Bernoulli **22**(2), 2016, 927–968
DOI: 10.3150/14-BEJ682

A dependent multiplier bootstrap for the sequential empirical copula process under strong mixing

AXEL BÜCHER[1] and IVAN KOJADINOVIC[2]

[1] *Fakultät für Mathematik, Ruhr-Universität Bochum, 44780 Bochum, Germany. E-mail: axel.buecher@ruhr-uni-bochum.de*

[2] *Laboratoire de mathématiques et de leurs applications, Université de Pau et des Pays de l'Adour, UMR CNRS 5142, B.P. 1155, 64013 Pau Cedex, France. E-mail: ivan.kojadinovic@univ-pau.fr*

Two key ingredients to carry out inference on the copula of multivariate observations are the empirical copula process and an appropriate resampling scheme for the latter. Among the existing techniques used for i.i.d. observations, the multiplier bootstrap of Rémillard and Scaillet (*J. Multivariate Anal.* **100** (2009) 377–386) frequently appears to lead to inference procedures with the best finite-sample properties. Bücher and Ruppert (*J. Multivariate Anal.* **116** (2013) 208–229) recently proposed an extension of this technique to strictly stationary strongly mixing observations by adapting the *dependent* multiplier bootstrap of Bühlmann (The blockwise bootstrap in time series and empirical processes (1993) ETH Zürich, Section 3.3) to the empirical copula process. The main contribution of this work is a generalization of the multiplier resampling scheme proposed by Bücher and Ruppert along two directions. First, the resampling scheme is now genuinely sequential, thereby allowing to transpose to the strongly mixing setting many of the existing multiplier tests on the unknown copula, including nonparametric tests for change-point detection. Second, the resampling scheme is now fully automatic as a data-adaptive procedure is proposed which can be used to estimate the bandwidth parameter. A simulation study is used to investigate the finite-sample performance of the resampling scheme and provides suggestions on how to choose several additional parameters. As by-products of this work, the validity of a sequential version of the dependent multiplier bootstrap for empirical processes of Bühlmann is obtained under weaker conditions on the strong mixing coefficients and the multipliers, and the weak convergence of the sequential empirical copula process is established under many serial dependence conditions.

Keywords: lag window estimator; multiplier central limit theorem; multivariate observations; partial-sum process; ranks; serial dependence

1. Introduction

Let $\mathbf{X}$ be a d-dimensional random vector with continuous marginal cumulative distribution functions (c.d.f.s) $F_1, \dots, F_d$. From the work of Sklar [45], the c.d.f. F of $\mathbf{X}$ can be written in a unique way as

$$F(\mathbf{x}) = C\{F_1(x_1), \dots, F_d(x_d)\}, \qquad \mathbf{x} \in \mathbb{R}^d,$$

where the function $C : [0, 1]^d \to [0, 1]$ is a copula and can be regarded as capturing the dependence among the components of $\mathbf{X}$. The above equation is at the origin of the increasing use of copulas for modeling multivariate distributions with continuous margins in many areas such as quantitative risk management (McNeil, Frey and Embrechts [31]), econometric modeling (Patton [35]), environmental modeling (Salvadori, De Michele and Kottegoda [41]), to name a very few.

Assume that C and $F_1, \dots, F_d$ are unknown and let $\mathbf{X}_1, \dots, \mathbf{X}_n$ be drawn from a strictly stationary sequence of continuous d-dimensional random vectors with c.d.f. F. For any $i \in \{1, \dots, n\}$ and $j \in \{1, \dots, d\}$, denote by $R_{ij}^{1:n}$ the (mid-)rank of X_{ij} among $X_{1j}, \dots, X_{nj}$ and let $\hat{U}_{ij}^{1:n} = R_{ij}^{1:n}/n$. The random vectors $\hat{\mathbf{U}}_i^{1:n} = (\hat{U}_{i1}^{1:n}, \dots, \hat{U}_{id}^{1:n})$, $i \in \{1, \dots, n\}$, are often referred to as *pseudo-observations* from the copula C, and a natural nonparametric estimator of C is the *empirical copula* of $\mathbf{X}_1, \dots, \mathbf{X}_n$ (Rüschendorf [40]; Deheuvels [19]), frequently defined as the empirical c.d.f. computed from the pseudo-observations, that is,

$$C_{1:n}(\mathbf{u}) = \frac{1}{n} \sum_{i=1}^{n} \mathbf{1}\big(\hat{\mathbf{U}}_i^{1:n} \le \mathbf{u}\big), \qquad \mathbf{u} \in [0,1]^d.$$

The empirical copula plays a key role in most nonparametric inference procedures on C. Examples of its use for parametric inference, nonparametric testing and goodness-of-fit testing can be found in Tsukahara [46], Rémillard and Scaillet [39], Genest, Rémillard and Beaudoin [23], respectively, among many others. The asymptotics of such procedures typically follow from the asymptotics of the *empirical copula process*. With applications to change-point detection in mind, a generalization of the latter process central to this work is the *two-sided sequential empirical copula process*. It is defined, for any $(s,t) \in \Delta = \{(s,t) \in [0,1]^2 : s \le t\}$ and $\mathbf{u} \in [0,1]^d$, by

$$\mathbb{C}_n(s,t,\mathbf{u}) = \frac{1}{\sqrt{n}} \sum_{i=\lfloor ns \rfloor+1}^{\lfloor nt \rfloor} \big\{\mathbf{1}\big(\hat{\mathbf{U}}_i^{\lfloor ns \rfloor+1:\lfloor nt \rfloor} \le u\big) - C(\mathbf{u})\big\}, \tag{1.1}$$

where, for any $y \ge 0$, $\lfloor y \rfloor$ is the greatest integer smaller or equal than y. The latter process can be rewritten in terms of the empirical copula $C_{\lfloor ns \rfloor+1:\lfloor nt \rfloor}$ of the sample $\mathbf{X}_{\lfloor ns \rfloor+1}, \dots, \mathbf{X}_{\lfloor nt \rfloor}$ as

$$\mathbb{C}_n(s,t,\mathbf{u}) = \sqrt{n}\lambda_n(s,t)\big\{C_{\lfloor ns \rfloor+1:\lfloor nt \rfloor}(\mathbf{u}) - C(\mathbf{u})\big\}, \qquad (s,t,\mathbf{u}) \in \Delta \times [0,1]^d,$$

where $\lambda_n(s,t) = (\lfloor nt \rfloor - \lfloor ns \rfloor)/n$ and with the convention that $C_{k:k-1}(\mathbf{u}) = 0$ for all $\mathbf{u} \in [0,1]^d$ and all $k \in \{1, \dots, n\}$.

The quantity $\mathbb{C}_n(0,1,\cdot,\cdot)$ is the standard empirical copula process which has been extensively studied in the literature (see, e.g., Rüschendorf [40]; Gaenssler and Stute [22]; Tsukahara [46]; van der Vaart and Wellner [49]; Segers [43]; Bücher and Volgushev [13]). Notice that the process $\mathbb{C}_n(0,\cdot,\cdot,\cdot)$ does not coincide with the sequential process initially studied by Rüschendorf [40] and defined by

$$\mathbb{C}_n^{\circ}(s,\mathbf{u}) = \frac{1}{\sqrt{n}} \sum_{i=1}^{\lfloor ns \rfloor} \big\{\mathbf{1}\big(\hat{\mathbf{U}}_i^{1:n} \le \mathbf{u}\big) - C(\mathbf{u})\big\}, \qquad (s,\mathbf{u}) \in [0,1]^{d+1}. \tag{1.2}$$

The above process, unlike $\mathbb{C}_n(0,\cdot,\cdot,\cdot)$, cannot be rewritten in terms of the empirical copula unless $s = 1$. Note that the weak convergence of $\mathbb{C}_n^{\circ}$ was further studied by Bücher and Volgushev

[13] under a large number of serial dependence scenarios and under mild smoothness conditions on the copula.

