Online Supplement to “Goodness-of-fit tests for multivariate copula-based time series models”

Betina Berghaus and Axel Bücher*

November 24, 2015

Abstract

This supplementary material contains a detailed description of a procedure to simulate a dependent multiplier sequence, a section illustrating the runtime performance of the proposed tests and a section containing the proofs of Lemma C.4, C.6 and C.7 from the main paper.

E Practical issues for the choice of the multipliers

As described in Section 3.1, the bootstrap version of the test statistic in Section 2 depends crucially on a sequence of dependent multipliers satisfying the conditions in Definition A.1 in Appendix A. Within this section, we present a slight modification of the approach in Bücher and Kojadinovic (2014), Section 5.2, originating from Bühlmann (1993) and Politis and White (2004), showing how to simulate such a sequence.

*Ruhr-Universität Bochum, Fakultät für Mathematik, Universitätsstr. 150, 44780 Bochum, Germany. E-mail: betina.berghaus@rub.de, axel.buecher@rub.de.

E.1 Simulating dependent multipliers

First of all, choose a bandwidth parameter ℓ_n (see Section E.2 below) and set $b_n = \text{round}\{(\ell_n + 1)/2\}$. Let κ be a positive, bounded function which is symmetric around zero such that $\kappa(x) > 0$ for all $x \in (-1, 1)$. Within the simulation study, we use the Parzen-kernel defined as

$$\kappa_P(x) = (1 - 6x^2 + 6|x|^3)\mathbb{1}(|x| \leq 1/2) + 2(1 - |x|)^3\mathbb{1}(1/2 < |x| \leq 1).$$

Now, for $j = 1, \ldots, \ell_n$, define weights $\tilde{\omega}_{jn} = \omega_{jn}(\sum_{i=1}^{\ell_n} \omega_{in}^2)^{-1/2}$ where $\omega_{jn} = \kappa\{(j - b_n)/b_n\}$. Finally, let $\xi_1, \ldots, \xi_{n+2b_n-2}$ be i.i.d. random variables which are independent from the sample $\boldsymbol{X}_1, \ldots, \boldsymbol{X}_n$ with $\mathbb{E}[\xi_1] = 0$, $\mathbb{E}[\xi_1^2] = 1$ and $\mathbb{E}[|\xi_1|^m] < \infty$, for any $m \geq 1$. Then, the sequence of random variables $Z_{1,n}, \ldots, Z_{n,n}$ defined as

$$Z_{i,n} = \sum_{j=1}^{\ell_n} \tilde{\omega}_{jn} \xi_{j+i-1}$$

asymptotically satisfies the conditions in Definition A.1, see Bücher and Kojadinovic (2014). For a detailed discussion on the properties of these multipliers and an alternative simulation procedure for the multpliers see Bücher and Kojadinovic (2014).

E.2 Choosing the bandwidth parameter ℓ_n

The simulation method described in the previous section depends on the choice of the 'bandwidth' parameter ℓ_n. Within the context of empirical copulas, Bücher and Kojadinovic (2014) derived a closed-form formula for a theoretically optimal choice, where optimality is to be understood as optimality with respect to a certain MSE-minimizing criterion. Moreover, they proposed a data-adaptive estimation procedure for this (theoretically) optimal bandwidth. In the following, we adapt their approach to the processes underlying the goodness-of-fit tests in Section 2.

For $\boldsymbol{u}, \boldsymbol{v} \in [0, 1]^d$, let $\sigma(\boldsymbol{u}, \boldsymbol{v})$ denote the characterizing covariance kernel of the process $\boldsymbol{u} \mapsto \mathbb{C}_C(\boldsymbol{u}) - \nabla C_{\theta_0}(\boldsymbol{u})\Theta$. Its bootstrap approximation, for a fixed $b \in \{1, \dots, B\}$, is given by $\mathbb{C}_n^{(b)}(\boldsymbol{u}) - \nabla C_{\theta_n}(\boldsymbol{u})\Theta_n^{(b)}$. The main idea of the subsequent developments is as follows: if $\hat{\sigma}_n$ denotes the covariance kernel of the bootstrap process, conditional on the data, then a theoretically optimal choice of ℓ_n is given by the minimizer of the integrated mean squared error of $\hat{\sigma}_n(\boldsymbol{u}, \boldsymbol{v})$, seen as an estimator for $\sigma(\boldsymbol{u}, \boldsymbol{v})$, with respect to ℓ_n. Unfortunately, necessary closed form expressions for the mean or the variance of $\hat{\sigma}_n$ are out of reach, whence we follow the proposal of Bücher and Kojadinovic (2014) and consider an asymptotically equivalent form of $\mathbb{C}_n^{(b)}(\boldsymbol{u}) - \nabla C_{\theta_n}(\boldsymbol{u})\Theta_n^{(b)}$ instead, for which calculations are feasible.

More precisely, define $\tilde{\mathbb{B}}_n^{(b)}(\boldsymbol{u}) = n^{-1/2} \sum_{i=1}^n Z_{i,n}^{(b)}\{\mathbb{1}(\boldsymbol{U}_i \leq \boldsymbol{u}) - C_{\theta_0}(\boldsymbol{u})\}$ and $\tilde{\Theta}_n^{(b)} = n^{-1/2} \sum_{i=1}^n Z_{i,n}^{(b)}\{J_{\theta_0}(\boldsymbol{U}_i) + K_{i,\theta_0}\}$, where K_{i,θ_0} is defined in Theorem 2.2, and let $\tilde{\mathbb{C}}_n^{(b)}(\boldsymbol{u}) = \tilde{\mathbb{B}}_n^{(b)}(\boldsymbol{u}) - \sum_{l=1}^d C_{\theta_0}^{[l]}(\boldsymbol{u})\tilde{\mathbb{B}}_n^{(b)}(\boldsymbol{u}^l)$. It follows from the arguments in the proof of Theorem 2.2 that

$$\sup_{\boldsymbol{u} \in [0,1]^d} \left| \mathbb{C}_n^{(b)}(\boldsymbol{u}) - \nabla C_{\theta_n}(\boldsymbol{u})\Theta_n^{(b)} - \{\tilde{\mathbb{C}}_n^{(b)}(\boldsymbol{u}) - \nabla C_{\theta_0}(\boldsymbol{u})\tilde{\Theta}_n^{(b)}\} \right| = o_P(1).$$

In contrast to $\hat{\sigma}_n$, the (conditional) covariance kernel of the (unobservable) process $\tilde{\mathbb{C}}_n^{(b)}(\boldsymbol{u}) - \nabla C_{\theta_0}(\boldsymbol{u})\tilde{\Theta}_n^{(b)}$ can be calculated explicitly, at least up to the first-order terms. We have

$$\begin{aligned}
\tilde{\sigma}_n(\boldsymbol{u}, \boldsymbol{v}) &= \mathrm{Cov}_Z\{\tilde{\mathbb{C}}_n^{(b)}(\boldsymbol{u}) - \nabla C_{\theta_0}(\boldsymbol{u})\tilde{\Theta}^{(b)}, \tilde{\mathbb{C}}_n^{(b)}(\boldsymbol{v}) - \nabla C_{\theta_0}(\boldsymbol{v})\tilde{\Theta}^{(b)}\} \\
&= \frac{1}{n} \sum_{i,j=1}^n \mathbb{E}_Z[Z_{i,n}^{(b)} Z_{j,n}^{(b)}] f(\boldsymbol{U}_i, \boldsymbol{u}) f(\boldsymbol{U}_j, \boldsymbol{v}) \\
&= \frac{1}{n} \sum_{i,j=1}^n \varphi\{(i-j)/\ell_n\} f(\boldsymbol{U}_i, \boldsymbol{u}) f(\boldsymbol{U}_j, \boldsymbol{v})
\end{aligned}$$

where Cov_Z and $\mathbb{E}_Z$ denote covariance and expectation conditional on the data, respectively, and where, for $i = 1, \dots, n$ and $\boldsymbol{u} \in [0,1]^d$,

$$\begin{aligned}
f(\boldsymbol{U}_i, \boldsymbol{u}) = \mathbb{1}(\boldsymbol{U}_i \leq \boldsymbol{u}) - C_{\theta_0}(\boldsymbol{u}) - \sum_{l=1}^d C_{\theta_0}^{[l]}(\boldsymbol{u})\{\mathbb{1}(U_{il} \leq u_l) - u_l\} \\
- \nabla C_{\theta_0}(\boldsymbol{u})\{J_{\theta_0}(\boldsymbol{U}_i) + K_{\theta_0,i}\}.
\end{aligned}$$

