SFB
823

Sequential detection of parameter changes in dynamic conditional correlation models

Katharina Pape, Pedro Galeano,
Dominik Wied

Nr. 7/2017

Discussion Paper

Sequential Detection of Parameter Changes in Dynamic Conditional Correlation Models

Katharina Pape, Pedro Galeano, Dominik Wied *

TU Dortmund, Universidad Carlos III de Madrid and Universität zu Köln

This Version: April 3, 2017

Abstract

A multivariate monitoring procedure is presented to detect changes in the parameter vector of the dynamic conditional correlation model proposed by Robert Engle in 2002. The benefit of the proposed procedure is that it can be used to detect changes in both the conditional and unconditional variance as well as in the correlation structure of the model. The detector is based on quasi log likelihood scores. More precisely, standardized derivations of quasi log likelihood contributions of points in the monitoring period are evaluated at parameter estimates calculated from a historical period. The null hypothesis of a constant parameter vector is rejected if these standardized terms differ too much from those that were expected under the assumption of a constant parameter vector. Under appropriate assumptions on moments and the structure of the parameter space, limit results are derived both under null hypothesis and alternatives. In a simulation study, size and power properties of the procedure are examined in various scenarios.

Keywords: Dynamic conditional correlation; Multivariate sequences; Online detection; Parameter changes; Threshold function.

JEL Classification: C12, C32

*K. Pape: TU Dortmund, Fakultät Statistik, 44221 Dortmund, Germany. E-Mail: pape@statistik.tu-dortmund.de, Phone: +49 231 755 3127. P. Galeano: Universidad Carlos III de Madrid, Departamento de Estadística and UC3M-BS Institute of Financial Big Data, 28903 Getafe, Madrid, Spain. E-mail: pedro.galeano@uc3m.es, Phone: +34 916 818 901. D. Wied: Universität zu Köln, Wirtschafts- und Sozialwissenschaftliche Fakultät, 50931 Köln, Germany. E-Mail: dwied@uni-koeln.de, Phone: +49 221 470 4514. Financial support by Deutsche Forschungsgemeinschaft (SFB 823, project A1) and Ministerio de Ciencia e Innovación grant ECO2015-66593-P is gratefully acknowledged.

1. Introduction

Recent years brought a lot of research in the fields of modelling volatility and correlation and testing for structural breaks, as well as in the intersection between both. In particular, the latter is motivated by the importance of information on structural changes in such parameters for financial applications. For instance, analysts need the aforementioned information for constructing optimal portfolios or for anticipating crises. A usual observation here is that volatilities and correlations increase in turbulent market phases.

While other articles often consider either variances or correlations, see, for instance, Wied and Galeano (2013) and Pape et al. (2016), among others, this paper aims at monitoring structural changes in both volatilities and correlations *jointly*. For that, we consider the well-known Dynamic Conditional Correlation (DCC) model by Engle (2002) and provide a method to monitor its parameters which steer the conditional volatilities and correlations. With constant parameters, the unconditional variances and correlations would be constant. Our method can be used to investigate if this assumption is realistic in empirical datasets. If the parameters are not constant in the observed period, then parameter estimates based on the constancy assumption are no longer reliable leading to biased volatilities and correlations forecasts.

We deal with monitoring parameter changes in dynamic conditional correlation models. Thus, based on an historical period of observations, we receive new data and the problem is to detect the presence of a changepoint in the model parameters as soon as possible once it has happened. Our approach is strongly motivated by those found in Chu et al. (1996) and Berkes et al. (2004). On the one hand, Chu et al. (1996) developed sequential tests of structural stability in linear models. Particularly, these authors considered two monitoring procedures: the first one is a fluctuation monitoring procedure based on recursive estimation of parameters, while the second one is a CUSUM monitoring procedure based on recursive residuals. On the other hand, Berkes et al. (2004) proposed a sequential monitoring scheme to detect changes in the parameters of a Generalized Autoregressive Conditional Heteroskedasticity (GARCH) model proposed by Bollerslev (1986). In this case, the sequential changepoint detector depends on quasi likelihood scores, thus not using residuals. Their idea was to use a historical data set to estimate the model parameters of the GARCH model, that are used to evaluate the contributions of the data from the monitoring period to the Gaussian quasi log likelihood function. Therefore, under the alternative of a parameter change in the monitoring period, it is expected that the absolute gradient contributions of post break observations tend to infinity.

The procedure proposed in this paper to monitor changes in the parameters of the DCC model can be seen as a multivariate extension of the monitoring scheme proposed by Berkes et al. (2004). Nevertheless, the extension is much more complex than it may seem. On the one hand, models that allow for dynamic modelling of both the variance and correlation possess a far more complex structure than other multivariate extensions of the univariate GARCH model. The challenge of handling the model and its quasi likelihood scores gets even more demanding if a multiplicative structure of the conditional covariance matrix is postulated as in the DCC model. On the other hand, the DCC models and their properties are far less well investigated than univariate GARCH models and especially the classical GARCH model considered by Berkes et al. (2004). For the models with dynamic conditional correlation, important results like conditions for the existence and uniqueness of a stationary solution or for the existence of unconditional moments of higher order have just been proposed recently, see Fermanian and Malongo (2016), or remain to be established which makes this type of model quite challenging in applications.

Even if we focus on the DCC model due to its enormous popularity for modeling multiple financial returns, the results of this paper may be extended to models with structure similar to the one of the DCC model of Engle (2002), e.g., the Conditional Correlation Model (CCC) of Bollerslev (1990), the DCC model of Tse and Tsui (2002) and the asymmetric generalized dynamic conditional correlation (AG-DCC) model of Capiello et al. (2006), among others. On the contrary, the extension to other popular multivariate volatility models, e.g., the multivariate extensions of the GARCH models as proposed by Bollerslev et al. (1988) or the BEKK model proposed by Engle and Kroner (1995), that ensures the nonnegative definiteness of the conditional covariance matrix under milder conditions on the parameters, is more complex as the structure of these models is quite different to the structure of dynamic correlation models.

The rest of the paper is organized as follows. Section 2 introduces the DCC model proposed by Engle (2002). Section 3 describes the monitoring problem and presents several assumptions needed to ensure the existence and uniqueness of a stationary solution of the model under consideration. Section 4 proposes our monitoring scheme and presents the asymptotic results. The performance of the procedure in finite samples is investigated by simulation and application to real data in Sections 5 and 6. Some concluding statements can be found in Section 7.

Additionally, the appendix contains several sections organized as follows. The first section provides a detailed presentation of the first and second order partial derivations of the contributions of individual observations to the quasi log likelihood function of the model. The second section contains the proofs of the theorems and propositions in Section 4. Finally, the last one provides the proofs of some additional calculation rules.

2. The Dynamic Conditional Correlation Model

2.1. The Model and Basic Assumptions

Let $\{y_t, t \in \mathbb{Z}\}$ be a sequence of p dimensional random vectors, $y_t = (y_{1t}, \ldots, y_{pt})'$, following a multivariate GARCH model given by

$$y_t = H_t^{1/2} \epsilon_t \tag{2.1}$$

where

$$H_t = Cov\left(y_t | \mathcal{F}_{t-1}\right) \tag{2.2}$$

is the positive definite conditional covariance matrix of y_t given the information set $\mathcal{F}_{t-1} = \sigma\left\{y_{t-1}, y_{t-2}, \ldots\right\}$ and $\{\epsilon_t, t \in \mathbb{Z}\}$ a Standard White Noise sequence in $\mathbb{R}^p$, i.e. $E\left(\epsilon_t\right) = \mathbf{0}_p$, $\mathsf{Cov}\left(\epsilon_t\right) = \mathbb{I}_p$, $\forall t \in \mathbb{Z}$, and the vectors ϵ_t are mutually independent. In the following, $\mathbf{0}_p$, $\mathbf{0}_{p\times p}$ and $\mathbb{I}_p$ denote the p dimensional vector of zeros, the $(p \times p)$ dimensional matrix of zeros, and the $(p \times p)$ dimensional identity matrix, respectively.

Among all available specifications of the conditional covariance matrix H_t, we focus on the one determined by the DCC model by Engle (2002), that assumes

$$H_t = D_t R_t D_t \tag{2.3}$$

where $D_t = \mathsf{diag}\left\{h_{1t}^{1/2}, \ldots, h_{pt}^{1/2}\right\}$, with h_{it}, $i = 1, \ldots, p$, the individual variances, that can be specified as univariate GARCH(1,1) models for instance:

$$h_{it} = \omega_i + \alpha_i y_{i,t-1}^2 + \beta_i h_{i,t-1}, \quad i = 1, \ldots, p. \tag{2.4}$$

Furthermore, $R_t := \mathsf{Cor}\left(y_t | \mathcal{F}_{t-1}\right)$ is the conditional correlation matrix of y_t which can be decomposed as

$$R_t = Q_t^* Q_t Q_t^* \tag{2.5}$$

where Q_t is a $(p \times p)$ matrix that is recursively determined as

$$Q_t = (1 - \alpha - \beta)\,\bar{Q} + \alpha z_{t-1} z_{t-1}' + \beta Q_{t-1} \tag{2.6}$$

with $z_t = D_t^{-1} y_t$ the 'standardized' vectors. The parameters α and β are nonnegative scalars which satisfy $\alpha + \beta < 1$ and $\bar{Q} = [\bar{q}_{ij}]_{i,j=1,\ldots,p}$ is both the unconditional covariance and correlation matrix of z_t. Since this implies that the main diagonal elements are one, the unknown parameters in the matrix $\bar{Q}$ are the entries of $\psi = \mathsf{vecl}\left(\bar{Q}\right) = (\bar{q}_{21}, \ldots, \bar{q}_{p,p-1})'$, where $\mathsf{vecl}(\cdot)$ is the operator that stacks the lower diagonal elements of a matrix into a vector. Finally, the normalizing matrix Q_t^* is given by

$$Q_t^* := \mathsf{diag}\left\{[Q_t]_{11}^{-1/2}, \ldots, [Q_t]_{pp}^{-1/2}\right\}$$

where $[Q_t]_{ii}$ denotes the i-th main diagonal entry of the matrix Q_t.

In summary, the vector of parameters in the DCC model is given as

$$\boldsymbol{\theta} = (\omega_1, \alpha_1, \beta_1, \ldots, \omega_p, \alpha_p, \beta_p, \alpha, \beta, \bar{q}_{21}, \ldots, \bar{q}_{p,p-1})'$$

which leads to a total number of $d := \frac{1}{2}(p+1)(p+4)$ unknown parameters in the model. Note that $\boldsymbol{\theta}$ can be decomposed as $\boldsymbol{\theta} = (\boldsymbol{\theta}_1', \boldsymbol{\theta}_2')'$ where

$$\boldsymbol{\theta}_1 = (\omega_1, \alpha_1, \beta_1, \ldots, \omega_p, \alpha_p, \beta_p)' = \left(\phi_1', \ldots, \phi_p'\right)'$$

with $\phi_j := (\omega_j, \alpha_j, \beta_j)'$, $j = 1, \ldots, p$, is the vector of variance parameters and

$$\boldsymbol{\theta}_2 = (\alpha, \beta, \bar{q}_{21}, \ldots, \bar{q}_{p,p-1})' = (\alpha, \beta, \psi')',$$

is the vector of correlation parameters.

An important issue in multivariate models with dynamic variance is that the positive definiteness of the conditional covariance matrix H_t has to be guaranteed for all $t \in \mathbb{Z}$ almost surely. Proposition 2 in Engle and Sheppard (2001) provides sufficient conditions for this property. Particularly, the matrix H_t as specified in (2.3), (2.4) and (2.6), is positive definite for all $t \in \mathbb{Z}$ almost surely, if Assumption 2.1 is satisfied:

Assumption 2.1. *1.* $\omega_i > 0, \ \forall i \in \{1, \ldots, p\}$.

2. $\alpha_i > 0$ *and* $\beta_i > 0$ *with* $\alpha_i + \beta_i < 1 \ \ \forall i \in \{1, \ldots, p\}$, *see also Nelson and Cao (1992).*

3. $h_{i0} > 0, \ \forall i \in \{1, \ldots, p\}$.

4. $\alpha > 0$ *and* $\beta > 0$ *with* $\alpha + \beta < 1$.

5. There exists $\delta_1 > 0$ *with* $\lambda_{\mathsf{min}}(\bar{Q}) > \delta_1$ *where* $\lambda_{\mathsf{min}}(\cdot)$ *is the smallest eigenvalue of a square matrix.*

6. There exists $\delta_2 > 0$ *with* $\lambda_{\mathsf{max}}(Q_t) < \delta_2 \ \forall t \in \mathbb{Z}$ *where* $\lambda_{\mathsf{max}}(\cdot)$ *is the largest eigenvalue of a square matrix.*

Note, that Assumption 2.1.6. is not necessary to verify the positive definiteness of R_t for all $t \in \mathbb{Z}$. This property is implied by the positive definiteness of Q_t and Proposition 1 in Engle and Sheppard (2001). The fact that Q_t is positive definite for all $t \in \mathbb{Z}$ is implied by the decomposition

$$Q_t = \frac{1-\alpha-\beta}{1-\beta}\bar{Q} + \alpha \sum_{n=0}^{\infty} \beta^n z_{t-n-1} z'_{t-n-1}$$

and 6.70.(a) in Seber (2008):

$$\begin{aligned} \lambda_{\mathsf{min}}(Q_t) &\overset{a.s.}{\geq} \lambda_{\mathsf{min}}\left(\frac{1-\alpha-\beta}{1-\beta}\bar{Q}\right) + \lambda_{\mathsf{min}}\left(\alpha \sum_{n=0}^{\infty} \beta^n z_{t-n-1} z'_{t-n-1}\right) \\ &\overset{a.s.}{\geq} \frac{1-\alpha-\beta}{1-\beta}\lambda_{\mathsf{min}}\left(\bar{Q}\right) > \frac{1-\alpha-\beta}{1-\beta}\delta_1 > 0. \end{aligned}$$

However, we use Assumption 2.1.6. to get fixed boundaries for the positive eigenvalues of R_t for all $t \in \mathbb{Z}$. Note that with 6.17.(a) in Seber (2008), we have

$$\max_{1\leq i\leq p} [Q_t]_{ii} < \sum_{i=1}^{p} [Q_t]_{ii} = \sum_{i=1}^{p} \lambda_i (Q_t) \leq p\lambda_{\mathsf{max}} (Q_t) \overset{a.s.}{<} p\delta_2 \tag{2.7}$$

where $\lambda_{\mathsf{max}} (Q_t) = \lambda_1 (Q_t), \ldots, \lambda_p (Q_t) = \lambda_{\mathsf{min}} (Q_t) \overset{a.s.}{>} 0$ are the ordered eigenvalues of Q_t. Furthermore, we have

$$\min_{1\leq i\leq p} [Q_t]_{ii} > \frac{1-\alpha-\beta}{1-\beta}. \tag{2.8}$$

Hence, (2.7), (2.8) and 6.95. in Seber (2008) imply boundaries for the eigenvalues of R_t:

$$\lambda_{\min} (R_t) \geq \lambda_{\min} (Q_t^*)^2 \lambda_{\min} (Q_t) \geq \left(\max_{1\leq i\leq p} [Q_t]_{ii} \right)^{-1} \frac{1-\alpha-\beta}{1-\beta} \delta_1 \geq \frac{1-\alpha-\beta}{1-\beta} \frac{\delta_1}{p\delta_2} \tag{2.9}$$

$$\lambda_{\mathsf{max}} (R_t) \leq \lambda_{\max} (Q_t^*)^2 \lambda_{\max} (Q_t) \leq \left(\min_{1\leq i\leq p} [Q_t]_{ii} \right)^{-1} \delta_2 \leq \frac{1-\beta}{1-\alpha-\beta} \delta_2 \tag{2.10}$$

The results (2.9) and (2.10) will be used extensively throughout the article and the proof section in the appendix.

2.2. The Estimation of the Model Parameters

Given an observed multivariate time series $y_1, \ldots, y_T$, the quasi maximum likelihood estimator (QMLE) of $\boldsymbol{\theta}$, denoted by $\hat{\theta}_T$, is obtained by maximizing the Gaussian quasi log likelihood (QLL) function

$$L_T (\boldsymbol{\theta}|y_1, \ldots, y_T) = \sum_{t=1}^{T} l_t (\boldsymbol{\theta}|y_1, \ldots, y_T), \tag{2.11}$$

with the individual QLL contributions

$$l_t (\boldsymbol{\theta}|y_1, \ldots, y_T) = -\frac{1}{2} \left(p \cdot \log 2\pi + \log \det (H_t) + y_t' H_t^{-1} y_t \right). \tag{2.12}$$

Both the one step and the two step quasi maximum likelihood estimator are consistent. The second one was proposed by Engle and Sheppard (2001) to reduce the calculation time, since the one step estimation gets computationally expensive even for moderate dimensions of y_t. However, preliminary simulations showed that the one step QMLE yields considerably better estimates. Hence, we use the latter one for our simulations and applications.

In the following, we denote the (one step) QMLE calculated from a sample of T observations as $\hat{\theta}_T$. The consistency proof of $\hat{\theta}_T$ won't be stated in detail, but is based on the unique optimization of $L(\cdot)$ which is the limit of the QLL function $L_T(\cdot)$ by the true parameter vector $\boldsymbol{\theta}$ and on the uniform convergence of $L_T(\cdot)$ to $L(\cdot)$ on the constrained parameter space U which will be defined in the next section. The first part can be proved along the lines of the proof of Lemma 5.5 in Berkes et al. (2003) while the uniform convergence can be shown by means of Theorem A.2.2 in White (1994) and analogously to the approach in Section B.2.3.

3. The Monitoring Problem and Associated Assumptions

Let $\theta_t \in \mathbb{R}^d$ be the parameter vector of the DCC model at time t. Assume a historical period of length m that is not affected by any structural change, i.e.

Assumption 3.1. *$\theta_1 = \ldots = \theta_m$ with m a positive integer.*

We are interested in testing the null hypothesis of a constant parameter vector

$$\mathsf{H}_0: \ \theta_t = \boldsymbol{\theta}, \quad t = 1, \ldots, m, m+1, \ldots$$

against the alternative of a change in the vector of parameters at an unknown point in the monitoring period

$$\mathsf{H}_1: \ \theta_t = \begin{cases} \boldsymbol{\theta}, & t = 1, \ldots, m, m+1, \ldots, m+k^*-1 \\ \boldsymbol{\theta}^*, & t = m+k^*, m+k^*+1, \ldots \end{cases}$$

with $\boldsymbol{\theta} = \left(\phi_1', \ldots, \phi_p', \alpha, \beta, \psi'\right)'$ the parameters before the change and $\boldsymbol{\theta}^* = \left(\phi_1^{*\prime}, \ldots, \phi_p^{*\prime}, \alpha^*, \beta^*, \psi^{*\prime}\right)'$ the parameter vector after the change where $\phi_i^* = (\omega_i^*, \alpha_i^*, \beta_i^*)'$, $i = 1, \ldots, p$, and $\psi^* = \mathsf{vecl}\left(\bar{Q}^*\right)$ with $\bar{Q}^* = \left[\bar{q}_{ij}^*\right]_{i,j=1,\ldots,p}$. Note that the change takes place at the k^*-th point of the monitoring period which is the $(m+k^*)$-th point in the whole time series. Denote the number of unknown parameters in the constant matrix $\bar{Q}$ as $p^- := \frac{1}{2}p\,(p-1)$.

Similarly as in Berkes et al. (2003), we assume that there exist constants $0 < \underline{u} < \overline{u}$ and $0 < \rho < 1$, such that the parameter space can be constrained to the set U:

$$U := \Big\{ u : \ \mathsf{max}\left\{t_1, \ldots, t_p, b, a+b, |q_1|, \ldots, |q_{p^-}|\right\} \leq \rho, \ \lambda_{\mathsf{min}}\left(\mathsf{F}_{\bar{Q}}(u)\right) > \delta_1,$$
$$\text{and} \ \ \underline{u} < \mathsf{min}\left\{x_1, s_1, t_1, \ldots, x_p, s_p, t_p, a, b\right\} \leq \mathsf{max}\left\{x_1, s_1, t_1, \ldots x_p, s_p, t_p, a, b\right\} \leq \overline{u} \Big\}$$

where $u = (x_1, s_1, t_1, \ldots, x_p, s_p, t_p, a, b, q_1, \ldots, q_{p^-})'$ is a generic element of the constrained parameter space U. The functions $\mathsf{F}_{\bar{Q}}(u)$ and $\mathsf{F}_{Q_t}(u)$ are defined as follows:

Definition 3.1. *Define for $i \in \{1, \ldots, p\}$, $t \in \mathbb{Z}$ and $u \in U$:*

$$(i)\ \mathsf{w}_{it}(u) := \frac{x_i}{1-t_i} + s_i \sum_{k=1}^{\infty} t_i^{k-1} y_{i,t-k}^2 = \frac{x_i}{1-t_i} + s_i \sum_{k=0}^{\infty} t_i^{k} y_{i,t-k-1}^2.$$

$$(ii)\ \mathsf{F}_{D_t}(u) := \mathsf{diag}\left\{\mathsf{w}_{1t}(u)^{1/2}, \ldots, \mathsf{w}_{pt}(u)^{1/2}\right\}.$$

$$(iii)\ \mathsf{F}_{R_t}(u) := \mathsf{F}_{Q_t^*}(u)\mathsf{F}_{Q_t}(u)\mathsf{F}_{Q_t^*}(u) \quad \textit{with} \quad \mathsf{F}_{Q_t^*}(u) := \mathsf{diag}\left\{[\mathsf{F}_{Q_t}(u)]_{11}^{-1/2}, \ldots, [\mathsf{F}_{Q_t}(u)]_{pp}^{-1/2}\right\}.$$

$$(iv)\ \mathsf{F}_{Q_t}(u) := \frac{1-a-b}{1-b}\mathsf{F}_{\bar{Q}}(u) + a\sum_{k=1}^{\infty} b^{k-1}\mathsf{z}_{t-k}(u)\mathsf{z}'_{t-k}(u).$$

$$(v)\ \mathsf{F}_{\bar{Q}}(u) := \begin{pmatrix} 1 & q_1 & q_2 & \cdots & q_{p-1} \\ q_1 & 1 & q_p & \cdots & \vdots \\ \vdots & q_p & \ddots & & q_{p^- -1} \\ \vdots & \vdots & & 1 & q_{p^-} \\ q_{p-1} & \cdots & q_{p^- -1} & q_{p^-} & 1 \end{pmatrix}.$$

Note that $\mathsf{z}_t(u) = \mathsf{F}_{D_t}(u)^{-1} y_t$ which implies $\mathsf{z}_{it}(u) = \frac{y_{it}}{\sqrt{\mathsf{w}_{it}(u)}}$, $\forall i = 1, \ldots, p$.
To enable consistent parameter estimation, we assume throughout the paper:

Assumption 3.2. *$\theta \in U$.*

The QLL function in (2.11) can be written as a function of an arbitrary element of the parameter space $u \in U$ and is given as $L_T(u) = \sum_{t=1}^{T} l_t(u)$ with

$$l_t(u) = -\frac{1}{2}\left(p \cdot \log 2\pi + \sum_{i=1}^{p} \log \mathsf{w}_{it}(u) + \log\det\left(F_{R_t}(u)\right) + \mathsf{z}'_t(u)\mathsf{F}_{R_t}(u)^{-1}\mathsf{z}_t(u)\right). \tag{3.1}$$

Note that the functions $\mathsf{w}_{it}(u)$ and $\mathsf{F}_{Q_t}(u)$ depend on an infinite past of observations. While the assumption of an infinite past may be adequate in the context of theoretical considerations, only finitely many past observations can be obtained in practice. Thus, we need the following terms:

Definition 3.2. *Define for $i \in \{1, \ldots, p\}$, $t \in \mathbb{Z}$ and $u \in U$:*

(i) $\widehat{\mathsf{w}}_{it}(u) := \dfrac{x_i}{1 - t_i} + s_i \sum_{k=1}^{t-1} t_i^{k-1} y_{i,t-k}^2 = \dfrac{x_i}{1-t_i} + s_i \sum_{k=0}^{t-2} t_i^k y_{i,t-k-1}^2.$

(ii) $\widehat{\mathsf{F}}_{D_t}(u) := \mathsf{diag}\left\{\widehat{\mathsf{w}}_{1t}(u)^{1/2}, \ldots, \widehat{\mathsf{w}}_{pt}(u)^{1/2}\right\}.$

(iii) $\widehat{\mathsf{F}}_{R_t}(u) := \widehat{\mathsf{F}}_{Q_t^*}(u)\widehat{\mathsf{F}}_{Q_t}(u)\widehat{\mathsf{F}}_{Q_t^*}(u)$ *with* $\widehat{\mathsf{F}}_{Q_t^*}(u) := \mathsf{diag}\left\{\left[\widehat{\mathsf{F}}_{Q_t}(u)\right]_{11}^{-1/2}, \ldots, \left[\widehat{\mathsf{F}}_{Q_t}(u)\right]_{pp}^{-1/2}\right\}.$

(iv) $\widehat{\mathsf{F}}_{Q_t}(u) := \dfrac{1-a-b}{1-b}\mathsf{F}_{\bar{Q}}(u) + a \sum_{k=1}^{t-1} b^{k-1} \widehat{\mathsf{z}}_{t-k}(u)\widehat{\mathsf{z}}'_{t-k}(u)$ *with* $\widehat{\mathsf{z}}_t(u) = \widehat{\mathsf{F}}_{D_t}(u)^{-1} y_t.$

(v) $\hat{L}_T(u) = \sum_{t=1}^{T} \hat{l}_t(u)$

$$\textit{with} \quad \hat{l}_t(u) = -\frac{1}{2}\left(p \cdot \log(2\pi) + \sum_{i=1}^{p} \log \widehat{\mathsf{w}}_{it}(u) + \log\det\left(\widehat{\mathsf{F}}_{R_t}(u)\right) + \widehat{\mathsf{z}}'_t(u)\widehat{\mathsf{F}}_{R_t}(u)^{-1}\widehat{\mathsf{z}}_t(u)\right).$$

3.1. The Existence of a Unique Stationary Solution

To verify the existence of a stationary and unique solution satisfying the DCC model, we have to impose some additional assumptions, see Fermanian and Malongo (2016):

Assumption 3.3. $\max_{1\le i\le p} \alpha_i + \max_{1\le i\le p} \beta_i < 1$ *and* $|\beta| < 1$.

Note that since $\mathsf{E}\left(z_t z_t' | \mathcal{F}_{t-1}\right) = R_t$, the sequence $\{\eta_t,\ t \in \mathbb{Z}\}$ with $\eta_t = R_t^{-1/2} z_t$ consists of independent random vectors with $\mathsf{E}\left(\eta_t | \mathcal{F}_{t-1}\right) = \mathbf{0}_p$ and $\mathsf{Cov}\left(\eta_t | \mathcal{F}_{t-1}\right) = \mathbb{I}_p$.

Assumption 3.4.

$$\mathsf{E}\left[ln\left(\beta^2 + \alpha^2 \frac{4(2p+1)\sqrt{p}}{\sqrt{C_\lambda} C_q} \|\eta_t\|_2^2\right)\right] < 0$$

$$\textit{with constants} \quad C_\lambda = \frac{\lambda_{\mathsf{min}}\left((1-\alpha-\beta)\,\bar{Q}\right)}{1-\beta^2} \quad \textit{and} \quad C_q = \frac{(1-\alpha-\beta)\min_{1\le i\le p} \bar{q}_{ii}}{1-\beta^2}.$$

Along the lines of Fermanian and Malongo (2016) and taking into account that the volatilities are modeled as GARCH(1,1) processes, the model equations (2.1)-(2.6) can be written equivalently as

$$X_t = T_t X_{t-1} + \zeta_t \tag{3.2}$$

with

$$X_t := \left(h_{1t}, \ldots, h_{pt}, y_{1t}^2, \ldots, y_{pt}^2, \mathsf{vecl}\,(Q_t)', \mathsf{vecl}\,(z_t z_t')'\right)',$$

$$\zeta_t := \left(\omega_1, \ldots, \omega_p, \omega_1 z_{1t}^2, \ldots, \omega_p z_{pt}^2, (1-\alpha-\beta)\,\mathsf{vecl}\left(\bar{Q}\right)', \mathsf{vecl}\,(z_t z_t')'\right)'$$

$$\text{and} \quad T_t = \left[\begin{array}{ccc|ccc|c|c}
\beta_1 & & 0 & \alpha_1 & & 0 & & \\
 & \ddots & & & \ddots & & \mathbf{0}_{p\times p^-} & \mathbf{0}_{p\times p^-} \\
0 & & \beta_p & 0 & & \alpha_p & & \\
\hline
\beta_1 z_{1t}^2 & & 0 & \alpha_1 z_{1t}^2 & & 0 & & \\
 & \ddots & & & \ddots & & \mathbf{0}_{p\times p^-} & \mathbf{0}_{p\times p^-} \\
0 & & \beta_p z_{pt}^2 & 0 & & \alpha_p z_{pt}^2 & & \\
\hline
 & \mathbf{0}_{p^-\times p} & & & \mathbf{0}_{p^-\times p} & & \beta\,\mathbb{I}_{p^-} & \alpha\,\mathbb{I}_{p^-} \\
\hline
 & \mathbf{0}_{p^-\times p} & & & \mathbf{0}_{p^-\times p} & & \mathbf{0}_{p^-\times p^-} & \mathbf{0}_{p^-\times p^-}
\end{array}\right].$$

According to Fermanian and Malongo (2016) and for the model defined by (2.1)-(2.6), a stationary solution of (3.2) exists if Assumption 3.3 is satisfied. If in addition Assumption 3.4 is satisfied, this stationary solution is unique.

4. The Monitoring Procedure

The following monitoring scheme works along the lines of the univariate approach proposed by Berkes et al. (2004). Denote by $l_t'(u)$ the gradient of the QLL contributions $l_t\,(u)$ with infinite past in (3.1) and $\hat{l}_t'(u)$ the gradient with finite past in Definition 3.2. The explicit form of the gradient is derived and presented in Section A of the appendix. Note that when the transpose of a vector and the gradient can be confused, we use ^{T} to denote the transpose of a vector.

Since the asymptotics are carried out under the assumption of a growing length of the historical period, it may be suitable to define some characteristics as the length of the monitoring period in terms of m. Besides, this allows for a more adequate comparison of simulation results for different lengths of the historical period. Hence, we denote the length of the monitoring period as mB. Thus, the variable B indicates how long the monitoring period is compared to the historical period. Furthermore, for any $u \in U$, define

$$\boldsymbol{D}(u) := \mathsf{E} l_t'(u) l_t'(u)^T$$

and assume

Assumption 4.1. $\boldsymbol{D} := \boldsymbol{D}(\boldsymbol{\theta})$ *is a finite and nonsingular matrix.*

For a finite and observed sample, we use the estimate

$$\widehat{\boldsymbol{D}}_m = \frac{1}{m} \sum_{t=1}^{m} \hat{l}_t'(\hat{\theta}_m) \hat{l}_t'(\hat{\theta}_m)^T$$

and the detector

$$V_k = \sum_{t=m+1}^{m+k} \widehat{\boldsymbol{D}}_m^{-\frac{1}{2}} \hat{l}_t'(\hat{\theta}_m)$$

with stopping rule

$$\tau_m = \min \left\{ k \leq mB : \ |V_k| > m^{\frac{1}{2}} \left(1 + \frac{k}{m}\right) \mathsf{b}\left(\frac{k}{m}\right) \right\}, \tag{4.1}$$

where $\mathsf{b}(\cdot)$ is a threshold function and $|\cdot|$ the norm that yields the maximum absolute entry of vectors and matrices. If $\tau_m < \infty$, a change in the parameters is indicated while $\tau_m = \infty$ signalizes that the detector did not cross the threshold function in the monitoring period and no changepoint could be detected. As in Berkes et al. (2004), some moderate conditions are imposed on the form of the threshold function $\mathsf{b}(\cdot)$:

Assumption 4.2. $\mathsf{b}(\cdot)$ *is continuous on* $(0, \infty)$ *and* $\inf_{0<t<\infty} \mathsf{b}(t) > 0$.

To avoid confusion with the model parameters, denote by $\tilde{\alpha} \in (0, 1)$ the significance level for testing the null hypothesis of no parameter change versus the alternative hypothesis of a change during the monitoring period.

Therefore, the threshold function $\mathsf{b}(\cdot)$ or at least the variable parts of the function should be chosen such that

$$\lim_{m\to\infty} \mathsf{P}_{\mathsf{H}_0}\{\tau_m < \infty\} = \tilde{\alpha} \quad \text{and} \quad \lim_{m\to\infty} \mathsf{P}_{\mathsf{H}_1}\{\tau_m < \infty\} = 1.$$

Berkes et al. (2004) choose the threshold function $\mathsf{b}(\cdot)$ as a constant that is obtained via simulation. Preliminary simulations suggested that the empirical size of the proposed multivariate procedure depends strongly on the length of the monitoring period, i.e. on the parameter B, just as in the univariate case presented by Berkes et al. (2004). To reduce this effect we include the length of the monitoring period into the stopping rule (4.1). Moreover, we prefer a curved threshold function to the linear one that results from choosing $\mathsf{b}(\cdot)$ as a constant function. More precisely, we use the threshold function, that was proposed by Horváth et al. (2004) and also used by Wied and Galeano (2013) among others, i.e.

$$\mathsf{b}(x) = \max\left\{\left(\frac{x}{1+x}\right)^{\gamma}, \varepsilon\right\}$$

where $\gamma \in [0, 1/2)$ is a tuning parameter and $\varepsilon > 0$ a constant that can be chosen arbitrarily small in applications. A larger value of γ results in a steeper threshold function that tends to detect early changes in the parameters with a higher probability while a smaller value of the tuning parameter leads to a lower slope of the threshold function which results in a higher probability to detect changes that arise later in the monitoring period.

We scale this threshold function by multiplying a constant $c = c(\tilde{\alpha})$ which is obtained via Monte Carlo simulations such that the probability that the detector crosses the threshold function in the monitoring period equals the theoretical size $\tilde{\alpha}$.

As a last preparation for our main results, the random vectors y_t need to posses eighth moments and cross moments. Since this property is partly determined by the behaviour of the innovation vectors, we assume

Assumption 4.3. *The innovation vectors ϵ_t possess absolute eighth moments and cross moments.*

The following proposition provides conditions that allow to pass the existence of these moments on to the outcome vectors y_t. We denote the vector of conditional variances as $h_t = (h_{1t}, \ldots, h_{pt})'$ and, adopting the notation of He and Teräsvirta (2004), the vector of squared outcomes as $y_t^{(2)} = \left(y_{1t}^2, \ldots, y_{pt}^2\right)'$. Recall the equity $y_t = D_t z_t$ with $z_t \sim (0, R_t)$ the 'standardised' vectors. Thus, a vector representation of (2.4) is given as

$$h_t = \boldsymbol{\omega} + \boldsymbol{A} y_{t-1}^{(2)} + \boldsymbol{B} h_{t-1} \tag{4.2}$$

with $\boldsymbol{\omega} := (\omega_1, \ldots, \omega_p)'$, $\boldsymbol{A} := \mathsf{diag}\{\alpha_1, \ldots, \alpha_p\}$ and $\boldsymbol{B} := \mathsf{diag}\{\beta_1, \ldots, \beta_p\}$.