As mentioned earlier, a first key ingredient of many of the existing inference procedures on the unknown copula C is the process $\mathbb{C}_n$ defined in (1.1). A second key ingredient is typically some resampling scheme allowing to obtain replicates of $\mathbb{C}_n$. When dealing with independent observations, several such resampling schemes for the empirical copula process $\mathbb{C}_n(0, 1, \cdot, \cdot)$ were proposed in the literature, ranging from the multinomial bootstrap of Fermanian, Radulović and Wegkamp [21] to the multiplier technique introduced in Scaillet [42] and investigated further in Rémillard and Scaillet [39]. Their finite-sample properties were compared in Bücher and Dette [8] who concluded that the multiplier bootstrap of Rémillard and Scaillet [39] has, overall, the best finite-sample behavior. In the case of strongly mixing observations, Bücher and Ruppert [11] recently proposed a similar resampling scheme by adapting the *dependent* multiplier bootstrap of Bühlmann ([15], Section 3.3) to the process $\mathbb{C}_n^\circ$ defined in (1.2). Their empirical investigations indicate that the latter outperforms in finite samples a block bootstrap based on the work of Künsch [30] and Bühlmann [14]. Note that the idea of *dependent multipliers* appearing in Bühlmann ([15], Section 3.3) can also be found in Chen and Fan ([16], Section 5.1) and was recently independently rediscovered by Shao [44] in the context of the *smooth function model* but not in the empirical process setting. For the sample mean as statistic of interest, the latter author connected this resampling technique to the *tapered block bootstrap* of Paparoditis and Politis [33].

The main aim of this work is to provide an extended version of the multiplier resampling scheme of Bücher and Ruppert [11] adapted to the two-sided sequential process $\mathbb{C}_n$ defined in (1.1). The influence of the parameters of the resulting bootstrap procedure is studied in detail, both theoretically and by means of extensive simulations. An important contribution of the paper is an approach for estimating the key bandwidth parameter which plays a role somehow analogous to that of the block length in the block bootstrap. As a practical consequence, the resulting dependent multiplier technique for $\mathbb{C}_n$ can be used in a fully automatic way and many of the existing multiplier tests on the unknown copula C derived in the case of i.i.d. observations can be transposed to the strongly mixing case. In addition, due to its sequential nature, the resampling scheme can be used to derive nonparametric tests for change-point detection particularly sensitive to changes in the copula. This last point will be discussed in more detail in Section 4, and is also the subject of a companion paper (Bücher *et al.* [10]). Finally, the obtained results could be used to develop statistical inference procedures for Markovian copula time series models as introduced in Darsow, Nguyen and Olsen [18]. Based on recent results from Beare [5] on the mixing properties of these time series, one could, for instance, apply the proposed multiplier bootstrap to derive uniform confidence bands for the empirical copula or to develop tests for simple goodness-of-fit hypotheses on the copula in theses models.

There are two important by-products of this work that can be of independent interest. First, the validity of a sequential version of the dependent multiplier bootstrap for empirical processes of Bühlmann ([15], Section 3.3) (which has also been considered in Bücher and Ruppert [11], proof of Proposition 2) is obtained under weaker conditions on the rate of decay of the strong mixing coefficients and the multipliers. The derived result is based on a sequential unconditional multiplier central limit theorem for the multivariate empirical process indexed by lower-left orthants that is adapted to the case of strongly mixing observations. Second, the weak convergence of the two-sided sequential empirical copula process $\mathbb{C}_n$ is established under many serial dependence scenarios, including mild strong mixing conditions.

The paper is organized as follows. The second section presents a sequential extension of the seminal work of Bühlmann ([15], Section 3.3). In the third section, the asymptotics of the two-sided sequential empirical copula process $\mathbb{C}_n$ are obtained under many serial dependence conditions. Based on the results of the second and third sections, a dependent multiplier bootstrap for $\mathbb{C}_n$ is derived next. In the fifth section, the practical steps necessary to carry out the derived bootstrap are examined. In particular, a procedure for estimating the key bandwidth parameter of the dependent multiplier bootstrap is proposed by adapting to the empirical process setting the approach put forward in Politis and White [38] and Patton, Politis and White [34], among others. In addition, two ways of generating dependent multiplier sequences central to this resampling technique are discussed. The last section partially reports the results of large-scale Monte Carlo experiments whose aim was to investigate the influence in finite samples of the various parameters involved in the dependent multiplier bootstrap for $\mathbb{C}_n$.

The following notation is used in the sequel. The arrow "$\rightsquigarrow$" denotes weak convergence in the sense of Definition 1.3.3 in van der Vaart and Wellner [48], and, given a set T, $\ell^\infty(T)$ (resp., $\mathcal{C}(T)$) represents the space of all bounded (resp., continuous) real-valued functions on T equipped with the uniform metric.

2. A dependent multiplier bootstrap for the multivariate empirical process under strong mixing

The multiplier bootstrap of Rémillard and Scaillet [39] that has been adopted as a resampling technique in the case of i.i.d. observations in many tests on the unknown copula C is a consequence of the multiplier central limit theorem for empirical processes (see, e.g., Kosorok [29], Theorem 10.1 and Corollary 10.3). A sequential version of the previous result can be proved (see Holmes, Kojadinovic and Quessy [24], Theorem 1) by using the method of proof adopted in van der Vaart and Wellner ([48], Theorem 2.12.1). While investigating the block bootstrap for empirical processes constructed from strongly mixing observations, Bühlmann ([15], Section 3.3) obtained what resembles to a conditional version of the multiplier central limit theorem, subsequently also referred to as a *dependent multiplier bootstrap* (note that a sequential version of this result appears in the proof of Proposition 2 of Bücher and Ruppert [11]). The main idea of Bühlmann is to replace i.i.d. multipliers by suitable serially dependent multipliers. In the rest of the paper, we say that a sequence of random variables $(\xi_{i,n})_{i\in\mathbb{Z}}$ is a *dependent multiplier sequence* if:

(M1) The sequence $(\xi_{i,n})_{i\in\mathbb{Z}}$ is strictly stationary with $\mathrm{E}(\xi_{0,n}) = 0$, $\mathrm{E}(\xi_{0,n}^2) = 1$ and $\sup_{n\geq 1} \mathrm{E}(|\xi_{0,n}|^\nu) < \infty$ for all $\nu \geq 1$, and is independent of the available sample $\mathbf{X}_1, \dots, \mathbf{X}_n$.

(M2) There exists a sequence $\ell_n \to \infty$ of strictly positive constants such that $\ell_n = \mathrm{o}(n)$ and the sequence $(\xi_{i,n})_{i\in\mathbb{Z}}$ is ℓ_n-dependent, that is, $\xi_{i,n}$ is independent of $\xi_{i+h,n}$ for all $h > \ell_n$ and $i \in \mathbb{N}$.