Mimicking the proofs of Proposition 5.1 and Proposition 5.2 in Bücher and Kojadinovic (2014), we obtain the following results regarding bias and variance of $\tilde{\sigma}_n$, seen as an estimator for σ.

Lemma E.1. *Additionally to the conditions assumed in (iii) of Theorem 2.2 suppose that a defined in Condition VI satisfies $a > 3(2+\nu)/\nu$ and that φ defined in Definition A.1 is twice continuously differentiable on $[-1,1]$ with $\varphi''(0) \neq 0$. Then, for any $\boldsymbol{u},\boldsymbol{v} \in [0,1]^d$,*

$$\mathbb{E}[\tilde{\sigma}_n(\boldsymbol{u},\boldsymbol{v})] - \sigma(\boldsymbol{u},\boldsymbol{v}) = \frac{\Gamma(\boldsymbol{u},\boldsymbol{v})}{\ell_n^2} + r_{n,1}(\boldsymbol{u},\boldsymbol{v}),$$

where $\sup_{\boldsymbol{u},\boldsymbol{v}\in[0,1]^d} |r_{1,n}(\boldsymbol{u},\boldsymbol{v})| = o(\ell_n^{-2})$ and $\Gamma(\boldsymbol{u},\boldsymbol{v}) = \frac{\varphi''(0)}{2}\sum_{k\in\mathbb{Z}} k^2\gamma(k,\boldsymbol{u},\boldsymbol{v})$ with $\gamma(k,\boldsymbol{u},\boldsymbol{v}) = \mathrm{Cov}\{f(\boldsymbol{U}_0,\boldsymbol{u}), f(\boldsymbol{U}_k,\boldsymbol{v})\}$.

Additionally, provided the function φ is Lipschitz-continuous and provided $\int_{[0,1]^d}\prod_{l=1}^d r_l(u_l)^{4+2\nu}\mathrm{d}C_{\theta_0}(\boldsymbol{u}) < \infty$, where $r_1,\ldots,r_d$ are defined in Condition V, then

$$\mathrm{Var}\{\tilde{\sigma}_n(\boldsymbol{u},\boldsymbol{v})\} = \frac{\ell_n}{n}\Delta(\boldsymbol{u},\boldsymbol{v}) + r_{n,2}(\boldsymbol{u},\boldsymbol{v}),$$

where $\Delta(\boldsymbol{u},\boldsymbol{v}) = \int_{-1}^1 \varphi(x)^2\mathrm{d}x\{\sigma(\boldsymbol{u},\boldsymbol{u})\sigma(\boldsymbol{v},\boldsymbol{v}) + \sigma(\boldsymbol{u},\boldsymbol{v})^2\}$ and where the remainder term satisfies $\sup_{\boldsymbol{u},\boldsymbol{v}\in[0,1]^d}|r_{n,2}(\boldsymbol{u},\boldsymbol{v})| = o(\ell_n/n)$.

As a consequence of Lemma E.1, the (pointwise) mean integrated squared error of $\tilde{\sigma}_n(\boldsymbol{u},\boldsymbol{v})$ can be written as

$$\mathrm{MSE}\{\tilde{\sigma}_n(\boldsymbol{u},\boldsymbol{v})\} = \frac{\{\Gamma(\boldsymbol{u},\boldsymbol{v})\}^2}{\ell_n^4} + \Delta(\boldsymbol{u},\boldsymbol{v})\frac{\ell_n}{n} + r_{n,1}^2(\boldsymbol{u},\boldsymbol{v}) + r_{n,2}(\boldsymbol{u},\boldsymbol{v}).$$

Furthermore, the integrated mean squared error is given by

$$\mathrm{IMSE}(\tilde{\sigma}_n) = \int_{[0,1]^{2d}} \mathrm{MSE}\{\tilde{\sigma}_n(\boldsymbol{u},\boldsymbol{v})\}\mathrm{d}(\boldsymbol{u},\boldsymbol{v}) = \frac{\bar{\Gamma}^2}{\ell_n^4} + \bar{\Delta}\frac{\ell_n}{n} + o(\ell_n^{-4}) + o(\ell_n/n),$$

where $\bar{\Gamma} = \int_{[0,1]^{2d}}\Gamma(\boldsymbol{u},\boldsymbol{v})\mathrm{d}(\boldsymbol{u},\boldsymbol{v})$ and $\bar{\Delta} = \int_{[0,1]^{2d}}\Delta(\boldsymbol{u},\boldsymbol{v})\mathrm{d}(\boldsymbol{u},\boldsymbol{v})$. Obviously, the function $\ell_n \mapsto \bar{\Gamma}/\ell_n^4 + \bar{\Delta}\ell_n/n$ is minimized for

$$\ell_{n,opt} = \left(\frac{4\bar{\Gamma}^2}{\bar{\Delta}}\right)^{1/5} n^{1/5},$$

which can be considered as a theoretically optimal choice for the bandwidth parameter ℓ_n.

In practice, the unknown quantities in $\ell_{n,opt}$ need to be estimated, namely $\sigma(\boldsymbol{u},\boldsymbol{v}) = \sum_{k\in\mathbb{Z}} \gamma(k,\boldsymbol{u},\boldsymbol{v})$ and $M(\boldsymbol{u},\boldsymbol{v}) = \sum_{k\in\mathbb{Z}} k^2\gamma(k,\boldsymbol{u},\boldsymbol{v})$. For that purpose, we can closely follow Bücher and Kojadinovic (2014) again. Let $L \in \{1,\dots,n\}$ be the smallest number such that the marginal autocorrelations at lag L appear to be negligible, see Bücher and Kojadinovic (2014) and Politis and White (2004) for details. Let K denote the trapezoidal kernel, defined as $K(x) = [\{2(1-|x|)\} \vee 0] \wedge 1$. Then, set

$$\hat{\sigma}_n(\boldsymbol{u},\boldsymbol{v}) = \sum_{k=-L}^{L} K(k/L)\hat{\gamma}_n(k,\boldsymbol{u},\boldsymbol{v}), \quad \hat{M}_n(\boldsymbol{u},\boldsymbol{v}) = \sum_{k=-L}^{L} K(k/L)k^2\hat{\gamma}_n(k,\boldsymbol{u},\boldsymbol{v}),$$

where, for $\boldsymbol{u},\boldsymbol{v} \in [0,1]^d$ and $k \in \{-L,\dots,L\}$,

$$\hat{\gamma}_n(k,\boldsymbol{u},\boldsymbol{v}) = \begin{cases} n^{-1}\sum_{i=1}^{n-k} \hat{f}(\hat{\boldsymbol{U}}_i,\boldsymbol{u})\hat{f}(\hat{\boldsymbol{U}}_{i+k},\boldsymbol{v}), & k \geq 0 \\ n^{-1}\sum_{i=1-k}^{n} \hat{f}(\hat{\boldsymbol{U}}_i,\boldsymbol{u})\hat{f}(\hat{\boldsymbol{U}}_{i+k},\boldsymbol{v}), & k < 0 \end{cases},$$

and where

$$\hat{f}(\hat{\boldsymbol{U}}_i,\boldsymbol{u}) = \mathbb{1}(\hat{\boldsymbol{U}}_i \leq \boldsymbol{u}) - C_n(\boldsymbol{u}) - \sum_{l=1}^{d} C_n^{[l]}(\boldsymbol{u})\{\mathbb{1}(\hat{U}_{il} \leq u_l) - u_l\} \\ - \nabla C_{\theta_n}(\boldsymbol{u})\{J_{\theta_0}(\hat{\boldsymbol{U}}_i) + \hat{K}_{i,n,\theta_n}\}.$$

The plug-in principle finally yields an estimator $\hat{\ell}_{n,opt}$ for $\ell_{n,opt}$.