By the use of $Z_t := \mathsf{diag}\{z_{1t}, \ldots, z_{pt}\}$, we have $y_t^{(2)} = Z_t^2 h_t$ which allows for an autoregressive representation of (4.2):

$$h_t = \boldsymbol{\omega} + \boldsymbol{C}_{t-1} h_{t-1}$$

with $\boldsymbol{C}_t := \boldsymbol{A} Z_t^2 + \boldsymbol{B}$. Note that the unconditional expectation of $\boldsymbol{C}_t$ is time independent under the stationary conditions 4.3 and 4.4. Denominate by $\bigotimes_k$ the k-fold Kronecker product of identical matrices and assume

Assumption 4.4. $\lambda_{max}(\mathsf{E}\left[\bigotimes_j \boldsymbol{C_0}\right]) < \infty, \quad \forall j \in \{1, \ldots, 4\}.$

Proposition 4.1. *Let $\{y_t\}$ be a sequence of random vectors that satisfy* (2.1)-(2.6) *and η_t the random vectors defined as $\eta_t := R_t^{-1/2} z_t$ with $z_t := D_t^{-1} y_t$. Under Assumptions 2.1, 3.2-3.4, 4.3 and 4.4, the random vectors y_t possess eighth moments and cross moments.*

The evidence of Proposition 4.1 is the last piece we need for the next proposition that forms the base for the following theorems:

Proposition 4.2. *Under Assumptions 2.1, 3.1-3.4, 4.1 and 4.3-4.4, we have*

$$\widehat{\boldsymbol{D}}_m \to \boldsymbol{D} \ a.s.$$

Theorem 4.1. *Under* H_0 *and Assumptions 2.1, 3.1-3.4 and 4.1-4.4, we have*

$$\lim_{m\to\infty} \mathsf{P}_{\mathsf{H}_0}\{\tau_m < \infty\} = \lim_{m\to\infty} \mathsf{P}_{\mathsf{H}_0}\left(\sup_{1<k\le mB} \frac{|V_k|}{m^{1/2}\left(1+\frac{k}{m}\right)\mathsf{b}\left(\frac{k}{m}\right)} \ge c\right) = \mathsf{P}_{\mathsf{H}_0}\left(\sup_{t\in(0,B]} \frac{|\mathcal{G}(t)|}{(1+t)\,\mathsf{b}\,(t)} \ge c\right)$$

where $\left\{\mathcal{G}(t) = (G_1(t), \ldots, G_d(t))',\ t \in [0, B]\right\}$ *is a d-variate stochastic process whose component* *processes are d independent mean zero Gaussian processes $\{G_j(t), t \in [0, B]\}$ with covariance function $\mathsf{E}\left(G_j(k)G_j(l)\right) = \mathsf{min}\{k, l\} + kl$, for $j = 1, \ldots, d$.*

Along the lines of Galeano and Wied (2014) or Berkes et al. (2004) and denoting $\{W_i(t),\ t \in [0, \infty)\}$ for $i = 1, \ldots, d$ as d independent one dimensional Standard Brownian Motions, we have that $|\mathcal{G}(t)|$ possesses the same distribution as $\max\limits_{1\le i\le p} \left|(1+t)W_i\left(\frac{t}{1+t}\right)\right|$ for all $t \in \mathbb{Z}$, which yields

$$\sup_{t\in(0,B]} \frac{|\mathcal{G}(t)|}{(1+t)\mathsf{b}(t)} \overset{d}{=} \sup_{t\in(0,B]} \max_{1\le i\le d} \frac{\left|W_i\left(\frac{t}{1+t}\right)\right|}{\mathsf{b}\left(t\right)} \overset{d}{=} \sup_{t\in(0,B]} \max_{1\le i\le d} \frac{\left|W_i\left(\frac{t}{1+t}\right)\right|}{\mathsf{max}\left\{\left(\frac{t}{1+t}\right)^{\gamma}, \varepsilon\right\}}$$
$$\overset{d}{=} \sup_{\tilde{s}\in\left(0, \frac{B}{1+B}\right]} \max_{1\le i\le d} \frac{|W_i\left(\tilde{s}\right)|}{\mathsf{max}\left\{\tilde{s}^{\gamma}, \varepsilon\right\}}$$

when $t = \frac{\tilde{s}}{1-\tilde{s}}$ is substituted. The notation $A_1 \overset{d}{=} A_2$ indicates that the random variables A_1 and A_2 possess the same distribution. Furthermore, choosing $\tilde{s} = \frac{s(1+B)}{B}$ yields

$$\sup_{t\in(0,B]} \frac{|\mathcal{G}(t)|}{(1+t)\mathsf{b}(t)} \overset{d}{=} \sup_{s\in(0,1]} \max_{1\le i\le d} \frac{\left|W_i\left(\frac{sB}{1+B}\right)\right|}{\mathsf{max}\left\{\left(\frac{sB}{1+B}\right)^{\gamma}, \varepsilon\right\}} \overset{d}{=} \left(\frac{B}{1+B}\right)^{\frac{1}{2}-\gamma} \sup_{s\in(0,1]} \max_{1\le i\le d} \frac{|W_i\left(s\right)|}{\mathsf{max}\left\{s^{\gamma}, \varepsilon\left(\frac{1+B}{B}\right)^{\gamma}\right\}}.$$

Thus, we can use Monte Carlo simulations to obtain critical values $c = c\left(\tilde{\alpha}\right)$ in dependence of the significance level $\tilde{\alpha}$ based on the equality

$$\mathsf{P}_{\mathsf{H}_0}\left(\left(\frac{B}{1+B}\right)^{1/2-\gamma} \sup_{s\in(0,1]} \max_{1\le i\le d} \frac{|W_i\left(s\right)|}{\mathsf{max}\left\{s^{\gamma}, \varepsilon\left(\frac{1+B}{B}\right)^{\gamma}\right\}} \ge c\left(\tilde{\alpha}\right)\right)$$
$$= 1 - \left[\mathsf{P}_{\mathsf{H}_0}\left(\sup_{s\in(0,1]} \frac{|W_1(s)|}{\mathsf{max}\left\{s^{\gamma}, \varepsilon\left(\frac{1+B}{B}\right)^{\gamma}\right\}} < \left(\frac{1+B}{B}\right)^{1/2-\gamma} c\left(\tilde{\alpha}\right)\right)\right]^d = \tilde{\alpha}$$

or alternatively

$$\mathsf{P}_{\mathsf{H}_0}\left(\sup_{s\in(0,1]} \frac{|W_1(s)|}{\mathsf{max}\left\{s^{\gamma}, \varepsilon\left(\frac{1+B}{B}\right)^{\gamma}\right\}} < \left(\frac{1+B}{B}\right)^{1/2-\gamma} c\left(\tilde{\alpha}\right)\right) = \left(1-\tilde{\alpha}\right)^{\frac{1}{d}}.$$

Simulations showed that the critical values obtained by using the limit distribution of the detector and simulating the maximum values of weighted Brownian Motions yield unfeasible high size distortions in finite samples for medium- and even for large-sized historical periods. As a consequence, the detector values tend to exceed the values of the scaled threshold function soon after the beginning of the monitoring period. To extenuate the resulting size distortions, the critical values can be obtained via Bootstrap type Monte Carlo simulations. Recall that $\hat{\theta}_m$ is the estimate of the parameter vector calculated from the historical data. We assume that the underlying DCC process features a similar behaviour as the process determined by the parameters estimated from the historical period if the latter one is sufficiently large. Hence, we simulate $b_{BT} = 199$ realisations of a DCC process whose structure is controlled by $\hat{\theta}_m$ and denote them as $Y^{*(i)} := \left\{ y_1^{*(i)}, \ldots, y_{m(B+1)}^{*(i)} \right\}$ for $i \in \{1, \ldots, b_{BT}\}$. An intuitive approach would be to calculate the detector values

$$\left| V_k^{*(i)} \right| = \left| \sum_{t=m+1}^{m+k} \left[\widehat{D}_m^{*(i)} \right]^{-1/2} \hat{l}_t^{*(i)\prime} \left(\hat{\theta}_m \right) \right|$$

from each sample $Y^{*(i)}$ with $\hat{l}_t^{*(i)\prime} \left(\hat{\theta}_m \right)$ the QLL contributions and

$$\widehat{D}_m^{*(i)} = \frac{1}{m} \sum_{t=1}^{m} \hat{l}_t^{*(i)\prime} \left(\hat{\theta}_m \right) \left[\hat{l}_t^{*(i)\prime} \left(\hat{\theta}_m \right) \right]^T$$

the estimate of D_p based on the first m observations of $Y^{*(i)}$. But since we are not interested in using the exact limit distribution of the detector, the matrix D_p is substituted by the identity matrix to avoid the additional uncertainty that goes along with the matrix estimation. Further simulations that are not part of this article showed that this approach yields a remarkable decrease of the size distortions compared to the use of an estimate of D_p. Denote the resulting detector as $\left| \tilde{V}_k^{*(i)} \right|$ and the maximum of the scaled detector values gained from sample $Y^{*(i)}$ as

$$T^{*(i)} := \max_{1 \le k \le [mB]} \frac{\left| V_k^{*(i)} \right|}{m^{1/2} \left(1 + \frac{k}{m} \right) \mathsf{b} \left(\frac{k}{m} \right)}, \quad \text{for } i \in \{1, \ldots, b_{BT}\}.$$

The $(1 - \tilde{\alpha})$ quantile of $\left\{ T^{*(1)}, \ldots, T^{*(b_{BT})} \right\}$ can be used as a critical value in finite sample applications. Although a detailed analysis of these critical values and their properties lies beyond the scope of this article, they show a satisfying behaviour in simulations.

Concludingly, we investigate the asymptotic distribution of the detector under a parameter change. Recall that under the alternative of a structural break at an unknown point in the monitoring period the parameter vector changes from $\boldsymbol{\theta}$ to $\boldsymbol{\theta}^*$ at the k^*-th point of the monitoring period. Therefore, we need to impose further assumptions on the parameters of the post break period that resemble Assumptions 3.3, 3.4 and 4.4:

Assumption 4.5. $\max\limits_{1\leq i\leq p} \alpha_i^* + \max\limits_{1\leq i\leq p} \beta_i^* < 1 \quad and \quad |\beta^*| < 1.$

Assumption 4.6.

$$E\left[ln\left([\beta^*]^2 + [\alpha^*]^2 \frac{4(2p+1)\sqrt{p}}{\sqrt{C_\lambda^* C_q^*}} \|\eta_t\|_2^2\right)\right] < 0$$

$$\textit{with constants} \quad C_\lambda^* = \frac{\lambda_{\mathsf{min}}\left((1-\alpha^*-\beta^*)\,\bar{Q}^*\right)}{1-[\beta^*]^2} \quad \textit{and} \quad C_q^* = \frac{(1-\alpha^*-\beta^*)\min\limits_{1\leq i\leq p}\bar{q}_{ii}^*}{1-[\beta^*]^2}.$$

Assumption 4.7. $\lambda_{max}(\mathsf{E}\left[\bigotimes_j \boldsymbol{C_0^*}\right]) < \infty, \quad \forall j \in \{1,\ldots,4\},$

where $\boldsymbol{C_t^*} := \boldsymbol{A}^* Z_t^2 + \boldsymbol{B}^*$ *with* $\boldsymbol{A}^* := \mathsf{diag}\left\{\alpha_1^*,\ldots,\alpha_p^*\right\}$ *and* $\boldsymbol{B}^* := \mathsf{diag}\left\{\beta_1^*,\ldots,\beta_p^*\right\}.$

Theorem 4.2. *Under the alternative of a structural break and Assumptions 3.1-3.4 and 4.1-4.7, we have*

$$\lim_{m\to\infty} \mathsf{P}_{\mathsf{H}_1}\{\tau_m < \infty\} = 1.$$

Since it takes some time until the influence of the post break observations on the detector is strong enough to indicate the presence of a changepoint, it has to be assumed that in general the changepoint location is not consistent with the first hitting time τ_m. Thus, once the sheer presence of a change in the parameter vector is signalized, the position of the changepoint has to be estimated. A possible estimator works analogously to the estimators in Wied et al. (2012) and in Wied and Galeano (2013) and is defined as

$$\hat{k} := \mathop{\mathsf{arg\,max}}_{1\leq k\leq \tau_m - 1} \frac{k}{\sqrt{\tau_m}} \left| \frac{1}{\tau_m - 1} \sum_{t=m+1}^{m+\tau_m-1} \hat{l}_t'\left(\hat{\theta}_m\right) - \frac{1}{k}\sum_{t=m+1}^{m+k} \hat{l}_t'\left(\hat{\theta}_m\right)\right|. \tag{4.3}$$

Though a detailed analysis of its properties lies beyond the scope of this paper, estimators of this type showed satisfactory behaviour in simulations and applications which is why we use (4.3) to estimate the location of the changepoint throughout the next sections.

5. Simulations

This section is devoted to the investigation of the procedure's performance under various simulation settings. Under the null as well as under alternative hypotheses, some parameters have to be specified. First, we choose the length of the historical period as $m \in \{500, 1.000, 2.000\}$. In terms of trading days this equals roughly 2, 4 and 8 years, respectively. We assume that the length of the monitoring period is considerably smaller than the length of the historical period with $B \in \{0.1, 0.2, \dots, 0.5\}$. The dimension of the random vectors is $p \in \{3, 5\}$ and the tuning parameter is chosen as $\gamma \in \{0, 0.2, 0.4\}$. These values support the ability of the monitoring procedure to detect early or later appearing structural breaks. In any case, we simulated 1.000 time series and applied our procedure to them. Note that all of the simulations are carried out for a significance level of $\tilde{\alpha} = 0.05$.

5.1. Simulations Under the Null Hypothesis

First, we investigate the behaviour under the null hypothesis of no structural break in the parameter vector. For each vector component, we choose all of the variance parameters either as $\phi_i = (0.01, 0.05, 0.9)'$ or as $\phi_i = (0.01, 0.2, 0.7)'$ for all $i \in \{1, \dots, p\}$ where the second case indicates a stronger effect of single shocks on the volatility of future observations. The correlation structure is determined by the parameters $(\alpha, \beta) = (0.05, 0.9)$ and the constant unconditional correlation matrix $\bar{Q}_p$ which is defined as

$$\bar{Q}_3 = \begin{bmatrix} 1 & 0.5 & 0.1 \\ 0.5 & 1 & 0.5 \\ 0.1 & 0.5 & 1 \end{bmatrix} \quad \text{and} \quad \bar{Q}_5 = \begin{bmatrix} 1 & 0.5 & 0.3 & 0.2 & 0.1 \\ 0.5 & 1 & 0.5 & 0.3 & 0.2 \\ 0.3 & 0.5 & 1 & 0.5 & 0.3 \\ 0.2 & 0.3 & 0.5 & 1 & 0.5 \\ 0.1 & 0.2 & 0.3 & 0.5 & 1 \end{bmatrix}.$$

The results from Table 1 suggest that the empirical size increases with B which is plausible since larger values of this parameter imply a growing length of the monitoring period and thus more uncertainty. While larger values of m and γ reduce the size distortions, higher dimensions tend to increase the probability to commit a type I error. Especially the influence of variations in the length of the historical period and the dimension are as expected. Furthermore, the empirical size is distinctly higher when the variance parameters are chosen as $\phi_i = (0.01, 0.05, 0.9)'$. This result was expectable, since the sum $\alpha_i + \beta_i$ is closer to one, i.e. we are closer to a unit root process than in the second scenario. Note that the direction of the effects may not be the decribed one for all of the scenarios since we only did a relatively small number of replications due to the high demand of computational ressources.

			$\phi_i = (0.01, 0.05, 0.9)'$			$\phi_i = (0.01, 0.2, 0.7)'$		
			$\boldsymbol{m} = \mathbf{500}$	**1.000**	**2.000**	$\boldsymbol{m} = \mathbf{500}$	**1.000**	**2.000**
$\boldsymbol{p} = \mathbf{3}$	$\boldsymbol{B} = \mathbf{0.1}$	$\boldsymbol{\gamma} = \mathbf{0}$	0.124	0.088	0.068	0.068	0.077	0.047
		$\boldsymbol{\gamma} = \mathbf{0.2}$	0.133	0.084	0.074	0.067	0.070	0.052
		$\boldsymbol{\gamma} = \mathbf{0.4}$	0.116	0.082	0.070	0.064	0.058	0.066
	$\boldsymbol{B} = \mathbf{0.2}$	$\boldsymbol{\gamma} = \mathbf{0}$	0.150	0.101	.0.094	0.069	0.057	0.088
		$\boldsymbol{\gamma} = \mathbf{0.2}$	0.151	0.086	0.091	0.069	0.077	0.087
		$\boldsymbol{\gamma} = \mathbf{0.4}$	0.118	0.080	0.083	0.065	0.060	0.060
	$\boldsymbol{B} = \mathbf{0.3}$	$\boldsymbol{\gamma} = \mathbf{0}$	0.177	0.111	0.089	0.120	0.087	0.081
		$\boldsymbol{\gamma} = \mathbf{0.2}$	0.151	0.094	0.073	0.095	0.075	0.083
		$\boldsymbol{\gamma} = \mathbf{0.4}$	0.143	0.084	0.086	0.071	0.054	0.069
	$\boldsymbol{B} = \mathbf{0.4}$	$\boldsymbol{\gamma} = \mathbf{0}$	0.213	0.113	0.118	0.106	0.109	0.104
		$\boldsymbol{\gamma} = \mathbf{0.2}$	0.193	0.120	0.105	0.110	0.098	0.106
		$\boldsymbol{\gamma} = \mathbf{0.4}$	0.147	0.094	0.072	0.077	0.066	0.063
	$\boldsymbol{B} = \mathbf{0.5}$	$\boldsymbol{\gamma} = \mathbf{0}$	0.197	0.118	0.141	0.129	0.110	0.122
		$\boldsymbol{\gamma} = \mathbf{0.2}$	0.179	0.135	0.097	0.100	0.109	0.115
		$\boldsymbol{\gamma} = \mathbf{0.4}$	0.166	0.111	0.090	0.086	0.073	0.073
$\boldsymbol{p} = \mathbf{5}$	$\boldsymbol{B} = \mathbf{0.1}$	$\boldsymbol{\gamma} = \mathbf{0}$	0.139	0.093	0.079	0.080	0.085	0.062
		$\boldsymbol{\gamma} = \mathbf{0.2}$	0.141	0.104	0.067	0.071	0.066	0.047
		$\boldsymbol{\gamma} = \mathbf{0.4}$	0.117	0.093	0.072	0.070	0.048	0.060
	$\boldsymbol{B} = \mathbf{0.2}$	$\boldsymbol{\gamma} = \mathbf{0}$	0.153	0.099	0.083	0.079	0.080	0.082
		$\boldsymbol{\gamma} = \mathbf{0.2}$	0.161	0.112	0.085	0.088	0.081	0.068
		$\boldsymbol{\gamma} = \mathbf{0.4}$	0.148	0.083	0.073	0.074	0.069	0.059
	$\boldsymbol{B} = \mathbf{0.3}$	$\boldsymbol{\gamma} = \mathbf{0}$	0.181	0.109	0.087	0.102	0.118	0.116
		$\boldsymbol{\gamma} = \mathbf{0.2}$	0.161	0.110	0.098	0.087	0.098	0.090
		$\boldsymbol{\gamma} = \mathbf{0.4}$	0.148	0.087	0.085	0.074	0.064	0.071
	$\boldsymbol{B} = \mathbf{0.4}$	$\boldsymbol{\gamma} = \mathbf{0}$	0.181	0.117	0.122	0.121	0.106	0.125
		$\boldsymbol{\gamma} = \mathbf{0.2}$	0.199	0.109	0.111	0.102	0.108	0.101
		$\boldsymbol{\gamma} = \mathbf{0.4}$	0.151	0.114	0.088	0.073	0.079	0.086
	$\boldsymbol{B} = \mathbf{0.5}$	$\boldsymbol{\gamma} = \mathbf{0}$	0.198	0.111	0.131	0.152	0.126	0.140
		$\boldsymbol{\gamma} = \mathbf{0.2}$	0.186	0.111	0.123	0.122	0.120	0.118
		$\boldsymbol{\gamma} = \mathbf{0.4}$	0.182	0.107	0.107	0.092	0.084	0.059

Table 1: Empirical size for various parameter combinations.

5.2. Simulations Under Various Alternatives

In this section we investigate the performance of the proposed procedure confronted with different types of structural breaks. More precisely, first we simulate changes in the variance parameters followed by changes in the unconditional correlation matrix $\bar{Q}$.

Since the results under the null showed a strong dependency on the length of the monitoring period, the simulations under alternative scenarios will be limited to the case of monitoring periods with length $0.2m$. This choice of B yields small deviations between the empirical and the theoretical size as the results from Table 1 suggest. Since the length of the monitoring period depends on m, we define the location of the changepoint k^* in terms of m as $k^* = [mB\lambda^*]$ where $[\cdot]$ is the largest integer smaller than a given real number and the fraction λ^* is chosen from $\{0.05, 0.3, 0.5\}$. This indicates changes located at the beginning or later in the monitoring period.

5.2.1. Changes in the Variance Parameters

We establish two settings of changes in the variance parameters. First of all, we assume that $\phi_i = (0.01, 0.05, 0.9)'$ changes to $\phi_i^* = (0.005, 0.2, 0.7)'$ followed by a change from $\phi_i = (0.01, 0.2, 0.7)'$ to $\phi^* = (0.05, 0.05, 0.9)'$ for all $i \in \{1, \dots, p\}$. These settings will be denoted as Scenario 1 and 2.

Note that next to the actual variation in the parameters which implies a change in the conditional variance structure, Scenario 1 contains a decrease in the unconditional variances of all componenents while Scenario 2 implies a variance increase. The results for Scenario 1 can be taken from Tables 2 and 3 and those for Scenario 2 from Tables 4 and 5.

The power depends positively on the length of the historical period and negatively on the dimension of the random vectors. While the first result was expectable the negative influence of p on the power may be explained by the fact that the share of the $3p$ variance parameters in the group of all parameters decreases with growing dimension. Thus, changes in the variance parameters might be harder to detect if p gets large. The ability to detect parameter changes is distinctly higher for changes located at the begin of the monitoring period than for later ones which is a typical property of sequential monitoring schemes, see e.g. Wied and Galeano (2013) or Pape et al. (2016). Furthermore, parameter changes that lead to decreased unconditional variance can be detected

				Empirical first hitting times					Estimated changepoint location				
$\boldsymbol{\lambda^*}$	$\boldsymbol{m}$	$\boldsymbol{\gamma}$	Power	Mean	SD	$Q_{0.25}$	$Q_{0.5}$	$Q_{0.75}$	Mean	SD	$Q_{0.25}$	$Q_{0.5}$	$Q_{0.75}$
0.05	**500** ($k^* = 5$)	**0**	0.973	44.87	14.49	35	42	51	18.05	10.39	11	16	23
		0.2	0.971	40.28	15.77	30	38	48	16.08	9.76	9	14	21
		0.4	0.971	40.24	18.28	29	38	50	16.03	10.32	9	14	21
	1.000 ($k^* = 10$)	**0**	1.000	60.40	15.46	51	57	68	24.40	11.71	17	22	29
		0.2	0.999	52.88	16.09	42	50	60	21.18	10.53	14	19	26
		0.4	1.000	47.09	19.83	36	45	56	19.20	11.89	12	17	24
0.3	**500** ($k^* = 30$)	**0**	0.916	66.47	15.56	58	66	77	31.62	9.50	27	31	36
		0.2	0.891	63.86	17.27	54	64	75	30.46	10.17	26	30	35
		0.4	0.849	64.25	22.60	54	68	79	31	12.26	26	31	38
	1.000 ($k^* = 60$)	**0**	0.998	110.78	20.19	97	109	122	60.59	12.06	55	60	66
		0.2	0.998	107.62	22.10	94	106	119	59.84	13.31	55	60	66
		0.4	0.994	107.05	29.89	94	108	123	58.73	17.64	54	60	66
0.5	**500** ($k^* = 50$)	**0**	0.724	77.82	17.47	70.75	81	91	43.54	13.39	38	47	51
		0.2	0.703	76.83	19.48	70	81	90	42.58	14.30	36	47	52
		0.4	0.558	71.13	28.28	66	80	91	39.39	17.75	31	46	52
	1.000 ($k^* = 100$)	**0**	0.965	148.56	22.61	135	149	163	92.07	20.31	87	97	103
		0.2	0.972	148.36	24.02	136	149	162	93.03	19.56	88	97	103
		0.4	0.931	150.02	36.68	140	156	172	92.43	24.40	90	98	104

Table 2: Power against changes in the parameters that imply a variance decrease (Scenario 1) and properties of the first hitting times τ_m and estimated changepoints $\hat{k}$ for $p = 3$.

				Empirical first hitting times					Estimated changepoint location				
$\boldsymbol{\lambda^*}$	$\boldsymbol{m}$	$\boldsymbol{\gamma}$	Power	Mean	SD	$Q_{0.25}$	$Q_{0.5}$	$Q_{0.75}$	Mean	SD	$Q_{0.25}$	$Q_{0.5}$	$Q_{0.75}$
0.05	**500** ($k^* = 5$)	**0**	0.927	44.46	13.82	36	42	50	17.23	9.07	11	15	22
		0.2	0.913	39.51	15.13	30	37	45	15.31	8.99	9	13	19
		0.4	0.892	38.88	16.83	29	37	48	15.46	10.01	9	13	20
	1.000 ($k^* = 10$)	**0**	0.997	60.62	14.83	51	58	67	23.61	11.11	16	21	28
		0.2	0.997	51.60	14.42	42	49	59	20.66	10.31	14	18	25
		0.4	0.996	47.04	18.90	37	45	57	18.85	11.01	12	17	24
0.3	**500** ($k^* = 30$)	**0**	0.859	66.82	15.25	58	67	77	32.15	9.05	27	31	37
		0.2	0.848	64.89	16.99	56	65	76	31.01	10.60	26	30	37
		0.4	0.806	64.09	22.07	55	66	78	30.23	11.81	25.25	30	37
	1.000 ($k^* = 60$)	**0**	0.993	108.95	17.86	97	107	119	59.93	12.89	54	60	65
		0.2	0.992	106.35	20.53	94	106	118	58.72	13.69	54	59	64
		0.4	0.992	107.68	28.43	97	109	123	58.45	16.38	54	60	66
0.5	**500** ($k^* = 50$)	**0**	0.680	79.04	16.92	73	82	91	43.53	12.98	37	46	51
		0.2	0.673	76.54	20.58	70	82	91	42.48	14.36	37	46	51
		0.4	0.553	70.67	30.36	65	82	93	38.90	18.00	31	46	51
	1.000 ($k^* = 100$)	**0**	0.973	147.58	22.32	136	148	161	91.81	19.18	86	96	102
		0.2	0.972	145.77	25.43	135	148	161	90.26	21.93	84	97	102
		0.4	0.947	148.30	38.44	139	155	169	90.00	25.49	87	97	103

Table 3: Power against changes in the parameters that imply a variance decrease (Scenario 1) and properties of the first hitting times τ_m and estimated changepoints $\hat{k}$ for $p = 5$.

much more reliably than changes that entail variance increases. This property underlines the results under the null which suggested a stronger tendency of the detector to cross the threshold function if

$\boldsymbol{\lambda^*}$	$\boldsymbol{m}$	γ	Power	Empirical first hitting times					Estimated changepoint location				
				Mean	SD	$Q_{0.25}$	$Q_{0.5}$	$Q_{0.75}$	Mean	SD	$Q_{0.25}$	$Q_{0.5}$	$Q_{0.75}$
0.05	**500** ($k^*=5$)	**0**	0.768	64.97	19.92	51	66	80	24.62	16.68	12	20	32
		0.2	0.702	56.90	23.55	39.25	57	75	21.57	15.39	10	17	29
		0.4	0.558	49.34	29.99	21.25	48	75.75	20.44	17.00	7	15	29
	1.000 ($k^*=10$)	**0**	0.974	116.05	32.61	95	114	138	39.92	30.82	16	29	55
		0.2	0.955	107.65	38.44	80	107	133	36.77	30.67	14	26	49
		0.4	0.810	107.81	49.94	72	109.5	146	39.06	35.14	14	25	53
0.3	**500** ($k^*=30$)	**0**	0.527	76.44	15.92	65	79	90	32.30	11.46	27	32	37
		0.2	0.482	73.28	18.45	61	76	88	31.38	12.84	25	31	36
		0.4	0.271	62.76	27.40	46	68	85	29.00	14.92	20	30	36
	1.000 ($k^*=60$)	**0**	0.809	151.87	28.70	130	155	176	58.77	15.35	52	60	65
		0.2	0.727	145.37	33.39	121	149	172	55.86	16.11	48	59	65
		0.4	0.457	142.46	47.80	122	153	179	54.78	22.12	46	59	65
0.5	**500** ($k^*=50$)	**0**	0.346	81.76	14.91	72	84	95	44.04	15.42	36	48.5	54
		0.2	0.292	76.91	19.49	70	81	92	40.72	16.24	29	46	52
		0.4	0.172	58.26	35.67	18	72	87.25	32.97	23.10	6.75	39.5	52
	1.000 ($k^*=100$)	**0**	0.505	167.32	24.59	153	172	187	86.40	22.86	74	94	102
		0.2	0.445	161.34	32.89	146	168	184	83.36	25.98	71	91	101
		0.4	0.221	138.81	61.53	122	164	182	74.26	37.55	52	90	101

Table 4: Power against changes in the parameters that imply a variance increase (Scenario 2) and properties of the first hitting times τ_m and estimated changepoints $\hat{k}$ for $p = 3$.

$\boldsymbol{\lambda^*}$	$\boldsymbol{m}$	γ	Power	Empirical first hitting times					Estimated changepoint location				
				Mean	SD	$Q_{0.25}$	$Q_{0.5}$	$Q_{0.75}$	Mean	SD	$Q_{0.25}$	$Q_{0.5}$	$Q_{0.75}$
0.05	**500** ($k^*=5$)	**0**	0.738	67.80	19.61	54	70.50	83	25.32	17.10	12.25	20	34
		0.2	0.644	60.38	23.72	43	62.5	80	23.15	17.21	11	18	31
		0.4	0.409	46.89	30.46	15	49	72	18.67	16.67	6	13	26
	1.000 ($k^*=10$)	**0**	0.964	127.63	32.68	105	127	151	43.43	31.97	18	33	62
		0.2	0.935	117.04	39.51	90	116	145	38.39	31.36	15	27	54
		0.4	0.704	112.02	55.44	73	120	157.25	41.77	37.69	14	27	59
0.3	**500** ($k^*=30$)	**0**	0.491	76.75	16.17	66	80	89	30.91	11.62	25.5	31	36
		0.2	0.394	72.05	20.39	59.25	73	89	30.59	12.72	24	31.5	36
		0.4	0.193	56.33	34.87	30	64	88	25.09	17.89	8	29	34
	1.000 ($k^*=60$)	**0**	0.742	155.38	29.41	137	159	178	58.46	15.89	53	60	65
		0.2	0.671	149.38	33.91	128	154	176	57.44	15.32	51	60	65
		0.4	0.363	135.31	58.76	114.5	151	181	53.02	26.20	43.5	59	65
0.5	**500** ($k^*=50$)	**0**	0.292	79.91	14.59	70	82	92	41.55	15.36	31	46.5	53
		0.2	0.265	76.12	19.48	65	81	91	39.81	15.82	29	46	52
		0.4	0.127	46.23	36.34	4.5	56	78	25.44	21.42	2	24	48
	1.000 ($k^*=100$)	**0**	0.461	166.82	23.85	150	171	186	83.79	21.66	71	91	100
		0.2	0.384	160.36	33.85	144.75	169	185	82.73	24.10	68.75	90.5	101
		0.4	0.175	125.46	71.73	64.5	156	182	65.53	39.96	31.5	80	99

Table 5: Power against changes in the parameters that imply a variance increase (Scenario 2) and properties of the first hitting times τ_m and estimated changepoints $\hat{k}$ for $p = 5$.

the initial variance parameters are chosen as $\phi_i = (0.01, 0.05, 0.9)'$ rather than $\phi_i = (0.01, 0.2, 0.7)'$. In addition and consistently with the results under the null, the power decreases with the tuning

parameter γ. This result occurs as well if the structural break is located at a later point in the monitoring period and should be detected with a higher value of γ more easily.
The results concerning the estimated changepoint locations in Tables 2-5 suggest that the position of changepoints located at a fraction of $\lambda^* = 0.3$ of the monitoring period can be estimated without large distortions while earlier and later changes respectively are systematically placed too early and late respectively in the dataset. Note that the results for the estimated changepoint locations depend strongly on the properties of the first hitting times since these define the length of the subsample that is used to locate the changepoint.

5.2.2. Changes in the Correlation Parameters

Another possible alternative scenario is a change in the correlation structure. We assume that the variance parameters as well as α and β stay constant while the unconditional covariance matrix changes from $\bar{Q} = \mathbb{I}_p$ to $\bar{Q}^*$ where the latter one is a matrix whose main diagonal entries are equal to one while all of the diagonal entries are Δ with $\Delta \in \{0.1, \ldots, 0.9\}$. The variance parameters and α and β are chosen as in Section 5.1.

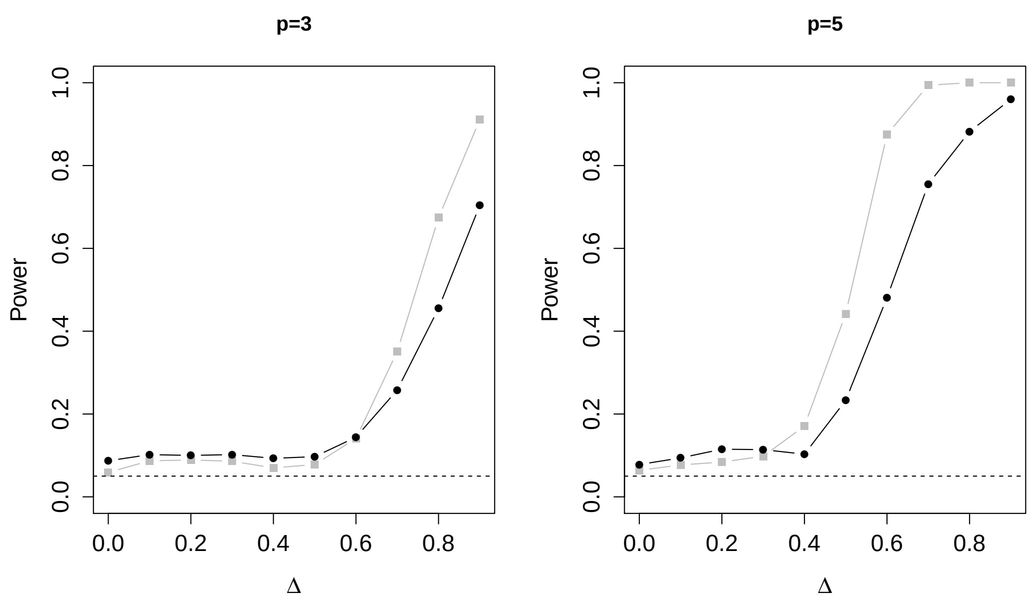

Figure 1: Power against correlation changes.
Black: $\phi_i = (0.01, 0.05, 0.9)'$, for $i = 1, \ldots, p$; Gray: $\phi_i = (0.01, 0.2, 0.7)'$, for $i = 1, \ldots, p$.

The power results for changes at fraction $\lambda^* = 0.05$ of the monitoring period are illustrated in Figure 1 for simulated time series of dimension 3 or 5, a historical period consisting of 1.000 data points and tuning parameter $\gamma = 0.2$. The results reveal problems to detect correlation changes of moderate magnitude for both choices of the vector of variance parameters. However, the power curve has a large slope for higher values and converges to one. While smaller changes in the correlation parameters can be detected more frequently if the variance parameters are chosen as $\phi_i = (0.01, 0.05, 0.9)'$ rather than $\phi_i = (0.01, 0.2, 0.7)'$, the opposite is true for larger values of Δ. The fact that some of the power results are quite low suggests that the QLL function seems to be very flat in some regions such that several kinds of parameter changes are hardly detectable.

6. Empirical Results

To investigate the performance of the proposed monitoring scheme under real conditions, we apply the procedure to a group of asset returns. We choose the assets of some insurance companies that are listed at different stock exchanges in Europe. More precisely, we use the log returns of the assets of *Allianz* (abbreviated by All), *AXA* , *Generali* (Gen), *ING* and *Munich Re* (MRe) due to the fact that a conjoint modeling seems to be adequate for the returns of assets from the same industrial sector and monetary area. As Engle (2002) argued, the DCC model is well-suited to model the typical features of multivariate time series of asset returns. Since Bollerslev (1986) stated that even GARCH models of order (1,1) are capable of explaining the behaviour of log returns very well, we use GARCH(1,1) models for the univariate conditional variance equations (2.4), which is in line with our approach in Section 2.