(M3) There exists a function $\varphi : \mathbb{R} \to [0, 1]$, symmetric around 0, continuous at 0, satisfying $\varphi(0) = 1$ and $\varphi(x) = 0$ for all $|x| > 1$ such that $\mathrm{E}(\xi_{0,n}\xi_{h,n}) = \varphi(h/\ell_n)$ for all $h \in \mathbb{Z}$.

To state the main result of this section, we need to introduce additional notation and definitions. Let $\mathbf{U}_1, \dots, \mathbf{U}_n$ be the unobservable sample obtained from $\mathbf{X}_1, \dots, \mathbf{X}_n$ by the probability integral transforms $U_{ij} = F_j(X_{ij})$, $i \in \{1, \dots, n\}$, $j \in \{1, \dots, d\}$. It follows that $\mathbf{U}_1, \dots, \mathbf{U}_n$ is a marginally uniform d-dimensional sample from the unknown c.d.f. C. The corresponding sequential empirical process is then defined as

$$\tilde{\mathbb{B}}_n(s, \mathbf{u}) = \frac{1}{\sqrt{n}} \sum_{i=1}^{\lfloor ns \rfloor} \{\mathbf{1}(\mathbf{U}_i \leq \mathbf{u}) - C(\mathbf{u})\}, \qquad (s, \mathbf{u}) \in [0, 1]^{d+1}. \tag{2.1}$$

Note that, in the rest of the paper, the notation of most of the quantities that are directly computed from the unobservable sample $\mathbf{U}_1, \dots, \mathbf{U}_n$ will involve the symbol "$\sim$."

Furthermore, let M be a large integer and let $(\xi_{i,n}^{(1)})_{i \in \mathbb{Z}}, \dots, (\xi_{i,n}^{(M)})_{i \in \mathbb{Z}}$ be M independent copies of the same dependent multiplier sequence. Then, for any $m \in \{1, \dots, M\}$ and $(s, \mathbf{u}) \in [0, 1]^{d+1}$, let

$$\tilde{\mathbb{B}}_n^{(m)}(s, \mathbf{u}) = \frac{1}{\sqrt{n}} \sum_{i=1}^{\lfloor ns \rfloor} \xi_{i,n}^{(m)} \{\mathbf{1}(\mathbf{U}_i \leq \mathbf{u}) - C(\mathbf{u})\}. \tag{2.2}$$

From the previous display, we see that the bandwidth sequence ℓ_n defined in assumption (M2) plays a role somehow analogous to that of the *block length* in the block bootstrap. Two ways of forming the dependent multiplier sequences $(\xi_{i,n}^{(m)})_{i \in \mathbb{Z}}$ will be presented in Section 5.2.

Finally, for the sake of completeness, let us recall the notion of *strongly mixing sequence*. For a sequence of d-dimensional random vectors $(\mathbf{Y}_i)_{i \in \mathbb{Z}}$, the σ-field generated by $(\mathbf{Y}_i)_{a \leq i \leq b}$, $a, b \in \mathbb{Z} \cup \{-\infty, +\infty\}$, is denoted by $\mathcal{F}_a^b$. The strong mixing coefficients corresponding to the sequence $(\mathbf{Y}_i)_{i \in \mathbb{Z}}$ are then defined by $\alpha_0 = 1/2$ and

$$\alpha_r = \sup_{p \in \mathbb{Z}} \sup_{A \in \mathcal{F}_{-\infty}^p, B \in \mathcal{F}_{p+r}^{+\infty}} \left|\mathrm{P}(A \cap B) - \mathrm{P}(A)\mathrm{P}(B)\right|, \qquad r \in \mathbb{N}, r > 0.$$

The sequence $(\mathbf{Y}_i)_{i \in \mathbb{Z}}$ is said to be *strongly mixing* if $\alpha_r \to 0$ as $r \to \infty$.

The following result, inspired by Bühlmann ([15], Section 3.3), could be regarded as an extension of the multiplier central limit theorem to the sequential and strongly mixing setting for empirical processes indexed by lower-left orthants. Its proof is given in Appendix A.

Theorem 2.1 (Dependent multiplier central limit theorem). *Assume that $\ell_n = \mathrm{O}(n^{1/2-\varepsilon})$ for some $0 < \varepsilon < 1/2$ and that $\mathbf{U}_1, \dots, \mathbf{U}_n$ is drawn from a strictly stationary sequence $(\mathbf{U}_i)_{i \in \mathbb{Z}}$ whose strong mixing coefficients satisfy $\alpha_r = \mathrm{O}(r^{-a})$, $a > 3 + 3d/2$. Then,*

$$\left(\tilde{\mathbb{B}}_n, \tilde{\mathbb{B}}_n^{(1)}, \dots, \tilde{\mathbb{B}}_n^{(M)}\right) \rightsquigarrow \left(\mathbb{B}_C, \mathbb{B}_C^{(1)}, \dots, \mathbb{B}_C^{(M)}\right)$$

in $\{\ell^\infty([0, 1]^{d+1})\}^{M+1}$, where $\mathbb{B}_C$ is the weak limit of the sequential empirical process $\tilde{\mathbb{B}}_n$ defined in (2.1), and $\mathbb{B}_C^{(1)}, \dots, \mathbb{B}_C^{(M)}$ are independent copies of $\mathbb{B}_C$.

Before commenting on the result and the assumptions of the above theorem, let us state a corollary that can be regarded as an unconditional and sequential analogue of Theorem 3.2 of Bühlmann [15], and may be of interest for applications of empirical processes outside the scope of copulas. Recall that $\mathbf{X}_1, \ldots, \mathbf{X}_n$ is drawn from a strictly stationary sequence of continuous d-dimensional random vectors with c.d.f. F and that the margins of F are denoted by $F_1, \ldots, F_d$. Then, let

$$\mathbb{Z}_n(s, \mathbf{x}) = \frac{1}{\sqrt{n}} \sum_{i=1}^{\lfloor ns \rfloor} \{\mathbf{1}(\mathbf{X}_i \le \mathbf{x}) - F(\mathbf{x})\}, \qquad (s, \mathbf{x}) \in [0, 1] \times \overline{\mathbb{R}}^d,$$

be the usual sequential empirical process based on the observed sequence $\mathbf{X}_1, \ldots, \mathbf{X}_n$ and, for any $m \in \{1, \ldots, M\}$, let

$$\hat{\mathbb{Z}}_n^{(m)}(s, \mathbf{x}) = \frac{1}{\sqrt{n}} \sum_{i=1}^{\lfloor ns \rfloor} \xi_{i,n}^{(m)} \{\mathbf{1}(\mathbf{X}_i \le \mathbf{x}) - F_n(\mathbf{x})\}, \qquad (s, \mathbf{x}) \in [0, 1] \times \overline{\mathbb{R}}^d,$$

where $\overline{\mathbb{R}} = [-\infty, \infty]$ and F_n is the empirical c.d.f. computed from $\mathbf{X}_1, \ldots, \mathbf{X}_n$. The following corollary is then a consequence of the fact that $\mathbb{Z}_n(s, \mathbf{x}) = \tilde{\mathbb{B}}_n\{s, F_1(x_1), \ldots, F_d(x_d)\}$ for all $(s, \mathbf{x}) \in [0, 1] \times \overline{\mathbb{R}}^d$ and that, under the conditions of Theorem 2.1, for all $m \in \{1, \ldots, M\}$,

$$\sup_{(s,\mathbf{x}) \in [0,1] \times \overline{\mathbb{R}}^d} \left| \hat{\mathbb{Z}}_n^{(m)}(s, \mathbf{x}) - \tilde{\mathbb{B}}_n^{(m)}\{s, F_1(x_1), \ldots, F_d(x_d)\} \right| \xrightarrow{\mathrm{P}} 0,$$

a proof of which follows from the proof of Lemma A.3 in the supplementary material (Bücher and Kojadinovic [9]).