F Runtime comparison

To compare the performance of our test to the test by Rémillard et al. (2012) in terms of runtime, we simulated $N = 100$ test runs of both tests under the null hypothesis for the Gaussian, the t_4 and the Clayton copula with $\tau = 0.2$. The results can be found in Figure 1, where we depict the relative runtimes r_{pb}/r_{mb} of the test by Rémillard et al. (2012) (r_{pb} for parametric bootstrap) compared to the runtime of the new test (r_{mb} for multiplier bootstrap) as a function of the sample size (with $B = 200$ bootstrap replications) and of the number of bootstrap replications (with sample size $n = 200$), respectively. In all considered scenarios the new test clearly outperforms the test by Rémillard et al. (2012) in terms of computational costs, up to a factor of almost 70 for the Clayton copula with $B = 500$ and $n = 200$.

Note that the function corresponding to the Clayton copula in the left panel of Figure 1 is the only decreasing function. This is due to the fact that, for the Clayton copula and with increasing sample size, the evaluation of the score function in the data points in the multiplier bootstrap requires comparatively more time than the simulation of a Markovian copula process in the parametric bootstrap (even though all involved formulas are explicit).

G Remaining proofs

Proof of Lemma C.4. As in the proof of Lemma C.1, we may assume without loss of generality that $\Omega = \Omega_{n,D}$ for some $D > 0$. Let us write $L_{n2}^{(b)} = A_n + B_n$, where

$$A_n = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} Z_{i,n}^{(b)} \{J_{\theta_0}(\hat{\boldsymbol{U}}_i) - J_{\theta_0}(\boldsymbol{U}_i)\}, \quad B_n = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} Z_{i,n}^{(b)} \{\hat{K}_{i,n,\theta_0} - K_{i,n,\theta_0}\},$$

and consider each term separately. For $\eta \in (0, 1/2)$, set $M_\eta = [\eta, 1-\eta]^d$. By the mean value theorem, there exist intermediate values $\tilde{\boldsymbol{U}}_i$ between $\hat{\boldsymbol{U}}_i$ and $\boldsymbol{U}_i$ such that we can write $A_n = \sum_{l=1}^{d} \{A_{n,l}(M_\eta) + A_{n,l}(M_\eta^C)\}$, where, for any

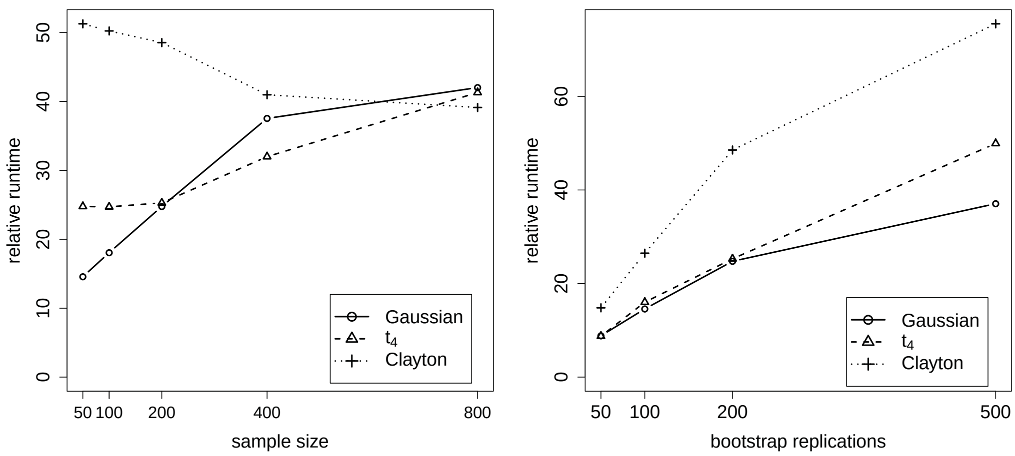

Figure 1: Relative runtimes of the test by Rémillard et al. (2012) compared to the runtimes of the new testing procedure as a function of the sample size (left) and the number of bootstrap replications (right) for the Gaussian, the t_4 and the Clayton copula.

$M \subset [0,1]^d$,

$$A_{n,l}(M) = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} Z_{i,n}^{(b)} J_{\theta_0}^{[l]}(\tilde{\boldsymbol{U}}_i)(\hat{U}_{i,l} - U_{i,l}) \mathbb{1}(\boldsymbol{U}_i \in M).$$

We begin with the treatment of $A_{n,l}(M_\eta^C)$, for a fixed l. Since we may assume $\Omega = \Omega_{n,D}$, we can bound $\|A_{n,l}(M_\eta^C)\|$ by

$$\frac{D}{n} \sum_{i=1}^{n} |Z_{i,n}^{(b)}| q_l(U_{i,l}) \sup_{\frac{|u - U_{i,l}|}{q_l(U_{i,l})} \leq \frac{B}{\sqrt{n}}} \tilde{r}_l(u) \prod_{l' \neq l} \sup_{\frac{|u - U_{i,l'}|}{q_{l'}(U_{i,l'})} \leq \frac{B}{\sqrt{n}}} r_{l'}(u) \mathbb{1}(U_i \in M_\eta^C)$$

As in the proof of Lemma C.1, by Condition V, for sufficiently large n, the expectation of the latter expression converges to zero as $\eta \to 0$. Therefore, it remains to consider $A_{n,l}(M_\eta)$ for fixed $\eta \in (0, 1/2)$.

Since $\max_{i=1}^{n} \|\hat{\boldsymbol{U}}_i - \boldsymbol{U}_i\| \leq D/\sqrt{n}$, we get that $A_{n,l}(M_\eta) = \bar{A}_{n,l}(M_\eta) + o_P(1)$

for $n \to \infty$, where

$$\bar{A}_{n,l}(M_\eta) = \frac{1}{\sqrt{n}} \sum_{i=1}^n Z_{i,n}^{(b)} J_{\theta_0}^{[l]}(\tilde{\boldsymbol{U}}_i)(\hat{U}_{i,l} - U_{i,l}) \mathbb{1}(\boldsymbol{U}_i \in M_\eta, \|\hat{\boldsymbol{U}}_i - \boldsymbol{U}_i\| \le \eta/2).$$

By the Cauchy-Schwarz inequality, we obtain

$$\begin{aligned} \mathbb{E}\big[\|\bar{A}_{n,l}(M_\eta)\|^2\big] &= \mathbb{E}\big[\{\bar{A}_{n,l}(M_\eta)\}'\{\bar{A}_{n,l}(M_\eta)\}\big] \\ &\le \sup_{\boldsymbol{u} \in M_{\eta/2}} \|J_{\theta_0}^{[l]}(\boldsymbol{u})\|^2 \times \mathbb{E}\big[\max_{i=1}^n |\hat{U}_{i,l} - U_{i,l}|^2\big] \times \frac{1}{n} \sum_{i,j=1}^n |\mathbb{E}[Z_{i,n}^{(b)} Z_{j,n}^{(b)}]|. \end{aligned}$$

The first factor on the right-hand side is bounded by Condition V. The second factor is bounded by D^2/n as we may assume that $\Omega = \Omega_{n,D}$. Regarding the third factor, note that $\frac{1}{n}\sum_{i,j=1}^n |\mathbb{E}[Z_{i,n}^{(b)} Z_{j,n}^{(b)}]| = O(\ell_n)$ as shown in (C.3). We can conclude that $\bar{A}_{n,l}(M_\eta)$ is of order $O_p(n^{-1/4-\kappa/2}) = o_P(1)$, and therefore also $A_n = o_P(1)$.