Since the results in Table 1 suggest that the size increases considerably with the length of the monitoring period and hence with B, we monitor the data under the use of a stepwise approach. In each step, we only use the first $[mB]$ observations following the historical period as monitoring period. If a changepoint is detected in this subsample, we cut off all of the observations before the estimated location of the structural break and use the following m data points as new historical period followed by a monitoring period of $[mB]$ points. If no changepoint can be detected in the monitoring period we cut off the first $[mB]$ observations of the previous historical period and use

$\gamma = 0.2, B = 0.2,\ m = 500$		$\gamma = 0.2, B = 0.2,\ m = 1000$	
τ_m	$\hat{k}$	τ_m	$\hat{k}$
2000/06/09	2000/05/22	2004/01/27	2003/12/22
2002/10/15	2002/10/10	2008/01/21	2008/01/04
2006/03/01	2005/12/05	2012/12/04	2012/10/15
2008/05/28	2008/05/22		
2010/08/04	2010/06/09		
2012/07/26	2012/07/05		
2014/07/09	2014/07/07		

Table 6: First hitting times and estimated changepoint locations.

the following m data points as new historical period. The results for $\gamma = 0.2$, $B = 0.2$ and different lengths of the historical period can be taken from Table 6. Additionally, the estimated changepoint locations for $m = 1.000$ are visualized in Figure 2 with two of the monitored time series. The figure shows that the time series of asset returns are splitted up into calm and more turbulent phases. The estimated changepoint locations in Table 6 can be linked to important economic events of the

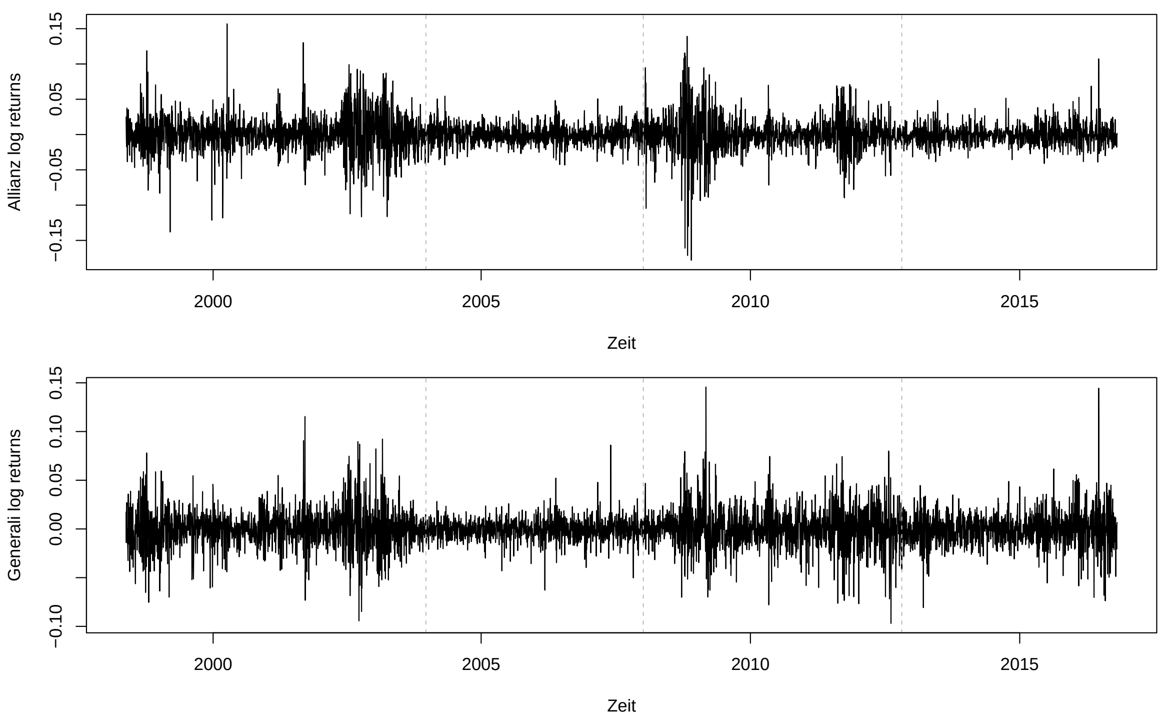

Figure 2: Log returns of the *Allianz* and *Generali* assets with the detected changepoints for $m = 1.000$ from Table 6 (dashed gray lines).

last two decades. The changepoint in 2002 marks the end of the crisis caused by the bursting of the dotcom bubble and the start of a calmer period that was interrupted in 2008 by the financial crisis followed by the debt crisis. The last changepoint might signalize the beginning of a recovery phase of the stock markets.

	1998/05/11 -2003/12/22	2003/12/23 -2008/01/04	2008/01/07 -2012/10/15	2012/10/16 -2016/10/25
$\hat{\omega}_{All}$	0.00002	0.00002	0.00001	<0.00001
$\hat{\omega}_{AXA}$	0.00001	0.00002	<0.00001	<0.00001
$\hat{\omega}_{Gen}$	0.00002	0.00003	<0.00001	<0.00001
$\hat{\omega}_{ING}$	0.00002	0.00001	0.00002	0.00001
$\hat{\omega}_{MRe}$	0.00001	0.00001	0.00002	0.00009
$\hat{\alpha}_{All}$	0.099	0.072	0.083	0.034
$\hat{\alpha}_{AXA}$	0.085	0.040	0.095	0.045
$\hat{\alpha}_{Gen}$	0.096	0.097	0.058	0.028
$\hat{\alpha}_{ING}$	0.082	0.113	0.093	0.125
$\hat{\alpha}_{MRe}$	0.120	0.063	0.117	0.072
$\hat{\beta}_{All}$	0.879	0.811	0.905	0.944
$\hat{\beta}_{AXA}$	0.887	0.920	0.895	0.936
$\hat{\beta}_{Gen}$	0.877	0.786	0.941	0.967
$\hat{\beta}_{ING}$	0.900	0.806	0.900	0.667
$\hat{\beta}_{MRe}$	0.853	0.703	0.883	0.924
$\left\Vert\hat{\boldsymbol{\theta}}_1\right\Vert_2$	0.182	0.171	0.167	0.140
$\hat{\alpha}$	0.011	0.022	0.032	0.013
$\hat{\beta}$	0.988	0.978	0.883	0.986
$\hat{q}_{All,AXA}$	0.422	0.597	0.779	0.778
$\hat{q}_{All,Gen}$	0.401	0.442	0.536	0.406
$\hat{q}_{All,ING}$	0.519	0.523	0.774	0.697
$\hat{q}_{All,MRe}$	0.651	0.505	0.745	0.704
$\hat{q}_{AXA,Gen}$	0.424	0.487	0.579	0.538
$\hat{q}_{AXA,ING}$	0.621	0.470	0.761	0.762
$\hat{q}_{AXA,MRe}$	0.457	0.510	0.666	0.607
$\hat{q}_{Gen,ING}$	0.468	0.245	0.511	0.442
$\hat{q}_{Gen,MRe}$	0.366	0.366	0.439	0.256
$\hat{q}_{ING,MRe}$	0.513	0.435	0.645	0.475
$\left\Vert\hat{\boldsymbol{\theta}}_2\right\Vert_2$	2.698	2.532	3.027	2.908
$\left\Vert\hat{\boldsymbol{\theta}}\right\Vert_2$	2.704	2.537	3.032	2.911

Table 7: Model parameters estimated from the data between successive detected changepoints for $m = 1.000$ from Table 6 and the euclidical norm of the estimated parameter vectors and the estimated vectors of variance and correlation parameters.

Estimates of the model parameters calculated from the data between two sucessive changepoints can be taken from Table 7. To measure the magnitude of the changes in the estimated parameters, the table also contains the euclidian norm of the parameter vectors estimated from the subsamples as well as the euclidian norm of the estimated vectors of variance and correlation parameters, respectively. The largest change in the parameters in terms of the euclidian norm can be found between the period before the financial crisis and the period of the crisis itself. A large part of this phenomenon seems to be caused by the fact that the correlation of asset returns tends to increase in times of crisis, see Sandoval Jr. and De Paula Franca (2012) among others.

7. Conclusion

We present a method to detect changes in the parameter vector of the DCC model proposed by Engle (2002) which is based on quasi log likelihood scores and allows to detect changes in the conditional and unconditional variance and covariance structure. We analyze the size and power properties of the presented procedure and apply it to a group of log returns that belong to the assets of several insurance companies. In applications it turns out as a heavy problem that the assumption of a historical period which is free from structural breaks cannot be checked with a known retrospective method. The search for a solution for this problem is left as a task for future research.

References

Alt, H. (2006): *Lineare Funktionalanalysis: Eine anwendungsorientierte Einführung*, Springer-Verlag Berlin Heidelberg.

Berkes, I., E. Gombay, L. Horváth, and P. Kokoszka (2004): "Sequential Change-point Detection in GARCH(p,q) models," *Econometric Theory*, 20, 1140–1167.

Berkes, I., L. Horváth, and P. Kokoszka (2003): "GARCH Processes: Structure and Estimation," *Bernoulli*, 9(2), 201–227.

Bollerslev, T. (1986): "Generalized Autoregressive Conditional Heteroscedasticity," *Journal of Econometrics*, 31(3), 307–327.

——— (1990): "Modeling the coherence in short-run nominal exchange rates: a multivariate generalized ARCH model," *Review of Economics and Statistics*, 72, 498–505.

Bollerslev, T., R. Engle, and J. Wooldridge (1988): "A Capital Asset Pricing Model with Time-varying Covariances," *Journal of Political Economy*, 96(1), 116–131.

Capiello, L., R. Engle, and K. Sheppard (2006): "Asymmetric Dynamics in the Correlations of Global Equity and Bond Returns," *Journal of Financial Econometrics*, 4(4), 537–572.

Chu, C.-S., M. Stinchcombe, and H. White (1996): "Monitoring structural change," *Econometrica*, 64(5), 1045–1065.

Engle, R. (2002): "Dynamic Conditional Correlation: A Simple Class of Multivariate Generalized Autoregressive Conditional Heteroscedasticity Moldels," *Journal of Business & Economic Statistics*, 20(3), 339–350.

Engle, R. and K. Kroner (1995): "Multivariate Simultaneous Generalized ARCH," *Econometric Theory*, 11(1), 122–150.

Engle, R. and K. Sheppard (2001): "Theoretical and Empirical Properties of Dynamic Conditional Correlation Multivariate GARCH," *NBER Working Paper No. 8554*.

Fermanian, J.-D. and H. Malongo (2016): "On the Stationarity of Dynamic Conditional Correlation Models," *Econometric Theory*.

Galeano, P. and D. Wied (2014): "Multiple Change Point Detection in the Correlation Structure of Financial Assets," *Computational Statistics and Data Analysis*, 76, 262–282.

Hafner, C. and H. Herwatz (2008): "Analytical Quasi Maximum Likelihood Inference in Multivariate Volatility Models," *Metrika*, 67, 219–239.

He, C. and T. Teräsvirta (2004): "An Extended Constant Conditional Correlation GARCH Model and its Fourth-Moment Structure," *Econometric Theory*, 20(5), 904–926.

Horváth, L., M. Hušková, P. Kokoszka, and J. Steinebach (2004): "Monitoring Changes in Linear Models," *Journal of Statistical Planning and Inference*, 126, 225–251.

Lütkepohl, H. (1996): *Handbook of Matrices*, John Wiley & Sons, New York, NY, USA.

NAKATANI, T. AND T. TERÄSVIRTA (2007): "Testing for Volatility Interactions in the Constant Conditional Correlation GARCH Model," *SSE/EFI Working Paper Series in Economics and Finance No. 649*.

NELSON, D. AND C. CAO (1992): "Inequality Constraints in the Univariate GARCH model," *Journal of Business & Economic Statistics*, 10(2), 229–235.

PAPE, K., D. WIED, AND P. GALEANO (2016): "Monitoring multivariate variance changes," *Journal of Empirical Finance*, 39(A), 54–68.

SANDOVAL JR., L. AND I. DE PAULA FRANCA (2012): "Correlation of Financial Markets in Times of Crisis," *Physica A: Statistical Mechanics and its Applications*, 391(1-2), 187–208.

SEBER, G. (2008): *A Matrix Handbook for Statisticians*, John Wiley & Sons, Hoboken, NJ, USA.

TSE, Y. AND K. TSUI (2002): "A Multivariate Generalized Autoregressive Conditional Heteroscedasticity Model with Time-varying Correlations," *Journal of Business & Economic Statistics*, 20(3), 351–362.

WHITE, H. (1994): *Estimation, Inference and Specification Analysis*, Cambridge University Press.

WIED, D. AND P. GALEANO (2013): "Monitoring Correlation Change in a Sequence of Random Variables," *Journal of Statistical Planning and Inference*, 143(1), 186–196.

WIED, D., W. KRÄMER, AND H. DEHLING (2012): "Testing for a Change in Correlation at an Unknown Point in Time Using an Extended Functional Delta Method," *Econometric Theory*, 68(3), 570–589.

APPENDIX

A. The Partial Derivations of the DCC QLL Function

A.1. Notation and Transformation Matrices

Recall that the QLL function of the DCC model introduced in Section 2.1is given as:

$$
\begin{aligned}
L_T(\theta) &= \sum_{t=1}^{T} l_t(\theta) \\
\text{with} \quad l_t(\theta) &= -\frac{1}{2}\left(p \cdot \log 2\pi + \log\det(H_t) + y_t' H_t^{-1} y_t\right) \\
&= -\frac{1}{2}\left(p \cdot \log 2\pi + 2\log\det(D_t) + \log\det(R_t) + z_t' R_t^{-1} z_t\right)
\end{aligned}
$$

As argued in Hafner and Herwatz (2008), it is sufficient to look at the lower diagonal entries of R_t in detail since the latter one is a symmetric matrix with ones on the main diagonal. Throughout the next sections, several matrices are used to interchange the position of the entries in vectors and matrices, see for instance Hafner and Herwatz (2008) or Lütkepohl (1996):

- $\mathsf{vec}(\cdot)$: the vec operator that stacks the entries of a matrix into a vector.
- $\mathsf{vech}(\cdot)$: the vech operator that stacks the diagonal and lower diagonal entries of a symmetric matrix into a vector.
- $\mathsf{vecl}(\cdot)$: the vecl operator that stacks the lower diagonal entries of a symmetric matrix into a vector.
- K_{mn}, the commutation matrix: $\mathsf{vec}(A') = \mathsf{K}_{mn} \cdot \mathsf{vec}(A)$ for a $(m \times n)$ matrix A.
- D_p, the duplication matrix: $\mathsf{D}_p \cdot \mathsf{vech}(A) = \mathsf{vec}(A)$ for a symmetric $(p \times p)$ matrix A.
- D_p^+, the Moore Penrose inverse of D_p. In general, this is not the elimination matrix L_p with $\mathsf{L}_p \cdot \mathsf{vec}(A) = \mathsf{vech}(A)$ which is some generalized inverse.
- $\mathsf{D}_{p,-}$, the matrix that results after deleting those columns from D_p that refer to the main diagonal entries of a symmetric $(p \times p)$ matrix A when D_p is multiplied by $\mathsf{vech}(A)$.

- $\mathsf{D}^+_{p,-}$, the Moore Penrose inverse of $\mathsf{D}_{p,-}$. This matrix is obtained when those rows that refer to the main diagonal elements of a symmetric $(p \times p)$ matrix A in vector $\mathsf{vech}(A)$ are deleted from D^+_p. Note, that for a symmetric $(p \times p)$ matrix A, we have $\mathsf{vecl}\,(A) = \mathsf{D}^+_{p,-}\mathsf{vec}\,(A)$.

- The number of lower diagonal elements of a $(k \times k)$ matrix is denoted by $p^+ = \frac{1}{2}(k+1)k$ if the main diagonal entries are included and by $p^- = \frac{1}{2}(k-1)k$ if they are excluded.

A.2. Some Calculation Rules for Matrices

The following rules will be needed throughout the next sections:

CR1 For the transformation matrices, we have

$$\mathsf{D}^+_{p,-} = \frac{1}{2}\left(\mathsf{D}_{p,-}\right)'.$$

The proof of this statement can be found in Section C.

CR2 Lütkepohl (1996), 10.4.2(1): $X \sim (m,m)$ symmetric:

$$\frac{\partial \mathsf{vec}(X)}{\partial \mathsf{vech}(X)'} = \mathsf{D}_m \quad \Rightarrow \quad \frac{\partial \mathsf{vec}(X)}{\partial \mathsf{vecl}(X)'} = \mathsf{D}_{m,-} \overset{\text{CR1}}{=} 2\left(\mathsf{D}^+_{m,-}\right)'.$$

CR3 For symmetric matrices $X \sim (m,m)$ and $Y(X) \sim (n,n)$, we have

$$\frac{\partial \mathsf{vech}\,(Y(X))}{\partial \mathsf{vecl}(X)'} = \frac{\partial \mathsf{L}_n \mathsf{vec}\,(Y(X))}{\partial \mathsf{vec}(X)'}\frac{\partial \mathsf{vec}(X)}{\partial \mathsf{vecl}(X)'} = \mathsf{L}_n \frac{\partial \mathsf{vec}\,(Y(X))}{\partial \mathsf{vec}(X)'}\mathsf{D}_{m,-}.$$

This is a direct consequence of CR2.

CR4 For symmetric $(n \times n)$ matrices X and $Y(X)$, we have

$$\frac{\partial \mathsf{vec}(XYX)}{\partial \mathsf{vec}(X)'} = (X \otimes X)\,\frac{\partial \mathsf{vec}\,(Y)}{\partial \mathsf{vec}\,(X)'} + (XY \otimes \mathbb{I}_n + \mathbb{I}_n \otimes XY)$$

The proof of this statement can be found in Section C.

A.3. The First Order Derivations with Respect to the Variance Parameters

Since the conditional correlation matrix R_t does not depend on the variance parameters, the partial derivations of the QLL contributions with respect to $\boldsymbol{\theta}_1$ are similar to (3.2) in Nakatani and Teräsvirta (2007):

$$\frac{\partial l_t(\theta)}{\partial \boldsymbol{\theta}_1} = -\frac{\partial \log \det D_t}{\partial \boldsymbol{\theta}_1} - \frac{1}{2}\frac{\partial y_t' H_t^{-1} y_t}{\partial \boldsymbol{\theta}_1} = -\frac{1}{2}\frac{\partial \mathsf{vec}\,(D_t)'}{\partial \boldsymbol{\theta}_1}\mathsf{vec}\left(2D_t^{-1} - D_t^{-1}R_t^{-1}z_t z_t' - z_t z_t' R_t^{-1} D_t^{-1}\right).$$

Using $v_{jt} = \left(1, y_{jt}^2, h_{jt}\right)'$ yields that the partial derivations of the non zero diagonal entries of D_t with respect to $\phi_i = (\omega_i, \alpha_i, \beta_i)'$ are given as:

$$g_{it} := \frac{\partial h_{it}^{1/2}}{\partial \phi_i} = \frac{1}{2} h_{jt}^{-1/2}\left(v_{j,t-1} + \beta_j \frac{\partial h_{j,t-1}}{\partial \phi_j}\right) \quad \text{and} \quad \frac{\partial h_{jt}^{1/2}}{\partial \phi_i} = \mathbf{0}_3, \text{ for } i \neq j.$$

Thus, we have

$$\frac{\partial \mathsf{vec}\,(D_t)'}{\partial \boldsymbol{\theta}_1} = \begin{bmatrix} g_{1t} & \mathbf{0}_{3\times p} & \mathbf{0}_3 & \mathbf{0}_{3\times p} & \dots & \mathbf{0}_3 & \mathbf{0}_{3\times p} & \mathbf{0}_3 \\ \mathbf{0}_3 & \mathbf{0}_{3\times p} & g_{2t} & \mathbf{0}_{3\times p} & \dots & \mathbf{0}_3 & \mathbf{0}_{3\times p} & \mathbf{0}_3 \\ \vdots & \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots \\ \mathbf{0}_3 & \mathbf{0}_{3\times p} & \mathbf{0}_3 & \mathbf{0}_{3\times p} & \dots & g_{p-1,t} & \mathbf{0}_{3\times p} & \mathbf{0}_3 \\ \mathbf{0}_3 & \mathbf{0}_{3\times p} & \mathbf{0}_3 & \mathbf{0}_{3\times p} & \dots & \mathbf{0}_3 & \mathbf{0}_{3\times p} & g_{pt} \end{bmatrix}.$$

Note that for theoretical considerations we use representations that depend on an infinite past of observations y_t while for simulation or parameter estimation we use the recursive form based on starting values for time $t = 0$.

- $\dfrac{\partial h_{it}}{\partial \omega_i} = 1 + \beta_i \dfrac{\partial h_{i,t-1}}{\partial \omega_i} = \dfrac{1}{1-\beta_i}$
- $\dfrac{\partial h_{it}}{\partial \alpha_i} = y_{i,t-1}^2 + \beta_i \dfrac{\partial h_{i,t-1}}{\partial \alpha_i} = \sum\limits_{n=0}^{\infty} \beta_i^n y_{i,t-n-1}^2$
- $\dfrac{\partial h_{it}}{\partial \beta_i} = h_{i,t-1} + \beta_i \dfrac{\partial h_{i,t-1}}{\partial \beta_i} = \sum\limits_{n=0}^{\infty} \beta_i^n h_{i,t-n-1} = \dfrac{\omega_i}{(1-\beta_i)^3} + \alpha_i \sum\limits_{n=0}^{\infty} \beta_i^n \sum\limits_{k=0}^{\infty} \beta_i^k y_{i,t-n-k-2}^2$

Starting values for estimation or simulation can be chosen as in the `ccgarch` package in `R` where $h_0 = (h_{10}, \dots, h_{p0})'$ is chosen as $\left(\mathsf{s}_1^2, \dots, \mathsf{s}_p^2\right)'$ with $\mathsf{s}_j^2 = \frac{1}{T}\sum\limits_{t=1}^{T} y_{jt}^2$, $j = 1, \dots, p$.
Hence, $v_{j0} = \left(1, \mathsf{s}_j^2, \mathsf{s}_j^2\right)'$, $j = 1, \dots, p$ and $\left(\frac{\partial h_{10}}{\partial \phi_1}, \dots, \frac{\partial h_{p0}}{\partial \phi_p}\right)' = \mathbf{0}_{p\times 3}$.

A.4. The First Order Derivations with Respect to the Correlation Parameters

Throughout this section, we consider the partial derivations of the QLL contributions with respect to those parameters that determine the correlation structure of the DCC model. Similar to (3.3) in Nakatani and Teräsvirta (2007), we have

$$\frac{\partial l_t(\theta)}{\partial \boldsymbol{\theta}_2} = -\frac{1}{2}\frac{\partial \log\det R_t}{\partial \boldsymbol{\theta}_2} - \frac{1}{2}\frac{\partial y_t' D_t^{-1} R_t^{-1} D_t^{-1} y_t}{\partial \boldsymbol{\theta}_2} = -\frac{1}{2}\frac{\partial \mathsf{vecl}\,(R_t)'}{\partial \boldsymbol{\theta}_2}\frac{\partial \mathsf{vec}\,(R_t)'}{\partial \mathsf{vecl}\,(R_t)}\frac{\partial \log\det R_t}{\partial \mathsf{vec}\,(R_t)}$$

$$= -\frac{1}{2}\frac{\partial \mathsf{vecl}\,(R_t)'}{\partial \boldsymbol{\theta}_2}\frac{\partial \mathsf{vec}\,(R_t)'}{\partial \mathsf{vecl}\,(R_t)}\frac{\partial \mathsf{vec}\left(R_t^{-1}\right)'}{\partial \mathsf{vec}\,(R_t)}\frac{\partial \mathsf{vec}\left(D_t^{-1}R_t^{-1}D_t^{-1}\right)'}{\partial \mathsf{vec}\left(R_t^{-1}\right)}\frac{\partial y_t' D_t^{-1} R_t^{-1} D_t^{-1} y_t}{\partial \mathsf{vec}\left(D_t^{-1}R_t^{-1}D_t^{-1}\right)}$$

$$= -\frac{1}{2}\frac{\partial \mathsf{vecl}\,(R_t)'}{\partial \boldsymbol{\theta}_2} 2\mathsf{D}_{p,-}^{+}\mathsf{vec}\left(R_t^{-1}\right) - \frac{1}{2}\frac{\partial \mathsf{vecl}\,(R_t)'}{\partial \boldsymbol{\theta}_2} 2\mathsf{D}_{p,-}^{+}(-1)\left(R_t^{-1}\otimes R_t^{-1}\right)\left(D_t^{-1}\otimes D_t^{-1}\right)\mathsf{vec}\left(y_t y_t'\right)$$

$$= -\frac{\partial \mathsf{vecl}\,(R_t)'}{\partial \boldsymbol{\theta}_2}\mathsf{D}_{p,-}^{+}\left[\mathsf{vec}\left(R_t^{-1}\right) - \left(R_t^{-1}\otimes R_t^{-1}\right)\left(D_t^{-1}\otimes D_t^{-1}\right)\left(y_t\otimes y_t\right)\right]$$

$$= -\frac{\partial \mathsf{vecl}\,(R_t)'}{\partial \boldsymbol{\theta}_2}\mathsf{vecl}\left(R_t^{-1}\left[\mathbb{I}_p - z_t z_t' R_t^{-1}\right]\right).$$

The second equality results from an extensive use of the chain rule, see 10.2.1(6) in Lütkepohl (1996). The third one uses 10.3.1(8), 10.3.3(10), 10.4.1(3), 10.4.2(1) and 10.6(1) in Lütkepohl (1996) and CR1 from Section A.2. The remaining equalities utilize CR2 and simple matrix computations. In the following, we consider the partial derivations of the lower diagonal entries of R_t with respect to the correlation parameters in detail. According to Hafner and Herwatz (2008), we have

$$\frac{\partial \mathsf{vecl}\,(R_t)}{\partial \boldsymbol{\theta}_2'} = \mathsf{D}_{p,-}^{+}\frac{\partial \mathsf{vec}\,(R_t)}{\partial \boldsymbol{\theta}_2'} = \mathsf{D}_{p,-}^{+}\frac{\partial \mathsf{vec}\,(R_t)}{\partial \mathsf{vech}\,(Q_t)'}\frac{\partial \mathsf{vech}\,(Q_t)}{\partial \boldsymbol{\theta}_2'} \tag{A.1}$$

where

$$\frac{\partial \mathsf{vec}\,(R_t)}{\partial \mathsf{vech}\,(Q_t)'} = \frac{\partial \mathsf{vec}\,(R_t)}{\partial \mathsf{vec}\,(Q_t^*)'}\frac{\partial \mathsf{vec}\,(Q_t^*)}{\partial \mathsf{vech}\,(Q_t^*)'}\frac{\partial \mathsf{vech}\,(Q_t^*)}{\partial \mathsf{vech}\,(Q_t)'} \overset{\text{CR2}}{=} \frac{\partial \mathsf{vec}\,(Q_t^* Q_t Q_t^*)}{\partial \mathsf{vec}\,(Q_t^*)'}\mathsf{D}_p\frac{\partial \mathsf{vech}\,(Q_t^*)}{\partial \mathsf{vech}\,(Q_t)'}. \tag{A.2}$$

With a slight difference to Section 4.3 in Hafner and Herwatz (2008), the derivation of $\mathsf{vec}\,(Q_t^* Q_t Q_t^*)$ with respect to $\mathsf{vec}\,(Q_t^*)$ is given as

$$\frac{\partial \mathsf{vec}(Q_t^* Q_t Q_t^*)}{\partial \mathsf{vec}(Q_t^*)'} \overset{\text{CR4}}{=} \left(Q_t^*\otimes Q_t^*\right)\frac{\partial \mathsf{vec}\,(Q_t)}{\partial \mathsf{vec}\,(Q_t^*)'} + \left(Q_t^* Q_t\otimes \mathbb{I}_p + \mathbb{I}_p\otimes Q_t^* Q_t\right) \tag{A.3}$$

and the derivation of $\mathsf{vech}\,(Q_t^*)$ with respect to $\mathsf{vech}\,(Q_t)$ as

$$\frac{\partial \mathsf{vech}\,(Q_t^*)}{\partial \mathsf{vech}\,(Q_t)'} = -\frac{1}{2}\mathsf{diag}\left\{\mathsf{vech}\left(Q_t^{-\frac{3}{2}}\right)\right\} \cdot \mathsf{diag}\left\{\mathsf{vech}\,(\mathbb{I}_p)\right\}$$

with $q_{ij,t} = [Q_t]_{ij}$ and $Q_t^{-3/2} := \left[q_{ij,t}^{-3/2}\right]_{i,j=1,\dots,p}$. Thus, (A.2) and (A.3) imply

$$\begin{aligned}\frac{\partial \mathsf{vecl}\,(R_t)}{\partial \boldsymbol{\theta}_2'} &= \mathsf{D}_{p,-}^{+}\left[(Q_t^* \otimes Q_t^*)\,\frac{\partial \mathsf{vec}\,(Q_t)}{\partial \mathsf{vec}\,(Q_t^*)'} + (Q_t^* Q_t \otimes \mathbb{I}_p + \mathbb{I}_p \otimes Q_t^* Q_t)\right]\mathsf{D}_p \frac{\partial \mathsf{vech}\,(Q_t^*)}{\partial \mathsf{vech}\,(Q_t)'}\frac{\partial \mathsf{vech}\,(Q_t)}{\partial \boldsymbol{\theta}_2'} \\ &= \left[\mathsf{D}_{p,-}^{+}\,(Q_t^* \otimes Q_t^*)\,\mathsf{D}_p + \mathsf{D}_{p,-}^{+}\,(Q_t^* Q_t \otimes \mathbb{I}_p + \mathbb{I}_p \otimes Q_t^* Q_t)\,\mathsf{D}_p \frac{\partial \mathsf{vech}\,(Q_t^*)}{\partial \mathsf{vech}\,(Q_t)'}\right]\frac{\partial \mathsf{vech}\,(Q_t)}{\partial \boldsymbol{\theta}_2'}.\end{aligned}$$

The derivation of $\mathsf{vech}\,(Q_t)$ with respect to $\boldsymbol{\theta}_2$ can be split up into:

$$\begin{aligned}\frac{\partial \mathsf{vech}\,(Q_t)}{\partial \alpha} &= -\mathsf{vech}(\bar{Q}) + \mathsf{vech}\left(z_{t-1}z_{t-1}'\right) + \beta\frac{\partial \mathsf{vech}\,(Q_{t-1})}{\partial \alpha} = -\frac{1}{1-\beta}\mathsf{vech}(\bar{Q}) + \sum_{n=0}^{\infty}\beta^n \mathsf{vech}\left(z_{t-n-1}z_{t-n-1}'\right) \\ \frac{\partial \mathsf{vech}\,(Q_t)}{\partial \beta} &= -\mathsf{vech}(\bar{Q}) + \mathsf{vech}\,(Q_{t-1}) + \beta\frac{\partial \mathsf{vech}\,(Q_{t-1})}{\partial \beta} = -\frac{1}{1-\beta}\mathsf{vech}(\bar{Q}) + \sum_{n=0}^{\infty}\beta^n \mathsf{vech}\,(Q_{t-n-1}) \\ \frac{\partial \mathsf{vech}\,(Q_t)}{\partial \mathsf{vecl}\,(\bar{Q})} &= (1-\alpha-\beta)\frac{\partial \mathsf{vech}(\bar{Q})}{\partial \mathsf{vecl}(\bar{Q})'} + \alpha\frac{\partial \mathsf{vech}\left(z_{t-1}z_{t-1}'\right)}{\partial \mathsf{vecl}(\bar{Q})'} + \beta\frac{\partial \mathsf{vech}\,(Q_{t-1})}{\partial \mathsf{vecl}(\bar{Q})'} \overset{\text{CR3}}{=} \frac{1-\alpha-\beta}{1-\beta}\mathsf{L}_p\mathsf{D}_{p,-}\,.\end{aligned}$$

As in the `ccgarch` package, we fix the initial values of $\left(\frac{\partial \mathsf{vech}(Q_0)}{\partial \alpha}, \frac{\partial \mathsf{vech}(Q_0)}{\partial \beta}\right)$ as $(0,0)$.

