi(\varepsilon_i(s))\Big|\Big)^2\Big\}^{1/2}ds$$
$$\le D\,(\log m)\sup_{s\in[0,1]}\|\psi(\varepsilon_0(s))\|_2\,\frac{1}{n}\sum_{\nu=1}^{n}\sup_{h\in[0,1/n]}\|r_{0M}(s_\nu)-r_{0M}(s_\nu-h)\|_2,$$

with some $D>0$, where in the last inequality we made use of the independence of the sequences $\{r_{iM}(\cdot)\}$ and $\{\varepsilon_i(\cdot)\}$. Since $\sup_{s\in[0,1]}\|\psi(\varepsilon_0(s))\|_2<\infty$, this proves (i).

(ii) Consider

$$
\begin{aligned}
&\max_{1\le k\le\lfloor mT\rfloor} \frac{1}{\sqrt{m}\,(k/m)^{\gamma}} \left| \sum_{i=m+1}^{m+k} \left(\frac{1}{n} \sum_{\nu=1}^{n} r_{iM}(s_\nu)\psi(\varepsilon_i(s_\nu)) - \int_0^1 r_{iM}(s)\psi(\varepsilon_i(s))ds \right) \right| \\
&= \max_{1\le k\le\lfloor mT\rfloor} \frac{1}{\sqrt{m}\,(k/m)^{\gamma}} \left| \sum_{i=m+1}^{m+k} \sum_{\nu=1}^{n} \int_{s_{\nu-1}}^{s_\nu} \big(r_{iM}(s_\nu)\psi(\varepsilon_i(s_\nu)) - r_{iM}(s)\psi(\varepsilon_i(s)) \big) ds \right| \\
&\le C \Bigg(\sum_{\nu=1}^{n} \int_{s_{\nu-1}}^{s_\nu} \max_{1\le k\le\lfloor mT\rfloor} \frac{1}{\sqrt{m}\,(k/m)^{\gamma}} \left| \sum_{i=m+1}^{m+k} \big(r_{iM}(s_\nu) - r_{iM}(s) \big) \psi(\varepsilon_i(s)) \right| ds \\
&\quad + \sum_{\nu=1}^{n} \int_{s_{\nu-1}}^{s_\nu} \max_{1\le k\le\lfloor mT\rfloor} \frac{1}{\sqrt{m}\,(k/m)^{\gamma}} \left| \sum_{i=m+1}^{m+k} r_{iM}(s_\nu) \big(\psi(\varepsilon_i(s_\nu)) - \psi(\varepsilon_i(s)) \big) \right| ds \Bigg) \\
&= \sum_{\nu=1}^{n} \int_{s_{\nu-1}}^{s_\nu} \big(I_{1,m}(s_\nu, s) + I_{2,m}(s_\nu, s) \big) ds.
\end{aligned}
\tag{5.11}
$$

According to (5.7) and Assumption (B.7a), as $m \to \infty$,

$$
E\Big(\sum_{\nu=1}^{n} \int_{s_{\nu-1}}^{s_\nu} I_{1,m}(s_\nu, s) ds \Big) \le C\,(\log m)\, \frac{1}{n} \sum_{\nu=1}^{n} \sup_{h\in[0,1/n]} \| r_{0M}(s_\nu) - r_{0M}(s_\nu - h) \|_2 = o(1).
\tag{5.12}
$$

By an analogous argument,

$$
\begin{aligned}
E\Big(\sum_{\nu=1}^{n} \int_{s_{\nu-1}}^{s_\nu} I_{2,m}(s_\nu, s) ds \Big) &\le \sum_{\nu=1}^{n} \int_{s_{\nu-1}}^{s_\nu} E \max_{1\le k\le\lfloor mT\rfloor} \frac{1}{\sqrt{m}\,(k/m)^{\gamma}} \left| \sum_{i=m+1}^{m+k} r_{iM}(s_\nu) \big(\psi(\varepsilon_i(s_\nu)) - \psi(\varepsilon_i(s)) \big) \right| ds \\
&\le C\,(\log m) \sup_{s\in[0,1]} \| r_{0M}(s) \|_2 \, \frac{1}{n} \sum_{\nu=1}^{n} \sup_{h\in[0,1/n]} \| \psi(\varepsilon_i(s_\nu)) - \psi(\varepsilon_i(s_\nu - h)) \|_2 = o(1),
\end{aligned}
\tag{5.13}
$$

where we have used the independence of the sequences $\{r_{iM}(\cdot)\}$ and $\{\varepsilon_i(\cdot)\}$ once again in combination with Assumptions (B.1) and (B.7b).