For the proof of the lemma, it remains to be shown that $B_n = o_P(1)$. We can decompose $B_n = \sum_{l=1}^d (B_{n,1,l} + B_{n,2,l} + B_{n,3,l})$, with

$$\begin{aligned} B_{n,1,l} &= \frac{1}{\sqrt{n}} \sum_{i=1}^n Z_{i,n}^{(b)} \frac{1}{n} \sum_{j=1}^n \{J_{\theta_0}^{[l]}(\hat{\boldsymbol{U}}_j) - J_{\theta_0}^{[l]}(\boldsymbol{U}_j)\}\{\mathbb{1}(U_{i,l} \le U_{j,l}) - U_{j,l}\} \\ B_{n,2,l} &= \frac{1}{\sqrt{n}} \sum_{i=1}^n Z_{i,n}^{(b)} \frac{1}{n} \sum_{j=1}^n \{J_{\theta_0}^{[l]}(\hat{\boldsymbol{U}}_j) - J_{\theta_0}^{[l]}(\boldsymbol{U}_j)\}(U_{j,l} - \hat{U}_{j,l}) \\ B_{n,3,l} &= \left\{\frac{1}{\sqrt{n}} \sum_{j=1}^n J_{\theta_0}^{[l]}(\boldsymbol{U}_j)(\hat{U}_{j,l} - U_{j,l})\right\} \times \left\{\frac{1}{n} \sum_{i=1}^n Z_{i,n}^{(b)}\right\}. \end{aligned}$$

$B_{n,3,l}$ converges in probability to zero: the first factor is of order $O_P(1)$ by a similar argumentation as before based on the fact that we may assume $\Omega = \Omega_{n,D}$, and the second factor is of order $O_P(\ell_n^{1/2}/n^{1/2})$ by (C.3).

Regarding $B_{n,1,l}$, we can bound

$$\|B_{n,1,l}\| \le \left\{\frac{1}{\sqrt{n}} \sup_{u \in [0,1]} \Big| \sum_{i=1}^n Z_{i,n}^{(b)} \frac{\mathbb{1}(U_{i,l} \le u) - u}{q_l(u)} \Big| \right\}$$

$$\times \left\{ \frac{1}{n} \sum_{j=1}^n \|J_{\theta_0}^{[l]}(\hat{\boldsymbol{U}}_j) - J_{\theta_0}^{[l]}(\boldsymbol{U}_j)\| q_l(U_{j,l}) \right\}.$$

The first factor on the right-hand side is of order $O_P(1)$ by Lemma D.1. The second factor converges to 0 in probability by the same argumentation as for the treatment of (C.2) in the proof Lemma C.1.

A similar argumentation also works for $B_{n,2,l}$: on the set $\Omega_{n,D}$, we have

$$\|B_{n,2,l}\| \le \left\{ \frac{1}{n} \sum_{i=1}^n |Z_{i,n}^{(b)}| \right\} \times \left\{ \frac{D}{n} \sum_{j=1}^n \|J_{\theta_0}^{[l]}(\hat{\boldsymbol{U}}_j) - J_{\theta_0}^{[l]}(\boldsymbol{U}_j)\| q_l(U_{j,l}) \right\},$$

and this expression is $o_P(1)$ since $\mathbb{E}[|Z_{i,n}^{(b)}|] < \infty$ and by the same reasons as in the proof of Lemma C.1. □

Proof of Lemma C.6. Using Condition III, tightness of the vector of processes follows from marginal tightness of $\tilde{\mathbb{B}}_n$ and $\tilde{\mathbb{B}}_n^{(b)}$, see Theorem 3.1 in Bücher and Kojadinovic (2014).

Regarding weak convergence of the finite dimensional distributions, we only consider $(\tilde{\mathbb{B}}_n, \tilde{\Theta}_n, \tilde{\mathbb{B}}_n^{(b)}, \tilde{\Theta}_n^{(b)})$ for the ease of reading. By the Cramér-Wold device, we have to show that, for any $q, q' \in \mathbb{N}$, any $c_1, \dots, c_q, \bar{c}_1, \dots, \bar{c}_{q'} \in \mathbb{R}$, $c, \bar{c} \in \mathbb{R}^p$ and any $\boldsymbol{u}_1, \dots, \boldsymbol{u}_q, \boldsymbol{v}_1, \dots, \boldsymbol{v}_{q'} \in [0,1]^d$,

$$V_n := \sum_{s=1}^q c_s \tilde{\mathbb{B}}_n(\boldsymbol{u}_s) + \sum_{s=1}^{q'} \bar{c}_s \tilde{\mathbb{B}}_n^{(b)}(\boldsymbol{v}_s) + c'\tilde{\Theta}_n + \bar{c}'\tilde{\Theta}_n^{(b)}$$
$$\rightsquigarrow V := \sum_{s=1}^q c_s \mathbb{B}_C(\boldsymbol{u}_s) + \sum_{s=1}^{q'} \bar{c}_s \mathbb{B}_C^{(b)}(\boldsymbol{v}_s) + c'\Theta + \bar{c}'\Theta^{(b)}.$$

First of all, we decompose

$$V_n = \frac{1}{\sqrt{n}} \sum_{i=1}^n W_i + W_i^{(b)} + T_i + T_i^{(b)}$$

with $W_i = \sum_{s=1}^q c_s \{\mathbb{1}(\boldsymbol{U}_i \le \boldsymbol{u}_s) - C_{\theta_0}(\boldsymbol{u}_s)\}, W_i^{(b)} = \sum_{s=1}^{q'} Z_{i,n}^{(b)} \bar{c}_s \{\mathbb{1}(\boldsymbol{U}_i \le \boldsymbol{v}_s) - C_{\theta_0}(\boldsymbol{v}_s)\}, T_i = c'\{J_{\theta_0}(\boldsymbol{U}_i) + K_{i,\theta_0}\}$ and $T_i^{(b)} = Z_{i,n}^{(b)} \bar{c}'\{J_{\theta_0}(\boldsymbol{U}_i) + K_{i,\theta_0}\}$.