A.5. The Second Order Partial Derivations with Respect to $\boldsymbol{\theta}$

Along the lines of Nakatani and Teräsvirta (2007), the Hessian of the QLL contributions can be split up into several blocks:

$$\frac{\partial^2 l_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}\partial \boldsymbol{\theta}'} = \begin{bmatrix} \frac{\partial^2 l_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}_1 \partial \boldsymbol{\theta}_1'} & \frac{\partial^2 l_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}_1 \partial \boldsymbol{\theta}_2'} \\ \frac{\partial^2 l_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}_2 \partial \boldsymbol{\theta}_1'} & \frac{\partial^2 l_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}_2 \partial \boldsymbol{\theta}_2'} \end{bmatrix}.$$

A.5.1. The Calculation of $\frac{\partial^2 l_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}_1 \partial \boldsymbol{\theta}_1'}$

Taking into account that the conditional correlation matrix R_t does not depend on the variance parameters, the upper left block of the Hessian is given analogously to (3.6) in Nakatani and Teräsvirta (2007):

$$\frac{\partial^2 l_t(\theta)}{\partial \boldsymbol{\theta}_1 \partial \boldsymbol{\theta}_1'} = -\frac{1}{2}\frac{\partial}{\partial \boldsymbol{\theta}_1'}\left(\frac{\partial \mathsf{vec}\left(D_t\right)'}{\partial \boldsymbol{\theta}_1}\mathsf{vec}\left(2D_t^{-1} - D_t^{-1}R_t^{-1}z_t z_t' - z_t z_t' R_t^{-1} D_t^{-1}\right)\right)$$

$$= -\frac{1}{2}\frac{\partial \mathsf{vec}\left(D_t\right)'}{\partial \boldsymbol{\theta}_1}\Big[-2\left(D_t^{-1}\otimes D_t^{-1}\right) + \left(z_t z_t' \otimes D_t^{-1}R_t^{-1}D_t^{-1}\right) + \left(D_t^{-1}R_t^{-1}D_t^{-1}\otimes z_t z_t'\right)$$

$$+ \left(D_t^{-1}\otimes D_t^{-1}R_t^{-1}z_t z_t'\right) + \left(D_t^{-1}R_t^{-1}z_t z_t' \otimes D_t^{-1}\right) + \left(z_t z_t' R_t^{-1}D_t^{-1}\otimes D_t^{-1}\right)$$

$$+ \left(D_t^{-1}\otimes z_t z_t' R_t^{-1} D_t^{-1}\right)\Big]\frac{\partial \mathsf{vec}\left(D_t\right)}{\partial \boldsymbol{\theta}_1'}$$

$$+ \left[\frac{1}{2}\left(\mathsf{vec}\left(D_t^{-1}R_t^{-1}z_t z_t'\right)\otimes \mathbb{I}_{3p}\right) + \frac{1}{2}\left(\mathsf{vec}\left(z_t z_t' R_t^{-1}D_t^{-1}\right)\otimes \mathbb{I}_{3p}\right) - \left(\mathsf{vec}\left(D_t^{-1}\right)\otimes \mathbb{I}_{3p}\right)\right]\frac{\partial^2 \mathsf{vec}\left(D_t\right)'}{\partial \boldsymbol{\theta}_1 \partial \boldsymbol{\theta}_1'}.$$

A closer look at the individual parts of the derivation yields

$$\frac{\partial^2 \mathsf{vec}\left(D_t\right)'}{\partial \boldsymbol{\theta}_1 \partial \boldsymbol{\theta}_1'} = \frac{\partial}{\partial \boldsymbol{\theta}_1'}\mathsf{vec}\left(\frac{\partial \mathsf{vech}\left(D_t\right)'}{\partial \boldsymbol{\theta}_1}\mathsf{D}_p'\right) = \left(\mathsf{D}_p \otimes \mathbb{I}_{3p}\right)\frac{\partial}{\partial \boldsymbol{\theta}_1'}\mathsf{vec}\left(\frac{\partial \mathsf{vech}\left(D_t\right)'}{\partial \boldsymbol{\theta}_1}\right).$$

Denote the non zero derivation blocks of the main diagonal entries of D_t as

$$g_{it}^{(2)} := \frac{\partial^2 h_{it}^{1/2}}{\partial \phi_i \partial \phi_i'}.$$

Thus, the second order derivations of $\mathsf{vech}\left(D_t\right)$ are given as

$$\frac{\partial}{\partial \boldsymbol{\theta}_1'}\mathsf{vec}\left(\frac{\partial \mathsf{vech}\left(D_t\right)'}{\partial \boldsymbol{\theta}_1}\right) = \begin{bmatrix} g_{1t}^{(2)} & \mathbf{0}_{3\times 3p^2} & \mathbf{0}_{3\times 3} & \mathbf{0}_{3\times 3p(p-1)} & \dots & \mathbf{0}_{3\times 3} & \mathbf{0}_{3\times 6p} & \mathbf{0}_{3\times 3} \\ \mathbf{0}_{3\times 3} & \mathbf{0}_{3\times 3p^2} & g_{2t}^{(2)} & \mathbf{0}_{3\times 3p(p-1)} & \dots & \mathbf{0}_{3\times 3} & \mathbf{0}_{3\times 6p} & \mathbf{0}_{3\times 3} \\ \vdots & \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots \\ \mathbf{0}_{3\times 3} & \mathbf{0}_{3\times 3p^2} & \mathbf{0}_{3\times 3} & \mathbf{0}_{3\times 3p(p-1)} & \dots & g_{p-1,t}^{(2)} & \mathbf{0}_{3\times 6p} & \mathbf{0}_{3\times 3} \\ \mathbf{0}_{3\times 3} & \mathbf{0}_{3\times 3p^2} & \mathbf{0}_{3\times 3} & \mathbf{0}_{3\times 3p(p-1)} & \dots & \mathbf{0}_{3\times 3} & \mathbf{0}_{3\times 6p} & g_{pt}^{(2)} \end{bmatrix}.$$

Note, that with $V_{it} := \frac{\partial v_{it}}{\partial \phi_i'} = \left(\mathbf{0}_3, \mathbf{0}_3, \frac{\partial h_{it}}{\partial \phi_i}\right)'$, we have

$$g_{it}^{(2)} = -\frac{1}{4}\frac{1}{h_{it}^{3/2}}\frac{\partial h_{it}}{\partial \phi_i'}\left(v_{i,t-1} + \beta_i \frac{\partial h_{i,t-1}}{\partial \phi_i}\right) + \frac{1}{2}\frac{1}{h_{it}^{1/2}}\left(V_{i,t-1} + \beta_i \frac{\partial^2 h_{i,t-1}}{\partial \phi_i \partial \phi_i'}\right)$$

with

- $\frac{\partial^2 h_{it}}{(\partial\omega_i)^2} = \beta_i \frac{\partial^2 h_{i,t-1}}{(\partial\omega_i)^2} = 0$
- $\frac{\partial^2 h_{it}}{\partial\omega_i \partial\alpha_i} = \beta_i \frac{\partial^2 h_{i,t-1}}{\partial\omega_i \partial\alpha_i} = 0$
- $\frac{\partial^2 h_{it}}{\partial\omega_i \partial\beta_i} = \frac{\partial h_{it}}{\partial\omega_i} + \beta_i \frac{\partial^2 h_{i,t-1}}{\partial\omega_i \partial\beta_i} = \frac{1}{(1-\beta_i)^2}$
- $\frac{\partial^2 h_{it}}{(\partial\alpha_i)^2} = \beta_i \frac{\partial^2 h_{i,t-1}}{(\partial\alpha_i)^2} = 0$
- $\frac{\partial^2 h_{it}}{\partial\alpha_i \partial\beta_i} = \frac{\partial h_{it}}{\partial\alpha_i} + \beta_i \frac{\partial^2 h_{i,t-1}}{\partial\alpha_i \partial\beta_i} = \sum_{n=0}^{\infty} n\beta_i^{n-1} y_{i,t-n-1}^2$
- $\frac{\partial^2 h_{it}}{(\partial\beta_i)^2} = 2\frac{\partial h_{i,t-1}}{\partial\beta_i} + \beta_i \frac{\partial^2 h_{i,t-1}}{(\partial\beta_i)^2} = 2\sum_{n=0}^{\infty} \beta_i^n \frac{\partial h_{i,t-n-2}}{\partial\beta_i} = \frac{2\omega_i}{(1-\beta_i)^4} + 2\alpha_i \sum_{n=0}^{\infty}\beta_i^n \sum_{k=0}^{\infty}\beta_i^k \sum_{l=0}^{\infty}\beta_i^l y_{i,t-n-k-l-2}^2.$

Analogously to the approach in Section A.3, a vector of zeros can be chosen as starting value for the recursive calculation of the second order partial derivations of the conditional variances with respect to the variance parameters.

A.5.2. The Calculation of $\frac{\partial^2 l_t(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}_1 \partial\boldsymbol{\theta}_2'}$ and $\frac{\partial^2 l_t(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}_2 \partial\boldsymbol{\theta}_1'}$

Analogously to (3.9) and (3.10) in Nakatani and Teräsvirta (2007) under limitation to the partial derivations of the lower diagonal entries of R_t with respect to the correlation parameters, the off diagonal blocks of the Hessian equal

$$\begin{aligned}
\frac{\partial^2 l_t(\theta)}{\partial\boldsymbol{\theta}_1 \partial\boldsymbol{\theta}_2'} &= -\frac{1}{2}\frac{\partial}{\partial\boldsymbol{\theta}_2'}\left(\frac{\partial \mathsf{vec}\,(D_t)'}{\partial\boldsymbol{\theta}_1}\mathsf{vec}\left(2D_t^{-1} - D_t^{-1}R_t^{-1}z_t z_t' - z_t z_t' R_t^{-1} D_t^{-1}\right)\right) \\
&= \frac{1}{2}\frac{\partial \mathsf{vec}\,(D_t)'}{\partial\boldsymbol{\theta}_1}\left[\frac{\partial \mathsf{vec}\left(D_t^{-1}R_t^{-1}z_t z_t'\right)}{\partial\boldsymbol{\theta}_2'} + \frac{\partial \mathsf{vec}\left(z_t z_t' R_t^{-1} D_t^{-1}\right)}{\partial\boldsymbol{\theta}_2'}\right] \\
&= \frac{1}{2}\frac{\partial \mathsf{vec}\,(D_t)'}{\partial\boldsymbol{\theta}_1}\left[\left(z_t z_t' \otimes D_t^{-1}\right) + \left(D_t^{-1} \otimes z_t z_t'\right)\right]\frac{\partial \mathsf{vec}\left(R_t^{-1}\right)}{\partial \mathsf{vec}\,(R_t)'}\frac{\partial \mathsf{vec}\,(R_t)}{\partial \mathsf{vecl}\,(R_t)'}\frac{\partial \mathsf{vecl}\,(R_t)}{\partial\boldsymbol{\theta}_2'} \\
&\overset{\text{CR2}}{=} -\frac{\partial \mathsf{vec}\,(D_t)'}{\partial\boldsymbol{\theta}_1}\left[\left(z_t z_t' R_t^{-1} \otimes D_t^{-1}R_t^{-1}\right) + \left(D_t^{-1}R_t^{-1} \otimes z_t z_t' R_t^{-1}\right)\right]\left(\mathsf{D}_{p,-}^{+}\right)'\frac{\partial \mathsf{vecl}\,(R_t)}{\partial\boldsymbol{\theta}_2'} \\
\frac{\partial^2 l_t(\theta)}{\partial\boldsymbol{\theta}_2 \partial\boldsymbol{\theta}_1'} &= \left[\frac{\partial^2 l_t(\theta)}{\partial\boldsymbol{\theta}_1 \partial\boldsymbol{\theta}_2'}\right]' = -\frac{\partial \mathsf{vecl}\,(R_t)'}{\partial\boldsymbol{\theta}_2}\mathsf{D}_{p,-}^{+}\left[\left(R_t^{-1}z_t z_t' \otimes R_t^{-1}D_t^{-1}\right) + \left(R_t^{-1}D_t^{-1} \otimes R_t^{-1}z_t z_t'\right)\right]\frac{\partial \mathsf{vec}\,(D_t)}{\partial\boldsymbol{\theta}_1'}.
\end{aligned}$$

A.5.3. The Calculation of $\frac{\partial^2 l_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}_2 \partial \boldsymbol{\theta}_2'}$

Finally, the lower right block of the Hessian is given as

$$\begin{aligned}
\frac{\partial^2 l_t(\theta)}{\partial \boldsymbol{\theta}_2 \partial \boldsymbol{\theta}_2'} &= -\frac{\partial}{\partial \boldsymbol{\theta}_2'}\left(\frac{\partial \mathsf{vecl}\,(R_t)'}{\partial \boldsymbol{\theta}_2}\mathsf{vecl}\left(R_t^{-1}\right)\right) + \frac{\partial}{\partial \boldsymbol{\theta}_2'}\left(\frac{\partial \mathsf{vecl}\,(R_t)'}{\partial \boldsymbol{\theta}_2}\mathsf{vecl}\left(R_t^{-1} z_t z_t' R_t^{-1}\right)\right) \\
&= -\frac{\partial \mathsf{vecl}\,(R_t)'}{\partial \boldsymbol{\theta}_2}\left[\frac{\partial \mathsf{vecl}\left(R_t^{-1}\right)}{\partial \boldsymbol{\theta}_2'} - \frac{\partial \mathsf{vecl}\left(R_t^{-1} z_t z_t' R_t^{-1}\right)}{\partial \boldsymbol{\theta}_2'}\right] \\
&\quad - \left[\left(\mathsf{vecl}\left(R_t^{-1}\right)' \otimes \mathbb{I}_{p^-+2}\right) - \left(\mathsf{vecl}\left(R_t^{-1} z_t z_t' R_t^{-1}\right)' \otimes \mathbb{I}_{p^-+2}\right)\right] \frac{\partial}{\partial \boldsymbol{\theta}_2'}\mathsf{vec}\left(\frac{\partial \mathsf{vecl}\,(R_t)'}{\partial \boldsymbol{\theta}_2}\right) \\
&= \frac{\partial \mathsf{vecl}\,(R_t)'}{\partial \boldsymbol{\theta}_2}\mathsf{D}_{p,-}^{+}\left[\left(R_t^{-1} \otimes R_t^{-1}\right) - \left(R_t^{-1} z_t z_t' R_t^{-1} \otimes R_t^{-1}\right) - \left(R_t^{-1} \otimes R_t^{-1} z_t z_t' R_t^{-1}\right)\right]\mathsf{D}_{p,-}\frac{\partial \mathsf{vecl}\,(R_t)}{\partial \boldsymbol{\theta}_2'} \\
&\quad + \left[\left(\mathsf{vecl}\left(R_t^{-1}\right)' \otimes \mathbb{I}_{p^-+2}\right) + \left(\mathsf{vecl}\left(R_t^{-1} z_t z_t' R_t^{-1}\right)' \otimes \mathbb{I}_{p^-+2}\right)\right] \frac{\partial}{\partial \boldsymbol{\theta}_2'}\mathsf{vec}\left(\frac{\partial \mathsf{vecl}\,(R_t)'}{\partial \boldsymbol{\theta}_2}\right).
\end{aligned}$$

A closer look at the second order partial derivations of the entries of the conditional correlation matrix with respect to the correlation parameters yields

$$\begin{aligned}
\frac{\partial}{\partial \boldsymbol{\theta}_2'}\mathsf{vec}\left(\frac{\partial \mathsf{vecl}\,(R_t)'}{\partial \boldsymbol{\theta}_2}\right) &= \frac{\partial}{\partial \boldsymbol{\theta}_2'}\mathsf{vec}\left(\frac{\partial \mathsf{vech}\,(Q_t)'}{\partial \boldsymbol{\theta}_2}\mathsf{D}_p'\left(Q_t^* \otimes Q_t^*\right)\left(\mathsf{D}_{p,-}^{+}\right)'\right) \\
&\quad + \frac{\partial}{\partial \boldsymbol{\theta}_2'}\mathsf{vec}\left(\frac{\partial \mathsf{vech}\,(Q_t)'}{\partial \boldsymbol{\theta}_2}\frac{\partial \mathsf{vech}\,(Q_t^*)'}{\partial \mathsf{vech}\,(Q_t)}\mathsf{D}_p'\left(Q_t Q_t^* \otimes \mathbb{I}_p + \mathbb{I}_p \otimes Q_t Q_t^*\right)\left(\mathsf{D}_{p,-}^{+}\right)'\right). \quad \text{(A.4)}
\end{aligned}$$

With 10.5.5(4) of Lütkepohl (1996) the first summand of (A.4) equals

$$\left(\mathsf{D}_{p,-}^{+} \otimes \frac{\partial \mathsf{vech}\,(Q_t)'}{\partial \boldsymbol{\theta}_2}\mathsf{D}_p'\right)\frac{\partial \mathsf{vec}\,(Q_t^* \otimes Q_t^*)}{\partial \boldsymbol{\theta}_2'} + \left(\mathsf{D}_{p,-}^{+}\left(Q_t^* \otimes Q_t^*\right)\mathsf{D}_p \otimes \mathbb{I}_{p^-+2}\right)\frac{\partial}{\partial \boldsymbol{\theta}_2'}\mathsf{vec}\left(\frac{\partial \mathsf{vech}\,(Q_t)'}{\partial \boldsymbol{\theta}_2}\right).$$

Furthermore, 10.5.5(7) of Lütkepohl (1996) gives

$$\begin{aligned}
&\frac{\partial \mathsf{vec}\,(Q_t^* \otimes Q_t^*)}{\partial \boldsymbol{\theta}_2'} = \left(\mathbb{I}_p \otimes \mathsf{K}_{pp} \otimes \mathbb{I}_p\right)\left[\frac{\partial \mathsf{vec}\,(Q_t^*)}{\partial \boldsymbol{\theta}_2'} \otimes \mathsf{vec}\,(Q_t^*) + \mathsf{vec}\,(Q_t^*) \otimes \frac{\partial \mathsf{vec}\,(Q_t^*)}{\partial \boldsymbol{\theta}_2'}\right] \\
&\text{with} \qquad \frac{\partial \mathsf{vec}\,(Q_t^*)}{\partial \boldsymbol{\theta}_2'} = \mathsf{D}_p\frac{\partial \mathsf{vech}\,(Q_t^*)}{\partial \mathsf{vech}\,(Q_t)'}\frac{\partial \mathsf{vech}\,(Q_t)}{\partial \boldsymbol{\theta}_2'}.
\end{aligned}$$

Note that

$$\frac{\partial}{\partial \boldsymbol{\theta}_2'}\mathsf{vec}\left(\frac{\partial \mathsf{vech}\left(Q_t\right)'}{\partial \boldsymbol{\theta}_2}\right) = \mathsf{K}_{p^+,p^-+2}\frac{\partial}{\partial \boldsymbol{\theta}_2'}\mathsf{vec}\left(\frac{\partial \mathsf{vech}\left(Q_t\right)}{\partial \boldsymbol{\theta}_2'}\right)$$

and that $\frac{\partial}{\partial \boldsymbol{\theta}_2'}\mathsf{vec}\left(\frac{\partial \mathsf{vech}(Q_t)}{\partial \boldsymbol{\theta}_2'}\right)$ can be split up into several block matrices:

$$\begin{bmatrix} \frac{\partial^2 \mathsf{vech}(Q_t)}{(\partial\alpha)^2} & \frac{\partial^2 \mathsf{vech}(Q_t)}{\partial\alpha\partial\beta} & \frac{\partial^2 \mathsf{vech}(Q_t)}{\partial\alpha\partial\mathsf{vecl}\left(\bar{Q}\right)'} \\ \frac{\partial^2 \mathsf{vech}(Q_t)}{\partial\alpha\partial\beta} & \frac{\partial^2 \mathsf{vech}(Q_t)}{(\partial\beta)^2} & \frac{\partial^2 \mathsf{vech}(Q_t)}{\partial\beta\partial\mathsf{vecl}\left(\bar{Q}\right)'} \\ \mathsf{vec}\left(\frac{\partial^2 \mathsf{vech}(Q_t)}{\partial\alpha\partial\mathsf{vecl}\left(\bar{Q}\right)'}\right) & \mathsf{vec}\left(\frac{\partial^2 \mathsf{vech}(Q_t)'}{\partial\beta\partial\mathsf{vecl}\left(\bar{Q}\right)'}\right) & \frac{\partial}{\partial\mathsf{vecl}\left(\bar{Q}\right)'}\mathsf{vec}\left(\frac{\partial \mathsf{vech}(Q_t)'}{\partial\mathsf{vecl}\left(\bar{Q}\right)'}\right) \end{bmatrix}$$

$$= \begin{bmatrix} \beta\frac{\partial^2 \mathsf{vech}(Q_{t-1})}{(\partial\alpha)^2} & \frac{\partial \mathsf{vech}(Q_{t-1})}{\partial\alpha} + \beta\frac{\partial^2 \mathsf{vech}(Q_{t-1})}{\partial\alpha\partial\beta} & -\frac{1}{1-\beta}\mathsf{L}_p\mathsf{D}_{p,-} \\ \frac{\partial \mathsf{vech}(Q_{t-1})}{\partial\alpha} + \beta\frac{\partial^2 \mathsf{vech}(Q_{t-1})}{\partial\alpha\partial\beta} & 2\frac{\partial \mathsf{vech}(Q_{t-1})}{\partial\beta} + \beta\frac{\partial^2 \mathsf{vech}(Q_{t-1})}{(\partial\beta)^2} & -\frac{2-\alpha-2\beta}{(1-\beta)^2}\mathsf{L}_p\mathsf{D}_{p,-} \\ -\frac{1}{1-\beta}\mathsf{vec}\left(\mathsf{L}_p\mathsf{D}_{p,-}\right) & -\frac{2-\alpha-2\beta}{(1-\beta)^2}\mathsf{vec}\left(\mathsf{L}_p\mathsf{D}_{p,-}\right) & \mathbf{0}_{p^+p^-\times p^-} \end{bmatrix}.$$

With the use of 10.5.5(4) in Lütkepohl (1996), the second summand of (A.4) equals

$$\left(\mathsf{D}_{p,-}^+ \otimes \frac{\partial \mathsf{vech}\left(Q_t\right)'}{\partial \boldsymbol{\theta}_2}\frac{\partial \mathsf{vech}\left(Q_t^*\right)'}{\partial \mathsf{vech}\left(Q_t\right)}\mathsf{D}_p'\right)\frac{\partial \mathsf{vec}\left(Q_tQ_t^* \otimes \mathbb{I}_p + \mathbb{I}_p \otimes Q_tQ_t^*\right)}{\partial \boldsymbol{\theta}_2'}$$
$$+\left(\mathsf{D}_{p,-}^+\left[Q_t^*Q_t \otimes \mathbb{I}_p + \mathbb{I}_p \otimes Q_t^*Q_t\right]\mathsf{D}_p \otimes \mathbb{I}_{p^-+2}\right)\frac{\partial}{\partial \boldsymbol{\theta}_2'}\mathsf{vec}\left(\frac{\partial \mathsf{vech}\left(Q_t^*\right)'}{\partial \boldsymbol{\theta}_2}\right)$$

with

$$\frac{\partial \mathsf{vec}\left(Q_tQ_t^* \otimes \mathbb{I}_p + \mathbb{I}_p \otimes Q_tQ_t^*\right)}{\partial \boldsymbol{\theta}_2'} = \left(\mathbb{I}_p \otimes \mathsf{K}_{pp} \otimes \mathbb{I}_p\right)\left[\frac{\partial \mathsf{vec}\left(Q_t^*Q_t\right)}{\partial \boldsymbol{\theta}_2'} \otimes \mathsf{vec}\left(\mathbb{I}_p\right) + \mathsf{vec}\left(\mathbb{I}_p\right) \otimes \frac{\partial \mathsf{vec}\left(Q_t^*Q_t\right)}{\partial \boldsymbol{\theta}_2'}\right]$$

and

$$\frac{\partial \mathsf{vec}\left(Q_t^*Q_t\right)}{\partial \boldsymbol{\theta}_2'} = \left(\mathbb{I}_p \otimes Q_t^*\right)\frac{\partial \mathsf{vec}\left(Q_t\right)}{\partial \boldsymbol{\theta}_2'} + \left(Q_t \otimes \mathbb{I}_p\right)\frac{\partial \mathsf{vec}\left(Q_t^*\right)}{\partial \boldsymbol{\theta}_2'}$$
$$= \left[\left(\mathbb{I}_p \otimes Q_t\right)\mathsf{D}_p\frac{\partial \mathsf{vech}\left(Q_t^*\right)}{\partial \mathsf{vech}\left(Q_t\right)'} + \left(Q_t^* \otimes \mathbb{I}_p\right)\mathsf{D}_p\right]\frac{\partial \mathsf{vech}\left(Q_t\right)}{\partial \boldsymbol{\theta}_2'}.$$

B.1. The Proof of Proposition 4.1

Denote $\mathbf{y}_t := \mathsf{vec}\left(y_t^{(2)} y_t^{(2)\prime}\right)$ with $y_t^{(2)} = \left(y_{1t}^2, \ldots, y_{pt}^2\right)'$, $\mathbf{h}_t := \mathsf{vec}\,(h_t h_t')$ and $\boldsymbol{Y}_t := \mathsf{E}\,[\mathbf{y}_t \mathbf{y}_t']$. Thus, the existence of the eighth moments and cross moments of y_t is implied by $\boldsymbol{Y}_t < \infty$.

Note, that with 7.2(7) in Lütkepohl (1996), we have

$$\mathbf{y}_t \mathbf{y}_t' = \mathsf{vec}\left(Z_t^2 h_t h_t' Z_t^2\right) \mathsf{vec}\left(Z_t^2 h_t h_t' Z_t^2\right)' = \left[\bigotimes\nolimits_2 Z_t^2\right] \mathbf{h}_t \mathbf{h}_t' \left[\bigotimes\nolimits_2 Z_t^2\right] \tag{B.1}$$

$$\Rightarrow\ \mathsf{vec}\,[\mathbf{y}_t \mathbf{y}_t'] = \left[\bigotimes\nolimits_4 Z_t^2\right] \mathsf{vec}\,[\mathbf{h}_t \mathbf{h}_t'] \tag{B.2}$$

where $\bigotimes_k$ denotes the k-fold Kronecker product of identical matrices. This yields

$$\mathsf{vec}\,(\boldsymbol{Y}_t) = \mathsf{E}\left(\left[\bigotimes\nolimits_4 Z_t^2\right] \mathsf{vec}\,[\mathbf{h}_t \mathbf{h}_t']\right). \tag{B.3}$$

In the (extended) CCC model that is considered by He and Teräsvirta (2004), the random vectors z_t are i.i.d. and independent of h_t which allows for a simple factorization of the expectation in (B.3). Unfortunately, in the model with dynamic conditional correlation this does not work anymore. To enable the factorization, we use the independent random vectors $\eta_t := R_t^{-1/2} z_t$ with $\eta_t \sim (0, \mathbb{I}_p)$. Denote $R_t^{1/2} := [\tilde{r}_{ij,t}]_{i,j=1,\ldots,p}$ and note that

$$z_{it} = \sum_{j=1}^{p} \tilde{r}_{ij,t} \eta_{jt} \quad \text{and} \quad z_{it}^2 = \sum_{j=1}^{p} \sum_{k=1}^{p} \tilde{r}_{ij,t} \tilde{r}_{ik,t} \eta_{jt} \eta_{kt}\ .$$

Note that all entries of $R_t^{1/2}$ are bounded by one in modulus since the i-th main diagonal of $R_t^{1/2} R_t^{1/2}$ is $\sum_{j=1}^{p} \tilde{r}_{ij,t}^2$ and equal one for all $i, j \in \{1, \ldots, p\}$. This property yields

$$z_{it}^2 < \sum_{j=1}^{p} \sum_{k=1}^{p} |\eta_{jt} \eta_{kt}| =: \tilde{z}_t^2 \quad \forall i \in \{1, \ldots, p\}\,.$$

Thus, using the independence of η_t and h_t , we have for (B.3):

$$\mathsf{vec}\,(\boldsymbol{Y}_t) \leq \mathsf{E}\left(\left[\tilde{z}_t^2\right]^4\right)\ \mathsf{E}\Big(\mathsf{vec}\,[\mathbf{h}_t \mathbf{h}_t']\Big). \tag{B.4}$$

First, we will argue why the first factor in B.4 is finite. Repeated substitution yields

$$\eta_t = R_t^{-1/2} z_t = R_t^{-1/2} D_t^{-1} y_t = R_t^{-1/2} D_t^{-1} \left[D_t R_t Dt \right]^{1/2} \epsilon_t.$$

Thus, the entries of η_t are weighted sums of the random variables $\epsilon_{1t}, \ldots, \epsilon_{pt}$ whose eighth moments and cross are finite by Assumption 4.3. This implies that $\mathsf{E}\left(\left[\tilde{z}_t^2\right]^4\right)$ is finite if this property applies to the weights. The latter one is implied by Lemmas B.7 and B.14 and with the arguments of the proof of Lemma B.13 in Section B.2

The finiteness of the second factor on the righthand side of (B.4) can be shown analogously to the approach in He and Teräsvirta (2004). If GARCH(1,1) models are postulated to explain the conditional variances, we have along the lines of the proof of Theorem 2 in He and Teräsvirta (2004):

- $$\mathbf{h}_t - \left[\bigotimes_2 \boldsymbol{C}_{t-1}\right] \mathbf{h}_{t-1} = \mathsf{vec}\left(\boldsymbol{\omega}\boldsymbol{\omega}'\right) + \sum_{k=1}^{2} \left[E_{k1} \otimes E_{k2}\right] h_{t-1} \quad \text{(B.5)}$$

  where $\left\{\left(E_{k1}, E_{k2}\right), k \in \{1, 2\}\right\}$ is the set of all permutations of $\left(\boldsymbol{\omega}, \boldsymbol{C}_{t-1}\right)$.

- $$\mathsf{vec}\left[\mathbf{h}_t h_t'\right] - \left[\bigotimes_3 \boldsymbol{C}_{t-1}\right] \mathsf{vec}\left[\mathbf{h}_{t-1} h_{t-1}'\right] \quad \text{(B.6)}$$

  $$= \mathsf{vec}\left[\mathsf{vec}\left(\boldsymbol{\omega}\boldsymbol{\omega}'\right)\boldsymbol{\omega}'\right] + \sum_{k=1}^{3} \left[E_{k1}^{(1)} \otimes E_{k2}^{(1)} \otimes E_{k3}^{(1)}\right] h_{t-1} + \sum_{k=1}^{3} \left[E_{k1}^{(2)} \otimes E_{k2}^{(2)} \otimes E_{k3}^{(2)}\right] \mathbf{h}_{t-1}$$

  where $\left\{\left(E_{k1}^{(1)}, \ldots, E_{k3}^{(1)}\right), k \in \{1, \ldots, 3\}\right\}$ is the set of all permutations of $\left(\boldsymbol{\omega}, \boldsymbol{\omega}, \boldsymbol{C}_{t-1}\right)$ and $\left\{\left(E_{k1}^{(2)}, \ldots, E_{k3}^{(2)}\right), k \in \{1, \ldots, 3\}\right\}$ is the set of all permutations of $\left(\boldsymbol{\omega}, \boldsymbol{C}_{t-1}, \boldsymbol{C}_{t-1}\right)$.

- $$\mathsf{vec}\left[\mathbf{h}_t \mathbf{h}_t'\right] - \left[\bigotimes_4 \boldsymbol{C}_{t-1}\right] \mathsf{vec}\left[\mathbf{h}_{t-1} \mathbf{h}_{t-1}'\right] \quad \text{(B.7)}$$

  $$= \mathsf{vec}\left[\mathsf{vec}\left(\boldsymbol{\omega}\boldsymbol{\omega}'\right)\mathsf{vec}\left(\boldsymbol{\omega}\boldsymbol{\omega}'\right)'\right] + \sum_{k=1}^{4} \left[E_{k1}^{(1)} \otimes E_{k2}^{(1)} \otimes E_{k3}^{(1)} \otimes E_{k4}^{(1)}\right] h_{t-1} + \sum_{k=1}^{6} \left[E_{k1}^{(2)} \otimes E_{k2}^{(2)} \otimes E_{k3}^{(2)} \otimes E_{k4}^{(2)}\right] \mathbf{h}_{t-1}$$

  $$+ \left(\left(\boldsymbol{C}_{t-1} \otimes \boldsymbol{\omega} + \boldsymbol{\omega} \otimes \boldsymbol{C}_{t-1}\right) \otimes \bigotimes_2 \boldsymbol{C}_{t-1} + \left[\bigotimes_2 \boldsymbol{C}_{t-1} \otimes \left(\boldsymbol{C}_{t-1} \otimes \boldsymbol{\omega} + \boldsymbol{\omega} \otimes \boldsymbol{C}_{t-1}\right)\right] \mathsf{K}_{p^2 p}\right) \mathsf{vec}\left(\mathbf{h}_{t-1} h_{t-1}'\right)$$

  where $\left\{\left(E_{k1}^{(1)}, \ldots, E_{k4}^{(1)}\right), k \in \{1, \ldots, 4\}\right\}$ is the set of all permutations of $\left(\boldsymbol{\omega}, \boldsymbol{\omega}, \boldsymbol{\omega}, \boldsymbol{C}_{t-1}\right)$ and $\left\{\left(E_{k1}^{(2)}, \ldots, E_{k4}^{(2)}\right), k \in \{1, \ldots, 6\}\right\}$ is the set of all permutations of $\left(\boldsymbol{\omega}, \boldsymbol{\omega}, \boldsymbol{C}_{t-1}, \boldsymbol{C}_{t-1}\right)$.

Consider the matrix polynomials $\Psi^{(j)}(L)$, $j = 1, 2, 3$, in the lag operator L with

$$\Psi^{(j)}(L) = \sum_{i=0}^{\infty} \psi_{i,t-1}^{(j)} L^i \quad \text{and} \quad \forall\ j \in \{1, 2, 3\}$$

with $\psi_{0,t-1}^{(j)} = \mathbb{I}_{p^{j+1}}$, $\psi_{1,t-1}^{(j)} = -\bigotimes_{j+1} \boldsymbol{C}_{t-1}$ and $\psi_{l,t-1}^{(j)} = 0$ for all $l \geq 2$ and $j \in \{1, 2, 3\}$. Along the lines of He and Teräsvirta (2004), the inverses of $\Psi^{(j)}(L)$, $j \in \{1, 2, 3\}$, exist if

$$\lambda_{\mathsf{max}} \left(\mathsf{E} \left[\bigotimes\nolimits_j \boldsymbol{C}_0 \right] \right) < \infty \quad \forall\ j \in \{1, \ldots, 4\}. \tag{B.8}$$

Thus, multiplying the matching inverses $\Psi^{(j)}(L)$ from the lefthand side to (B.5)-(B.7), the property (B.8) allows a filter representation of $\mathbf{h}_t$, $\mathsf{vec}\left(\mathbf{h}_t h_t'\right)$ and $\mathsf{vec}\left(\mathbf{h}_t \mathbf{h}_t'\right)$ on the process $\{h_t\}$. Since $h_t < \infty$ for all $t \in \mathbb{Z}$ almost surely, (B.8) yields the finity of the righthand side of (B.4). This implies the existence of the eighth moments and cross moments of y_t which completes the proof.

■

B.2. The Proof of Proposition 4.2

The proof of Proposition 4.2 is organized as follows: First of all, we show that the finite past variation matrix $\widehat{\boldsymbol{D}}_m(\boldsymbol{\theta}) = \frac{1}{m} \sum_{t=1}^{m} \hat{l}_t'(\boldsymbol{\theta}) \hat{l}_t'(\boldsymbol{\theta})^T$ is a suitable substitute for the matrix with infinite past $\boldsymbol{D}_m(\boldsymbol{\theta}) = \frac{1}{m} \sum_{t=1}^{m} l_t'(\boldsymbol{\theta}) l_t'(\boldsymbol{\theta})^T$, i.e.

$$\sup_{u \in U} \left| \hat{\boldsymbol{D}}_m(u) - \boldsymbol{D}_m(u) \right| \stackrel{a.s.}{\to} 0. \tag{B.9}$$

To ensure the existence of the limit matrix $\boldsymbol{D}(u) = \mathsf{E}\left(l_0'(u) l_0'(u)^T\right)$ for all $u \in U$ in the first place, it has to be shown that

$$\mathsf{E}\left(\sup_{u \in U} \left| l_0'(u) l_0'(u)^T \right| \right) < \infty. \tag{B.10}$$

Since the QMLE $\hat{\theta}_m$ is a strongly consistent estimator of $\boldsymbol{\theta}$ it remains to verify the uniform convergence of $\boldsymbol{D}_m(\cdot)$ to $\boldsymbol{D}(\cdot)$. The latter properties imply

$$\hat{\boldsymbol{D}}_m \left(\hat{\theta}_m \right) \stackrel{a.s.}{\to} \boldsymbol{D}(\theta) = \boldsymbol{D}. \tag{B.11}$$

B.2.1. The Proof of $\sup_{u\in U}\left|\hat{\boldsymbol{D}}_m(u)-\boldsymbol{D}_m(u)\right| \overset{a.s.}{\to} 0$

Along the lines of Berkes et al. (2003), we use a multivariate version of the classic mean value theorem (MVT) throughout the proof section:

Lemma B.1. *Let $f:\mathbb{R}^l \to \mathbb{R}$ be a function that is continuous in all of its arguments. Furthermore, let $\boldsymbol{x}$ and $\boldsymbol{y}$ be some l dimensional real valued vectors. Then, there exists a vector $\xi \in \mathbb{R}^l$ with $|\xi-\boldsymbol{x}| \le |\boldsymbol{x}-\boldsymbol{y}|$ and $|\xi-\boldsymbol{y}| \le |\boldsymbol{x}-\boldsymbol{y}|$, such that*

$$|f(\boldsymbol{x})-f(\boldsymbol{y})| = \left|\left.\frac{\partial f(\boldsymbol{x})}{\partial \boldsymbol{x}}\right|_{\boldsymbol{x}=\xi}\right| |\boldsymbol{x}-\boldsymbol{y}|. \tag{B.12}$$

Proof: Denote $\boldsymbol{x} := (x_1,\dots,x_l)'$, $\boldsymbol{y} := (y_1,\dots,y_l)'$ and $\boldsymbol{y}_i := (x_1,\dots,x_{i-1},y_i,x_{i+1},\dots,x_l)'$. Note that $|\boldsymbol{x}-\boldsymbol{y}_i| = |x_i-y_i|$. A componentwise application of the univariate MVT implies that for all $i\in\{1,\dots,l\}$ there exist $\xi_i \in [x_i,y_i]$ and $\boldsymbol{\xi}_i =: (x_1,\dots,x_{i-1},\xi_i,x_{i+1},\dots,x_l)'$ with

$$|f(\boldsymbol{x})-f(\boldsymbol{y}_i)| = \left|\left.\frac{\partial f(\boldsymbol{x})}{\partial x_i}\right|_{\boldsymbol{x}=\boldsymbol{\xi}_i}\right| |x_i-y_i|.$$

Thus, choosing $\xi := (\xi_1,\dots,\xi_l)'$ yields $|\boldsymbol{\xi}_i-\boldsymbol{x}| \le |\xi_i-x_i| \le |x_i-y_i|$ and $|\boldsymbol{\xi}_i-\boldsymbol{y}| \le |x_i-y_i|$ for all $i\in\{1,\dots,p\}$ as well as (B.12) which completes the proof. ■

With Lemma B.1, we have

$$\begin{aligned}\sup_{u\in U}\left|\widehat{\boldsymbol{D}}_m(u)-\boldsymbol{D}_m(u)\right| &= \frac{1}{m}\sup_{u\in U}\left|\sum_{t=1}^m\left[\hat{l}_t'(u)\hat{l}_t'(u)^T - l_t'(u)l_t'(u)^T\right]\right| \\ &\le \frac{2d}{m}\sup_{u\in U,t\in\mathbb{Z}}|v_t(u)|\sup_{u\in U}\left|\sum_{t=1}^m\left[\hat{l}_t'(u)-l_t'(u)\right]\right| \end{aligned}\tag{B.13}$$

where $v_t(u)\in\mathbb{R}^d$ is such that $|v_t(u)-l_t'(u)| \le \left|\hat{l}_t'(u)-l_t'(u)\right|$ and $\left|v_t(u)-\hat{l}_t'(u)\right| \le \left|\hat{l}_t'(u)-l_t'(u)\right|$. For the sum in (B.13), we have

$$\sup_{u\in U}\left|\sum_{t=1}^m\left[\hat{l}_t'(u)-l_t'(u)\right]\right| = \max\left\{\max_{1\le j\le p}\sup_{u\in U}\left|\sum_{t=1}^m\left(\frac{\partial\hat{l}_t(u)}{\partial \mathsf{r}_j}-\frac{\partial l_t(u)}{\partial \mathsf{r}_j}\right)\right|, \sup_{u\in U}\left|\sum_{t=1}^m\left(\frac{\partial\hat{l}_t(u)}{\partial u_2}-\frac{\partial l_t(u)}{\partial u_2}\right)\right|\right\}$$

with $u_1 := \left(\mathsf{r}_1',\dots,\mathsf{r}_p'\right)'$ where $\mathsf{r}_j := (x_j,s_j,t_j)'$, $j=1,\dots,p$, and $u_2 := (a,b,q_1,\dots,q_{p^-})'$.