Corollary 2.2 (Dependent multiplier bootstrap for $\mathbb{Z}_n$). *Assume that $\ell_n = \mathrm{O}(n^{1/2-\varepsilon})$ for some $0 < \varepsilon < 1/2$ and that $\mathbf{X}_1, \ldots, \mathbf{X}_n$ is drawn from a strictly stationary sequence $(\mathbf{X}_i)_{i \in \mathbb{Z}}$ of continuous d-dimensional random vectors whose strong mixing coefficients satisfy $\alpha_r = \mathrm{O}(r^{-a})$, $a > 3 + 3d/2$. Then,*

$$\left(\mathbb{Z}_n, \hat{\mathbb{Z}}_n^{(1)}, \ldots, \hat{\mathbb{Z}}_n^{(M)}\right) \rightsquigarrow \left(\mathbb{Z}_F, \mathbb{Z}_F^{(1)}, \ldots, \mathbb{Z}_F^{(M)}\right)$$

in $\{\ell^\infty([0, 1] \times \overline{\mathbb{R}}^d)\}^{M+1}$, where $\mathbb{Z}_F$ is the weak limit of $\mathbb{Z}_n$, and $\mathbb{Z}_F^{(1)}, \ldots, \mathbb{Z}_F^{(M)}$ are independent copies of $\mathbb{Z}_F$.

Remark 2.3. In the literature, the "validity" (or "consistency") of a bootstrap procedure is often shown by establishing weak convergence of conditional laws (see, e.g., van der Vaart [47], Chapter 23). In most theoretical developments of this type, the necessary additional step of approximating conditional laws by simulation from the random resampling mechanism sufficiently many times is typically omitted (van der Vaart [47], page 329). An appropriate unconditional weak convergence result of the form of the one established in Corollary 2.2 (see also Segers [43], and references therein for other examples) already includes the repetition of the random resampling mechanism and can be used to deduce consistency of a bootstrap procedure in many situations of practical interest. A rather general result in that direction is provided in Proposition F.1 of the supplementary material (Bücher and Kojadinovic [9]). As an important consequence, in many situations of practical interest, both paradigms (conditional and unconditional) can be used, and one can choose the approach that appears to be easiest for the particular problem at hand. In the empirical process setting, we tend to favor the unconditional paradigm as the usual workhorses for empirical process theory, the (extended) continuous mapping theorem and the functional delta method, appear to be applicable under less restrictive conditions in an unconditional setting (see, e.g., Kosorok [29], Section 10.1.4).

From a practical perspective, Corollary 2.2 is, for instance, a first necessary step to transpose to the strongly mixing setting the goodness-of-fit and nonparametric change-point tests based on empirical c.d.f.s considered in Kojadinovic and Yan [28] and Holmes, Kojadinovic and Quessy [24], respectively.

We end this section by a few comments on the assumptions of Theorem 2.1 and Corollary 2.2:

- The requirement that $\ell_n = \mathrm{O}(n^{1/2-\varepsilon})$ for some $0 < \varepsilon < 1/2$ is used for proving the finite-dimensional convergence involved in Theorem 2.1, while the condition $\alpha_r = \mathrm{O}(r^{-a})$, $a > 3 + 3d/2$, is needed for the proof of the asymptotic equicontinuity.
- Theorem 3.2 of Bühlmann [15] can be regarded as a nonsequential conditional analogue of Corollary 2.2 with slightly more constrained multiplier random variables. The condition on the rate of decay of the strong mixing coefficients in that result is $\sum_{r=0}^{\infty}(r+1)^p \alpha_r^{1/2} < \infty$ with $p = \max\{8d + 12, \lfloor 2/\varepsilon \rfloor + 1\}$ and is therefore stronger than the condition involved in Theorem 2.1.
- The condition on the strong mixing coefficients in Theorem 2.1 and Corollary 2.2 is clearly satisfied if $\mathbf{X}_1, \ldots, \mathbf{X}_n$ are i.i.d., so that the above unconditional resampling scheme remains valid for independent observations. In the latter case however, the Monte Carlo experiments carried out in Bücher and Ruppert [11] suggest that a simpler scheme with i.i.d. multipliers (based, e.g., on Theorem 1 of Holmes, Kojadinovic and Quessy [24]) will lead to better finite-sample performance. As noted by a referee, this was to be expected since the use of a resampling scheme designed to capture dependence for observations that are serially independent should naturally result in an efficiency loss, especially if the tuning parameter is estimated.

3. Asymptotics of the sequential empirical copula process under serial dependence

In the case of i.i.d. observations, the classical empirical copula process turns out to be asymptotically equivalent to a linear functional of the multivariate sequential empirical process $\tilde{\mathbb{B}}_n$ defined in (2.1) (see Segers [43], Proposition 4.3). This representation is at the heart of the multiplier bootstrap of Rémillard and Scaillet [39]. Obtaining such an asymptotic representation for the two-sided sequential empirical copula process $\mathbb{C}_n$ defined in (1.1) is therefore a preliminary step before a *dependent* multiplier bootstrap for $\mathbb{C}_n$ under strong mixing can be derived as a consequence of Theorem 2.1. The desired result is actually a corollary of a more general result. Indeed, in this section, the asymptotics of $\mathbb{C}_n$ are established under many serial dependence scenarios as a consequence of the weak convergence of the multivariate sequential empirical process $\tilde{\mathbb{B}}_n$. More specifically, the following condition is considered.

Condition 3.1. *The sample* $\mathbf{U}_1, \dots, \mathbf{U}_n$ *is drawn from a strictly stationary sequence* $(\mathbf{U}_i)_{i\in\mathbb{Z}}$ *such that* $\tilde{\mathbb{B}}_n$ *converges weakly in* $\ell^\infty([0,1]^{d+1})$ *to a tight centered Gaussian process* $\mathbb{B}_C$ *concentrated on*

$$\big\{\alpha^\star \in \mathcal{C}\big([0,1]^{d+1}\big) : \alpha^\star(s,\mathbf{u}) = 0 \textit{ if one of the components of } (s,\mathbf{u}) \textit{ is } 0 \textit{ and}$$
$$\alpha^\star(s,1,\dots,1) = 0 \textit{ for all } s \in (0,1]\big\}.$$

Note that, in the case of serial independence, the above condition is an immediate consequence of Theorem 2.12.1 of van der Vaart and Wellner [48]. As shall be discussed below, it is also met under strong mixing.

We also consider the following smoothness condition on C proposed by Segers [43]. As explained by the latter author, this condition is nonrestrictive in the sense that it is necessary for the candidate weak limit of $\mathbb{C}_n$ to exist pointwise and have continuous sample paths.

Condition 3.2. *For any* $j \in \{1,\dots,d\}$, *the partial derivatives* $\dot{C}_j = \partial C/\partial u_j$ *exist and are continuous on* $\{\mathbf{u} \in [0,1]^d : u_j \in (0,1)\}$.