The subsequent proof is based on the 'big block-small block'-technique. The assumption on a in Condition VI is equivalent to $\frac{1}{2(1+\nu)} < \frac{1}{2} - \frac{2+\nu}{a\nu}$ whence, noting that also $\kappa > \frac{1}{2(1+\nu)}$ by assumption, we may choose $0 < \eta_1 < \eta_2 < \kappa$ such that $\frac{1}{2(1+\nu)} < \eta_1 < \eta_2 < \frac{1}{2} - \frac{2+\nu}{a\nu}$. Now, set $b_n = \lfloor n^{1/2-\eta_1} \rfloor$ (the length of the big blocks), $s_n = \lfloor n^{1/2-\eta_2} \rfloor$ (the length of the small blocks) and $k_n = \lfloor n/(b_n + s_n) \rfloor$ (the number of big or small blocks). Notice, that $k_n = O(n^{1/2+\eta_1})$. For $j = 1, \dots, k_n$, set

$$B_{jn} = \sum_{i=(j-1)(b_n+s_n)+1}^{(j-1)(b_n+s_n)+b_n} W_i + W_i^{(b)} + T_i + T_i^{(b)},$$

$$S_{jn} = \sum_{i=(j-1)(b_n+s_n)+b_n+1}^{j(b_n+s_n)} W_i + W_i^{(b)} + T_i + T_i^{(b)},$$

such that we can write

$$V_n = \frac{1}{\sqrt{n}} \sum_{j=1}^{k_n} B_{jn} + \frac{1}{\sqrt{n}} \sum_{j=1}^{k_n} S_{jn} + \frac{1}{\sqrt{n}} R_n,$$

where $R_n = \sum_{i=k_n(b_n+s_n)+1}^{n} W_i + W_i^{(b)} + T_i + T_i^{(b)}$ is the sum over the remaining indices that are not part of a big or a small block.

First of all, let us show that the variance of V_n is equal to $\mathrm{Var}(\frac{1}{\sqrt{n}} \sum_{j=1}^{k_n} B_{jn}) + o(1)$ as $n \to \infty$, i.e., that

$$\frac{1}{n} \mathrm{Var}\Big(\sum_{j=1}^{k_n} S_{jn}\Big) + \frac{1}{n} \mathrm{Var}(R_n) + \frac{2}{n} \sum_{j,j'=1}^{k_n} \mathrm{Cov}(B_{jn}, S_{j'n}) + \frac{2}{n} \sum_{j=1}^{k_n} \mathrm{Cov}(B_{jn}, R_n) + \frac{2}{n} \sum_{j=1}^{k_n} \mathrm{Cov}(S_{jn}, R_n) \quad \text{(F.1)}$$

vanishes as $n \to \infty$. To this end, we will frequently exploit the following bounds which are consequences of Lemma 3.9 and Lemma 3.11 in Dehling and Philipp (2002)

$$
\begin{aligned}
|\mathbb{E}[W_i W_{i'}]| &= \Big| \sum_{s,s'=1}^{q} c_s c_{s'} \mathbb{E}[\{\mathbb{1}(\boldsymbol{U}_i \le \boldsymbol{u}_s) - C_{\theta_0}(\boldsymbol{u}_s)\}\{\mathbb{1}(\boldsymbol{U}_i \le \boldsymbol{u}_{s'}) - C_{\theta_0}(\boldsymbol{u}_{s'})\}]\Big| \\
&\le 4 \sum_{s,s'=1}^{q} |c_s c_{s'}| \alpha(|i-i'|) \le \text{const} \times \alpha(|i-i'|) \\
|\mathbb{E}[W_i T_{i'}]| &= \Big| \sum_{s=1}^{q} c_s c' \mathbb{E}[\{\mathbb{1}(\boldsymbol{U}_i \le \boldsymbol{u}_s) - C_{\theta}(\boldsymbol{u}_s)\} J_{\theta_0}(\boldsymbol{U}_{i'})]\Big| \\
&\le 20\|c\| \sum_{s=1}^{q} |c_s| \alpha(|i-i'|)^{\frac{1+\nu}{2+\nu}} \mathbb{E}[\|J_{\theta_0}\|^{2+\nu}]^{\frac{1}{2+\nu}} \le \text{const} \times \alpha(|i-i'|)^{\frac{1+\nu}{2+\nu}} \\
|\mathbb{E}[T_i T_{i'}]| &= |c' \mathbb{E}[J_{\theta_0}(\boldsymbol{U}_i) J_{\theta_0}(\boldsymbol{U}_{i'})'] c| \le 40\|c\|^2 \alpha(|i-i'|)^{\frac{\nu}{2+\nu}} \mathbb{E}[\|J_{\theta_0}\|^{2+\nu}]^{\frac{2}{2+\nu}} \\
&\le \text{const} \times \alpha(|i-i'|)^{\frac{\nu}{2+\nu}}.
\end{aligned}
$$

Analogously, $|\mathbb{E}[W_i^{(b)} W_{i'}^{(b)}]| \le \text{const} \times \alpha(|i-i'|))$, $|\mathbb{E}[W_i^{(b)} T_{i'}^{(b)}]| \le \text{const} \times \alpha(|i-i'|)^{\frac{1+\nu}{2+\nu}}$ and $|\mathbb{E}[T_i^{(b)} T_{i'}^{(b)}]| \le \text{const} \times \alpha(|i-i'|)^{\frac{\nu}{2+\nu}}$. Notice that all the other pairs of random variables are uncorrelated and that the largest bound is a constant multiple of $\alpha(|i-i'|)^{\frac{\nu}{2+\nu}}$. Now, we can begin with the discussion of the first summand in (F.1). We have

$$
\frac{1}{n} \operatorname{Var}\Big(\sum_{j=1}^{k_n} S_{jn}\Big) = \frac{1}{n} \sum_{j=1}^{k_n} \operatorname{Var}(S_{jn}) + \frac{2}{n} \sum_{j \ne j'} \operatorname{Cov}(S_{jn}, S_{j'n}). \tag{F.2}
$$

Since the distance between any two summands in S_{jn} and $S_{j'n}$ for $j \ne j'$ is at least b_n, their covariance is of order $\alpha(b_n)^{\frac{\nu}{2+\nu}} = O(b_n^{-a\nu/(2+\nu)})$. Observing that S_{jn} consists of s_n summands, we obtain that the second term in the last display is of order $O(k_n^2 s_n^2 b_n^{-a\nu/(2+\nu)} n^{-1}) = O(n^{1-a\nu/(4+2\nu)+\eta_1(2+a\nu/(2+\nu))-2\eta_2}) = O(n^{1-a\nu/(4+2\nu)+\eta_2 a\nu/(2+\nu)}) = o(1)$ since, by construction, $\eta_1 < \eta_2 < \frac{1}{2} - \frac{2+\nu}{a\nu}$. For the first sum on the right-hand side of (F.2), we have, by dominated convergence,

$$
\frac{1}{n} \sum_{j=1}^{k_n} \operatorname{Var}(S_{nj}) = \frac{1}{n} \sum_{j=1}^{k_n} \mathbb{E}[S_{nj}^2] \le \text{ const} \times \frac{1}{n} \sum_{j=1}^{k_n} \sum_{i=-s_n}^{s_n} (s_n - |i|) \alpha(|i|)^{\frac{\nu}{2+\nu}}
$$

$$= O(k_n s_n n^{-1}) = O(n^{\eta_1 - \eta_2}) = o(1).$$

For the second term of (F.1), we have

$$\begin{aligned}
\frac{1}{n}\operatorname{Var}(R_n) &\leq \text{ const} \times \frac{1}{n} \sum_{i,i'=k_n(b_n+s_n)+1}^{n} \alpha(|i-i'|)^{\frac{\nu}{2+\nu}} \\
&\leq \text{ const} \times \frac{1}{n} \sum_{i=-\{n-k_n(b_n+s_n)\}}^{n-k_n(b_n+s_n)} (n - k_n(s_n+b_n) - |i|)\alpha(|i|)^{\frac{\nu}{2+\nu}} \\
&= O(\{n - k_n(b_n+s_n)\}/n) = O((b_n+s_n)/n) = o(1),
\end{aligned}$$

where we used that $k_n \geq n/(b_n+s_n) - 1$.