(I) The Proof of $\sup_{u\in U}\left|\sum_{t=1}^{m}\left(\frac{\partial\hat{l}_t(u)}{\partial u_1}-\frac{\partial l_t(u)}{\partial u_1}\right)\right| \overset{a.s.}{=} \mathcal{O}(1)$**.**

Throughout the next sections, we will use the following statement repeatedly. The lemma is a generalisation of Lemma 2.2 in Berkes et al. (2003). Adopting their notation let

$$\log^+ x := \begin{cases} \log x, & x > 1 \\ 0, & \text{else} \end{cases}.$$

Lemma B.2. *Let* $\{X_t,\ t \in \mathbb{N}_0\}$ *be a sequence of identically distributed but not necessarily independent random variables satisfying*

$$\mathsf{E} \log^+ |X_0| < \infty. \tag{B.14}$$

Then, we have that $\sum_{k=0}^{\infty} k^j a^k X_k$ *converges with probability one for any* $a \in \mathbb{R}$ *with* $|a| < 1$ *and any fixed* $j \in \mathbb{N}_0$*.*

Proof: Analogously to the proof of Lemma 2.2 in Berkes et al. (2003), it suffices to show that the conditions for the Borel-Cantelli Lemma are satisfied for all $\zeta > 1$. Note, that the function $f : \mathbb{R}_+ \to \mathbb{R}_+$ with $f(x) = \frac{\zeta^x}{x^j}$ has a minimum located at $x_{\mathsf{min}} = -\frac{j}{\log \zeta}$ and is strictly monotonic increasing for larger values of x. Thus, there exists some constant $\zeta_0 \in (1, \zeta)$ with $\frac{\zeta^k}{k^j} > \zeta_0^k$ for any integer $k \geq k_0$ where k_0 is the smallest integer that is larger than x_{min}. Thus, for our counterpart of (2.5) in Berkes et al. (2003), we have

$$\sum_{k=0}^{\infty} \mathsf{P}\left(|X_k| > \frac{\zeta^k}{k^j}\right) \leq \sum_{k=0}^{k_0-1} \mathsf{P}\left(|X_k| > \frac{\zeta^k}{k^j}\right) + \sum_{k=k_0}^{\infty} \mathsf{P}\left(|X_k| > \frac{\zeta^k}{k^j}\right) \leq k_0 + \sum_{k=0}^{\infty} \mathsf{P}\left(|X_k| > \zeta_0^k\right). \tag{B.15}$$

Along the lines of Berkes et al. (2003), the righthand side of (B.15) is finite if (B.14) is satisfied. Hence, the Borel-Cantelli Lemma yields the almost sure convergence for any nonnegative integer j which completes the proof.

■

Lemma B.3. *Denote* $T_{jt} := \sum_{k=t-1}^{\infty} \rho^k y_{j,t-k-1}^2$*, for* $j = 1, \ldots, p$ *and* $t = 1, \ldots, m$*. Under Assumptions 2.1, 3.1-3.4, 4.3 and 4.4, we have for all* $i, j \in \{1, \ldots, p\}$ *and for* $m \to \infty$*:*

- $\sum_{t=1}^{m} T_{jt} \overset{a.s.}{=} \mathcal{O}(1)$*; and*
- $\left| \sum_{t=1}^{m} T_{it} y_{it} y_{jt} \right| \overset{a.s.}{=} \mathcal{O}(1)$.

Proof: Note that $\left\{ y_{jt}^2, t \in \mathbb{Z} \right\}$ and $\{y_{it} y_{jt}, t \in \mathbb{Z}\}$ are sequences of unconditionally identically distributed random variables. With Assumptions 4.3 and 4.4, we have for all $i, j \in \{1, \ldots, p\}$:

$$\mathsf{E} \log^+ \left(y_{jt}^2 \right) \leq \mathsf{E} \left(y_{jt}^2 \right) < \infty \quad \text{and} \quad \mathsf{E} \log^+ \left(y_{it} y_{jt} \right) \leq \mathsf{E} \log^+ \left(y_{it}^2 y_{jt}^2 \right) \leq \mathsf{E} \left(y_{it}^2 y_{jt}^2 \right) < \infty. \tag{B.16}$$

Hence, Lemma B.2 yields $T_{jt} \overset{a.s.}{=} \mathcal{O}(1)$, $\forall\ j \in \{1, \ldots, p\}$ and $t \in \mathbb{Z}$.

Additionally, for any $i, j \in \{1, \ldots, p\}$ and for $m \to \infty$, we have

$$\sum_{t=1}^{m} T_{jt} = \sum_{t=1}^{m} \sum_{k=t-1}^{\infty} \rho^k y_{j,t-k-1}^2 = \sum_{t=1}^{m} \sum_{l=0}^{\infty} \rho^{l+t-1} y_{jl}^2 = \sum_{t=0}^{m-1} \rho^t \sum_{l=0}^{\infty} \rho^l y_{jl}^2 = \frac{1-\rho^m}{1-\rho} \sum_{l=0}^{\infty} \rho^l y_{jl}^2 \overset{a.s.}{=} \mathcal{O}(1)$$

and

$$\left| \sum_{t=1}^{m} T_{it} y_{it} y_{jt} \right| = \left| \sum_{t=1}^{m} y_{it} y_{jt} \sum_{l=0}^{\infty} \rho^{l+t-1} y_{il}^2 \right| \leq \left| \sum_{t=0}^{m-1} \rho^t y_{i,t+1} y_{j,t+1} \right| \cdot \left| \sum_{l=0}^{\infty} \rho^l y_{il}^2 \right| \overset{a.s.}{=} \mathcal{O}(1)$$

where the last equality is implied by Lemma B.2 in both cases. ■

Lemma B.4. *Under Assumptions 2.1, 3.1-3.4, 4.3 and 4.4, we have for all* $j \in \{1, \ldots, p\}$ *and for* $m \to \infty$*:*

- $\sup_{u \in U} \left| \sum_{t=1}^{m} \left(\widehat{\mathsf{w}}_{jt}(u)^{\frac{1}{2}} - \mathsf{w}_{jt}(u)^{\frac{1}{2}} \right) \right| \overset{a.s}{=} \mathcal{O}(1)$*;*
- $\sup_{u \in U} \left| \sum_{t=1}^{m} \left(\widehat{\mathsf{w}}_{jt}(u)^{-\frac{1}{2}} - \mathsf{w}_{jt}(u)^{-\frac{1}{2}} \right) \right| \overset{a.s}{=} \mathcal{O}(1)$*; and*
- $\sup_{u \in U} \left| \sum_{t=1}^{m} \left[\left(\widehat{\mathsf{w}}_{it}(u) \widehat{\mathsf{w}}_{jt}(u) \right)^{-\frac{1}{2}} - \left(\mathsf{w}_{it}(u) \mathsf{w}_{jt}(u) \right)^{-\frac{1}{2}} \right] \right| \overset{a.s}{=} \mathcal{O}(1)$.

Proof: Analogously to the arguments in the proof of Lemma 5.4 in Berkes et al. (2003), there exist positive constants $C_1 := \frac{u}{1-\underline{u}}$ and $C_2 := \frac{u}{1-\rho}$ with

$$0 < C_1 \leq \widehat{\mathsf{w}}_{jt}(u) \leq \mathsf{w}_{jt}(u) \leq C_2 \left(1 + \sum_{k=0}^{\infty} \rho^k y_{j,t-1}^2\right) \overset{a.s.}{=} \mathcal{O}(1) \quad \forall\, j \in \{1,\dots,p\}\,. \tag{B.17}$$

Since $\widehat{\mathsf{w}}_{jt}(u) = \min\left[\widehat{\mathsf{w}}_{jt}(u), \mathsf{w}_{jt}(u)\right]$, the MVT yields for any $j \in \{1,\dots,p\}$

$$\sup_{u\in U}\left|\sum_{t=1}^{m}\left(\widehat{\mathsf{w}}_{jt}(u)^{\frac{1}{2}} - \mathsf{w}_{jt}(u)^{\frac{1}{2}}\right)\right| \leq \sup_{u\in U}\sum_{t=1}^{m}\frac{1}{2\widehat{\mathsf{w}}_{jt}(u)}\left(\mathsf{w}_{jt}(u) - \widehat{\mathsf{w}}_{jt}(u)\right) \leq \frac{1}{2C_1}\sum_{t=1}^{m} T_{jt} \overset{a.s.}{=} \mathcal{O}(1). \tag{B.18}$$

For the second statement, (B.17) and (B.18) yield

$$\sup_{u\in U}\left|\sum_{t=1}^{m}\left(\widehat{\mathsf{w}}_{jt}(u)^{-\frac{1}{2}} - \mathsf{w}_{jt}(u)^{-\frac{1}{2}}\right)\right| \leq \sup_{u\in U}\sum_{t=1}^{m}\frac{1}{\widehat{\mathsf{w}}_{jt}(u)}\left(\mathsf{w}_{jt}(u)^{\frac{1}{2}} - \widehat{\mathsf{w}}_{jt}(u)^{\frac{1}{2}}\right) \overset{a.s.}{=} \mathcal{O}(1).$$

Finally, with the MVT and (B.17), we have for any $i, j \in \{1,\dots,p\}$

$$\begin{aligned}
&\sup_{u\in U}\sum_{t=1}^{m}\left[\left(\widehat{\mathsf{w}}_{it}(u)\widehat{\mathsf{w}}_{jt}(u)\right)^{-\frac{1}{2}} - \left(\mathsf{w}_{it}(u)\mathsf{w}_{jt}(u)\right)^{-\frac{1}{2}}\right] \\
&= \sup_{u\in U}\sum_{t=1}^{m}\left(\widehat{\mathsf{w}}_{it}(u)\widehat{\mathsf{w}}_{jt}(u)\mathsf{w}_{it}(u)\mathsf{w}_{jt}(u)\right)^{-\frac{1}{2}}\left[\left(\mathsf{w}_{it}(u)\mathsf{w}_{jt}(u)\right)^{\frac{1}{2}} - \left(\widehat{\mathsf{w}}_{it}(u)\widehat{\mathsf{w}}_{jt}(u)\right)^{\frac{1}{2}}\right] \\
&\leq \frac{1}{C_1^2}\sup_{u\in U}\sum_{t=1}^{m}\frac{1}{2}\left(\widehat{\mathsf{w}}_{it}(u)\widehat{\mathsf{w}}_{jt}(u)\right)^{-\frac{1}{2}}\left|\mathsf{w}_{it}(u)\mathsf{w}_{jt}(u) - \widehat{\mathsf{w}}_{it}(u)\widehat{\mathsf{w}}_{jt}(u)\right| \\
&\leq \frac{1}{2C_1^3}\sup_{u\in U}\sum_{t=1}^{m}\left(\left|\mathsf{w}_{it}(u) - \widehat{\mathsf{w}}_{it}(u)\right|\widehat{\mathsf{w}}_{jt}(u) + \left|\mathsf{w}_{jt}(u) - \widehat{\mathsf{w}}_{jt}(u)\right|\mathsf{w}_{it}(u)\right) \\
&\leq \frac{\rho}{2C_1^3}\sup_{u\in U}\sum_{t=1}^{m}\left(T_{it}\widehat{\mathsf{w}}_{jt}(u) + T_{jt}\mathsf{w}_{it}(u)\right) \leq \frac{\rho}{C_1^3}\max_{1\leq j\leq p}\sup_{u\in U, t\in\mathbb{Z}}\mathsf{w}_{jt}(u)\sum_{t=1}^{m} T_{it} \overset{a.s.}{=} \mathcal{O}(1).
\end{aligned}$$

■

Lemma B.5. *Under Assumptions 2.1, 3.1-3.4, 4.3 and 4.4, we have for any* $j \in \{1,\dots,p\}$*:*

- $\sup_{u\in U, t\in\mathbb{Z}}\left|\frac{\partial \mathsf{vec}\left(\mathsf{F}_{D_t}(u)\right)'}{\partial \mathsf{r}_j}\right| \overset{a.s.}{=} \mathcal{O}(1)$*; and*
- $\sup_{u\in U}\left|\sum_{t=1}^{m}\left(\frac{\partial \mathsf{vec}\left(\widehat{\mathsf{F}}_{D_t}(u)\right)'}{\partial \mathsf{r}_j} - \frac{\partial \mathsf{vec}\left(\mathsf{F}_{D_t}(u)\right)'}{\partial \mathsf{r}_j}\right)\right| \overset{a.s.}{=} \mathcal{O}(1).$

Proof: Using (B.17), (B.16) and Lemma B.2, we have for any $j \in \{1, \dots, p\}$ and $m \to \infty$:

- $$\sup_{u\in U, t\in\mathbb{Z}} \left| \frac{\partial \mathsf{vec}\left(\mathsf{F}_{D_t}(u)\right)'}{\partial x_j} \right| \leq \frac{1}{2}\frac{1}{1-\rho} \sup_{u\in U, t\in\mathbb{Z}} \mathsf{w}_{jt}(u)^{-\frac{1}{2}} \overset{a.s.}{=} \mathcal{O}(1) \tag{B.19}$$
- $$\sup_{u\in U, t\in\mathbb{Z}} \left| \frac{\partial \mathsf{vec}\left(\mathsf{F}_{D_t}(u)\right)'}{\partial s_j} \right| \leq \frac{1}{2} \sup_{u\in U, t\in\mathbb{Z}} \mathsf{w}_{jt}(u)^{-\frac{1}{2}} \sup_{u\in U, t\in\mathbb{Z}} \sum_{k=0}^{\infty} \rho^k y_{j,t-k-1}^2 \overset{a.s.}{=} \mathcal{O}(1) \tag{B.20}$$
- $$\sup_{u\in U, t\in\mathbb{Z}} \left| \frac{\partial \mathsf{vec}\left(\mathsf{F}_{D_t}(u)\right)'}{\partial t_j} \right| \leq \frac{1}{2}\frac{\rho}{\underline{u}} \sup_{u\in U, t\in\mathbb{Z}} \mathsf{w}_{jt}(u)^{-\frac{1}{2}} \sup_{u\in U, t\in\mathbb{Z}} \sum_{k=0}^{\infty} k\rho^k y_{j,t-k-1}^2 \overset{a.s.}{=} \mathcal{O}(1). \tag{B.21}$$

Thus, the first statement is an immediate consequence of (B.19)-(B.21).

The second statement follows from (B.22)-(B.24) below. With Lemma B.4, we have

$$\sup_{u\in U} \left| \sum_{t=1}^{m} \left(\frac{\partial \mathsf{vec}\left(\widehat{\mathsf{F}}_{D_t}(u)\right)'}{\partial x_j} - \frac{\partial \mathsf{vec}\left(\mathsf{F}_{D_t}(u)\right)'}{\partial x_j} \right) \right| \leq \frac{1}{2}\frac{1}{1-\rho} \sup_{u\in U} \left| \sum_{t=1}^{m} \left(\widehat{\mathsf{w}}_{jt}(u)^{-\frac{1}{2}} - \mathsf{w}_{jt}(u)^{-\frac{1}{2}} \right) \right| \overset{a.s.}{=} \mathcal{O}(1) \tag{B.22}$$

Furthermore, (B.17) and Lemmas B.3 and B.4 yield

$$\begin{aligned} &\sup_{u\in U} \left| \sum_{t=1}^{m} \left(\frac{\partial \mathsf{vec}\left(\widehat{\mathsf{F}}_{D_t}(u)\right)'}{\partial s_j} - \frac{\partial \mathsf{vec}\left(\mathsf{F}_{D_t}(u)\right)'}{\partial s_j} \right) \right| \\ &\leq \frac{1}{2} \sup_{u\in U} \left| \sum_{t=1}^{m} \left(\widehat{\mathsf{w}}_{jt}(u)^{-\frac{1}{2}} - \mathsf{w}_{jt}(u)^{-\frac{1}{2}} \right) \right| \sup_{u\in U, t\in\mathbb{Z}} \sum_{k=0}^{t-2} t_j^k y_{j,t-k-1}^2 + \frac{1}{2} C_1^{-1/2} \sum_{t=1}^{m} T_{jt} \overset{a.s.}{=} \mathcal{O}(1). \end{aligned} \tag{B.23}$$

Finally, we have

$$\begin{aligned} &\sup_{u\in U} \left| \sum_{t=1}^{m} \left(\frac{\partial \mathsf{vec}\left(\widehat{\mathsf{F}}_{D_t}(u)\right)'}{\partial t_j} - \frac{\partial \mathsf{vec}\left(\mathsf{F}_{D_t}(u)\right)'}{\partial t_j} \right) \right| \\ &\leq \frac{1}{2} C_1^{-3/2} \frac{\rho}{\underline{u}} \sum_{t=1}^{m} \left[\sum_{k=t-1}^{\infty} k\rho^k y_{j,t-k-1}^2 \right] \left[\sum_{k=0}^{t-2} k\rho^k y_{j,t-k-1}^2 \right] + \frac{1}{2} C_1^{-1/2} \frac{\rho}{\underline{u}} \sum_{t=1}^{m} \sum_{k=t-1}^{\infty} k\rho^k y_{j,t-k-1}^2 \end{aligned} \tag{B.24}$$

with

$$\sum_{t=1}^{m} \sum_{k=t-1}^{\infty} k\rho^k y_{j,t-k-1}^2 = \sum_{t=1}^{m} \sum_{l=0}^{\infty} (l+t-1)\rho^{l+t-1} y_{jl}^2 = \sum_{t=0}^{m} \rho^t \sum_{l=0}^{\infty} l\rho^l y_{jl}^2 + \sum_{t=0}^{m} t\rho^t \sum_{l=0}^{\infty} \rho^l y_{jl}^2. \tag{B.25}$$

Since (B.25) is $\mathcal{O}(1)$ almost surely with (B.16) and Lemma B.2, the same applies to (B.24) which completes the proof. ■

Lemma B.6. *Under Assumptions 2.1, 3.1-3.4, 4.3 and 4.4, we have for $m \to \infty$:*

- $\sup_{u \in U} \left| \sum_{t=1}^{m} \left(\widehat{\mathsf{F}}_{D_t}(u) - \mathsf{F}_{D_t}(u) \right) \right| \overset{a.s.}{=} \mathcal{O}(1)$; *and*
- $\sup_{u \in U} \left| \sum_{t=1}^{m} \left(\widehat{\mathsf{F}}_{Q_t}(u) - \mathsf{F}_{Q_t}(u) \right) \right| \overset{a.s.}{=} \mathcal{O}(1)$.

Proof: For the first statement, Lemma B.4 implies

$$\sup_{u \in U} \left| \sum_{t=1}^{m} \left(\widehat{\mathsf{F}}_{D_t}(u) - \mathsf{F}_{D_t}(u) \right) \right| = \max_{1 \le j \le p} \left| \sum_{t=1}^{m} \left(\widehat{\mathsf{w}}_{jt}(u)^{\frac{1}{2}} - \mathsf{w}_{jt}(u)^{\frac{1}{2}} \right) \right| \overset{a.s.}{=} \mathcal{O}(1).$$

For the second statement (B.16), (B.17), and Lemmas B.2, B.3 and B.4 yield

$$\begin{aligned}
&\sup_{u \in U} \left| \sum_{t=1}^{m} \left(\widehat{\mathsf{F}}_{Q_t}(u) - \mathsf{F}_{Q_t}(u) \right) \right| = \sup_{u \in U} a \left| \sum_{t=1}^{m} \sum_{k=1}^{\infty} b^{k-1} \left(\widehat{\mathsf{z}}_t(u) \widehat{\mathsf{z}}_t'(u) - \mathsf{z}_t(u) \mathsf{z}_t'(u) \right) \right| \\
&\le \sup_{u \in U} \max_{1 \le i,j \le p} \rho \left| \sum_{t=1}^{m} \sum_{k=1}^{\infty} \rho^{k-1} y_{i,t-k} y_{j,t-k} \left[(\widehat{\mathsf{w}}_{i,t-k}(u) \widehat{\mathsf{w}}_{j,t-k}(u))^{-\frac{1}{2}} - (\mathsf{w}_{i,t-k}(u) \mathsf{w}_{j,t-k}(u))^{-\frac{1}{2}} \right] \right| \\
&\le \sup_{t \in \mathbb{Z}} \max_{1 \le i,j \le p} \left| \sum_{k=1}^{\infty} \rho^k y_{i,t-k} y_{j,t-k} \right| \sup_{u \in U} \left| \sum_{t=1}^{m} \left[(\widehat{\mathsf{w}}_{i,t-k}(u) \widehat{\mathsf{w}}_{j,t-k}(u))^{-\frac{1}{2}} - (\mathsf{w}_{i,t-k}(u) \mathsf{w}_{j,t-k}(u))^{-\frac{1}{2}} \right] \right| \overset{a.s.}{=} \mathcal{O}(1).
\end{aligned}$$

■

In the following, the proofs for terms with finite and infinite past work analogously and will be omitted for one of these cases.

Lemma B.7. *Under Assumptions 2.1, 3.1-3.4, 4.3 and 4.4, we have for $m \to \infty$:*

- $\sup_{u \in U, t \in \mathbb{Z}} |\mathsf{F}_{D_t}(u)| \overset{a.s.}{=} \mathcal{O}(1)$, $\sup_{u \in U, t \in \mathbb{Z}} |\mathsf{F}_{Q_t}(u)| \overset{a.s.}{=} \mathcal{O}(1)$, *and* $\sup_{u \in U, t \in \mathbb{Z}} |\mathsf{F}_{R_t}(u)| \overset{a.s.}{=} \mathcal{O}(1)$; *and*
- $\sup_{u \in U, t \in \mathbb{Z}} \left| \widehat{\mathsf{F}}_{D_t}(u) \right| \overset{a.s.}{=} \mathcal{O}(1)$, $\sup_{u \in U, t \in \mathbb{Z}} \left| \widehat{\mathsf{F}}_{Q_t}(u) \right| \overset{a.s.}{=} \mathcal{O}(1)$, *and* $\sup_{u \in U, t \in \mathbb{Z}} \left| \widehat{\mathsf{F}}_{R_t}(u) \right| \overset{a.s.}{=} \mathcal{O}(1)$.

Proof: Statement (B.17) and Lemma 2.2 in Berkes et al. (2003) yield

$$\sup_{u \in U, t \in \mathbb{Z}} |\mathsf{F}_{D_t}(u)| = \sup_{u \in U, t \in \mathbb{Z}} \max_{1 \le i \le p} \mathsf{w}_{it}(u)^{1/2} \overset{a.s.}{=} \mathcal{O}(1)$$

and

$$\begin{aligned}
\sup_{u \in U, t \in \mathbb{Z}} |\mathsf{F}_{Q_t}(u)| &\le \sup_{u \in U, t \in \mathbb{Z}} \max_{1 \le i \le p^-} \frac{1-a-b}{1-b} q_i + a \sup_{u \in U, t \in \mathbb{Z}} \left| \sum_{k=0}^{\infty} b^k \mathsf{z}_{t-k-1}(u) \mathsf{z}_{t-k-1}'(u) \right| \\
&\le \frac{1-2\underline{u}}{1-\rho} + \rho \sup_{u \in U, t \in \mathbb{Z}} \max_{1 \le i,j \le p} \left| \sum_{k=0}^{\infty} \rho^k \frac{y_{i,t-k-1}}{\mathsf{w}_{i,t-k-1}(u)^{1/2}} \frac{y_{j,t-k-1}}{\mathsf{w}_{j,t-k-1}(u)^{1/2}} \right| \overset{a.s.}{=} \mathcal{O}(1).
\end{aligned}$$

The last equality is a result of Lemma B.2 and

$$\mathsf{E}\log^+\left(\frac{y_{it}}{\mathsf{w}_{it}(u)^{1/2}}\frac{y_{jt}}{\mathsf{w}_{jt}(u)^{1/2}}\right) \le \mathsf{E}\log^+\left(\frac{y_{it}^2}{\mathsf{w}_{it}(u)^{1/2}}\frac{y_{jt}^2}{\mathsf{w}_{jt}(u)^{1/2}}\right) \le \mathsf{E}\left(\frac{y_{it}^2}{\mathsf{w}_{it}(u)^{1/2}}\frac{y_{jt}^2}{\mathsf{w}_{jt}(u)^{1/2}}\right)$$
$$\le C_1^{-1}\mathsf{E}\left(y_{it}^2 y_{jt}^2\right) < \infty.$$

Concluding, $\mathsf{F}_{R_t}(u)$ is a correlation matrix for all $t \in \mathbb{Z}$ almost surely. Hence, the absolute entries are bounded by 1 almost surely which completes the proof. ■

In the following, denote by S_n the set of all $n!$ permutations of $\{1,\dots,n\}$ and $sgn(\sigma)$ the sign of the permutation σ that indicates whether an even or odd number of pairwise interchanges of neighbouring entries in $(1,\dots,n)$ is necessary to obtain the permutation σ.

Lemma B.8. *Under Assumptions 2.1, 3.1-3.4, 4.3 and 4.4, we have for $m \to \infty$:*

- $\sup_{u\in U, t\in\mathbb{Z}} |\det \mathsf{F}_{Q_t}(u)| \overset{a.s.}{=} \mathcal{O}(1)$;
- $\sup_{u\in U, t\in\mathbb{Z}} \left|\det \widehat{\mathsf{F}}_{Q_t}(u)\right| \overset{a.s.}{=} \mathcal{O}(1)$; *and*
- $\sup_{u\in U, t\in\mathbb{Z}} \left|\sum_{t=1}^{m}\left(\det \widehat{\mathsf{F}}_{Q_t}(u) - \det \mathsf{F}_{Q_t}(u)\right)\right| \overset{a.s.}{=} \mathcal{O}(1)$.

Proof: Note that the determinant of a matrix equals the products of its eigenvalues. Hence, Assumption 2.1-6. implies

$$\sup_{u\in U, t\in\mathbb{Z}} |\det \mathsf{F}_{Q_t}(u)| \le \left[\sup_{u\in U, t\in\mathbb{Z}} \lambda_{\max}\left(\mathsf{F}_{Q_t}(u)\right)\right]^p = \delta_2^p.$$

For the second statement, Lemmas B.6 and B.7 yield

$$\begin{aligned}
&\sup_{u\in U}\left|\sum_{t=1}^{m}\left(\det \widehat{\mathsf{F}}_{Q_t}(u) - \det \mathsf{F}_{Q_t}(u)\right)\right| \\
&= \sup_{u\in U}\left|\sum_{t=1}^{m}\sum_{\sigma\in S_p}\left(\left[\widehat{\mathsf{F}}_{Q_t}(u)\right]_{1\sigma(1)}\cdot\ldots\cdot\left[\widehat{\mathsf{F}}_{Q_t}(u)\right]_{p\sigma(p)} - [\mathsf{F}_{Q_t}(u)]_{1\sigma(1)}\cdot\ldots\cdot[\mathsf{F}_{Q_t}(u)]_{p\sigma(p)}\right)\right| \\
&= \sup_{u\in U}\left|\sum_{t=1}^{m}\sum_{\sigma\in S_p}\sum_{j=1}^{p-1}\left([\mathsf{F}_{Q_t}(u)]_{j\sigma(j)} - \left[\widehat{\mathsf{F}}_{Q_t}(u)\right]_{j\sigma(j)}\right)\prod_{i=1}^{j-1}[\mathsf{F}_{Q_t}(u)]_{i\sigma(i)}\prod_{k=j+1}^{p}\left[\widehat{\mathsf{F}}_{Q_t}(u)\right]_{k\sigma(k)}\right|
\end{aligned}$$

$$\leq p!\,(p-1) \sup_{u\in U, t\in\mathbb{Z}} \left|\sum_{t=1}^{m}\left(\mathsf{F}_{Q_t}(u)-\widehat{\mathsf{F}}_{Q_t}(u)\right)\right| \max_{1\leq j\leq p}\left[\sup_{u\in U,t\in\mathbb{Z}}\left|\mathsf{F}_{Q_t}(u)\right|\right]^{j-1} \max_{1\leq j\leq p}\left[\sup_{u\in U,t\in\mathbb{Z}}\left|\widehat{\mathsf{F}}_{Q_t}(u)\right|\right]^{p-j}$$

$$\overset{a.s.}{=} \mathcal{O}(1).$$

■

Lemma B.9. *Under Assumptions 2.1, 3.1-3.4, 4.3 and 4.4, we have for $m \to \infty$:*

- $\sup_{u\in U,t\in\mathbb{Z}} \det \widehat{\mathsf{F}}_{Q_t^*}(u) \overset{a.s.}{=} \mathcal{O}(1)$; *and*
- $\sup_{u\in U}\left|\sum_{t=1}^{m}\left[\left(\det \widehat{\mathsf{F}}_{Q_t^*}(u)\right)^2-\left(\det \mathsf{F}_{Q_t^*}(u)\right)^2\right]\right| \overset{a.s.}{=} \mathcal{O}(1).$

Proof: For the first statement, we have

$$\sup_{u\in U,t\in\mathbb{Z}} \det \widehat{\mathsf{F}}_{Q_t^*}(u) = \sup_{u\in U,t\in\mathbb{Z}} \prod_{i=1}^{p}\left[\widehat{\mathsf{F}}_{Q_t}(u)\right]_{ii}^{-1/2} \leq \sup_{u\in U}\left(\frac{1-a-b}{1-b}\right)^{-p/2} \leq \left(\frac{1-\underline{u}}{1-\rho}\right)^{p/2} = \mathcal{O}(1).$$

For the second statement, consider the interval $\mathrm{I}_{Q_t^*}(u) := \left[\underline{\xi_t}^*(u), \bar{\xi}_t^*(u)\right]$ with

$$\underline{\xi_t}^*(u) := \min\left\{\det \widehat{\mathsf{F}}_{Q_t^*}(u), \det \mathsf{F}_{Q_t^*}(u)\right\} \quad \text{and} \quad \bar{\xi}_t^*(u) := \max\left\{\det \widehat{F}_{Q_t^*}(u), \det \mathsf{F}_{Q_t^*}(u)\right\}.$$

Hence, with the MVT and for any $t\in\mathbb{Z}$ and $u \in U$, there exists a $\xi_t^*(u)\in \mathrm{I}_{Q_t^*}(u)$ such that

$$\sup_{u\in U}\left|\sum_{t=1}^{m}\left[\left(\det \widehat{\mathsf{F}}_{Q_t^*}(u)\right)^2-\left(\det \mathsf{F}_{Q_t^*}(u)\right)^2\right]\right| = \sup_{u\in U}\left|\sum_{t=1}^{m}2\xi_t^*(u)\left[\det \widehat{\mathsf{F}}_{Q_t^*}(u)-\det \mathsf{F}_{Q_t^*}(u)\right]\right|$$

$$\leq 2\sup_{u\in U,t\in\mathbb{Z}}\bar{\xi}_t^*(u)\sup_{u\in U}\left|\sum_{t=1}^{m}\left(\prod_{i=1}^{p}\left[\widehat{\mathsf{F}}_{Q_t}(u)\right]_{ii}^{-\frac{1}{2}}-\prod_{i=1}^{p}\left[\mathsf{F}_{Q_t}(u)\right]_{ii}^{-\frac{1}{2}}\right)\right|$$

$$= 2\sup_{u\in U,t\in\mathbb{Z}}\bar{\xi}_t^*(u)\sup_{u\in U}\left|\sum_{t=1}^{m}\sum_{j=1}^{p}\left(\left[\widehat{\mathsf{F}}_{Q_t}(u)\right]_{jj}^{-\frac{1}{2}}-\left[\mathsf{F}_{Q_t}(u)\right]_{jj}^{-\frac{1}{2}}\right)\prod_{i=1}^{j-1}\left[\mathsf{F}_{Q_t}(u)\right]_{ii}^{-\frac{1}{2}}\prod_{k=j+1}^{p}\left[\widehat{\mathsf{F}}_{Q_t}(u)\right]_{kk}^{-\frac{1}{2}}\right|$$

$$\overset{MVT}{\leq} 2\sup_{u\in U,t\in\mathbb{Z}}\bar{\xi}_t^*(u)\sup_{u\in U}\sum_{j=1}^{p}\left|\sum_{t=1}^{m}\frac{1}{2}\left[\mathsf{F}_{\bar{Q}}(u)\right]_{jj}^{-\frac{3}{2}}\left(\left[\widehat{\mathsf{F}}_{Q_t}(u)\right]_{jj}-\left[\mathsf{F}_{Q_t}(u)\right]_{jj}\right)\right|\sup_{u\in U}\left(\frac{1-a-b}{1-b}\right)^{-\frac{p-1}{2}}$$

$$\leq \left(\frac{1-\underline{u}}{1-\rho}\right)^{\frac{p-1}{2}}\sup_{u\in U,t\in\mathbb{Z}}\bar{\xi}_t^*(u)\; p\sup_{u\in U}\left|\sum_{t=1}^{m}\left(\widehat{\mathsf{F}}_{Q_t}(u)-\mathsf{F}_{Q_t}(u)\right)\right| \overset{a.s.}{=} \mathcal{O}(1)$$

with Lemmas B.6, B.9 and (2.8).

■

Lemma B.10. *Under Assumptions 2.1, 3.1-3.4, 4.3 and 4.4, we have for $m \to \infty$:*

- $\sup\limits_{u \in U, t \in \mathbb{Z}} \frac{1}{\det \mathsf{F}_{R_t}(u)} \overset{a.s.}{=} \mathcal{O}(1)$,
- $\sup\limits_{u \in U, t \in \mathbb{Z}} \frac{1}{\det \widehat{\mathsf{F}}_{R_t}(u)} \overset{a.s.}{=} \mathcal{O}(1)$; *and*
- $\sup\limits_{u \in U, t \in \mathbb{Z}} \left| \sum\limits_{t=1}^{m} \left(\frac{1}{\det \widehat{\mathsf{F}}_{R_t}(u)} - \frac{1}{\det \mathsf{F}_{R_t}(u)} \right) \right| \overset{a.s.}{=} \mathcal{O}(1)$.