As we continue, for any $j \in \{1,\dots,d\}$, we define $\dot{C}_j$ to be zero on the set $\{\mathbf{u} \in [0,1]^d : u_j \in \{0,1\}\}$ (see also Segers [43]; Bücher and Volgushev [13]). It then follows that, under Condition 3.2, $\dot{C}_j$ is defined on the whole of $[0,1]^d$. Also, for any $j \in \{1,\dots,d\}$ and any $\mathbf{u} \in [0,1]^d$, $\mathbf{u}^{(j)}$ is the vector of $[0,1]^d$ defined by $u_i^{(j)} = u_j$ if $i = j$ and 1 otherwise.

Finally, in order to study $\mathbb{C}_n$, we need to be able to easily go back and forth between normalized ranks and empirical quantile functions. To this end, ties must not occur. In the case of serial independence, it is sufficient to assume that the marginal distributions are continuous. However, in the case of serial dependence, continuity of the marginal distributions is *not* sufficient to guarantee the absence of ties (see, e.g., Bücher and Segers [12], Example 4.2). This leads to a last condition.

Condition 3.3. *For any* $j \in \{1,\dots,d\}$, *there are no ties in the component series* $X_{1j},\dots,X_{nj}$ *with probability one.*

The following theorem is the main result of this section. It is proved in Appendix B.

Theorem 3.4 (Asymptotics of the sequential empirical copula process). *Under Conditions* 3.1, 3.2 *and* 3.3,

$$\sup_{(s,t,\mathbf{u})\in\Delta\times[0,1]^d} \big|\mathbb{C}_n(s,t,\mathbf{u}) - \tilde{\mathbb{C}}_n(s,t,\mathbf{u})\big| \xrightarrow{\mathrm{P}} 0,$$

where

$$\tilde{\mathbb{C}}_n(s,t,\mathbf{u}) = \big\{\tilde{\mathbb{B}}_n(t,\mathbf{u}) - \tilde{\mathbb{B}}_n(s,\mathbf{u})\big\} - \sum_{j=1}^d \dot{C}_j(\mathbf{u})\big\{\tilde{\mathbb{B}}_n\big(t,\mathbf{u}^{(j)}\big) - \tilde{\mathbb{B}}_n\big(s,\mathbf{u}^{(j)}\big)\big\}. \tag{3.1}$$

Consequently, $\mathbb{C}_n \rightsquigarrow \mathbb{C}_C$ *in* $\ell^\infty(\Delta \times [0,1]^d)$, *where, for* $(s,t,\mathbf{u}) \in \Delta \times [0,1]^d$,

$$\mathbb{C}_C(s,t,\mathbf{u}) = \{\mathbb{B}_C(t,\mathbf{u}) - \mathbb{B}_C(s,\mathbf{u})\} - \sum_{j=1}^d \dot{C}_j(\mathbf{u})\{\mathbb{B}_C(t,\mathbf{u}^{(j)}) - \mathbb{B}_C(s,\mathbf{u}^{(j)})\}. \tag{3.2}$$

The asymptotics of $\mathbb{C}_n$ under strong mixing immediately follow from the previous theorem. The necessary tool is Theorem 1 of Bücher [7], which states that, if $\mathbf{U}_1,\dots,\mathbf{U}_n$ is drawn from a strictly stationary sequence $(\mathbf{U}_i)_{i\in\mathbb{Z}}$ whose strong mixing coefficients satisfy $\alpha_r = \mathrm{O}(r^{-a})$, $a>1$, then $\tilde{\mathbb{B}}_n \rightsquigarrow \mathbb{B}_C$ in $\ell^\infty([0,1]^{d+1})$. In other words, $\mathbf{U}_1,\dots,\mathbf{U}_n$ satisfies Condition 3.1.

Corollary 3.5. *Assume that* $\mathbf{X}_1,\dots,\mathbf{X}_n$ *is drawn from a strictly stationary sequence* $(\mathbf{X}_i)_{i\in\mathbb{Z}}$ *whose strong mixing coefficients satisfy* $\alpha_r = \mathrm{O}(r^{-a})$, $a>1$. *Then, under Conditions* 3.2 *and* 3.3,

$$\sup_{(s,t,\mathbf{u})\in\Delta\times[0,1]^d} \left|\mathbb{C}_n(s,t,\mathbf{u}) - \tilde{\mathbb{C}}_n(s,t,\mathbf{u})\right| \xrightarrow{\mathrm{P}} 0,$$

where $\tilde{\mathbb{C}}_n$ *is defined in* (3.1).

The conditions of the above corollary are, for instance, satisfied (with much to spare) when $\mathbf{X}_1,\dots,\mathbf{X}_n$ is drawn from a stationary vector ARMA process with absolutely continuous innovations (see Mokkadem [32]).

4. A dependent multiplier bootstrap for $\mathbb{C}_n$ under strong mixing

Analogously to the approach adopted in Rémillard and Scaillet [39] (see also Segers [43]), we shall now combine the asymptotic representation for $\mathbb{C}_n$ stated in Corollary 3.5 with Theorem 2.1 to show the validity of a dependent multiplier bootstrap for $\mathbb{C}_n$ under strong mixing. The corresponding result, stated in Proposition 4.2 below, can be regarded as an extension of Proposition 2 in Bücher and Ruppert [11], where a similar but conditional result was established for the process $\mathbb{C}_n^\circ$ defined in (1.2) under stricter conditions on the mixing rate and the multipliers.

The underlying idea is as follows: the fact that the limiting vector of processes in Theorem 2.1 has independent components suggests regarding $\tilde{\mathbb{B}}_n^{(1)},\dots,\tilde{\mathbb{B}}_n^{(M)}$ as "almost" independent copies of $\tilde{\mathbb{B}}_n$ when n is large. Unfortunately, the $\tilde{\mathbb{B}}_n^{(m)}$ cannot be computed because C is unknown and the sample $\mathbf{U}_1,\dots,\mathbf{U}_n$ is unobservable. Estimating C by the empirical copula $C_{1:n}$ and $\mathbf{U}_1,\dots,\mathbf{U}_n$ by the pseudo-observations $\hat{\mathbf{U}}_1^{1:n},\dots,\hat{\mathbf{U}}_n^{1:n}$, we obtain the following computable version of $\tilde{\mathbb{B}}_n^{(m)}$ defined, for any $(s,\mathbf{u}) \in [0,1]^{d+1}$, by

$$\hat{\mathbb{B}}_n^{(m)}(s,\mathbf{u}) = \frac{1}{\sqrt{n}} \sum_{i=1}^{\lfloor ns\rfloor} \xi_{i,n}^{(m)}\{\mathbf{1}(\hat{\mathbf{U}}_i^{1:n} \le \mathbf{u}) - C_{1:n}(\mathbf{u})\}. \tag{4.1}$$

Starting from the asymptotic representation of $\mathbb{C}_n$ in terms of $\tilde{\mathbb{B}}_n$ stated in Corollary 3.5, we see that, to obtain "almost" independent copies of $\mathbb{C}_n$ for large n in the spirit of Rémillard and Scaillet [39], we additionally need to estimate the partial derivatives $\dot{C}_j$, $j \in \{1,\dots,d\}$, appearing in (3.2). As we continue, we consider estimators $\dot{C}_{j,n}$ of $\dot{C}_j$ satisfying the following condition put forward in Segers [43].