Now, let us bound the third term in (F.1). First we notice that, if $j = j'$ or $j' = j - 1$, we have $|\mathbb{E}[B_{jn}S_{j'n}]| \leq \text{const} \sum_{i=1}^{b_n} \sum_{i'=b_n+1}^{s_n+b_n} \alpha(|i-i'|)^{\frac{\nu}{2+\nu}} \leq \text{const} \sum_{i=1}^{s_n+b_n} i\alpha(i)^{\frac{\nu}{2+\nu}} \leq \text{const} < \infty$, since $a\nu/(2+\nu) > 2 + 1/\nu$ by Condition VI. In the other cases the distance between the blocks B_{jn} and $S_{j'n}$ is at least b_n, such that $|\mathbb{E}[B_{jn}S_{j'n}]| = O(b_n s_n \alpha(b_n)^{\frac{\nu}{2+\nu}})$. Together, this yields

$$\begin{aligned}
\frac{2}{n}\sum_{j,j'=1}^{k_n} \mathbb{E}[B_{jn}S_{j'n}] &= O(n^{-1}k_n) + O(k_n^2 b_n s_n b_n^{-a\frac{\nu}{2+\nu}} n^{-1}) \\
&= O(n^{-1/2+\eta_1}) + O(n^{1-\frac{a}{2}\frac{\nu}{2+\nu}+\eta_1(1+a\frac{\nu}{2+\nu})-\eta_2}) = o(1),
\end{aligned}$$

where the last equality follows exactly as above for the treatment of the first summand in (F.1). In the same manner, we get $\frac{2}{n}\sum_{j=1}^{k_n} \mathbb{E}[B_{jn}R_n] = O(n^{\frac{1}{2}-\frac{a}{2}\frac{\nu}{2+\nu}-\eta_1+a\frac{\nu}{2+\nu}\eta_2}) = o(1)$ and $\frac{2}{n}\sum_{j=1}^{k_n} \mathbb{E}[S_{nj}R_n] = O(n^{\frac{1}{2}-\frac{a}{2}\frac{\nu}{2+\nu}-\eta_2+a\frac{\nu}{2+\nu}\eta_1}) + O(\{b_n+s_n\}/n) = o(1)$.

For the next step of the proof, let $B'_{jn}, j = 1, \ldots, k_n$ denote independent random variables such that each B'_{jn} has the same distribution as B_{jn}. We will show that the characteristic function of $n^{-1/2}\sum_{j=1}^{k_n} B_{jn}$ is asymptotically equivalent to the characteristic function of $n^{-1/2}\sum_{j=1}^{k_n} B'_{jn}$. For $t \in \mathbb{R}$, define $\Psi_{jn}(t) = \exp(\mathrm{i}tn^{-1/2}B_{jn})$ and notice that $\mathbb{E}[\prod_{j=1}^{k_n}\Psi_{jn}(t)]$ and $\prod_{j=1}^{k_n}\mathbb{E}[\Psi_{jn}(t)]$ are the characteristic functions of $n^{-1/2}\sum_{j=1}^{k_n} B_{jn}$ and $n^{-1/2}\sum_{j=1}^{k_n} B'_{jn}$, respectively. The difference of the two characteristic functions can be decomposed as follows

$$\Big|\mathbb{E}\Big[\prod_{j=1}^{k_n}\Psi_{jn}(t)\Big]-\prod_{j=1}^{k_n}\mathbb{E}[\Psi_{jn}(t)]\Big| \leq \Big|\mathbb{E}\Big[\prod_{j=1}^{k_n}\Psi_{jn}(t)\Big]-\mathbb{E}[\Psi_{1n}(t)]\mathbb{E}[\prod_{j=2}^{k_n}\Psi_{jn}(t)]\Big|$$
$$+|\mathbb{E}[\Psi_{1n}(t)]|\times\Big|\mathbb{E}\Big[\prod_{j=2}^{k_n}\Psi_{jn}(t)\Big]-\mathbb{E}[\Psi_{2n}(t)]\mathbb{E}[\prod_{j=3}^{k_n}\Psi_{jn}(t)]\Big|$$
$$+\cdots+\Big|\prod_{j=1}^{k_n-2}\mathbb{E}[\Psi_{jn}(t)]\Big|\times\Big|\mathbb{E}\Big[\prod_{k_n-1}^{k_n}\Psi_{jn}(t)\Big]-\prod_{j=k_n-1}^{k_n}\mathbb{E}[\Psi_{jn}(t)]\Big|$$

Applying Lemma 3.9 in Dehling and Philipp (2002) (k_n-1) times, we get

$$\Big|\mathbb{E}\Big[\prod_{j=1}^{k_n}\Psi_{jn}(t)\Big]-\prod_{j=1}^{k_n}\mathbb{E}[\Psi_{jn}(t)]\Big|$$
$$\leq 2\pi\times(k_n-1)\max_{i=1,\ldots,k_n-1}\alpha(\sigma(\Psi_{in}),\sigma\{\prod_{j=i+1}^{k_n}\Psi_{jn}(t)\}),$$

which is of order $O(k_n\alpha(s_n))=O(n^{1/2-a/2+a\eta_2+\eta_1})$. Using that $\eta_1<\eta_2$, this expression converges to 0 by the choice of η_2 and the fact that $(2+v)/(a\nu)>1/a>1/(a+1)$. As a consequence, provided $n^{-1/2}\sum_{j=1}^{k_n}B'_{jn}$ converges weakly, then so does $n^{-1/2}\sum_{j=1}^{k_n}B_{jn}$ with the same limiting distribution.

Therefore, in order to finalize the proof, it remains to be shown that $n^{-1/2}\sum_{j=1}^{k_n}B'_{jn}$ converges weakly to V. This will be accomplished by proving that the variance of $n^{-1/2}\sum_{j=1}^{k_n}B'_{jn}$ converges to $\operatorname{Var}(V)$ as $n\to\infty$ and that the Lindeberg-condition from the Lindeberg-Feller central limit theorem for independent triangular arrays is met. We begin with the convergence of the variance and note that

$$\operatorname{Var}(V)=\sum_{s,s'=1}^{q}c_sc_{s'}\sum_{i\in\mathbb{Z}}\gamma(i,\boldsymbol{u}_s,\boldsymbol{u}_{s'})+\sum_{s,s'=1}^{q'}\bar{c}_s\bar{c}_{s'}\sum_{i\in\mathbb{Z}}\gamma(i,\boldsymbol{v}_s,\boldsymbol{v}_{s'})$$
$$+\sum_{s=1}^{q}c_sc'\sum_{\imath\in\mathbb{Z}}\bar{\gamma}(i,\boldsymbol{u}_s)+\sum_{s=1}^{q'}\bar{c}_s\bar{c}'\sum_{i\in\mathbb{Z}}\bar{\gamma}(i,\boldsymbol{v}_s)$$

$$+ c' \sum_{i\in\mathbb{Z}} \tilde{\gamma}(i) c + \bar{c}' \sum_{i\in\mathbb{Z}} \tilde{\gamma}(i)\bar{c},$$

where $\gamma(i, \boldsymbol{u}, \boldsymbol{v}) = \mathrm{Cov}\{\mathbb{1}(\boldsymbol{U}_1 \le \boldsymbol{u}), \mathbb{1}(\boldsymbol{U}_{1+i} \le \boldsymbol{v})\}$, $\bar{\gamma}(i, \boldsymbol{u}) = \mathrm{Cov}\{J_{\theta_0}(\boldsymbol{U}_1) + K_{1,\theta_0}, \mathbb{1}(\boldsymbol{U}_{1+i} \le \boldsymbol{u})\}$ and $\tilde{\gamma}(i) = \mathrm{Cov}\{J_{\theta_0}(\boldsymbol{U}_1) + K_{1,\theta_0}, J_{\theta_0}(\boldsymbol{U}_{1+i}) + K_{1+i,\theta_0}\}$. Now, let us show that $\mathrm{Var}(\frac{1}{\sqrt{n}} \sum_{j=1}^{k_n} B'_{jn}) = \mathrm{Var}(V_n) + o(1)$. To this end, note that

$$\mathrm{Var}\left(n^{-1/2} \sum_{j=1}^{k_n} B'_{jn}\right) = \frac{1}{n} \sum_{j=1}^{k_n} \mathrm{Var}(B'_{jn}) = \frac{1}{n} \sum_{j=1}^{k_n} \mathrm{Var}(B_{jn})$$
$$= \mathrm{Var}\left(n^{-1/2} \sum_{j=1}^{k_n} B_{jn}\right) - \frac{1}{n} \sum_{j\neq j'} \mathrm{Cov}(B_{jn}, B_{j'n}).$$