Proof: For the inverted determinants, statement (2.9) yields

$$\begin{aligned}
\sup_{u \in U, t \in \mathbb{Z}} \det \mathsf{F}_{R_t}(u) \geq \left(\sup_{u \in U, t \in \mathbb{Z}} \lambda_{\min} \left(\mathsf{F}_{R_t}(u) \right) \right)^p &\overset{a.s.}{>} \left[\frac{1-a-b}{1-b} \frac{\delta_1}{p\delta_2} \right]^p \\
\Rightarrow \sup_{u \in U, t \in \mathbb{Z}} \frac{1}{\det \mathsf{F}_{R_t}(u)} &\overset{a.s.}{<} \left[\frac{1-b}{1-a-b} \frac{p\delta_2}{\delta_1} \right]^p = \mathcal{O}(1).
\end{aligned}$$

For the last statement of the lemma, consider $\mathrm{I}_{R_t}(u) := \left[\underline{\xi_t}(u), \bar{\xi}_t(u) \right]$ with

$$\underline{\xi_t}(u) := \min \left\{ \det \widehat{\mathsf{F}}_{R_t}(u), \det \mathsf{F}_{R_t}(u) \right\} \quad \text{and} \quad \bar{\xi}_t(u) := \max \left\{ \det \widehat{\mathsf{F}}_{R_t}(u), \det \mathsf{F}_{R_t}(u) \right\}.$$

Thus, with the MVT and for any $t \in \mathbb{Z}$ and $u \in U$, there exists a $\xi_t(u) \in \mathbb{I}_{R_t}(u)$ such that

$$\begin{aligned}
&\sup_{u \in U} \left| \sum_{t=1}^{m} \left(\frac{1}{\det \widehat{\mathsf{F}}_{R_t}(u)} - \frac{1}{\det \mathsf{F}_{R_t}(u)} \right) \right| \leq \sup_{u \in U, t \in \mathbb{Z}} \left| - \frac{1}{\xi_t^2(u)} \right| \sup_{u \in U} \left| \sum_{t=1}^{m} \left(\det \widehat{\mathsf{F}}_{R_t}(u) - \det \mathsf{F}_{R_t}(u) \right) \right| \\
&\leq \sup_{u \in U, t \in \mathbb{Z}} \frac{1}{\underline{\xi_t}^2(u)} \sup_{u \in U} \left| \sum_{t=1}^{m} \left(\left[\det \widehat{\mathsf{F}}_{Q_t^*}(u) \right]^2 \det \widehat{\mathsf{F}}_{Q_t}(u) - \left[\det \mathsf{F}_{Q_t^*}(u) \right]^2 \det \mathsf{F}_{Q_t}(u) \right) \right| \\
&\leq \sup_{u \in U, t \in \mathbb{Z}} \frac{1}{\underline{\xi_t}^2(u)} \sup_{u \in U, t \in \mathbb{Z}} \det \widehat{\mathsf{F}}_{Q_t}(u) \sup_{u \in U} \left| \sum_{t=1}^{m} \left(\left[\det \widehat{\mathsf{F}}_{Q_t^*}(u) \right]^2 - \left[\det \mathsf{F}_{Q_t^*}(u) \right]^2 \right) \right| \\
&\quad + \sup_{u \in U, t \in \mathbb{Z}} \frac{1}{\underline{\xi_t}^2(u)} \left[\sup_{u \in U, t \in \mathbb{Z}} \det \mathsf{F}_{Q_t^*}(u) \right]^2 \sup_{u \in U} \left| \sum_{t=1}^{m} \left(\det \widehat{\mathsf{F}}_{Q_t}(u) - \det \mathsf{F}_{Q_t}(u) \right) \right| \overset{a.s.}{=} \mathcal{O}(1)
\end{aligned}$$

with Lemmas B.8 and B.9. ■

Denote by $X^{(i,j)}$ the matrix that results from $X \sim (n \times n)$ by omitting the i-th row and the j-th column with $1 \leq i, j \leq n$.

Lemma B.11. *Under Assumptions 2.1, 3.1-3.4, 4.3 and 4.4, we have for $m \to \infty$:*

- $\sup_{u \in U, t \in \mathbb{Z}} \mathsf{det}\ \mathsf{F}_{R_t}(u)^{(i,j)} \overset{a.s.}{=} \mathcal{O}(1)$; *and*
- $\sup_{u \in U} \left| \sum_{t=1}^{m} \left(\mathsf{det}\ \widehat{\mathsf{F}}_{R_t}(u)^{(i,j)} - \mathsf{det}\ \mathsf{F}_{R_t}(u)^{(i,j)} \right) \right| \overset{a.s.}{=} \mathcal{O}(1)$.

Proof: The proof works analogously to the proof of Lemma B.8 and uses the arguments in the proof of Lemma B.10.

Lemma B.12. *Under Assumptions 2.1, 3.1-3.4, 4.3 and 4.4, we have for $m \to \infty$:*

- $\sup_{u \in U, t \in \mathbb{Z}} \left| \sum_{t=1}^{m} \left(\widehat{\mathsf{F}}_{D_t}(u)^{-1} - \mathsf{F}_{D_t}(u)^{-1} \right) \right| \overset{a.s.}{=} \mathcal{O}(1)$; *and*
- $\sup_{u \in U, t \in \mathbb{Z}} \left| \sum_{t=1}^{m} \left(\widehat{\mathsf{F}}_{R_t}(u)^{-1} - \mathsf{F}_{R_t}(u)^{-1} \right) \right| \overset{a.s.}{=} \mathcal{O}(1)$.

Proof: With Lemmas B.4, B.10 and B.11, we have:

$$\sup_{u \in U, t \in \mathbb{Z}} \left| \sum_{t=1}^{m} \left(\widehat{\mathsf{F}}_{D_t}(u)^{-1} - \mathsf{F}_{D_t}(u)^{-1} \right) \right| = \sup_{u \in U, t \in \mathbb{Z}} \max_{1 \le j \le p} \left| \sum_{t=1}^{m} \left(\widehat{\mathsf{w}}_{jt}(u)^{-\frac{1}{2}} - \mathsf{w}_{jt}(u)^{-\frac{1}{2}} \right) \right| \overset{a.s.}{=} \mathcal{O}(1)$$

and

$$\begin{aligned} &\sup_{u \in U, t \in \mathbb{Z}} \left| \sum_{t=1}^{m} \left(\widehat{\mathsf{F}}_{R_t}(u)^{-1} - \mathsf{F}_{R_t}(u)^{-1} \right) \right| \\ &= \sup_{u \in U, t \in \mathbb{Z}} \max_{1 \le i,j \le p} \left| \sum_{t=1}^{m} \left(\frac{1}{\mathsf{det}\ \widehat{\mathsf{F}}_{R_t}(u)} \mathsf{det}\ \widehat{\mathsf{F}}_{R_t}(u)^{(i,j)} - \frac{1}{\mathsf{det}\ \mathsf{F}_{R_t}(u)} \mathsf{det}\ \mathsf{F}_{R_t}(u)^{(i,j)} \right) \right| \\ &\le \sup_{u \in U, t \in \mathbb{Z}} \left| \sum_{t=1}^{m} \left(\frac{1}{\mathsf{det}\ \widehat{\mathsf{F}}_{R_t}(u)} - \frac{1}{\mathsf{det}\ \mathsf{F}_{R_t}(u)} \right) \right| \max_{1 \le i,j \le p} \sup_{u \in U, t \in \mathbb{Z}} \left| \mathsf{det}\ \widehat{\mathsf{F}}_{R_t}(u)^{(i,j)} \right| \\ &\quad + \sup_{u \in U, t \in \mathbb{Z}} \left| \frac{1}{\mathsf{det}\ \mathsf{F}_{R_t}(u)} \right| \max_{1 \le i,j \le p} \sup_{u \in U} \left| \sum_{t=1}^{m} \left(\mathsf{det}\ \widehat{\mathsf{F}}_{R_t}(u)^{(i,j)} - \mathsf{det}\ \mathsf{F}_{R_t}(u)^{(i,j)} \right) \right| \overset{a.s.}{=} \mathcal{O}(1) \end{aligned}$$

■

Lemma B.13. *Under Assumptions 2.1, 3.1-3.4, 4.3 and 4.4, we have for $m \to \infty$:*

- $\sup_{u \in U, t \in \mathbb{Z}} \left| \mathsf{F}_{D_t}(u)^{1/2} \right| \overset{a.s.}{=} \mathcal{O}(1)$, $\sup_{u \in U, t \in \mathbb{Z}} \left| \mathsf{F}_{Q_t}(u)^{1/2} \right| \overset{a.s.}{=} \mathcal{O}(1)$, *and* $\sup_{u \in U, t \in \mathbb{Z}} \left| \mathsf{F}_{Q_t^*}(u)^{1/2} \right| \overset{a.s.}{=} \mathcal{O}(1)$;
- $\sup_{u \in U, t \in \mathbb{Z}} \left| \widehat{\mathsf{F}}_{D_t}(u)^{1/2} \right| \overset{a.s.}{=} \mathcal{O}(1)$, $\sup_{u \in U, t \in \mathbb{Z}} \left| \widehat{\mathsf{F}}_{Q_t}(u)^{1/2} \right| \overset{a.s.}{=} \mathcal{O}(1)$, *and* $\sup_{u \in U, t \in \mathbb{Z}} \left| \widehat{\mathsf{F}}_{Q_t^*}(u)^{1/2} \right| \overset{a.s.}{=} \mathcal{O}(1)$.

Proof: First of all, (B.17) implies $\sup_{u \in U, t \in \mathbb{Z}} \left| \mathsf{F}_{D_t}(u)^{1/2} \right| = \sup_{u \in U, t \in \mathbb{Z}} \max_{1 \le i \le p} \mathsf{w}_{it}(u)^{1/4} \overset{a.s.}{=} \mathcal{O}(1)$.

Furthermore, consider the eigenvalue decomposition of $\mathsf{F}_{Q_t}(u)^{1/2}$ with $U_t(u)$ the matrix whose columns are the orthonormalized eigenvectors that belong to the ordered eigenvalues of $\mathsf{F}_{Q_t}(u)$ which form the main diagonal of the diagonal matrix $\Lambda_t(u)$. Note that due to the normalization we have $\sup_{u\in U, t\in\mathbb{Z}} |U_t(u)| \le 1$ and $\sup_{u\in U, t\in\mathbb{Z}} |\mathsf{F}_{Q_t}(u)| \overset{a.s.}{=} \mathcal{O}(1)$ which implies $\lambda_{\max}(\mathsf{F}_{Q_t}(u)) \overset{a.s.}{=} \mathcal{O}(1)$. These statements yield

$$\sup_{u\in U,t\in\mathbb{Z}} \left|\mathsf{F}_{Q_t}(u)^{\frac{1}{2}}\right| = \sup_{u\in U,t\in\mathbb{Z}} \left|U_t(u)\Lambda_t(u)^{\frac{1}{2}}U_t(u)'\right| \le 2p^2 \sup_{u\in U,t\in\mathbb{Z}} |U_t(u)| \sup_{u\in U,t\in\mathbb{Z}} \lambda_{\max}(\mathsf{F}_{Q_t}(u))^{\frac{1}{2}} = \mathcal{O}(1)$$

and $$\sup_{u\in U,t\in\mathbb{Z}} \left|\mathsf{F}_{Q_t^*}(u)^{\frac{1}{2}}\right| \overset{a.s.}{\le} \sup_{u\in U,t\in\mathbb{Z}} \left|\frac{1-a-b}{1-b}\min_{1\le i\le p}\left[\mathsf{F}_{\bar{Q}}(u)\right]_{ii}\right|^{-\frac{1}{4}} \overset{a.s.}{\le} \left(\frac{1-\underline{u}}{1-2\rho}\right)^{\frac{1}{4}} = \mathcal{O}(1).$$ ■

Lemma B.14. *Under Assumptions 2.1, 3.1-3.4, 4.3 and 4.4, we have for $m\to\infty$:*

- $\sup_{u\in U,t\in\mathbb{Z}} \left|\mathsf{F}_{D_t}(u)^{-1}\right| \overset{a.s.}{=} \mathcal{O}(1)$ *and* $\sup_{u\in U,t\in\mathbb{Z}} \left|\mathsf{F}_{R_t}(u)^{-1}\right| \overset{a.s.}{=} \mathcal{O}(1)$*; and*
- $\sup_{u\in U,t\in\mathbb{Z}} \left|\widehat{\mathsf{F}}_{D_t}(u)^{-1}\right| \overset{a.s.}{=} \mathcal{O}(1)$ *and* $\sup_{u\in U,t\in\mathbb{Z}} \left|\widehat{\mathsf{F}}_{R_t}(u)^{-1}\right| \overset{a.s.}{=} \mathcal{O}(1)$.

Proof: Note that (B.17) implies $\sup_{u\in U,t\in\mathbb{Z}} \left|\mathsf{F}_{D_t}(u)^{-1}\right| = \sup_{u\in U,t\in\mathbb{Z}} \max_{1\le j\le p} \mathsf{w}_{jt}(u)^{-1/2} \le C_1^{-1/2} \overset{a.s.}{=} \mathcal{O}(1)$. To prove $\sup_{u\in U,t\in\mathbb{Z}} \left|\mathsf{F}_{R_t}(u)^{-1}\right| \overset{a.s.}{=} \mathcal{O}(1)$, we investigate the matrix $\mathsf{F}_{R_t}(u)^{-1}$ in detail. For this purpose, keep in mind that $\mathsf{F}_{R_t}(u)^{-1} = [\det \mathsf{F}_{R_t}(u)]^{-1}\,\mathsf{A}_t(u)$ with $\mathsf{A}_t(u) := (\mathsf{a}_{ij,t}(u))_{i,j=1,\dots,p}$ the adjoint matrix and $\mathsf{a}_{ij,t}(u)$ the cofactor of $[\mathsf{F}_{R_t}(u)]_{ij}$ which is defined as $\mathsf{a}_{ij,t}(u) := (-1)^{i+j}M_{ij,t}(u)$ with $M_{ij,t}(u) := \det \mathsf{F}_{R_t}(u)^{(i,j)}$ the minor of $[\mathsf{F}_{R_t}(u)]_{ij}$. Since the entries of $\mathsf{F}_{R_t}(u)$ do not exceed one in modulus, $M_{ij,t}(u)$ is bounded by a constant:

$$\sup_{u\in U}\max_{1\le i,j\le p} |M_{ij,t}(u)|\,| \le \sup_{u\in U}\max_{1\le i,j\le p} \sum_{\sigma\in S_{p-1}} \prod_{k=1}^{p-1} \left|\left[\mathsf{F}_{R_t}(u)^{(i,j)}\right]_{k\sigma(k)}\right| < (p-1)!\,. \tag{B.26}$$

Thus, analogously to the argumentation in Lemma B.10, (2.9) and (B.26) yield

$$\sup_{u\in U,t\in\mathbb{Z}} \left|\mathsf{F}_{R_t}(u)^{-1}\right| \le \sup_{u\in U,t\in\mathbb{Z}} \left|\frac{1}{\det \mathsf{F}_{R_t}(u)}\right| \max_{1\le i,j\le p} \sup_{u\in U,t\in\mathbb{Z}} \left|\det \mathsf{F}_{R_t}(u)^{(i,j)}\right| \overset{a.s.}{<} \left[\frac{1-b}{1-a-b}\frac{p\delta_2}{\delta_1}\right]^p (p-1)!$$

which completes the proof. ■

Lemma B.15. *Under Assumptions 2.1, 3.1-3.4, 4.3 and 4.4, we have for $m \to \infty$:*

- $\sup_{u\in U, t\in\mathbb{Z}} \left|\widehat{\mathsf{z}}_t(u)\widehat{\mathsf{z}}_t'(u)\right| \overset{a.s.}{=} \mathcal{O}(1)$; *and*

- $\sup_{u\in U} \left|\sum_{t=1}^m \left(\widehat{\mathsf{z}}_t(u)\widehat{\mathsf{z}}_t'(u) - \mathsf{z}_t(u)\mathsf{z}_t'(u)\right)\right| \overset{a.s.}{=} \mathcal{O}(1)$.

Proof: With $\widehat{\mathsf{F}}_{H_t}(u) := \widehat{\mathsf{F}}_{D_t}(u)\widehat{\mathsf{F}}_{R_t}(u)\widehat{\mathsf{F}}_{D_t}(u)$ and Assumption 4.3, the Lemmas B.13 and B.14 yield

$$\begin{aligned}
&\sup_{u\in U,t\in\mathbb{Z}} \left|\widehat{\mathsf{z}}_t(u)\widehat{\mathsf{z}}_t'(u)\right| \le p^2 \sup_{u\in U,t\in\mathbb{Z}} \left|\widehat{\mathsf{F}}_{D_t}(u)^{-1}\right| \sup_{u\in U,t\in\mathbb{Z}} \left|\widehat{\mathsf{F}}_{H_t}(u)^{1/2}\epsilon_t\epsilon_t'\widehat{\mathsf{F}}_{H_t}(u)^{1/2}\right| \sup_{u\in U,t\in\mathbb{Z}} \left|\widehat{\mathsf{F}}_{D_t}(u)^{-1}\right| \\
&\le p^4 \left(\sup_{u\in U,t\in\mathbb{Z}} \left|\widehat{\mathsf{F}}_{D_t}(u)^{-1}\right|\right)^2 \left(\sup_{u\in U,t\in\mathbb{Z}} \left|\widehat{\mathsf{F}}_{D_t}(u)^{1/2}\right|\right)^4 \left(\sup_{u\in U,t\in\mathbb{Z}} \left|\widehat{\mathsf{F}}_{R_t}(u)^{1/2}\right|\right)^2 \sup_{t\in\mathbb{Z}} \left|\epsilon_t\epsilon_t'\right| \\
&\le p^6 \left(\sup_{u\in U,t\in\mathbb{Z}} \left|\widehat{\mathsf{F}}_{D_t}(u)^{-1}\right|\right)^2 \left(\sup_{u\in U,t\in\mathbb{Z}} \left|\widehat{\mathsf{F}}_{D_t}(u)^{1/2}\right|\right)^4 \left(\sup_{u\in U} \left|\widehat{\mathsf{F}}_{Q_t^*}(u)^{1/2}\right|\right)^4 \left(\sup_{u\in U,t\in\mathbb{Z}} \left|\widehat{\mathsf{F}}_{Q_t}(u)^{1/2}\right|\right)^2 \\
&\qquad \times \sup_{t\in\mathbb{Z}} \left|\epsilon_t\epsilon_t'\right| \overset{a.s.}{=} \mathcal{O}(1).
\end{aligned}$$

Furthermore, with (B.17) and Lemma B.3, we have

$$\begin{aligned}
&\sup_{u\in U} \left|\sum_{t=1}^m \left(\widehat{\mathsf{z}}_t(u)\widehat{\mathsf{z}}_t'(u) - \mathsf{z}_t(u)\mathsf{z}_t'(u)\right)\right| = \sup_{u\in U} \max_{1\le i,j\le p} \left|\sum_{t=1}^m y_{it}y_{jt} \left[(\widehat{\mathsf{w}}_{it}(u)\widehat{\mathsf{w}}_{jt}(u))^{-\frac{1}{2}} - (\mathsf{w}_{it}(u)\mathsf{w}_{jt}(u))^{-\frac{1}{2}}\right]\right| \\
&\le \frac{\rho}{2C_1^3} \sup_{u\in U} \max_{1\le i,j\le p} \left|\sum_{t=1}^m \left(T_{it}\widehat{\mathsf{w}}_{jt}(u) + T_{jt}\mathsf{w}_{it}(u)\right) y_{it}y_{jt}\right| \\
&\le \frac{\rho}{2C_1^3} \max_{1\le i\le p} \sup_{u\in U,t\in\mathbb{Z}} \mathsf{w}_{it}(u) \max_{1\le i,j\le p} \left(\left|\sum_{t=1}^m T_{it}y_{it}y_{jt}\right| + \left|\sum_{t=1}^m T_{jt}y_{it}y_{jt}\right|\right) \overset{a.s.}{=} \mathcal{O}(1).
\end{aligned}$$

■

Denote $K_{1t}(u) := \widehat{\mathsf{F}}_{D_t}(u)^{-1} + \widehat{\mathsf{F}}_{D_t}(u)^{-1}\widehat{\mathsf{F}}_{R_t}(u)^{-1}\widehat{\mathsf{z}}_t(u)\widehat{\mathsf{z}}_t'(u)$ and analogously $\widehat{K_{1t}}(u)$ in dependence of a finite past of observations.

Lemma B.16. *Under Assumptions 2.1, 3.1-3.4, 4.3 and 4.4, we have for $m \to \infty$:*

- $\sup_{u\in U, t\in\mathbb{Z}} \left|\widehat{K_{1t}}(u)\right| \overset{a.s.}{=} \mathcal{O}(1)$; *and*

- $\sup_{u\in U} \left|\sum_{t=1}^m \left(\widehat{K_{1t}}(u) - K_{1t}(u)\right)\right| \overset{a.s.}{=} \mathcal{O}(1)$.

Proof: The Lemmas B.12, B.14 and B.15 imply

$$\sup_{u\in U,t\in\mathbb{Z}}\left|\widehat{K}_{1t}(u)\right|$$

$$\leq \sup_{u\in U,t\in\mathbb{Z}}\left|\widehat{\mathsf{F}}_{D_t}(u)^{-1}\right| + p^2 \sup_{u\in U,t\in\mathbb{Z}}\left|\widehat{\mathsf{F}}_{D_t}(u)^{-1}\right| \sup_{u\in U,t\in\mathbb{Z}}\left|\widehat{\mathsf{F}}_{R_t}(u)^{-1}\right| \sup_{u\in U,t\in\mathbb{Z}}\left|\widehat{\mathsf{z}}_t(u)\widehat{\mathsf{z}}_t'(u)\right| \overset{a.s.}{=} \mathcal{O}(1)$$

$$\text{and} \quad \sup_{u\in U}\left|\sum_{t=1}^{m}\left(\widehat{K}_{1t}(u) - K_{1t}(u)\right)\right|$$

$$\leq p^2 \sup_{u\in U,t\in\mathbb{Z}}\left|\widehat{\mathsf{F}}_{R_t}(u)^{-1}\right| \sup_{u\in U,t\in\mathbb{Z}}\left|\widehat{\mathsf{z}}_t(u)\widehat{\mathsf{z}}_t'(u)\right| \sup_{u\in U}\left|\sum_{t=1}^{m}\left(\widehat{\mathsf{F}}_{D_t}(u)^{-1} - \mathsf{F}_{D_t}(u)^{-1}\right)\right|$$

$$+ p^2 \sup_{u\in U,t\in\mathbb{Z}}\left|\mathsf{F}_{D_t}(u)^{-1}\right| \sup_{u\in U,t\in\mathbb{Z}}\left|\widehat{\mathsf{z}}_t(u)\widehat{\mathsf{z}}_t'(u)\right| \sup_{u\in U}\left|\sum_{t=1}^{m}\left(\widehat{\mathsf{F}}_{R_t}(u)^{-1} - \mathsf{F}_{R_t}(u)^{-1}\right)\right|$$

$$+ p^2 \sup_{u\in U,t\in\mathbb{Z}}\left|\mathsf{F}_{D_t}(u)^{-1}\right| \sup_{u\in U,t\in\mathbb{Z}}\left|\mathsf{F}_{R_t}(u)^{-1}\right| \sup_{u\in U}\left|\sum_{t=1}^{m}\left(\widehat{\mathsf{z}}_t(u)\widehat{\mathsf{z}}_t'(u) - \mathsf{z}_t(u)\mathsf{z}_t'(u)\right)\right| \overset{a.s.}{=} \mathcal{O}(1).$$

■

Lemma B.17. *Under Assumptions 2.1, 3.1-3.4, 4.3 and 4.4, we have for all* $j \in \{1,\ldots,p\}$ *and* $m \to \infty$*:*

- $\sup_{u\in U,t\in\mathbb{Z}}\left|\frac{\partial l_t(u)}{\partial \mathsf{r}_j}\right| \overset{a.s.}{=} \mathcal{O}(1)$*; and*
- $\sup_{u\in U}\left|\sum_{t=1}^{m}\left(\frac{\partial \hat{l}_t(u)}{\partial \mathsf{r}_j} - \frac{\partial l_t(u)}{\partial \mathsf{r}_j}\right)\right| \overset{a.s.}{=} \mathcal{O}(1).$

Proof: The Lemmas B.5, B.12 and B.16 imply

$$\sup_{u\in U,t\in\mathbb{Z}}\left|\frac{\partial l_t(u)}{\partial \mathsf{r}_j}\right| \leq \sup_{u\in U,t\in\mathbb{Z}} \frac{\partial \mathsf{vec}\left(\mathsf{F}_{D_t}(u)\right)'}{\partial \mathsf{r}_j} \sup_{u\in U,t\in\mathbb{Z}} K_{1t}(u) \overset{a.s.}{=} \mathcal{O}(1)$$

$$\text{and} \quad \sup_{u\in U}\left|\sum_{t=1}^{m}\left(\frac{\partial \hat{l}_t(u)}{\partial \mathsf{r}_j} - \frac{\partial l_t(u)}{\partial \mathsf{r}_j}\right)\right| = \sup_{u\in U}\left|\sum_{t=1}^{m}\left(\frac{\partial \mathsf{vec}\left(\widehat{\mathsf{F}}_{D_t}(u)\right)'}{\partial \mathsf{r}_j}\widehat{K}_{1t}(u) - \frac{\partial \mathsf{vec}\left(\mathsf{F}_{D_t}(u)\right)'}{\partial \mathsf{r}_j}K_{1t}(u)\right)\right|$$

$$\leq p \sup_{u\in U,t\in\mathbb{Z}}\left|\widehat{K}_{1t}(u)\right| \sup_{u\in U}\left|\sum_{t=1}^{m}\left(\frac{\partial \mathsf{vec}\left(\widehat{\mathsf{F}}_{D_t}(u)\right)'}{\partial \mathsf{r}_j} - \frac{\partial \mathsf{vec}\left(\mathsf{F}_{D_t}(u)\right)'}{\partial \mathsf{r}_j}\right)\right|$$

$$+ p \sup_{u\in U,t\in\mathbb{Z}}\left|\frac{\partial \mathsf{vec}\left(\mathsf{F}_{D_t}(u)\right)'}{\partial \mathsf{r}_j}\right| \sup_{u\in U}\left|\sum_{t=1}^{m}\left(\widehat{K}_{1t}(u) - K_{1t}(u)\right)\right|$$

$$+ p \sup_{u\in U,t\in\mathbb{Z}}\left|\frac{\partial \mathsf{vec}\left(\mathsf{F}_{D_t}(u)\right)'}{\partial \mathsf{r}_j}\right| \sup_{u\in U}\left|\sum_{t=1}^{m}\left(\widehat{\mathsf{F}}_{D_t}(u)^{-1} - \mathsf{F}_{D_t}(u)^{-1}\right)\right| \overset{a.s.}{=} \mathcal{O}(1).$$

$\blacksquare$

(II) The proof of $\sup_{u\in U}\left|\sum_{t=1}^{m}\left(\frac{\partial \hat{l}_t(u)}{\partial u_2}-\frac{\partial l_t(u)}{\partial u_2}\right)\right| \overset{a.s.}{=} \mathcal{O}(1)$.

Denote $K_{2t}(u) := \mathsf{F}_{R_t}(u)^{-1} - \mathsf{F}_{R_t}(u)^{-1}\mathsf{z}_t(u)\mathsf{z}_t'(u)\mathsf{F}_{R_t}(u)^{-1}$ and analogously $\widehat{K}_{2t}(u)$ in dependence of a finite past of observations.

Lemma B.18. *Under Assumptions 2.1, 3.1-3.4, 4.3 and 4.4, we have for $m\to\infty$:*

- $\sup_{u\in U,t\in\mathbb{Z}} |K_{2t}(u)| \overset{a.s.}{=} \mathcal{O}(1)$*; and*
- $\sup_{u\in U}\left|\sum_{t=1}^{m}\left(\widehat{K}_{2t}(u)-K_{2t}(u)\right)\right| \overset{a.s.}{=} \mathcal{O}(1).$

Proof: With Lemmas B.12, B.14 and B.15, we have

$$\begin{aligned}
&\sup_{u\in U,t\in\mathbb{Z}} |K_{2t}(u)| \\
&\quad\le \sup_{u\in U,t\in\mathbb{Z}}\left|\mathsf{F}_{R_t}(u)^{-1}\right| - p^2 \sup_{u\in U,t\in\mathbb{Z}}\left|\mathsf{F}_{R_t}(u)^{-1}\right| \sup_{u\in U,t\in\mathbb{Z}}\left|\mathsf{z}_t(u)\mathsf{z}_t'(u)\right| \sup_{u\in U,t\in\mathbb{Z}}\left|\mathsf{F}_{R_t}(u)^{-1}\right| \overset{a.s.}{=} \mathcal{O}(1). \\
&\text{and}\quad \sup_{u\in U}\left|\sum_{t=1}^{m}\left(\widehat{K}_{2t}(u)-K_{2t}(u)\right)\right| \\
&\quad\le \sup_{u\in U,t\in\mathbb{Z}}\left|\widehat{\mathsf{z}}_t(u)\widehat{\mathsf{z}}_t'(u)\right|\left(\sup_{u\in U,t\in\mathbb{Z}}\left|\widehat{\mathsf{F}}_{R_t}(u)^{-1}\right| + \sup_{u\in U,t\in\mathbb{Z}}\left|\mathsf{F}_{R_t}(u)^{-1}\right|\right)\sup_{u\in U}\left|\sum_{t=1}^{m}\left(\widehat{\mathsf{F}}_{R_t}(u)^{-1}-\mathsf{F}_{R_t}(u)^{-1}\right)\right| \\
&\qquad + \sup_{u\in U,t\in\mathbb{Z}}\left|\widehat{\mathsf{F}}_{R_t}(u)^{-1}\right| \sup_{u\in U,t\in\mathbb{Z}}\left|\mathsf{F}_{R_t}(u)^{-1}\right| \sup_{u\in U}\left|\sum_{t=1}^{m}\left(\widehat{\mathsf{z}}_t(u)\widehat{\mathsf{z}}_t'(u)-\mathsf{z}_t(u)\mathsf{z}_t'(u)\right)\right| \overset{a.s.}{=} \mathcal{O}(1).
\end{aligned}$$

$\blacksquare$

Lemma B.19. *Under Assumptions 2.1, 3.1-3.4, 4.3 and 4.4, we have for $m\to\infty$:*

- $\sup_{u\in U,t\in\mathbb{Z}}\left|\frac{\partial \mathsf{vech}\left(\mathsf{F}_{Q_t}(u)\right)'}{\partial u_2}\right| \overset{a.s.}{=} \mathcal{O}(1)$*; and*
- $\sup_{u\in U}\left|\sum_{t=1}^{m}\left(\frac{\partial \mathsf{vech}\left(\mathsf{F}_{Q_t}(u)\right)'}{\partial u_2}-\frac{\partial \mathsf{vech}\left(\widehat{\mathsf{F}}_{Q_t}(u)\right)'}{\partial u_2}\right)\right| \overset{a.s.}{=} \mathcal{O}(1).$

Proof: With $\sup_{u\in U}\left|\mathsf{F}_{\bar{Q}}(u)\right| \overset{a.s.}{\le} 1$, we have

- $$\sup_{u\in U,t\in\mathbb{Z}}\left|\frac{\partial \mathsf{vech}\left(\mathsf{F}_{Q_t}(u)\right)'}{\partial \mathsf{vecl}\left(\mathsf{F}_{\bar{Q}}(u)\right)}\right| = \frac{1-a-b}{1-b} = \mathcal{O}(1)$$

- $\sup_{u\in U,t\in\mathbb{Z}} \left| \frac{\partial \mathsf{vech}\left(\mathsf{F}_{Q_t}(u)\right)'}{\partial a} \right| \leq \frac{1}{1-\rho} + \sup_{u\in U,t\in\mathbb{Z}} \left| \sum_{n=0}^{\infty} \rho^n \mathsf{z}_{t-n-1}(u)\mathsf{z}'_{t-n-1}(u) \right| \overset{a.s.}{=} \mathcal{O}(1)$
- $\sup_{u\in U,t\in\mathbb{Z}} \left| \frac{\partial \mathsf{vech}\left(\mathsf{F}_{Q_t}(u)\right)'}{\partial b} \right| \leq \frac{\overline{u}}{(1-\rho)^2} + \sup_{u\in U,t\in\mathbb{Z}} \left| \sum_{n=0}^{\infty} n\rho^{n-1} \mathsf{z}_{t-n-1}(u)\mathsf{z}'_{t-n-1}(u) \right| \overset{a.s.}{=} \mathcal{O}(1)$

where the validity of the second and third statement is implied by $(*)$ from Section B.3.1 and Lemma B.2. Thus, the first part of the Lemma holds with Lemmas B.7 and B.15. Furthermore, with Lemmas B.6 and B.15, we have

- $\sup_{u\in U} \left| \sum_{t=1}^{m} \left(\frac{\partial \mathsf{vech}\left(\mathsf{F}_{Q_t}(u)\right)'}{\partial \mathsf{vecl}\left(\mathsf{F}_{\bar{Q}}(u)\right)} - \frac{\partial \mathsf{vech}\left(\widehat{\mathsf{F}}_{Q_t}(u)\right)'}{\partial \mathsf{vecl}\left(\mathsf{F}_{\bar{Q}}(u)\right)} \right) \right| = 0$
- $\sup_{u\in U} \left| \sum_{t=1}^{m} \left(\frac{\partial \mathsf{vech}\left(\mathsf{F}_{Q_t}(u)\right)'}{\partial a} - \frac{\partial \mathsf{vech}\left(\widehat{\mathsf{F}}_{Q_t}(u)\right)'}{\partial a} \right) \right| = \sup_{u\in U} \left| \sum_{t=1}^{m} \sum_{k=0}^{\infty} b^k \left(\mathsf{z}_{t-k}(u)\mathsf{z}'_{t-k}(u) - \widehat{\mathsf{z}}_{t-k}(u)\widehat{\mathsf{z}}'_{t-k}(u) \right) \right|$
$\leq \frac{1}{1-b} \sup_{u\in U,t\in\mathbb{Z}} \left| \sum_{t=1}^{m} \left(\mathsf{z}_{t-k}(u)\mathsf{z}'_{t-k}(u) - \widehat{\mathsf{z}}_{t-k}(u)\widehat{\mathsf{z}}'_{t-k}(u) \right) \right| \overset{a.s.}{=} \mathcal{O}(1)$
- $\sup_{u\in U} \left| \sum_{t=1}^{m} \left(\frac{\partial \mathsf{vech}\left(\mathsf{F}_{Q_t}(u)\right)'}{\partial b} - \frac{\partial \mathsf{vech}\left(\widehat{\mathsf{F}}_{Q_t}(u)\right)'}{\partial b} \right) \right| = \sup_{u\in U} \left| \sum_{t=1}^{m} \sum_{k=0}^{\infty} b^k \left(\mathsf{vech}\left(\mathsf{F}_{Q_t}(u)\right) - \mathsf{vech}\left(\widehat{\mathsf{F}}_{Q_t}(u)\right) \right) \right|$
$\leq \frac{1}{1-b} \sup_{u\in U,t\in\mathbb{Z}} \left| \sum_{t=1}^{m} \left(\mathsf{F}_{Q_t}(u) - \widehat{\mathsf{F}}_{Q_t}(u) \right) \right| \overset{a.s.}{=} \mathcal{O}(1).$

■

Lemma B.20. *Under Assumptions 2.1, 3.1-3.4, 4.3 and 4.4, we have for $m \to \infty$:*

- $\sup_{u\in U} \left| \sum_{t=1}^{m} \left(\mathsf{F}_{Q_t^*}(u) \otimes \mathsf{F}_{Q_t^*}(u) - \widehat{\mathsf{F}}_{Q_t^*}(u) \otimes \widehat{\mathsf{F}}_{Q_t^*}(u) \right) \right| \overset{a.s.}{=} \mathcal{O}(1)$; *and*
- $\sup_{u\in U} \left| \sum_{t=1}^{m} \left(\mathsf{F}_{Q_t^*}(u)\mathsf{F}_{Q_t}(u) - \widehat{\mathsf{F}}_{Q_t^*}(u)\widehat{\mathsf{F}}_{Q_t}(u) \right) \right| \overset{a.s.}{=} \mathcal{O}(1).$

Proof: With Lemmas B.6 and B.7 and the univariate MVT, we have

$$\begin{aligned}
&\sup_{u\in U} \left| \sum_{t=1}^{m} \left(\mathsf{F}_{Q_t^*}(u) \otimes \mathsf{F}_{Q_t^*}(u) - \widehat{\mathsf{F}}_{Q_t^*}(u) \otimes \widehat{\mathsf{F}}_{Q_t^*}(u) \right) \right| \\
&= \sup_{u\in U} \max_{1\leq i,j\leq p} \left| \sum_{t=1}^{m} \left(\left[\mathsf{F}_{Q_t}(u)\right]_{ii}^{-1/2} \left[\mathsf{F}_{Q_t}(u)\right]_{jj}^{-1/2} - \left[\widehat{\mathsf{F}}_{Q_t}(u)\right]_{ii}^{-1/2} \left[\widehat{\mathsf{F}}_{Q_t}(u)\right]_{jj}^{-1/2} \right) \right| \\
&\leq 2 \sup_{u\in U} \max_{1\leq j\leq p} \left| \left[\mathsf{F}_{Q_t}(u)\right]_{jj}^{-1/2} \right| \sup_{u\in U} \max_{1\leq i\leq p} \left| \sum_{t=1}^{m} \left(\left[\mathsf{F}_{Q_t}(u)\right]_{ii}^{-1/2} - \left[\widehat{\mathsf{F}}_{Q_t}(u)\right]_{ii}^{-1/2} \right) \right|
\end{aligned}$$

$$\leq 2\sup_{u\in U}\left|\min_{1\leq j\leq p}[\mathsf{F}_{Q_t}(u)]_{jj}\right|^{-1/2}\frac{1}{2}\sup_{u\in U}\left|\min_{1\leq j\leq p}[\mathsf{F}_{Q_t}(u)]_{jj}\right|^{-3/2}\sup_{u\in U}\left|\sum_{t=1}^{m}\Big(\mathsf{F}_{Q_t}(u)-\widehat{\mathsf{F}}_{Q_t}(u)\Big)\right|$$

$$\leq \left(\frac{1-\underline{u}}{1-\rho}\right)^2\sup_{u\in U}\left|\sum_{t=1}^{m}\Big(\mathsf{F}_{Q_t}(u)-\widehat{\mathsf{F}}_{Q_t}(u)\Big)\right|\overset{a.s.}{=}\mathcal{O}(1)$$

$$\text{and}\quad \sup_{u\in U}\left|\sum_{t=1}^{m}\Big(\mathsf{F}_{Q_t^*}(u)\mathsf{F}_{Q_t}(u)-\widehat{\mathsf{F}}_{Q_t^*}(u)\widehat{\mathsf{F}}_{Q_t}(u)\Big)\right|$$

$$\leq \sup_{u\in U,t\in\mathbb{Z}}|\mathsf{F}_{Q_t}(u)|\sup_{u\in U}\left|\sum_{t=1}^{m}\Big(\mathsf{F}_{Q_t^*}(u)-\widehat{\mathsf{F}}_{Q_t^*}(u)\Big)\right|+\sup_{u\in U,t\in\mathbb{Z}}\left|\widehat{\mathsf{F}}_{Q_t^*}(u)\right|\sup_{u\in U}\left|\sum_{t=1}^{m}\Big(\mathsf{F}_{Q_t}(u)-\widehat{\mathsf{F}}_{Q_t}(u)\Big)\right|$$

$$=\left(\sup_{u\in U,t\in\mathbb{Z}}|\mathsf{F}_{Q_t}(u)|+\sup_{u\in U,t\in\mathbb{Z}}\left|\widehat{\mathsf{F}}_{Q_t^*}(u)\right|\right)\sup_{u\in U}\left|\sum_{t=1}^{m}\Big(\mathsf{F}_{Q_t}(u)-\widehat{\mathsf{F}}_{Q_t}(u)\Big)\right|\overset{a.s.}{=}\mathcal{O}(1)$$

■

Denote

$$K_{3t}(u):=\mathsf{D}_{p,-}^{+}\Big(\mathsf{F}_{Q_t^*}(u)\otimes\mathsf{F}_{Q_t^*}(u)\Big)\mathsf{D}_p+\mathsf{D}_{p,-}^{+}\Big(\mathsf{F}_{Q_t^*}(u)\mathsf{F}_{Q_t}(u)\otimes\mathbb{I}_p+\mathbb{I}_p\otimes\mathsf{F}_{Q_t^*}(u)\mathsf{F}_{Q_t}(u)\Big)\mathsf{D}_p\frac{\partial\mathsf{vech}\left(\mathsf{F}_{Q_t^*}(u)\right)}{\partial\mathsf{vech}\left(\mathsf{F}_{Q_t}(u)\right)'}$$

and analogously $\widehat{K}_{3t}(u)$ in dependence of a finite past of observations.