(iii) In a first step, we replace $\frac{1}{n}\sum_{\nu=1}^{n} \lambda'(0, s_\nu)\big(r_{iM}^{q}(s_\nu) - Er_{0M}^{q}(s_\nu)\big)$ in (5.9)–(5.10) with

$$
x_i = \int_0^1 \lambda'(0, s)\big(r_{iM}^{q}(s) - Er_{0M}^{q}(s)\big) ds, \quad i = 1, 2, \ldots;\ q = 1, 2.
$$

This can be done in a similar way as in the proof of (5.8). We even have an additional multiplication by $1/\sqrt{m}$ here. Note that the sequence $\{x_i\}_{i=1,2\ldots}$ is again strictly stationary and ergodic with $Ex_0 = 0$.

Now, it suffices to prove (5.9) and (5.10) with $\max_{K<k\le\lfloor mT\rfloor}$ instead of $\max_{1\le k\le\lfloor mT\rfloor}$ and $\sum_{i=K+1}^{m+k}$ replacing $\sum_{i=m+1}^{m+k}$, where $K = K_m$ is such that $K \to \infty$, but $K/m^{1-\gamma} \to 0$, e.g., for $K = \log m$.

In view of the strict stationarity and $Ex_0 = 0$, the ergodic theorem gives, as $m \to \infty$,

$$
\max_{K<k\le\lfloor mT\rfloor} \frac{1}{k} \left| \sum_{i=m+1}^{m+k} x_i \right| = o_P(1).
$$

On observing that

$$
\max_{K<k\le\lfloor mT\rfloor} \frac{1}{m\,(k/m)^{\gamma}} \left| \sum_{i=m+1}^{m+k} x_i \right| \le T^{1-\gamma} \max_{K<k\le\lfloor mT\rfloor} \frac{1}{k} \left| \sum_{i=m+1}^{m+k} x_i \right|,
$$

this completes the proof of (5.9) and (5.10), respectively. □

Acknowledgments

The authors thank two anonymous referees for their constructive suggestions, which led to improve the presentation of their results. They also gratefully acknowledge support by grant GAČR P201/12/1277 (O. Chochola, M. Hušková and Z. Prášková) and DFG grant STE 306/22-1 (J.G. Steinebach).

References

[1] T.G. Andersen, T. Bollerslev, F.X. Diebold, J. Wu, Realized beta: Persistence and predictability, in: T. Fomby, D. Terrell (Eds.), Advances in Econometrics: Econometric Analysis of Economic and Financial Time Series in Honor of R.F. Engle and C.W.J. Granger vol. B, Elsevier, Amsterdam, 2006, pp. 1–40.
[2] D.W.K. Andrews, Heteroskedasticity and autocorrelation consistent covariance matrix estimation, Econometrica 59 (1991) 817–858.

[3] A. Aue, S. Hörmann, L. Horváth, M. Hušková, J.G. Steinebach, Sequential testing for the stability of high-frequency portfolio betas, Econom. Theory 28 (2012) 804–837.
[4] I. Berkes, S. Hörmann, J. Schauer, Split invariance principles for stationary processes, Ann. Probab. 39 (2011) 2441–2473.
[5] P. Billingsley, Convergence of Probability Measures, Wiley, New York, 1968.
[6] O. Chochola, Robust monitoring procedures for dependent data (Ph.D. thesis), Charles University of Prague, 2013.
[7] O. Chochola, M. Hušková, Z. Prášková, J.G. Steinebach, Robust monitoring of CAPM portfolio betas, J. Multivariate Anal. 118 (2013) 374–395.
[8] I. Fazekas, O. Klesov, A general approach to the strong law of large numbers, Theory Probab. Appl. 45 (2000) 436–448.
[9] E. Ghysels, On stable factor structures in the pricing of risk: Do time-varying betas help or hurt? J. Finance 53 (1998) 549–573.
[10] S. Hörmann, P. Kokoszka, Weakly dependent functional data, Ann. Statist. 38 (2010) 1845–1884.
[11] L. Horváth, P. Kokoszka, Inference for Functioal Data with Applications, Springer, New York, 2012.
[12] P.J. Huber, Robust Statistics, Wiley, New York, 1981.
[13] C. Kirch, Resampling methods for the change analysis of dependent data (Ph.D. thesis), University of Cologne, 2006.
[14] J. Lintner, The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets, Review Econom. Statist. 47 (1965) 13–37.
[15] R. Merton, An intertemporal asset pricing model, Econometrica 41 (1973) 867–880.
[16] F. Móricz, Moment inequalities and the strong laws of large numbers, Z. Wahrscheinlichkeitstheor. Verwandte Geb. 35 (1976) 299–314.
[17] F. Móricz, R. Serfling, W. Stout, Moment and probability bounds with quasi-superadditive structure for the maximum partial sums, Ann. Probab. 10 (1982) 1032–1040.
[18] W.F. Sharpe, Capital asset prices: A theory of market equilibrium under conditions of risk, J. Finance 19 (1964) 424–442.
[19] A. Zeileis, Econometric computing with HC and HAC covariance matrix estimators, J. Statist. Software 11 (2004) 1–17.