Since we have already shown in the beginning of the proof that the variance on the right-hand side equals $\mathrm{Var}(V_n) + o(1)$, it remains to be shown that $\frac{1}{n}\sum_{j\neq j'} \mathrm{Cov}(B_{jn}, B_{j'n}) = o(1)$. Since the distance between the random variables within the two blocks is at least s_n, we have $\mathbb{E}[B_{jn}B_{j'n}] = O(b_n^2 \alpha(s_n)^{\frac{\nu}{2+\nu}})$ for $j \neq j'$. Therefore, $n^{-1}\sum_{j\neq j'} \mathrm{Cov}(B_{jn}, B_{j'n}) = O(n^{-1} k_n^2 b_n^2 s_n^{-a\frac{\nu}{2+\nu}}) = O(n^{1-\frac{a}{2}\frac{\nu}{2+\nu}+a\frac{\nu}{2+\nu}\eta_2})$, which is $o(1)$ as shown above.

Now, let us show that $\mathrm{Var}(V_n) \to \mathrm{Var}(V)$ as $n \to \infty$. We can write $\mathrm{Var}(V_n)$ as

$$\frac{1}{n} \sum_{i,i'=1}^{n} \Bigg[\Bigg\{ \sum_{s,s'=1}^{q} c_s c_{s'} \gamma(i-i', \boldsymbol{u}_s, \boldsymbol{u}_{s'}) + \sum_{s,s'=1}^{q'} \bar{c}_s \bar{c}_{s'} \varphi\{(i-i')/\ell_n\} \gamma(i-i', \boldsymbol{v}_s, \boldsymbol{v}_{s'}) \Bigg\}$$
$$+ \Bigg\{ \sum_{s=1}^{q} c' c_s \bar{\gamma}(i-i', \boldsymbol{u}_s) + \sum_{s=1}^{q'} \bar{c}' \bar{c}_s \varphi\{(i-i')/\ell_n\} \bar{\gamma}(i-i', \boldsymbol{v}_s) \Bigg\}$$
$$+ \Big\{ c' \tilde{\gamma}(i-i') c + \varphi\{(i-i')/\ell_n\} \bar{c}' \tilde{\gamma}(i-i') \bar{c} \Big\} \Bigg]$$
$$= \sum_{i=-n}^{n} \frac{n-|i|}{n} \Bigg[\Bigg\{ \sum_{s,s'=1}^{q} c_s c_{s'} \gamma(i, \boldsymbol{u}_s, \boldsymbol{u}_{s'}) + \sum_{s,s'=1}^{q'} c_s c_{s'} \varphi(i/\ell_n) \gamma(i, \boldsymbol{v}_s, \boldsymbol{v}_{s'}) \Bigg\}$$
$$+ \Bigg\{ \sum_{s=1}^{q} c' c_s \bar{\gamma}(i, \boldsymbol{u}_s) + \sum_{s=1}^{q'} \bar{c}' \bar{c}_s \varphi(i/\ell_n) \bar{\gamma}(i, \boldsymbol{v}_s) \Bigg\}$$

$$+\left\{c'\tilde{\gamma}(i)c+\varphi(i/\ell_n)\bar{c}'\tilde{\gamma}(i)\bar{c}\right\}\Bigg]. \tag{F.3}$$

For the sake of brevity, we will only show convergence of the terms in the first curly brackets on the right-hand side of the last display to the respective terms in $\mathrm{Var}(V)$. We have

$$\sum_{i=-n}^{n}\frac{n-|i|}{n}\sum_{s,s'=1}^{q}c_s c_{s'}\gamma(i,\boldsymbol{u}_s,\boldsymbol{u}_{s'})$$
$$=\sum_{s,s'=1}^{q}c_s c_{s'}\sum_{i=-n}^{n}\gamma(i,\boldsymbol{u}_s,\boldsymbol{u}_{s'})-\sum_{s,s'=1}^{q}c_s c_{s'}\frac{1}{n}\sum_{i=-n}^{n}|i|\gamma(i,\boldsymbol{u}_s,\boldsymbol{u}_{s'}).$$

Since $|\gamma(i,\boldsymbol{u},\boldsymbol{v})|\le \mathrm{const}\times\alpha(|i|)$ and $\sum_{i=1}^{\infty}|i|\alpha(i)<\infty$, the second term on the right-hand side vanishes as $n\to\infty$, whereas the first term converges to $\sum_{s,s'=1}^{q}c_s c_{s'}\sum_{i\in\mathbb{Z}}\gamma(i,\boldsymbol{u}_s,\boldsymbol{u}_{s'})$.

Moreover, as $\varphi(h)=0$ for $|h|>1$ and $\ell_n=o(n)$, we have

$$\sum_{i=-n}^{n}\frac{n-|i|}{n}\sum_{s,s'=1}^{q'}\bar{c}_s\bar{c}_{s'}\varphi(i/\ell_n)\gamma(i,\boldsymbol{v}_s,\boldsymbol{v}_{s'})$$
$$=\sum_{s,s'=1}^{q'}\bar{c}_s\bar{c}_{s'}\sum_{i\in\mathbb{Z}}\frac{n-|i|}{n}\varphi(i/\ell_n)\gamma(i,\boldsymbol{v}_s,\boldsymbol{v}_{s'}),$$

By continuity of φ in 0, we have $\frac{n-|i|}{n}\varphi(i/\ell_n)\to 1$ as $n\to\infty$ for any fixed $i\in\mathbb{Z}$. Moreover, $|\frac{n-|i|}{n}\varphi(i/\ell_n)\gamma(i,\boldsymbol{v}_s,\boldsymbol{v}_{s'})|\le \mathrm{const}\times\alpha(i)$ for all $i\in\mathbb{Z}$ and all $s,s'=1,\dots,q'$. Therefore, by dominated convergence, as $n\to\infty$,

$$\sum_{s,s'=1}^{q'}\bar{c}_s\bar{c}_{s'}\sum_{i=-n}^{n}\frac{n-|i|}{n}\varphi(i/\ell_n)\gamma(i,\boldsymbol{v}_s,\boldsymbol{v}_{s'})\to\sum_{s,s'=1}^{q'}\bar{c}_s\bar{c}_{s'}\sum_{i\in\mathbb{Z}}\gamma(i,\boldsymbol{v}_s,\boldsymbol{v}_{s'}).$$

The convergence of the remaining summands in (F.3) follows along similar lines, exploiting that $\|\bar{\gamma}(i,\boldsymbol{u})\|\le \mathrm{const}\times\alpha(i)^{\frac{1+\nu}{2+\nu}}$, $\|\tilde{\gamma}(i)\|_{op}\le \mathrm{const}\times\alpha(i)^{\frac{\nu}{2+\nu}}$ and that $\sum_{i=1}^{\infty}|i|\alpha(i)^{\frac{\nu}{2+\nu}}<\infty$.