Condition 4.1. *There exists a constant $K > 0$ such that $|\dot{C}_{j,n}(\mathbf{u})| \le K$ for all $j \in \{1, \ldots, d\}$, $n \ge 1$ and $\mathbf{u} \in [0,1]^d$, and, for any $\delta \in (0, 1/2)$ and $j \in \{1, \ldots, d\}$,*

$$\sup_{\substack{\mathbf{u}\in[0,1]^d \\ u_j\in[\delta,1-\delta]}} \left|\dot{C}_{j,n}(\mathbf{u}) - \dot{C}_j(\mathbf{u})\right| \xrightarrow{\mathrm{P}} 0.$$

Three estimators of the partial derivatives satisfying Condition 4.1 are discussed in Section 5.3.

We can now define empirical processes that can be fully computed and that, under appropriate conditions, can be regarded as "almost" independent copies of $\mathbb{C}_n$ for large n. For any $m \in \{1, \ldots, M\}$ and $(s, t, \mathbf{u}) \in \Delta \times [0, 1]^d$, let

$$\begin{aligned}\hat{\mathbb{C}}_n^{(m)}(s,t,\mathbf{u}) &= \left\{\hat{\mathbb{B}}_n^{(m)}(t,\mathbf{u}) - \hat{\mathbb{B}}_n^{(m)}(s,\mathbf{u})\right\} \\ &\quad - \sum_{j=1}^d \dot{C}_{j,n}(\mathbf{u})\left\{\hat{\mathbb{B}}_n^{(m)}\left(t,\mathbf{u}^{(j)}\right) - \hat{\mathbb{B}}_n^{(m)}\left(s,\mathbf{u}^{(j)}\right)\right\}.\end{aligned} \tag{4.2}$$

The following proposition is a consequence of Corollary 3.5 and Theorem 2.1 and can be proved by adapting the arguments of Segers ([43], proof of Proposition 4.3) to the current sequential and strongly mixing setting. Its proof can be found in Section D of the supplementary material (Bücher and Kojadinovic [9]).

Proposition 4.2 (Dependent multiplier bootstrap for $\mathbb{C}_n$). *Assume that $\ell_n = \mathrm{O}(n^{1/2-\varepsilon})$ for some $0 < \varepsilon < 1/2$ and that $\mathbf{X}_1, \ldots, \mathbf{X}_n$ is drawn from a strictly stationary sequence $(\mathbf{X}_i)_{i\in\mathbb{Z}}$ whose strong mixing coefficients satisfy $\alpha_r = \mathrm{O}(r^{-a})$, $a > 3 + 3d/2$. Then, under Conditions 3.2, 3.3 and 4.1,*

$$\left(\mathbb{C}_n, \hat{\mathbb{C}}_n^{(1)}, \ldots, \hat{\mathbb{C}}_n^{(M)}\right) \rightsquigarrow \left(\mathbb{C}_C, \mathbb{C}_C^{(1)}, \ldots, \mathbb{C}_C^{(M)}\right)$$

in $\{\ell^\infty(\Delta \times [0,1]^d)\}^{M+1}$, where $\mathbb{C}_C$ is the weak limit of the two-sided sequential empirical copula process $\mathbb{C}_n$ defined in (3.2), and $\mathbb{C}_C^{(1)}, \ldots, \mathbb{C}_C^{(M)}$ are independent copies of $\mathbb{C}_C$.

We end this section by briefly illustrating how Proposition 4.2 can be used in the context of change-point detection. As discussed in Bücher *et al.* [10], a broad class of nonparametric tests for change-point detection particularly sensitive to changes in the copula can be derived from the process

$$\mathbb{D}_n(s,\mathbf{u}) = \sqrt{n}\lambda_n(0,s)\lambda_n(s,1)\left\{C_{1:\lfloor ns\rfloor}(\mathbf{u}) - C_{\lfloor ns\rfloor+1:n}(\mathbf{u})\right\}, \qquad (s,\mathbf{u}) \in [0,1]^{d+1}.$$

The above definition is a mere transposition to the copula context of the "classical construction" adopted, for instance, in Csörgő and Horváth ([17], Section 2.6). Under the null hypothesis of no change in the distribution, the process $\mathbb{D}_n$ can be simply rewritten as

$$\mathbb{D}_n(s,\mathbf{u}) = \lambda_n(s,1)\mathbb{C}_n(0,s,\mathbf{u}) - \lambda_n(0,s)\mathbb{C}_n(s,1,\mathbf{u}), \qquad (s,\mathbf{u}) \in [0,1]^{d+1}.$$

To be able to compute approximate p-values for statistics derived from $\mathbb{D}_n$ (given the unwieldy nature of the weak limit of $\mathbb{D}_n$), it is then natural to define the processes

$$\hat{\mathbb{D}}_n^{(m)}(s,\mathbf{u}) = \lambda_n(s,1)\hat{\mathbb{C}}_n^{(m)}(0,s,\mathbf{u}) - \lambda_n(0,s)\hat{\mathbb{C}}_n^{(m)}(s,1,\mathbf{u}), \qquad (s,\mathbf{u}) \in [0,1]^{d+1},$$

$m \in \{1,\dots,M\}$, which could be thought of as "almost" independent copies of $\mathbb{D}_n$ under the null hypothesis of no change in the distribution. Under the null and the conditions of Proposition 4.2, we immediately obtain from Proposition 4.2 that $\mathbb{D}_n, \hat{\mathbb{D}}_n^{(1)},\dots,\hat{\mathbb{D}}_n^{(M)}$ weakly converge jointly to independent copies of the same limit. As discussed in Remark 2.3, the latter result is the key step for establishing that classical tests based on $\mathbb{D}_n$ hold their level asymptotically. To illustrate this point further, let us focus on the Kolmogorov–Smirnov statistic $W_n = \sup_{(s,\mathbf{u})\in[0,1]^{d+1}} |\mathbb{D}_n(s,\mathbf{u})|$ and let $\hat{W}_n^{(m)} = \sup_{(s,\mathbf{u})\in[0,1]^{d+1}} |\hat{\mathbb{D}}_n^{(m)}(s,\mathbf{u})|$, $m \in \{1,\dots,M\}$. The continuous mapping theorem then implies that, under the null and the conditions of Proposition 4.2, $(W_n, W_n^{(1)},\dots,W_n^{(M)}) \rightsquigarrow (W, W^{(1)},\dots,W^{(M)})$, where W, the weak limit of W_n, is a continuous random variable, and $W^{(1)},\dots,W^{(M)}$ are independent copies of W. The above unconditional result ensures that the conclusion of Proposition F.1 in Section F of the supplementary material (Bücher and Kojadinovic [9]) holds, which implies that a test based on W_n whose approximate p-value is computed as $M^{-1}\sum_{m=1}^M \mathbf{1}(\hat{W}_n^{(m)} \geq W_n)$ will hold its level asymptotically as $n \to \infty$ followed by $M \to \infty$. To show that such a test is consistent under the alternative of changes in the copula only, one typically needs to prove that $n^{-1/2}W_n \xrightarrow{\mathrm{P}} c > 0$ and that, for any $m \in \{1,\dots,M\}$, $W_n^{(m)} = O_{\mathrm{P}}(\ell_n^{1/2})$, also under the alternative (see, e.g., Inoue [25], Theorem 2.5 for related results in the context of nonparametric change-point detection in multivariate c.d.f.s).