Lemma B.21. *Under Assumptions 2.1, 3.1-3.4, 4.3 and 4.4, we have for $m\to\infty$:*

- $\sup_{u\in U,t\in\mathbb{Z}}|K_{3t}(u)|\overset{a.s.}{=}\mathcal{O}(1)$*; and*
- $\sup_{u\in U}\left|\sum_{t=1}^{m}\Big(K_{3t}(u)-\widehat{K}_{3t}(u)\Big)\right|\overset{a.s.}{=}\mathcal{O}(1).$

Proof: First of all, we have

$$\sup_{u\in U,t\in\mathbb{Z}}\left|\frac{\partial\mathsf{vech}\left(\mathsf{F}_{Q_t^*}(u)\right)}{\partial\mathsf{vech}\left(\mathsf{F}_{Q_t}(u)\right)'}\right|\leq\frac{1}{2}\sup_{u\in U,t\in\mathbb{Z}}\left(\min_{1\leq j\leq p}\left|[F_{Q_t}(u)]_{jj}\right|\right)^{-3/2}\overset{a.s.}{\leq}\left(\frac{1-\underline{u}}{1-\rho}\right)^{3/2}=\mathcal{O}(1).\tag{B.27}$$

Furthermore, the MVT and Lemma B.6 yield

$$\begin{aligned}\sup_{u\in U}\left|\sum_{t=1}^{m}\left(\frac{\partial\mathsf{vech}\left(\mathsf{F}_{Q_t^*}(u)\right)}{\partial\mathsf{vech}\left(\mathsf{F}_{Q_t}(u)\right)'}-\frac{\partial\mathsf{vech}\left(\widehat{\mathsf{F}}_{Q_t^*}(u)\right)}{\partial\mathsf{vech}\left(\mathsf{F}_{Q_t}(u)\right)'}\right)\right|&=\frac{1}{2}\max_{1\leq j\leq p}\sup_{u\in U}\left|\sum_{t=1}^{m}\left([\mathsf{F}_{Q_t}(u)]_{jj}^{-3/2}-\left[\widehat{\mathsf{F}}_{Q_t}(u)\right]_{jj}^{-3/2}\right)\right|\\&\leq\frac{3}{4}\sup_{u\in U,t\in\mathbb{Z}}\left(\min_{1\leq j\leq p}[\mathsf{F}_{Q_t}(u)]_{jj}\right)^{-5/2}\sup_{u\in U}\left|\sum_{t=1}^{m}\Big(\mathsf{F}_{Q_t}(u)-\widehat{\mathsf{F}}_{Q_t}(u)\Big)\right|\\&\leq\frac{3}{4}\left(\frac{1-\underline{u}}{1-\rho}\right)^{5/2}\sup_{u\in U}\left|\sum_{t=1}^{m}\Big(\mathsf{F}_{Q_t}(u)-\widehat{\mathsf{F}}_{Q_t}(u)\Big)\right|\overset{a.s.}{=}\mathcal{O}(1).\end{aligned}\tag{B.28}$$

In combination with the Lemmas B.7 and B.20, (B.27) and (B.28) imply

$$\sup_{u\in U,t\in\mathbb{Z}}\left|K_{3t}(u)\right| \leq \sup_{u\in U,t\in\mathbb{Z}}\left|\mathsf{F}_{Q_t^*}(u)\otimes\mathsf{F}_{Q_t^*}(u)\right| + \sup_{u\in U,t\in\mathbb{Z}}\left|\mathsf{F}_{Q_t^*}(u)\mathsf{F}_{Q_t}(u)\right| \sup_{u\in U,t\in\mathbb{Z}}\left|\frac{\partial \mathsf{vech}\left(\mathsf{F}_{Q_t^*}(u)\right)}{\partial \mathsf{vech}\left(\mathsf{F}_{Q_t}(u)\right)'}\right|$$

$$\leq \left(\sup_{u\in U,t\in\mathbb{Z}}\left|\mathsf{F}_{Q_t^*}(u)\right|\right)^2 + p\sup_{u\in U,t\in\mathbb{Z}}\left|\mathsf{F}_{Q_t^*}(u)\right|\sup_{u\in U,t\in\mathbb{Z}}\left|\mathsf{F}_{Q_t}(u)\right|\left(\frac{1-\underline{u}}{1-\rho}\right)^{3/2} \overset{a.s.}{=} \mathcal{O}(1)$$

and

$$\sup_{u\in U}\left|\sum_{t=1}^m\left(K_{3t}(u)-\widehat{K}_{3t}(u)\right)\right| \leq \sup_{u\in U}\left|\sum_{t=1}^m\left(\mathsf{F}_{Q_t^*}(u)\otimes\mathsf{F}_{Q_t^*}(u)-\widehat{\mathsf{F}}_{Q_t^*}(u)\otimes\widehat{\mathsf{F}}_{Q_t^*}(u)\right)\right|$$

$$+2\sup_{u\in U,t\in\mathbb{Z}}\left|\frac{\partial \mathsf{vech}\left(\widehat{\mathsf{F}}_{Q_t^*}(u)\right)}{\partial \mathsf{vech}\left(\mathsf{F}_{Q_t}(u)\right)'}\right|\sup_{u\in U}\left|\sum_{t=1}^m\left(\mathsf{F}_{Q_t^*}(u)\mathsf{F}_{Q_t}(u)-\widehat{\mathsf{F}}_{Q_t^*}(u)\widehat{\mathsf{F}}_{Q_t}(u)\right)\right|$$

$$+2\sup_{u\in U,t\in\mathbb{Z}}\left|\widehat{\mathsf{F}}_{Q_t^*}(u)\widehat{\mathsf{F}}_{Q_t}(u)\right|\sup_{u\in U}\left|\sum_{t=1}^m\left(\frac{\partial \mathsf{vech}\left(\mathsf{F}_{Q_t^*}(u)\right)}{\partial \mathsf{vech}\left(\mathsf{F}_{Q_t}(u)\right)'}-\frac{\partial \mathsf{vech}\left(\widehat{\mathsf{F}}_{Q_t^*}(u)\right)}{\partial \mathsf{vech}\left(\mathsf{F}_{Q_t}(u)\right)'}\right)\right| \overset{a.s.}{=} \mathcal{O}(1)$$

■

Lemma B.22. *Under Assumptions 2.1, 3.1-3.4, 4.3 and 4.4, we have for $m\to\infty$:*

- $\sup_{u\in U,t\in\mathbb{Z}}\left|\frac{\partial \mathsf{vecl}\left(\widehat{\mathsf{F}}_{R_t}(u)\right)'}{\partial u_2}\right| \overset{a.s.}{=} \mathcal{O}(1)$*; and*
- $\sup_{u\in U}\left|\sum_{t=1}^m\left(\frac{\partial \mathsf{vecl}\left(\mathsf{F}_{R_t}(u)\right)'}{\partial u_2}-\frac{\partial \mathsf{vecl}\left(\widehat{\mathsf{F}}_{R_t}(u)\right)'}{\partial u_2}\right)\right| \overset{a.s.}{=} \mathcal{O}(1)$.

Proof: With Lemmas B.19 and B.21, we have

- $\sup_{u\in U,t\in\mathbb{Z}}\left|\frac{\partial \mathsf{vecl}\left(\widehat{\mathsf{F}}_{R_t}(u)\right)'}{\partial u_2}\right| \leq \sup_{u\in U,t\in\mathbb{Z}}\left|K_{3t}(u)\right|\left|\sup_{u\in U,t\in\mathbb{Z}}\frac{\partial \mathsf{vech}\left(\mathsf{F}_{Q_t}(u)\right)'}{\partial u_2}\right| \overset{a.s.}{=} \mathcal{O}(1)$
- $\sup_{u\in U}\left|\sum_{t=1}^m\left(\frac{\partial \mathsf{vecl}\left(\mathsf{F}_{R_t}(u)\right)'}{\partial u_2}-\frac{\partial \mathsf{vecl}\left(\widehat{\mathsf{F}}_{R_t}(u)\right)'}{\partial u_2}\right)\right|$

$$\leq \sup_{u\in U,t\in\mathbb{Z}}\left|K_{3t}(u)\right|\sup_{u\in U}\left|\sum_{t=1}^m\left(\frac{\partial \mathsf{vech}\left(\mathsf{F}_{Q_t}(u)\right)'}{\partial u_2}-\frac{\partial \mathsf{vech}\left(\mathsf{F}_{Q_t}(u)\right)'}{\partial u_2}\right)\right|$$

$$+\sup_{u\in U,t\in\mathbb{Z}}\left|\frac{\partial \mathsf{vech}\left(\mathsf{F}_{Q_t}(u)\right)'}{\partial u_2}\right|\sup_{u\in U}\left|\sum_{t=1}^m\left(K_{3t}(u)-\widehat{K}_{3t}(u)\right)\right| \overset{a.s.}{=} \mathcal{O}(1).$$

■

Lemma B.23. *Under Assumptions 2.1, 3.1-3.4, 4.3 and 4.4, we have for $m\to\infty$:*

- $\sup_{u\in U,t\in\mathbb{Z}}\left|\frac{\partial l_t(u)}{\partial u_2}\right| \overset{a.s.}{=} \mathcal{O}(1)$*; and*
- $\sup_{u\in U}\left|\sum_{t=1}^m\left(\frac{\partial \hat{l}_t(u)}{\partial u_2}-\frac{\partial l_t(u)}{\partial u_2}\right)\right| \overset{a.s.}{=} \mathcal{O}(1)$.

Proof: With Lemmas B.19 and B.18, we have

- $\sup\limits_{u\in U, t\in\mathbb{Z}} \left|\frac{\partial l_t(u)}{\partial u_2}\right| \leq \sup\limits_{u\in U, t\in\mathbb{Z}} \left|\frac{\partial \mathsf{vecl}\left(\mathsf{F}_{R_t}(u)\right)'}{\partial u_2}\right| + \sup\limits_{u\in U, t\in\mathbb{Z}} |K_{2t}(u)|$
- $$\sup_{u\in U} \left|\sum_{t=1}^{m}\left(\frac{\partial \hat{l}_t(u)}{\partial u_2} - \frac{\partial l_t(u)}{\partial u_2}\right)\right| \leq \sup_{u\in U, t\in\mathbb{Z}} |K_{2t}(u)| \sup_{u\in U}\left|\sum_{t=1}^{m}\left(\frac{\partial \mathsf{vecl}\left(\mathsf{F}_{R_t}(u)\right)'}{\partial u_2} - \frac{\partial \mathsf{vecl}\left(\widehat{\mathsf{F}}_{R_t}(u)\right)'}{\partial u_2}\right)\right|$$
$$+ \sup_{u\in U, t\in\mathbb{Z}} \left|\frac{\partial \mathsf{vecl}\left(\widehat{\mathsf{F}}_{R_t}(u)\right)'}{\partial u_2}\right| \sup_{u\in U}\left|\sum_{t=1}^{m}\left(\mathsf{F}_{R_t}(u)^{-1} - \widehat{\mathsf{F}}_{R_t}(u)^{-1}\right)\right|$$
$$+ \sup_{u\in U, t\in\mathbb{Z}} \left|\frac{\partial \mathsf{vecl}\left(\widehat{\mathsf{F}}_{R_t}(u)\right)'}{\partial u_2}\right| \sup_{u\in U}\left|\sum_{t=1}^{m}\left(\widehat{K_{2t}}(u) - K_{2t}(u)\right)\right|$$

Thus, with Lemmas B.17 and B.23, we have $\sup\limits_{u\in U}\left|\widehat{D}_m(u) - D_m(u)\right| \overset{a.s.}{=} \mathcal{O}\left(\frac{1}{m}\right)$, which completes the proof of (B.9). ■

B.2.2. The Proof of $\mathsf{E}\sup\limits_{u\in U}\left|l_0'(u)l_0'(u)^T\right| < \infty$

Along the lines of Berkes et al. (2003), we have

$$\sup_{u\in U}\left|l_0'(u)l_0'(u)^T\right| \leq \left(\sup_{u\in U}\left|\frac{\partial l_0(u)}{\partial u}\right|\right)^2 \overset{a.s.}{=} \mathcal{O}(1)$$

with Lemmas B.17 and B.23. This implies

$$\mathsf{E}\sup_{u\in U}\left|l_0'(u)l_0'(u)^T\right| < \infty.$$

B.2.3. The Proof of the Uniform Convergence of $\boldsymbol{D}_m(\cdot)$ to $\boldsymbol{D}(\cdot)$

We use Theorem A.2.2 in White (1994). The requirements are satisfied since U is a compact set and $l_t'(u)l_t'(u)^T$ is ergodic, continuous in u for all y_t and measurable in y_t for all $u \in U$. We choose the dominiating function as $\sup\limits_{u\in U}\left|l_t'(u)l_t'(u)^T\right|$. Thus, the finiteness of the expectation is implied by (B.10) which yields the uniform convergence of $\boldsymbol{D}_m(\cdot)$ to $\boldsymbol{D}(\cdot)$. ■

B.2.4. The Proof of Proposition 4.2.

Since $\boldsymbol{D}_m(\cdot)$ converges uniformly to $\boldsymbol{D}(\cdot)$, (B.11) follows directly from the consistency of the estimator $\hat{\theta}_m$ and the positive definiteness of the variation matrix $\boldsymbol{D}$, i.e. Assumption 4.1, which completes the proof of Proposition 4.2. ■

B.3. The Proof that the Second Order Derivations are of Finite Expectation

B.3.1. Notation

In the following, for $\varphi \in (0,1)$ and $i,j \in \{1,\dots,p\}$, we adopt the notation

- $\mathsf{G}_0^y(i,j,\varphi) := \sum_{n=0}^{\infty} \varphi^n y_{i,-n-1} y_{j,-n-1}$
- $\mathsf{G}_1^y(i,j,\varphi) := \sum_{n=0}^{\infty} n\varphi^{n-1} y_{i,-n-1} y_{j,-n-1}$
- $\mathsf{G}_2^y(i,j,\varphi) := \sum_{n=0}^{\infty} n(n-1)\varphi^{n-2} y_{i,-n-1} y_{j,-n-1}$
- $\mathsf{G}_0^{\mathsf{z}}(i,j,u) := \sum_{n=0}^{\infty} b^n \mathsf{z}_{i,-n-1}(u) \mathsf{z}_{j,-n-1}(u)$
- $\mathsf{G}_1^{\mathsf{z}}(i,j,u) := \sum_{n=0}^{\infty} n b^{n-1} \mathsf{z}_{i,-n-1}(u) \mathsf{z}_{j,-n-1}(u)$
- $\mathsf{G}_2^{\mathsf{z}}(i,j,u) := \sum_{n=0}^{\infty} n(n-1) b^{n-2} \mathsf{z}_{i,-n-1}(u) \mathsf{z}_{j,-n-1}(u)$
- $\mathsf{Z}(u) := \left(\sum_{s=1}^{p} \mathsf{z}_{s0}(u)\right)^2 = \sum_{s=1}^{p}\sum_{t=1}^{p} \mathsf{z}_{s0}(u)\mathsf{z}_{t0}(u).$

Note that $\mathsf{G}^{\mathsf{z}}(i,j,u) \overset{a.s.}{\leq} \frac{1}{C_1}\mathsf{G}^y(i,j,b)$ for all $u \in U$.

Furthermore, with $\bar{\varphi} := \max\{\varphi_1,\varphi_2\}$ and $\varphi_1,\varphi_2 \in (0,1)$, we have for $i,j,k,l \in \{1,\dots,p\}$:

$$\begin{aligned} \mathsf{G}_0^y(i,j,\varphi_1)\,\mathsf{G}_0^y(k,l,\varphi_2) &= \sum_{m=0}^{\infty}\sum_{n=0}^{m} \varphi_1^n \varphi_2^{m-n} y_{i,-n-1} y_{j,-n-1} y_{k,-m+n-1} y_{l,-m+n-1} \\ &\leq \sum_{m=0}^{\infty} \left(\bar{\varphi}^{\frac{1}{2}}\right)^m \sum_{n=0}^{m} \left(\bar{\varphi}^{\frac{1}{2}}\right)^n y_{i,-n-1} y_{j,-n-1} y_{k,-m+n-1} y_{l,-m+n-1}. \end{aligned} \tag{B.29}$$

With Assumptions 4.3 and 4.4 the double sum in (B.29) is stochastically bounded, since under the use of the Cauchy Schwarz inequality, we have $\mathsf{E}|y_{is}y_{js}y_{kt}y_{lt}| \leq \left[\mathsf{E}\left(y_{is}^2 y_{js}^2\right)\right]^{1/2} \left[\mathsf{E}\left(y_{kt}^2 y_{lt}^2\right)\right]^{1/2} < \infty$ which implies $\mathsf{E}\log^+(y_{is}y_{js}y_{kt}y_{lt}) < \infty$ for all $s,t \in \mathbb{Z}$ and all $i,j,k,l \in \{1,\dots,p\}$.

Thus, Lemma B.2 can be applied to the sum on the righthand side of (B.29).

Analogously, all products $\mathsf{G}_{m_1}^y(i,j,\varphi_1)\,\mathsf{G}_{m_2}^y(k,l,\varphi_2)$ or $\mathsf{G}_{m_1}^{\mathsf{z}}(i,j,u)\,\mathsf{G}_{m_2}^{\mathsf{z}}(k,l,u)$ with $\varphi_1,\varphi_2 \in (0,1)$ and $m_1,m_2 \in \{0,1,2\}$ are of finite expectation. This property can be expanded to products of four terms of type $\mathsf{G}_m^y(i,j,\varphi)$ or $\mathsf{G}_m^z(i,j,u)$, $i,j \in \{1,\dots,p\}$, $m \in \{0,1,2\}$, $\varphi \in (0,1)$, which can be shown by the use of Assumptions 4.3 and 4.4 and an application of the generalized Hölder inequality, that is Lemma 1.16 in Alt (2006) with $m = 8$, $q = 1$ and $p_i = 8$, $i = 1,\dots,8$. While considering products of terms of type $\mathsf{G}_m^{\mathsf{z}}(i,j,u)$, one or several factors can be substituted by $\mathsf{Z}(u)$ which leads to a decomposition into finitely many summands that are of finite expectation each. We denote the property of finite expectations of all of these products as $(*)$.

Furthermore, denote $\mathsf{F}_{R_0}(u)^{-1} := \left[r_{ij,0}^{-}(u)\right]_{i,j=1,\ldots,p}$ and recall that the arguments in the proof of Lemma B.14 imply

$$\sup_{u\in U,t\in\mathbb{Z}}\left|r_{ij,t}^{-}(u)\right| \overset{a.s.}{<} \left[\frac{1-b}{1-a-b}\frac{p\delta_2}{\delta_1}\right]^p (p-1)! =: \delta_* \quad \forall i,j\in\{1,\ldots,p\}\,. \tag{B.30}$$

B.3.2. The Partial Derivations of $\mathsf{w}_{i0}(u)$

The following statements on the magnitude of the partial derivations of $w_{i0}(u)$ with respect to the variance parameters will be used in the next sections:

$$\bullet\ \left(\frac{\partial \mathsf{w}_{i0}(u)}{\partial x_i}\right)^2 = \frac{1}{4}\frac{1}{(1-t_i)^2} \qquad \bullet\ \left(\frac{\partial \mathsf{w}_{i0}(u)}{\partial s_i}\right)^2 = \frac{1}{4}\mathsf{G}_0^y(i,i,t_i)^2 \tag{B.31}$$

$$\bullet\ \left(\frac{\partial \mathsf{w}_{i0}(u)}{\partial t_i}\right)^2 = \frac{1}{4}\frac{x_i^2}{(1-t_i)^4} + \frac{1}{2}\frac{x_i}{(1-t_i)^2}\mathsf{G}_1^y(i,i,t_i) + \frac{1}{4}\mathsf{G}_1^y(i,i,t_i)^2 \tag{B.32}$$

$$\bullet\ \frac{\partial \mathsf{w}_{i0}(u)}{\partial x_i}\frac{\partial \mathsf{w}_{i0}(u)}{\partial s_i} = \frac{1}{4}\frac{1}{1-t_i}\mathsf{G}_0^y(i,i,t_i) \qquad \bullet\ \frac{\partial \mathsf{w}_{i0}(u)}{\partial x_i}\frac{\partial \mathsf{w}_{i0}(u)}{\partial t_i} = \frac{1}{4}\frac{x_i}{(1-t_i)^3} + \frac{1}{4}\frac{1}{1-t_i}\mathsf{G}_1^y(i,i,t_i) \tag{B.33}$$

$$\bullet\ \frac{\partial \mathsf{w}_{i0}(u)}{\partial s_i}\frac{\partial \mathsf{w}_{i0}(u)}{\partial t_i} = \frac{1}{4}\frac{x_i}{(1-t_i)^2}\mathsf{G}_0^y(i,i,t_i) + \frac{1}{4}\mathsf{G}_0^y(i,i,t_i)\,\mathsf{G}_1^y(i,i,t_i)\,. \tag{B.34}$$

Furthermore, the following statements on derivations of the derivations of the roots of $\mathsf{w}_{i0}(u)$ will be useful:

$$\bullet\ \frac{\partial \mathsf{w}_{j0}(u)^{\frac{1}{2}}}{\partial x_j} \le \frac{1}{2C_1^{\frac{1}{2}}(1-t_j)} \qquad \bullet\ \frac{\partial \mathsf{w}_{j0}(u)^{\frac{1}{2}}}{\partial s_j} \le \frac{\mathsf{G}_0^y(j,j,t_j)}{2C_1^{\frac{1}{2}}} \qquad \bullet\ \frac{\partial \mathsf{w}_{j0}(u)^{\frac{1}{2}}}{\partial t_j} \le \frac{\mathsf{G}_1^y(j,j,t_j)}{2C_1^{\frac{1}{2}}} \tag{B.35}$$

$$\bullet\ \left|\frac{\partial^2 \mathsf{w}_{j0}(u)^{\frac{1}{2}}}{(\partial x_j)^2}\right| \le \frac{1}{4C_1^{\frac{3}{2}}(1-t_j)^2} \qquad \bullet\ \left|\frac{\partial^2 \mathsf{w}_{j0}(u)^{\frac{1}{2}}}{(\partial s_j)^2}\right| \le \frac{\mathsf{G}_0^y(j,j,t_j)^2}{4C_1^{\frac{3}{2}}} \tag{B.36}$$

$$\bullet\ \left|\frac{\partial^2 \mathsf{w}_{j0}(u)^{\frac{1}{2}}}{(\partial t_j)^2}\right| \le \frac{\mathsf{G}_1^y(j,j,t_j)}{4C_1^{\frac{3}{2}}} + \frac{\mathsf{G}_2^y(j,j,t_j)}{2C_1^{\frac{1}{2}}} \qquad \bullet\ \left|\frac{\partial^2 \mathsf{w}_{j0}(u)^{\frac{1}{2}}}{\partial x_j \partial s_j}\right| \le \frac{\mathsf{G}_0^y(j,j,t_j)}{4C_1^{\frac{3}{2}}(1-t_j)} \tag{B.37}$$

$$\bullet\ \left|\frac{\partial^2 \mathsf{w}_{j0}(u)^{\frac{1}{2}}}{\partial x_j \partial t_j}\right| \le \frac{1}{2C_1^{\frac{1}{2}}(1-t_j)^2} + \frac{\mathsf{G}_1^y(j,j,t_j)}{4C_1^{\frac{3}{2}}(1-t_j)} \qquad \bullet\ \left|\frac{\partial^2 \mathsf{w}_{j0}(u)^{\frac{1}{2}}}{\partial s_j \partial t_j}\right| \le \frac{\mathsf{G}_0^y(j,j,t_j)^2}{4C_1^{\frac{3}{2}}} + \frac{\mathsf{G}_1^y(j,j,t_j)}{2C_1^{\frac{1}{2}}} \tag{B.38}$$

B.3.3. The Partial Derivations of $\mathsf{F}_{Q_0}(u)$ and $\mathsf{F}_{Q_0^*}(u)$

In the following, we will have a closer look at the first and second order partial derivations of the (i,j)-th entry of $\mathsf{F}_{Q_0}(u)$:

- $\dfrac{\partial\left[\mathsf{F}_{Q_0}(u)\right]_{ij}}{\partial a} = -\dfrac{1}{1-b} + \mathsf{G}_0^z(i,j,u)$
- $\dfrac{\partial\left[\mathsf{F}_{Q_0}(u)\right]_{ij}}{\partial b} = -\dfrac{a}{(1-b)^2} + a\mathsf{G}_1^z(i,j,u)$ (B.39)
- $\dfrac{\partial^2\left[\mathsf{F}_{Q_0}(u)\right]_{ij}}{(\partial a)^2} = 0$
- $\dfrac{\partial^2\left[\mathsf{F}_{Q_0}(u)\right]_{ij}}{(\partial b)^2} = -\dfrac{2a}{(1-b)^3} + a\mathsf{G}_2^z(i,j,u)$ (B.40)
- $\dfrac{\partial^2\left[\mathsf{F}_{Q_0}(u)\right]_{ij}}{\partial a\partial b} = -\dfrac{1}{(1-b)^2} + \mathsf{G}_1^z(i,j,u)\,.$ (B.41)

The entries of $\mathsf{F}_{Q_0^*}(u)$ are either 0 or $[\mathsf{F}_{Q_0}(u)]_{ii}^{-1/2}$, for $i \in \{1,\dots,p\}$. Hence, only the derivations of the main diagonal entries will be considered.

- $$\frac{\partial^2\left[\mathsf{F}_{Q_0^*}(u)\right]_{ii}}{(\partial a)^2} = \frac{\partial^2\left[\mathsf{F}_{Q_0}(u)\right]_{ii}^{-\frac{1}{2}}}{(\partial a)^2} = \frac{3}{4}\left[\mathsf{F}_{Q_0}(u)\right]_{ii}^{-\frac{5}{2}}\left(\frac{\partial\left[\mathsf{F}_{Q_0}(u)\right]_{ii}}{\partial a}\right)^2 < \frac{3a^2}{4}\left(\frac{1-b}{1-a-b}\right)^{\frac{5}{2}}\mathsf{G}_0^z(i,i,u)^2 \quad \text{(B.42)}$$
- $$\begin{aligned}\frac{\partial^2\left[F_{Q_0^*}(u)\right]_{ii}}{(\partial b)^2} &= \frac{3}{4}\left[F_{Q_0}(u)\right]_{ii}^{-\frac{5}{2}}\left(\frac{\partial\left[F_{Q_0}(u)\right]_{ii}}{\partial b}\right)^2 - \frac{1}{2}\left[F_{Q_0}(u)\right]_{ii}^{-\frac{3}{2}}\frac{\partial^2\left[F_{Q_0}(u)\right]_{ii}}{(\partial b)^2}\\ &< \frac{3a^2}{4}\left(\frac{1-b}{1-a-b}\right)^{\frac{5}{2}}\mathsf{G}_1^z(i,i,u)^2 + \frac{1}{2}\left(\frac{1-b}{1-a-b}\right)^{\frac{3}{2}}\frac{2}{(1-\rho)^3} \quad \text{(B.43)}\end{aligned}$$
- $$\begin{aligned}\frac{\partial^2\left[F_{Q_0^*}(u)\right]_{ii}}{\partial a\partial b} &= \frac{3}{4}\left[F_{Q_0}(u)\right]_{ii}^{-\frac{5}{2}}\frac{\partial\left[F_{Q_0}(u)\right]_{ii}}{\partial a}\frac{\partial\left[F_{Q_0}(u)\right]_{ii}}{\partial b} - \frac{1}{2}\left[F_{Q_0}(u)\right]_{kk}^{-\frac{3}{2}}\frac{\partial^2\left[F_{Q_0}(u)\right]_{ii}}{\partial a\partial b}\\ &< \frac{3a^2}{4}\left(\frac{1-b}{1-a-b}\right)^{\frac{5}{2}}\mathsf{G}_0^z(i,i,u)\mathsf{G}_1^z(i,i,u) + \frac{1}{2}\left(\frac{1-b}{1-a-b}\right)^{\frac{3}{2}}\frac{1}{(1-\rho)^2}. \quad \text{(B.44)}\end{aligned}$$

B.3.4. The Proof that the Expectation of $\frac{\partial^2 l_0(u)}{\partial u_1\partial u_1'}$ is Finite

$$\begin{aligned}\mathsf{E}\left(\frac{\partial^2 l_0(u)}{\partial u_1\partial u_1'}\right) &= -\frac{1}{2}\mathsf{E}\left(\frac{\partial\mathsf{vec}\left(\mathsf{F}_{D_0}(u)\right)'}{\partial u_1}\Big[2\Big(\mathsf{F}_{D_0}(u)^{-1}\otimes\mathsf{F}_{D_0}(u)^{-1}\Big)\right. && \text{(B.45)}\\
&+\Big(\mathsf{z}_0(u)\mathsf{z}_0'(u)\otimes\mathsf{F}_{D_0}(u)^{-1}\mathsf{F}_{R_0}(u)^{-1}\mathsf{F}_{D_0}(u)^{-1}\Big)+\Big(\mathsf{F}_{D_0}(u)^{-1}\mathsf{F}_{R_0}(u)^{-1}\mathsf{F}_{D_0}(u)^{-1}\otimes\mathsf{z}_0(u)\mathsf{z}_0'(u)\Big)\\
&+\Big(\mathsf{F}_{D_0}(u)^{-1}\otimes\mathsf{F}_{D_0}(u)^{-1}\mathsf{F}_{R_0}(u)^{-1}\mathsf{z}_0(u)\mathsf{z}_0'(u)\Big)+\Big(\mathsf{F}_{D_0}(u)^{-1}\mathsf{F}_{R_0}(u)^{-1}\mathsf{z}_0(u)\mathsf{z}_0'(u)\otimes\mathsf{F}_{D_0}(u)^{-1}\Big)\\
&+\Big(\mathsf{z}_0(u)\mathsf{z}_0'(u)\mathsf{F}_{R_0}(u)^{-1}\mathsf{F}_{D_0}(u)^{-1}\otimes\mathsf{F}_{D_0}(u)^{-1}\Big)\\
&\left.+\Big(\mathsf{F}_{D_0}(u)^{-1}\otimes\mathsf{z}_0(u)\mathsf{z}_0'(u)\mathsf{F}_{R_0}(u)^{-1}\mathsf{F}_{D_0}(u)^{-1}\Big)\Big]\right)\frac{\partial\mathsf{vec}\left(\mathsf{F}_{D_0}(u)\right)}{\partial u_1'} && \text{(B.46)}\\
&+\mathsf{E}\left(\left[\frac{1}{2}\left(\mathsf{vec}\Big(\mathsf{F}_{D_0}(u)^{-1}\mathsf{F}_{R_0}(u)^{-1}\mathsf{z}_0(u)\mathsf{z}_0'(u)\Big)\otimes\mathbb{I}_{3p}\right)\right.\right.\\
&\left.\left.+\frac{1}{2}\left(\mathsf{vec}\Big(\mathsf{z}_0(u)\mathsf{z}_0'(u)\mathsf{F}_{R_0}(u)^{-1}\mathsf{F}_{D_0}(u)^{-1}\Big)\otimes\mathbb{I}_{3p}\right)-\left(\mathsf{vec}\Big(\mathsf{F}_{D_0}(u)^{-1}\Big)\otimes\mathbb{I}_{3p}\right)\right]\frac{\partial^2\mathsf{vec}\left(\mathsf{F}_{D_0}(u)\right)'}{\partial u_1\partial u_1'}\right) && \text{(B.47)}\end{aligned}$$

First of all, we have a closer look at the summands in (B.46). We use different matrices $\mathsf{V}^{(l)} := \left[\mathsf{v}_{ij}^{(l)}\right]_{i,j=1,\ldots,p}$, $l \in \{1,2\}$, that have to be specified appropriately to obtain the terms in (B.46). This general approach yields that

$$\frac{\partial \mathsf{vec}\left(\mathsf{F}_{D_0}(u)\right)'}{\partial u_1}\left(\mathsf{V}^{(1)} \otimes \mathsf{V}^{(2)}\right)\frac{\partial \mathsf{vec}\left(\mathsf{F}_{D_0}(u)\right)}{\partial u_1'}$$

is a block diagonal matrix with (3×3) blocks

$$\frac{\mathsf{v}_{kk}^{(1)}\mathsf{v}_{kk}^{(2)}}{\mathsf{w}_{k0}(u)}\frac{\partial \mathsf{w}_{k0}(u)}{\partial \mathsf{r}_k}\frac{\partial \mathsf{w}_{k0}(u)}{\partial \mathsf{r}_k'} \qquad \text{for } k \in \{1,\ldots,p\}$$

on its main diagonal. Note, that with $(*)$ the terms (B.31)-(B.34) are stochastically bounded. For the first summand in (B.46), we have $\frac{\mathsf{v}_{ii}^{(1)}\mathsf{v}_{ii}^{(2)}}{\mathsf{w}_{i0}(u)} = \mathsf{w}_{i0}(u)^{-2} \leq C_1^{-2}$. Thus, the expectation of this summand is finite. For the second and third summand in (B.46), we have

$$\mathsf{E}\left(\frac{\mathsf{v}_{ii}^{(1)}\mathsf{v}_{ii}^{(2)}}{\mathsf{w}_{i0}(u)}\right) \leq C_1^{-3}\mathsf{E}\left(y_{i0}^2\left|r_{ii,0}^-(u)\right|\right) \leq \delta_* C_1^{-3}\mathsf{E}\left(y_{i0}^2\right) \tag{B.48}$$

and for the fourth to seventh summand, we have

$$\mathsf{E}\left(\frac{\mathsf{v}_{ii}^{(1)}\mathsf{v}_{ii}^{(2)}}{\mathsf{w}_{i0}(u)}\right) = \mathsf{E}\left(\frac{y_{i0}}{\mathsf{w}_{i0}(u)^2}\sum_{k=1}^{p}\frac{y_{k0}}{\mathsf{w}_{k0}(u)}\left|r_{ik,0}^-(u)\right|\right) \leq \delta_* C_1^{-3}\sum_{k=1}^{p}\mathsf{E}\left(y_{i0}y_{k0}\right). \tag{B.49}$$

Hence, with $(*)$ it is obvious that products of one of the terms (B.31)-(B.34) and (B.48) or (B.49) are of finite expectation. Furthermore, note that (B.47) is finite if the following expectations are finite for all $u \in U$, $i_1,\ldots,i_4 \in \{1,\ldots,p\}$ and $x, y \in \{x_{i_4}, s_{i_4}, t_{i_4}\}$:

$$\mathsf{E}\left(w_{i_1 0}(u)^{-1/2}\left|r_{i_2 i_3,0}^-(u)\right|\mathsf{Z}(u)\frac{\partial^2 w_{i_4 0}(u)^{1/2}}{\partial x \partial y}\right) \leq \delta_* C_1^{-1/2}\mathsf{E}\left(\mathsf{Z}(u)\frac{\partial^2 w_{i_4 0}(u)^{1/2}}{\partial x \partial y}\right) \tag{B.50}$$

$$\mathsf{E}\left(w_{i_1 0}(u)^{-1/2}\frac{\partial^2 w_{i_1 0}(u)^{1/2}}{\partial x \partial y}\right) \leq C_1^{-1/2}\mathsf{E}\left(\frac{\partial^2 w_{i_1 0}(u)^{1/2}}{\partial x \partial y}\right). \tag{B.51}$$

The finity of (B.50) and (B.51) is a direct consequence of (B.30), $(*)$ and (B.35)-(B.38). Therefore, combining the previous results finally yields $\frac{\partial^2 l_0(u)}{\partial u_1 \partial u_1'} < \infty$.