Finally, let us prove the Lindeberg condition, i.e., that, for any $\varepsilon > 0$,

$$\frac{1}{n}\sum_{j=1}^{k_n}\mathbb{E}[B'^2_{jn}\mathbb{1}(|B'_{jn}| > \sqrt{n}\varepsilon)] = \frac{1}{n}\sum_{j=1}^{k_n}\mathbb{E}[B^2_{jn}\mathbb{1}(|B_{jn}| > \sqrt{n}\varepsilon)] \to 0$$

as $n \to \infty$. To bound the former expression we use Hölder's inequality with $p = (2+\nu)/2$ and $q = (2+\nu)/\nu$ and Markov's inequality to obtain

$$\begin{aligned}
\frac{1}{n}\sum_{j=1}^{k_n}\mathbb{E}[B^2_{jn}\mathbb{1}(|B_{jn}| \leq \sqrt{n}\varepsilon)] &\leq \frac{1}{n}\sum_{j=1}^{k_n}\mathbb{E}[|B_{jn}|^{2+\nu}]^{\frac{2}{2+\nu}}\mathbb{E}[\mathbb{1}(|B_{jn}| \leq \sqrt{n}\varepsilon)^{\frac{2+\nu}{\nu}}]^{\frac{\nu}{2+\nu}} \\
&= \frac{1}{n}\sum_{j=1}^{k_n}\mathbb{E}[|B_{jn}|^{2+\nu}]^{\frac{2}{2+\nu}}\mathbb{P}(|B_{jn}| \leq \sqrt{n}\varepsilon)^{\frac{\nu}{2+\nu}} \\
&\leq \frac{1}{n}\sum_{j=1}^{k_n}\mathbb{E}[|B_{jn}|^{2+\nu}]^{\frac{2}{2+\nu}}\mathbb{E}[|B_{jn}|^{2+\nu}]^{\frac{\nu}{2+\nu}}(\sqrt{n}\varepsilon)^{-\nu} \\
&= \frac{1}{n}\sum_{j=1}^{k_n}\mathbb{E}[|B_{jn}|^{2+\nu}](\sqrt{n}\varepsilon)^{-\nu}
\end{aligned}$$

By Minkowski's inequality, we can bound $\mathbb{E}[|B_{jn}|^{2+\nu}]^{\frac{1}{2+\nu}}$ by a sum over b_n summands of the form $\mathbb{E}[|W_i|^{2+\nu}]^{\frac{1}{2+\nu}} + \mathbb{E}[|W_i^{(b)}|^{2+\nu}]^{\frac{1}{2+\nu}} + \mathbb{E}[|T_i|^{2+\nu}]^{\frac{1}{2+\nu}} + \mathbb{E}[|T_i^{(b)}|^{2+\nu}]^{\frac{1}{2+\nu}}$, whence $\mathbb{E}[|B_{jn}|^{2+\nu}] = O(b_n^{2+\nu})$. This finally implies

$$\frac{1}{n}\sum_{j=1}^{k_n}\mathbb{E}[B^2_{jn}\mathbb{1}(|B_{jn}| \leq \sqrt{n}\varepsilon)] = O(b_n^{2+\nu}k_n n^{-\nu/2-1}) = O(n^{1/2-\eta_1(1+\nu)}) = o(1),$$

by the definition of η_1 and the Lemma is proved. □

Proof of Lemma C.7. It follows from the mean value theorem that

$$\sqrt{n}\{C_{\theta_n}(\boldsymbol{u}) - C_{\theta_0}(\boldsymbol{u})\} = \nabla C_{\theta_0}(\boldsymbol{u})\Theta_n + R_n(\boldsymbol{u}),$$

where $R_n(\boldsymbol{u}) = \sqrt{n}\sum_{s=1}^{p}\left\{\frac{\partial}{\partial\theta_s}C_{\tilde{\theta}}(\boldsymbol{u}) - \frac{\partial}{\partial\theta_s}C_{\theta_0}(\boldsymbol{u})\right\}(\theta_{ns} - \theta_{0s})$ and where $\tilde{\theta}$ denotes an intermediate point lying between θ_0 and θ_n. The Cauchy-Schwarz inequality allows to estimate

$$\sup_{\boldsymbol{u}\in[0,1]^d} \|R_n(\boldsymbol{u})\| \leq \sup_{\|\theta'-\theta_0\|<\|\theta_n-\theta_0\|} \sup_{\boldsymbol{u}\in[0,1]^d} \|\nabla C_{\theta'}(\boldsymbol{u}) - \nabla C_{\theta_0}(\boldsymbol{u})\| \times \|\Theta_n\|.$$

Since $\Theta_n = \tilde{\Theta}_n + o_p(1)$ for $n \to \infty$ as a consequence of Lemma C.1 and C.2, we obtain that $\|\Theta_n\| = O_P(1)$ from Lemma C.6 (note that Condition III was used in Lemma C.6 only for the tightness part, whence we do not need to assume it here). Fix $\eta > 0$. By Condition IV, we may choose $\delta > 0$ such that $\sup_{\|\theta'-\theta_0\|<\delta} \sup_{\boldsymbol{u}\in[0,1]^d} \|\nabla C_{\theta'}(\boldsymbol{u}) - \nabla C_{\theta_0}(\boldsymbol{u})\| \leq \eta$. Therefore,

$$\mathbb{P}(\sup_{\boldsymbol{u}\in[0,1]^d} \sup_{\|\theta'-\theta_0\|<\|\theta_n-\theta\|} \|\nabla C_{\theta'}(\boldsymbol{u}) - \nabla C_{\theta_0}(\boldsymbol{u})\| > \eta) \leq \mathbb{P}(\|\theta_n - \theta_0\| > \delta),$$

which converges to 0 as $n \to \infty$, since θ_n is a consistent estimator of θ_0. □

Acknowledgements. The authors would like to thank three unknown referees and two editors for their constructive comments on an earlier version of this manuscript, which led to a substantial improvement of the paper. The authors are also thankful to Brendan K. Beare for kindly providing us with the gasoline price markup data and Ivan Kojadinovic for helpful discussions regarding this manuscript.

Parts of this paper were written when Axel Bücher was a post-doctoral researcher at Université catholique de Louvain, Belgium, a visiting professor at Universität Heidelberg, Germany, and a visiting professor at Technische Universität Dortmund, Germany.

This work has been supported in parts by the Collaborative Research Center "Statistical modeling of nonlinear dynamic processes" (SFB 823) of the German Research Foundation (DFG) and by the IAP research network Grant P7/06 of the Belgian government (Belgian Science Policy), which is gratefully acknowledged.

References

Bücher, A. and I. Kojadinovic (2014). A dependent multiplier bootstrap for the sequential empirical copula process under strong mixing. *Bernoulli*, in press.

Bühlmann, P. L. (1993). *The blockwise bootstrap in time series and empirical processes*. ProQuest LLC, Ann Arbor, MI. Thesis (Dr.Sc.Math)–Eidgenoessische Technische Hochschule Zürich (Switzerland).

Dehling, H. and W. Philipp (2002). Empirical process techniques for dependent data. In *Empirical process techniques for dependent data*, pp. 3–113. Boston, MA: Birkhäuser Boston.

Politis, D. N. and H. White (2004). Automatic block-length selection for the dependent bootstrap. *Econometric Rev. 23*(1), 53–70.

Rémillard, B., N. Papageorgiou, and F. Soustra (2012). Copula-based semiparametric models for multivariate time series. *J. Multivariate Anal. 110*, 30–42.