Additional details, simulation results as well as illustrations on financial data can be found in Bücher *et al.* [10] for tests based on maximally selected Cramér–von Mises statistics.

5. Practical issues

The practical use of the derived dependent multiplier bootstrap for $\mathbb{C}_n$ requires the generation of dependent multiplier sequences and the estimation of the partial derivatives of the copula. These two practical issues are discussed in the second and third subsection below, while the first subsection addresses the key choice of the bandwidth parameter ℓ_n involved in the definition of dependent multiplier sequences.

5.1. Estimation of the bandwidth parameter ℓ_n

The bandwidth parameter ℓ_n defined in assumption (M2) plays a role somehow similar to that of the block length in the block bootstrap of Künsch [30]. Its value is therefore expected to have a crucial influence on the finite-sample performance of the dependent multiplier bootstrap for $\mathbb{C}_n$.

The choice of a similar bandwidth parameter is discussed, for instance, in Paparoditis and Politis [33] for the tapered block bootstrap using results from Künsch [30]. Related results are presented in Bühlmann ([15], Lemmas 3.12 and 3.13) and Shao ([44], Proposition 2.1) for the dependent multiplier bootstrap when the statistic of interest is the sample mean. The aim of this section is to extend the aforementioned results to the dependent multiplier bootstrap for $\mathbb{C}_n$ and propose an estimator of ℓ_n in the spirit of those investigated in Paparoditis and Politis [33], Politis and White [38] and Patton, Politis and White [34] for other resampling schemes. Since the dependent multiplier bootstrap for $\mathbb{C}_n$ is based on the corresponding bootstrap approximation for $\tilde{\mathbb{B}}_n$, we propose to base our estimator of the bandwidth parameter on the accuracy of the latter technique.

Let E_ξ and Cov_ξ denote the expectation and covariance, respectively, conditional on the data $\mathbf{X}_1, \dots, \mathbf{X}_n$, and, for any $\mathbf{u}, \mathbf{v} \in [0,1]^d$, let $\sigma_C(\mathbf{u}, \mathbf{v}) = \mathrm{Cov}\{\mathbb{B}_C(1,\mathbf{u}), \mathbb{B}_C(1,\mathbf{v})\}$. Now, fix $m \in \{1, \dots, M\}$ and, for any $\mathbf{u}, \mathbf{v} \in [0,1]^d$, let

$$
\begin{aligned}
\tilde{\sigma}_n(\mathbf{u}, \mathbf{v}) &= \mathrm{Cov}_\xi\big\{\tilde{\mathbb{B}}_n^{(m)}(1,\mathbf{u}), \tilde{\mathbb{B}}_n^{(m)}(1,\mathbf{v})\big\} \\
&= \mathrm{E}_\xi\big\{\tilde{\mathbb{B}}_n^{(m)}(1,\mathbf{u})\tilde{\mathbb{B}}_n^{(m)}(1,\mathbf{v})\big\} \\
&= \frac{1}{n}\sum_{i,j=1}^{n} \mathrm{E}_\xi\big(\xi_{i,n}^{(m)}\xi_{j,n}^{(m)}\big)\big\{\mathbf{1}(\mathbf{U}_i \le \mathbf{u}) - C(\mathbf{u})\big\}\big\{\mathbf{1}(\mathbf{U}_j \le \mathbf{v}) - C(\mathbf{v})\big\} \\
&= \frac{1}{n}\sum_{i,j=1}^{n} \varphi\big\{(i-j)/\ell_n\big\}\big\{\mathbf{1}(\mathbf{U}_i \le \mathbf{u}) - C(\mathbf{u})\big\}\big\{\mathbf{1}(\mathbf{U}_j \le \mathbf{v}) - C(\mathbf{v})\big\},
\end{aligned}
\tag{5.1}
$$

where $\tilde{\mathbb{B}}_n^{(m)}$ is defined in (2.2). For the moment, although it is based on the unobservable sample $\mathbf{U}_1, \dots, \mathbf{U}_n$ and the unknown copula C, we shall regard $\tilde{\sigma}_n(\mathbf{u}, \mathbf{v})$ as an estimator of $\sigma_C(\mathbf{u}, \mathbf{v})$.

The following two results extend Lemmas 3.12 and 3.13 of Bühlmann [15] and Proposition 2.1 of Shao [44]. They can be proved by adapting the arguments used in the proofs of Lemmas 3.12 and 3.13 of Bühlmann [15]. The resulting proofs are given in the supplementary material (Bücher and Kojadinovic [9]) for completeness.

Proposition 5.1. *Assume that* $\ell_n = \mathrm{O}(n^{1/2-\varepsilon})$ *for some* $0 < \varepsilon < 1/2$, *that* $\mathbf{U}_1, \dots, \mathbf{U}_n$ *is drawn from a strictly stationary sequence* $(\mathbf{U}_i)_{i\in\mathbb{Z}}$ *whose strong mixing coefficients satisfy* $\alpha_r = \mathrm{O}(r^{-a})$, $a > 3$, *and that* φ *defined in assumption* (M3) *is additionally twice continuously differentiable on* $[-1,1]$ *with* $\varphi''(0) \ne 0$. *Then, for any* $\mathbf{u}, \mathbf{v} \in [0,1]^d$,

$$
\mathrm{E}\big\{\tilde{\sigma}_n(\mathbf{u}, \mathbf{v})\big\} - \sigma_C(\mathbf{u}, \mathbf{v}) = \frac{\Gamma(\mathbf{u}, \mathbf{v})}{\ell_n^2} + r_{n,1}(\mathbf{u}, \mathbf{v}),
$$

where $\sup_{\mathbf{u},\mathbf{v}\in[0,1]^d} |r_{n,1}(\mathbf{u}, \mathbf{v})| = \mathrm{o}(\ell_n^{-2})$ *and*

$$
\Gamma(\mathbf{u}, \mathbf{v}) = \frac{\varphi''(0)}{2}\sum_{k=-\infty}^{\infty} k^2 \gamma(k, \mathbf{u}, \mathbf{v}) \qquad \text{with } \gamma(k, \mathbf{u}, \mathbf{v}) = \mathrm{Cov}\big\{\mathbf{1}(\mathbf{U}_0 \le \mathbf{u}), \mathbf{1}(\mathbf{U}_k \le \mathbf{v})\big\}.
$$

Proposition 5.2. *Assume that* $\mathbf{U}_1, \ldots, \mathbf{U}_n$ *is drawn from a strictly stationary sequence* $(\mathbf{U}_i)_{i\in\mathbb{Z}}$ *whose strong mixing coefficients satisfy* $\alpha_r = \mathrm{O}(r^{-a})$, $a > 3$, *and that there exists* $\lambda > 0$ *such that* φ *defined in assumption* (M3) *additionally satisfies* $|\varphi(x) - \varphi(y)| \le \lambda|x - y|$ *for all* $x, y \in \mathbb{R}$. *Then, for any* $\mathbf{u}, \mathbf{v} \in [0, 1]^d$,

$$\mathrm{Var}\big\{\tilde\sigma_n(\mathbf{u},\mathbf{v})\big\} = \frac{\ell_n}{n}\Delta(\mathbf{u},\mathbf{v}) + r_{n,2}(\mathbf{u},\mathbf{v}),$$

where

$$\Delta(\mathbf{u},\mathbf{v}) = \left\{\int_{-1}^{1}\varphi(x)^2\,\mathrm{d}x\right\}\left[\sigma_C(\mathbf{u},\mathbf{u})\sigma_C(\mathbf{v},\mathbf{v}) + \big\{\sigma_C(\mathbf{u},\mathbf{v})\big\}^2\right]$$

and $\sup_{\mathbf{u},\mathbf{v}\in[0,1]^d}|r_{n,2}(\mathbf{u},\mathbf{v})| = \mathrm{o}(\ell_n/n)$.