B.3.5. The Proof that the Expectations of $\frac{\partial^2 l_0(u)}{\partial u_1 \partial u_2'}$ and $\frac{\partial^2 l_0(u)}{\partial u_2 \partial u_1'}$ are Finite

$$\mathsf{E}\left(\frac{\partial^2 l_0(u)}{\partial u_1 \partial u_2'}\right) = -\frac{1}{4}\mathsf{E}\left(\frac{\partial \mathsf{vec}\left(\mathsf{F}_{D_0}(u)\right)'}{\partial u_1}\left[\left(\mathsf{z}_0(u)\mathsf{z}_0'(u)\mathsf{F}_{R_0}(u)^{-1} \otimes \mathsf{F}_{D_0}(u)^{-1}\mathsf{F}_{R_0}(u)^{-1}\right)\right.\right.$$
$$\left.\left. + \left(\mathsf{F}_{D_0}(u)^{-1}\mathsf{F}_{R_0}(u)^{-1} \otimes \mathsf{z}_0(u)\mathsf{z}_0'(u)\mathsf{F}_{R_0}(u)^{-1}\right)\right] 2\left(\mathsf{D}_{p,-}^{+}\right)' \frac{\partial \mathsf{vecl}\left(\mathsf{F}_{R_0}(u)\right)}{\partial u_2'}\right) \quad \text{(B.52)}$$

The cross derivations are of finite expectation if the following expectation is finite for all $u \in U$, $i_1, \ldots, i_5 \in \{1, \ldots, p\}$, $x \in \{x_{i_1}, s_{i_1}, t_{i_1}\}$ and $y \in \{a, b, q_1, \ldots, q_{p^-}\}$:

$$\mathsf{E}\left(\frac{\partial \mathsf{w}_{i_1 0}^{1/2}}{\partial x}\mathsf{Z}(u)\left|r_{i_2 i_3,0}^{-}(u)\right|^2 \mathsf{w}_{i_4 0}^{-1/2}\frac{\partial\left[\mathsf{F}_{Q_0^*}(u)\right]_{i_5 i_5}}{\partial y}\right) \leq \delta_*^2 C_1^{-1/2}\mathsf{E}\left(\frac{\partial \mathsf{w}_{i_1 0}^{1/2}}{\partial x}\mathsf{Z}(u)\frac{\partial\left[\mathsf{F}_{Q_0^*}(u)\right]_{i_5 i_5}}{\partial y}\right) \quad \text{(B.53)}$$

The finity of (B.53) is implied by (B.30), (∗) and the statements in Sections B.3.2 and B.3.3 which completes the proof.

B.3.6. The Proof that the Expectation of $\frac{\partial^2 l_0(u)}{\partial u_2 \partial u_2'}$ is Finite

It has to be shown that the following expectations are finite:

$$\mathsf{E}\left(\frac{\partial \mathsf{vecl}\left(\mathsf{F}_{R_0}(u)\right)'}{\partial u_2}\mathsf{D}_{p,-}^{+}\left(\mathsf{F}_{R_0}(u)^{-1} \otimes \mathsf{F}_{R_0}(u)^{-1}\right)\mathsf{D}_{p,-}\frac{\partial \mathsf{vecl}\left(\mathsf{F}_{R_0}(u)\right)}{\partial u_2'}\right) \quad \text{(B.54)}$$

$$\mathsf{E}\left(\frac{\partial \mathsf{vecl}\left(\mathsf{F}_{R_0}(u)\right)'}{\partial u_2}\mathsf{D}_{p,-}^{+}\left(\mathsf{F}_{R_0}(u)^{-1}\mathsf{z}_t(u)\mathsf{z}_t'(u)\mathsf{F}_{R_0}(u)^{-1} \otimes \mathsf{F}_{R_0}(u)^{-1}\right)\mathsf{D}_{p,-}\frac{\partial \mathsf{vecl}\left(\mathsf{F}_{R_0}(u)\right)}{\partial u_2'}\right) \quad \text{(B.55)}$$

$$\mathsf{E}\left(\frac{\partial \mathsf{vecl}\left(\mathsf{F}_{R_0}(u)\right)'}{\partial u_2}\mathsf{D}_{p,-}^{+}\left(\mathsf{F}_{R_0}(u)^{-1} \otimes \mathsf{F}_{R_0}(u)^{-1}\mathsf{z}_t(u)\mathsf{z}_t'(u)\mathsf{F}_{R_0}(u)^{-1}\right)\mathsf{D}_{p,-}\frac{\partial \mathsf{vecl}\left(\mathsf{F}_{R_0}(u)\right)}{\partial u_2'}\right) \quad \text{(B.56)}$$

$$\mathsf{E}\left[\left(\mathsf{vecl}\left(\mathsf{F}_{R_0}(u)^{-1}\left[\mathbb{I}_p - \mathsf{z}_t(u)\mathsf{z}_t'(u)\mathsf{F}_{R_0}(u)^{-1}\right]\right)'\mathsf{D}_{p,-}^{+} \otimes \frac{\partial \mathsf{vech}\left(\mathsf{F}_{Q_0}(u)\right)'}{\partial u_2}\mathsf{D}_p'\right)\frac{\partial \mathsf{vecl}\left(\mathsf{F}_{Q_0^*}(u) \otimes \mathsf{F}_{Q_0^*}(u)\right)}{\partial u_2'}\right] \quad \text{(B.57)}$$

$$\mathsf{E}\left[\left(\mathsf{vecl}\left(\mathsf{F}_{R_0}(u)^{-1}\left[\mathbb{I}_p - \mathsf{z}_t(u)\mathsf{z}_t'(u)\mathsf{F}_{R_0}(u)^{-1}\right]\right)'\mathsf{D}_{p,-}^{+} \otimes \frac{\partial \mathsf{vech}\left(\mathsf{F}_{Q_0^*}(u)\right)'}{\partial u_2}\mathsf{D}_p'\right)\left(\mathbb{I}_p \otimes \mathsf{K}_{pp} \otimes \mathbb{I}_p\right)\right.$$
$$\left. \times \left[\frac{\partial \mathsf{vec}\left(\mathsf{F}_{Q_0^*}(u)\mathsf{F}_{Q_0}(u)\right)}{\partial u_2'} \otimes \mathsf{vec}\left(\mathbb{I}_p\right) + \mathsf{vec}\left(\mathbb{I}_p\right) \otimes \frac{\partial \mathsf{vec}\left(\mathsf{F}_{Q_0^*}(u)\mathsf{F}_{Q_0}(u)\right)}{\partial u_2'}\right]\right] \quad \text{(B.58)}$$

$$\mathsf{E}\left[\left(\mathsf{vecl}\left(\mathsf{F}_{R_0}(u)^{-1}\left[\mathbb{I}_p - \mathsf{z}_t(u)\mathsf{z}_t'(u)\mathsf{F}_{R_0}(u)^{-1}\right]\right)'\mathsf{D}_{p,-}^{+}\left(\mathsf{F}_{Q_0}(u)^* \otimes \mathsf{F}_{Q_0}(u)^*\right)\mathsf{D}_p \otimes \mathbb{I}_{p^-+2}\right)\right.$$
$$\left. \times \frac{\partial}{\partial u_2'}\mathsf{vec}\left(\frac{\partial \mathsf{vech}\left(\mathsf{F}_{Q_0}(u)\right)'}{\partial u_2}\right)\right] \quad \text{(B.59)}$$

$$\mathsf{E}\left[\left(\mathsf{vecl}\Big(\mathsf{F}_{R_0}(u)^{-1}\left[\mathbb{I}_p-\mathsf{z}_t(u)\mathsf{z}_t'(u)\mathsf{F}_{R_0}(u)^{-1}\right]\Big)'\mathsf{D}_{p,-}^{+}\Big(\mathsf{F}_{Q_0^*}(u)\mathsf{F}_{Q_0}(u)\otimes\mathbb{I}_p+\mathbb{I}_p\otimes\mathsf{F}_{Q_0^*}(u)\mathsf{F}_{Q_0}(u)\Big)\mathsf{D}_p\otimes\mathbb{I}_{p^-+2}\right)\right.$$
$$\left.\times\ \frac{\partial}{\partial u_2'}\mathsf{vec}\left(\frac{\partial\mathsf{vech}\left(\mathsf{F}_{Q_0^*}(u)\right)'}{\partial u_2}\right)\right]. \tag{B.60}$$

The following statements will be useful for the next parts of this section:

- $$\mathsf{d}(i,j):=\Big([\mathsf{F}_{Q_0}(u)]_{ii}\,[\mathsf{F}_{Q_0}(u)]_{jj}\Big)^{-1/2}\overset{a.s.}{<}\frac{1-b}{1-a-b} \tag{B.61}$$
- $$\frac{\mathsf{G}_0^z(k,k,u)}{[\mathsf{F}_{Q_0}(u)]_{kk}}\le\frac{\mathsf{G}_0^z(k,k,u)}{a\mathsf{G}_0^z(k,k,u)}=\frac{1}{a} \tag{B.62}$$
- $$[\mathsf{F}_{Q_0}(u)]_{ij}\,\mathsf{d}(i,j)\le 1\quad\Rightarrow\quad\frac{[\mathsf{F}_{Q_0}(u)]_{ij}}{[\mathsf{F}_{Q_0}(u)]_{ii}^{1/2}}\le[\mathsf{F}_{Q_0}(u)]_{jj}^{1/2}<[1+a\mathsf{G}_0^z(j,j,u)]^{1/2}. \tag{B.63}$$

The finity of (B.55) and (B.56) is a direct consequence of the finity of the following expectations for $u\in U$, $x\in\{a,b,q_1,\dots,q_{p^-}\}$ and for all $i_1,\dots,i_5\in\{1,\dots,p\}$:

$$\mathsf{E}\left(\left[\frac{\partial\left[\mathsf{F}_{Q_t^*}(u)\right]_{i_1i_1}}{\partial x}\right]^2\left[\frac{\partial\left[\mathsf{F}_{Q_t}(u)\right]_{i_2i_3}}{\partial\left[\mathsf{F}_{Q_t}(u)\right]_{i_2i_2}^{1/2}}\right]^2\left|r_{i_4i_5,0}^-(u)\right|^3\mathsf{Z}(u)\right)\le\delta_*^3\mathsf{E}\left(\left[\frac{\partial\left[\mathsf{F}_{Q_t^*}(u)\right]_{i_1i_1}}{\partial x}\right]^2[1+a\mathsf{G}_0^z(i_3,i_3,u)]\,\mathsf{Z}(u)\right)$$
$$\mathsf{E}\left(\left[\frac{\partial\left[\mathsf{F}_{Q_t^*}(u)\right]_{i_1i_1}}{\partial x}\right]^2\left[\mathsf{F}_{Q_0^*}(u)\right]^4\left|r_{i_4i_5,0}^-(u)\right|^3\mathsf{Z}(u)\right)\le\delta_*^3\mathsf{E}\left(\left[\frac{\partial\left[\mathsf{F}_{Q_t^*}(u)\right]_{i_1i_1}}{\partial x}\right]^2\left(\frac{1-b}{1-a-b}\right)^2\mathsf{Z}(u)\right)$$

which is implied by (B.30), $(*)$ and the statements in Section B.3.3. Analogously, the finity of (B.54) is obtained by substituting $\left|r_{i_4i_5,0}^-(u)\right|^3\mathsf{Z}(u)$ by $\left|r_{i_4i_5,0}^-(u)\right|^2\overset{a.s.}{<}\delta_*^2$.

Furthermore, the finity of (B.57) and (B.58) is a consequence of the finity of the following expectations for all $u\in U$, $i_1,\dots,i_5\in\{1,\dots,p\}$ and $x,y\in\{a,b,q_1,\dots,q_{p^-}\}$:

$$\mathsf{E}\left(\left|r_{i_1i_2,0}^-(u)\right|^2\mathsf{Z}(u)\frac{\partial\left[\mathsf{F}_{Q_0}(u)\right]_{i_3i_4}}{\partial x}\frac{\partial\left[\mathsf{F}_{Q_0^*}(u)\right]_{i_5i_5}}{\partial y}\left[\mathsf{F}_{Q_t}(u)\right]_{i_6i_6}^{-\frac{1}{2}}\right)$$
$$\le\delta_*^2\left(\tfrac{1-b}{1-a-b}\right)^{\frac{1}{2}}\mathsf{E}\left(\mathsf{Z}(u)\frac{\partial\left[\mathsf{F}_{Q_0}(u)\right]_{i_3i_4}}{\partial x}\frac{\partial\left[\mathsf{F}_{Q_0^*}(u)\right]_{i_5i_5}}{\partial y}\right)$$
$$\mathsf{E}\left(\left|r_{i_1i_2,0}^-(u)\right|^2\mathsf{Z}(u)\left[\frac{\partial\left[\mathsf{F}_{Q_0^*}(u)\right]_{i_3i_3}}{\partial x}\right]^2\left[\mathsf{F}_{Q_0}(u)\right]_{i_3i_3}\right)\le\delta_*^2\mathsf{E}\left(\mathsf{Z}(u)\left[\frac{\partial\left[\mathsf{F}_{Q_0}(u)\right]_{i_3i_3}}{\partial x}\right]^2\right)$$
$$\mathsf{E}\left(\left|r_{i_1i_2,0}^-(u)\right|^2\mathsf{Z}(u)\frac{\partial\left[\mathsf{F}_{Q_0^*}(u)\right]_{i_3i_3}}{\partial x}\frac{\partial\left[\mathsf{F}_{Q_0^*}(u)\right]_{i_3i_3}}{\partial y}\left[\mathsf{F}_{Q_0^*}(u)\right]_{i_3i_3}\right)\le\delta_*^2\left(\tfrac{1-b}{1-a-b}\right)^{1/2}\mathsf{E}\left(\mathsf{Z}(u)\left[\frac{\partial\left[\mathsf{F}_{Q_0^*}(u)\right]_{i_3i_3}}{\partial x}\right]^2\right).$$

Finally, with

$$\frac{\partial}{\partial u_2'}\mathsf{vec}\left(\frac{\partial \mathsf{vech}\left(\mathsf{F}_{Q_0^*}(u)\right)}{\partial u_2'}\right) = \begin{bmatrix} \frac{\partial^2 \mathsf{vech}\left(F_{Q_0^*}(u)\right)}{(\partial a)^2} & \frac{\partial^2 \mathsf{vech}\left(F_{Q_0^*}(u)\right)}{\partial a \partial b} & \mathbf{0}_{p^+ \times p^-} \\ \frac{\partial^2 \mathsf{vech}\left(F_{Q_0^*}(u)\right)}{\partial a \partial b} & \frac{\partial^2 \mathsf{vech}\left(F_{Q_0^*}(u)\right)}{(\partial b)^2} & \mathbf{0}_{p^* \times p^-} \\ \mathbf{0}_{p^+ p^-} & \mathbf{0}_{p^+ p^-} & \mathbf{0}_{p^+ p^- \times p^-} \end{bmatrix}$$

the terms (B.59) and (B.60) are finite if the same applies to the following expectations for all $u \in U$, $i_1, \dots, i_6 \in \{1, \dots, p\}$ and $x, y \in \{a, b, q_1, \dots, q_{p^-}\}$:

$$\mathsf{E}\left(\left|r^-_{i_1 i_2, 0}(u)\right|^2 \mathsf{Z}(u)\left[\mathsf{F}_{Q_0}(u)\right]^{-1}_{i_3 i_3} \frac{\partial^2 \left[\mathsf{F}_{Q_0}(u)\right]_{i_4 i_5}}{\partial x \partial y}\right) \leq \delta_*^2 \frac{1-b}{1-a-b}\mathsf{E}\left(\mathsf{Z}(u)\frac{\partial^2 \left[\mathsf{F}_{Q_0}(u)\right]_{i_4 i_5}}{\partial x \partial y}\right) \tag{B.64}$$

$$\mathsf{E}\left(\left|r^-_{i_1 i_2, 0}(u)\right|^2 \mathsf{Z}(u)\frac{\left[\mathsf{F}_{Q_0}(u)\right]_{i_3 i_4}}{\left[\mathsf{F}_{Q_0}(u)\right]^{1/2}_{i_3 i_3}} \frac{\partial^2 \left[\mathsf{F}_{Q_0}(u)\right]_{i_5 i_6}}{\partial x \partial y}\right) \leq \delta_*^2 \mathsf{E}\left(\mathsf{Z}(u)\left[1 + a\mathsf{G}_0^z(i_4, i_4, u)\right]^{1/2} \frac{\partial^2 \left[\mathsf{F}_{Q_0}(u)\right]_{i_5 i_6}}{\partial x \partial y}\right). \tag{B.65}$$

With (B.30), $(*)$ and the statements from Section B.3.3, the terms (B.64) and (B.65) are finite wich completes the proof and yields $\mathsf{E}\left(\frac{\partial^2 l_0(u)}{\partial u_2 \partial u_2'}\right) < \infty$.

B.4. The Proof of Theorem 4.1

We use Theorem A.2.2 in White (1994). Since U is a compact set and $l_t''(u)$ is continuous in u for all y_t and measurable in y_t for all $u \in U$, it remains to verify the dominance condition. Choosing the dominating function as $\sup_{u \in U} |l_t''(u)|$ and using the results of Section B.3 implies $\mathsf{E} \sup_{u \in U} |l_t''(u)| < \infty$. This yields

$$\sup_{u \in U}\left|\frac{1}{m}\sum_{i=1}^m l_i''(u) - \mathsf{E}\left(l_0''(u)\right)\right| = \frac{1}{m}\sup_{u \in U}\left|\sum_{i=1}^m \left[l_i''(u) - \mathsf{E}\left(l_0''(u)\right)\right]\right| \overset{a.s.}{\to} 0. \tag{B.66}$$

Recall that with Lemmas B.17 and B.23, we have

$$\sup_{u \in U}\left|\sum_{i=1}^n \left[\hat{l}_i'(u) - l_i'(u)\right]\right| \overset{a.s.}{=} \mathcal{O}(1). \tag{B.67}$$

Using the consistency of the QMLE as well as (B.66) and (B.67), we can argue along the lines of Berkes et al. (2004) that

$$\sup_{1\leq k<mB} \frac{\left|\sum_{i=m+1}^{m+k} \widehat{D}_m^{-1/2} \hat{l}_i'\left(\hat{\theta}_m\right)\right|}{m^{1/2}\left(1+\frac{k}{m}\right)\mathsf{b}\left(\frac{k}{m}\right)} \xrightarrow{d} \sup_{t\in(0,B]} \frac{\left|\left[W_{\boldsymbol{D}}(1+t)-(1+t)W_{\boldsymbol{D}}(1)\right]\boldsymbol{D}^{-1/2}\right|}{(1+t)\mathsf{b}\left(t\right)}.$$

where the limit process $\{W_{\boldsymbol{D}}(s),\ s\in[0,\infty)\}$ is a d-variate mean zero Gaussian process with covariance function $\mathsf{E}\left(W_{\boldsymbol{D}}^T(k)W_{\boldsymbol{D}}(l)\right) = \min\{k,l\}\,\boldsymbol{D}$. A simple recalculation of the properties of the processes yieds that $\left\{\left[W_{\boldsymbol{D}}(1+t)-(1+t)W_{\boldsymbol{D}}(1)\right]\boldsymbol{D}^{-1/2},\ t\in[0,\infty)\right\}$ and $\{\mathcal{G}(t),\ t\in[0,\infty)\}$ possess the same distribution which completes the proof. ■

B.5. The Proof of Theorem 4.2

Assume that the vector of parameters $\boldsymbol{\theta}$ changes to $\boldsymbol{\theta}^*$ at the k^*-th point of the monitoring period and there is a positive constant λ^* that determines the fraction of the monitoring period where the changepoint is located, i.e. $\lambda^* := \frac{k^*}{mB}$. To avoid that λ^* is shrinked towards zero with m tending to infinity we assume λ^* as constant over time. In the following, we consider the decomposition for $k>k^*$:

$$\frac{\sum_{i=m+1}^{m+k} l_i'\left(\hat{\theta}_m\right)}{m^{1/2}\left(1+\frac{k}{m}\right)\mathsf{b}\left(\frac{k}{m}\right)} = \frac{\sum_{i=m+1}^{m+k^*-1} l_i'\left(\hat{\theta}_m\right)}{m^{1/2}\left(1+\frac{k}{m}\right)\mathsf{b}\left(\frac{k}{m}\right)} + \frac{\sum_{i=m+k^*}^{m+k} l_i'\left(\hat{\theta}_m\right)}{m^{1/2}\left(1+\frac{k}{m}\right)\mathsf{b}\left(\frac{k}{m}\right)}. \tag{B.68}$$

For the first summand in (B.68) that sums the gradient contributions of pre break observations, we obtain along the lines of the proof of Theorem 4.1 or the proof of Theorem 3.1. in Berkes et al. (2004):

$$\sup_{k^*\leq k<\infty} \frac{\left|\sum_{i=m+1}^{m+k^*-1} l_i'\left(\hat{\theta}_m\right)\right|}{m^{1/2}\left(1+\frac{k}{m}\right)\mathsf{b}\left(\frac{k}{m}\right)} \xrightarrow{d} \sup_{t\in[\lambda^*B,\infty)} \frac{\left|W_{\boldsymbol{D}}\left(1+\lambda^*B\right)-\left(1+\lambda^*B\right)W_{\boldsymbol{D}}\left(1\right)\right|}{(1+t)\mathsf{b}(t)}. \tag{B.69}$$

The second summand in (B.68) contains the gradient contributions of the post break observations that are determined by parameter vector $\boldsymbol{\theta}^*$. Thus, an expansion of $l_i'\left(\hat{\theta}_m\right)$ into a Taylor series at $\boldsymbol{\theta}^*$ yields

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

as $m \to \infty$. Along the lines of Theorem 4.4 in Berkes et al. (2004), we can deduce

$$\left(\hat{\theta}_m - \boldsymbol{\theta}^* \right) = \left(\hat{\theta}_m - \boldsymbol{\theta} \right) + \left(\boldsymbol{\theta} - \boldsymbol{\theta}^* \right) = -\frac{1}{m} \sum_{i=1}^{m} \left[\mathsf{E} \left(l_0''(\boldsymbol{\theta}) \right) \right]^{-1} l_i'(\boldsymbol{\theta}) \left[1 + o_P(1) \right] + \left(\boldsymbol{\theta} - \boldsymbol{\theta}^* \right).$$

Note that since $\boldsymbol{\theta}^*$ is in the interior of U, there exists U_2 a neighbourhood of $\boldsymbol{\theta}^*$ inside of which $\frac{1}{m} \sum_{i=1}^{m} l_i''(u)$ converges uniformly to its expectation with Theorem A.2.2 in White (1994) and the results of Section B.3 that imply the finity of the expectation of the dominating function $\sup_{u \in U} |l_i''(u)|$. This property implies the convergence in probability of

$$\sup_{k^* \le k < \infty} \frac{\left| \left(\hat{\theta}_m - \boldsymbol{\theta}^* \right)' \sum_{i=m+k^*}^{m+k} l_i'' \left(\boldsymbol{\theta}^* \right) - \left[-\frac{k-k^*+1}{m} \left[\mathsf{E} \left(l_0''(\boldsymbol{\theta}) \right) \right]^{-1} \mathsf{E} \left(l_0'' \left(\boldsymbol{\theta}^* \right) \right) \sum_{i=1}^{m} l_i'(\boldsymbol{\theta}) + \left(k - k^* + 1 \right) \left(\boldsymbol{\theta} - \boldsymbol{\theta}^* \right)' \mathsf{E} \left(l_0'' \left(\boldsymbol{\theta}^* \right) \right) \right] \right|}{m^{1/2} \left(1 + \frac{k}{m} \right) \mathsf{b} \left(\frac{k}{m} \right)}$$

to zero. Furthermore, the triangle inequality yields

$$\sup_{k^* \le k < \infty} \frac{\left| \left[\sum_{i=m+k^*}^{m+k} l_i' \left(\boldsymbol{\theta}^* \right) - \frac{k-k^*+1}{m} \left[\mathsf{E} \left(l_0''(\boldsymbol{\theta}) \right) \right]^{-1} \mathsf{E} \left(l_0'' \left(\boldsymbol{\theta}^* \right) \right) \sum_{i=1}^{m} l_i'(\boldsymbol{\theta}) \right] - \left(k - k^* + 1 \right) \left(\boldsymbol{\theta} - \boldsymbol{\theta}^* \right)' \mathsf{E} \left(l_0'' \left(\boldsymbol{\theta}^* \right) \right) \right|}{m^{1/2} \left(1 + \frac{k}{m} \right) \mathsf{b} \left(\frac{k}{m} \right)}$$

$$\ge \left| \sup_{k^* \le k < \infty} \frac{\left| \sum_{i=m+k^*}^{m+k} l_i' \left(\boldsymbol{\theta}^* \right) - \frac{k-k^*+1}{m} \left[\mathsf{E} \left(l_0''(\boldsymbol{\theta}) \right) \right]^{-1} \mathsf{E} \left(l_0'' \left(\boldsymbol{\theta}^* \right) \right) \sum_{i=1}^{m} l_i'(\boldsymbol{\theta}) \right|}{m^{1/2} \left(1 + \frac{k}{m} \right) \mathsf{b} \left(\frac{k}{m} \right)} - \sup_{k^* \le k < \infty} \frac{\left| \left(k - k^* + 1 \right) \left(\boldsymbol{\theta} - \boldsymbol{\theta}^* \right)' \mathsf{E} \left(l_0'' \left(\boldsymbol{\theta}^* \right) \right) \right|}{m^{1/2} \left(1 + \frac{k}{m} \right) \mathsf{b} \left(\frac{k}{m} \right)} \right|$$

with

$$\sup_{k^* \le k < \infty} \frac{\left| \sum_{i=m+k^*}^{m+k} l_i' \left(\boldsymbol{\theta}^* \right) - \frac{k-k^*+1}{m} \left[\mathsf{E} \left(l_0''(\boldsymbol{\theta}) \right) \right]^{-1} \mathsf{E} \left(l_0'' \left(\boldsymbol{\theta}^* \right) \right) \sum_{i=1}^{m} l_i'(\boldsymbol{\theta}) \right|}{m^{1/2} \left(1 + \frac{k}{m} \right) \mathsf{b} \left(\frac{k}{m} \right)}$$

$$\xrightarrow{d} \sup_{t \in (\lambda^*, \infty)} \frac{\left| W_{\boldsymbol{D}^*}(1+t) - W_{\boldsymbol{D}^*}(1+\lambda^*) - (t - \lambda^*) \left[\mathsf{E} \left(l_0''(\boldsymbol{\theta}) \right) \right]^{-1} \mathsf{E} \left(l_0'' \left(\boldsymbol{\theta}^* \right) \right) W_{\boldsymbol{D}}(1) \right|}{(1+t) \, \mathsf{b} \, (t)} \tag{B.70}$$

where $\boldsymbol{D}^* := \mathsf{Cov} \left[l_0' \left(\boldsymbol{\theta}^* \right) \right]$ and $\{ W_{\boldsymbol{D}^*}(t), \ t \in [0, \infty) \}$ a zero mean Gaussian process with covariance function $K_{\boldsymbol{D}^*}(s,t) = \min \{ s, t \} \, \boldsymbol{D}^*$.

In addition, we have

$$\sup_{k^* \le k < \infty} \frac{\left|(k - k^* + 1)(\boldsymbol{\theta} - \boldsymbol{\theta}^*)' \, \mathsf{E}\left(l_0''(\boldsymbol{\theta}^*)\right)\right|}{m^{1/2}\left(1 + \frac{k}{m}\right) \mathsf{b}\left(\frac{k}{m}\right)} = \sqrt{m} \sup_{k^* \le k < \infty} \frac{\left|\frac{k - k^* + 1}{m}(\boldsymbol{\theta} - \boldsymbol{\theta}^*)' \, \mathsf{E}\left(l_0''(\boldsymbol{\theta}^*)\right)\right|}{\left(1 + \frac{k}{m}\right) \mathsf{b}\left(\frac{k}{m}\right)} \to \infty$$

as $m \to \infty$ almost surely. Since the thresholds (B.69) and (B.70) are stochastically bounded if the variable parts in $\mathsf{b}(\cdot)$ are chosen such that the procedure keeps its size under the null, this completes the proof. ■

C. The Proofs of the Calculation Rules in Section A.2

C.1. The Proof of CR1:

Recall that $\mathsf{D}_{p,-}$ is a $(p^2 \times p^-)$ matrix while $\mathsf{D}_{p,-}^+$ is of dimension $(p^- \times p^2)$. Furthermore, $\mathsf{D}_{p,-}$ is a matrix whose entries are zero or one and whose column sums are all equal to 2 whereas the row sums are 1, i.e. every column has two ones but none of them is in the same row as any of the ones in a different column.

$$\mathsf{D}_{p,-} := \left(\mathsf{d}_{\cdot 1}, \mathsf{d}_{\cdot 2} \ldots, \mathsf{d}_{\cdot p^+}\right) \;\Rightarrow\; \mathsf{d}_{\cdot i}'\mathsf{d}_{\cdot j} = \begin{cases} 2, & i = j \\ 0 & i \neq j \end{cases} \;\Rightarrow\; \mathsf{D}_{p,-}'\mathsf{D}_{p,-} = 2\mathbb{I}_{p^+}$$

Thus, the eigenvalues of $\mathsf{D}_{p,-}'\mathsf{D}_{p,-}$ are all equal to 2 with multiplicity p^-. According to page 335 in Seber (2008), this implies that the singular values of $\mathsf{D}_{p,-}$ are $\sqrt{2}$ and the singular value decomposition of $\mathsf{D}_{p,-}$ is given as

$$\mathsf{D}_{p,-} = U \begin{bmatrix} \sqrt{2} \cdot \mathbb{I}_{p^-} \\ \mathbf{0}_{p^+ \times p^-} \end{bmatrix} V'$$

with orthogonal matrices $U \sim (p^2 \times p^2)$ and $V \sim (p^- \times p^-)$ that contain the orthonormalized eigenvectors of $\mathsf{D}_{p,-}\mathsf{D}_{p,-}'$ and $\mathsf{D}_{p,-}'\mathsf{D}_{p,-}$, respectively. Hence, with 5.5.1(9) of Lütkepohl (1996), the singular value decomposition of the Moore Penrose inverse of $D_{p,-}$ is given as

$$\mathsf{D}_{p,-}^+ = V \begin{bmatrix} \frac{1}{\sqrt{2}} \cdot \mathbb{I}_{p^-} \\ \mathbf{0}_{p^+ \times p^-} \end{bmatrix} U' = \frac{1}{2} V \begin{bmatrix} \sqrt{2} \cdot \mathbb{I}_{p^-} \\ \mathbf{0}_{p^+ \times p^-} \end{bmatrix} U' = \frac{1}{2}\mathsf{D}_{p,-}'.$$

C.2. The Proof of CR4:

We use 17.30(h) in Seber (2008):

$$\begin{aligned} &Z \sim (l \times m),\ U(Z) \sim (q \times r),\ V(Z) \sim (r \times t) \\ &\Rightarrow \frac{\partial \mathsf{vec}\,(UV)}{\partial \mathsf{vec}\,(Z)'} = (V \otimes \mathbb{I}_q)' \frac{\partial \mathsf{vec}\,(U)}{\partial \mathsf{vec}\,(Z)'} + (\mathbb{I}_t \otimes U) \frac{\partial \mathsf{vec}\,(V)}{\partial \mathsf{vec}\,(X)'}. \end{aligned}$$

We consider symmetric $(n \times n)$ matrices X and $Y(X)$ and choose $U := XY$ and $V := X$.

This yields $q = r = t = l = m = n$ and

$$\begin{aligned} \frac{\partial \mathsf{vec}(XY(X)X)}{\partial \mathsf{vec}(X)'} &= (X \otimes \mathbb{I}_n)' \frac{\partial \mathsf{vec}\,(XY)}{\partial \mathsf{vec}\,(X)'} + (\mathbb{I}_n \otimes XY) \frac{\partial \mathsf{vec}\,(X)}{\partial \mathsf{vec}\,(X)'} \\ &\overset{U=X,\ V=Y}{=} (X \otimes \mathbb{I}_n) \left[(Y \otimes \mathbb{I}_n)' \frac{\partial \mathsf{vec}\,(X)}{\partial \mathsf{vec}\,(X)'} + (\mathbb{I}_n \otimes X) \frac{\partial \mathsf{vec}\,(Y)}{\partial \mathsf{vec}\,(X)'} \right] + (\mathbb{I}_n \otimes XY) \\ &= (X \otimes \mathbb{I}_n)(Y \otimes \mathbb{I}_n) + (X \otimes \mathbb{I}_n)(\mathbb{I}_p \otimes X) \frac{\partial \mathsf{vec}\,(Y)}{\partial \mathsf{vec}\,(X)'} + (\mathbb{I}_n \otimes XY) \\ &= (X \otimes X) \frac{\partial \mathsf{vec}\,(Y)}{\partial \mathsf{vec}\,(X)'} + (XY \otimes \mathbb{I}_n + \mathbb{I}_n \otimes XY)\,. \end{aligned}$$

■