Under the combined conditions of Propositions 5.1 and 5.2, we have that, for any $\mathbf{u}, \mathbf{v} \in [0, 1]^2$, the mean squared error of $\tilde\sigma_n(\mathbf{u},\mathbf{v})$ is

$$\mathrm{MSE}\big\{\tilde\sigma_n(\mathbf{u},\mathbf{v})\big\} = \frac{\{\Gamma(\mathbf{u},\mathbf{v})\}^2}{\ell_n^4} + \Delta(\mathbf{u},\mathbf{v})\frac{\ell_n}{n} + r_n(\mathbf{u},\mathbf{v}),$$

where $r_n(\mathbf{u},\mathbf{v}) = \{r_{n,1}(\mathbf{u},\mathbf{v})\}^2 + 2\Gamma(\mathbf{u},\mathbf{v})r_{n,1}(\mathbf{u},\mathbf{v})/\ell_n^2 + r_{n,2}(\mathbf{u},\mathbf{v})$. This allows us to define the integrated mean squared error

$$\mathrm{IMSE}_n = \int_{[0,1]^{2d}}\mathrm{MSE}\big\{\tilde\sigma_n(\mathbf{u},\mathbf{v})\big\}\,\mathrm{d}\mathbf{u}\,\mathrm{d}\mathbf{v} \sim \frac{\bar\Gamma^2}{\ell_n^4} + \bar\Delta\frac{\ell_n}{n}, \tag{5.2}$$

where

$$\bar\Gamma^2 = \int_{[0,1]^{2d}}\big\{\Gamma(\mathbf{u},\mathbf{v})\big\}^2\,\mathrm{d}\mathbf{u}\,\mathrm{d}\mathbf{v} \quad \text{and} \quad \bar\Delta = \int_{[0,1]^{2d}}\Delta(\mathbf{u},\mathbf{v})\,\mathrm{d}\mathbf{u}\,\mathrm{d}\mathbf{v}. \tag{5.3}$$

Notice that $\bar\Delta$ can be rewritten as

$$\bar\Delta = \left\{\int_{-1}^{1}\varphi(x)^2\,\mathrm{d}x\right\}\left[\left\{\int_{[0,1]^d}\sigma_C(\mathbf{u},\mathbf{u})\,\mathrm{d}\mathbf{u}\right\}^2 + \int_{[0,1]^{2d}}\big\{\sigma_C(\mathbf{u},\mathbf{v})\big\}^2\,\mathrm{d}\mathbf{u}\,\mathrm{d}\mathbf{v}\right].$$

Differentiating the function $x \mapsto \bar\Gamma^2/x^4 + \bar\Delta x/n$ and equating the derivative to zero, we obtain that the value of ℓ_n that minimizes IMSE_n is, asymptotically,

$$\ell_n^{\mathrm{opt}} = \left(\frac{4\bar\Gamma^2}{\bar\Delta}\right)^{1/5} n^{1/5}. \tag{5.4}$$

From (5.4), we see that, to estimate ℓ_n^{opt}, we need to estimate the infinite sums $K(\mathbf{u},\mathbf{v}) = \sum_{k\in\mathbb{Z}}k^2\gamma(k,\mathbf{u},\mathbf{v})$ and $\sigma_C(\mathbf{u},\mathbf{v}) = \sum_{k\in\mathbb{Z}}\gamma(k,\mathbf{u},\mathbf{v})$ for all $\mathbf{u},\mathbf{v} \in [0, 1]^d$. Should $\mathbf{U}_1, \ldots, \mathbf{U}_n$ be observable, this could be done by adapting the procedures described in Paparoditis and Politis

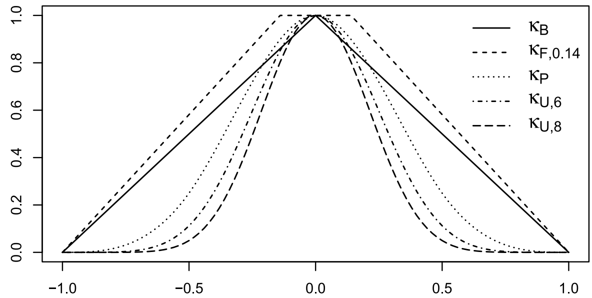

Figure 1. Graphs of the functions κ_B, $\kappa_{F,0.14}$ and κ_P, as well as $\kappa_{U,6}$ and $\kappa_{U,8}$ defined in Section 5.2.2.

([33], page 1111) or Politis and White ([38], Section 3) to the current empirical process setting. Let $L \geq 1$ be an integer to be determined from $\mathbf{X}_1, \ldots, \mathbf{X}_n$ later and fix $\mathbf{u}, \mathbf{v} \in [0,1]^d$. Proceeding in the spirit of Politis and Romano [37] and Politis [36], the quantity $K(\mathbf{u}, \mathbf{v})$ could be estimated by $\check{K}_n(\mathbf{u}, \mathbf{v}) = \sum_{k=-L}^{L} \kappa_{F,0.5}(k/L) k^2 \check{\gamma}_n(k, \mathbf{u}, \mathbf{v})$, where

$$\kappa_{F,c}(x) = \big[\big\{\big(1 - |x|\big)/(1-c)\big\} \vee 0\big] \wedge 1, \qquad c \in [0,1], \tag{5.5}$$

is the "flat top" (trapezoidal) kernel parametrized by $c \in [0,1]$ (see Figure 1), and $\check{\gamma}_n(k, \mathbf{u}, \mathbf{v})$ is the estimated cross-covariance at lag $k \in \{-(n-1), \ldots, n-1\}$, computed from the sequences $\{\mathbf{1}(\mathbf{U}_i \leq \mathbf{u})\}_{i \in \{1,\ldots,n\}}$ and $\{\mathbf{1}(\mathbf{U}_i \leq \mathbf{v})\}_{i \in \{1,\ldots,n\}}$, that is,

$$\check{\gamma}_n(k, \mathbf{u}, \mathbf{v}) = \begin{cases} n^{-1} \displaystyle\sum_{i=1}^{n-k} \big\{\mathbf{1}(\mathbf{U}_i \leq \mathbf{u}) - \tilde{H}_n(\mathbf{u})\big\}\big\{\mathbf{1}(\mathbf{U}_{i+k} \leq \mathbf{v}) - \tilde{H}_n(\mathbf{v})\big\}, & k \geq 0, \\ n^{-1} \displaystyle\sum_{i=1-k}^{n} \big\{\mathbf{1}(\mathbf{U}_i \leq \mathbf{u}) - \tilde{H}_n(\mathbf{u})\big\}\big\{\mathbf{1}(\mathbf{U}_{i+k} \leq \mathbf{v}) - \tilde{H}_n(\mathbf{v})\big\}, & k \leq 0, \end{cases}$$

with $\tilde{H}_n$ being the empirical c.d.f. computed from $\mathbf{U}_1, \ldots, \mathbf{U}_n$. Similarly, $\sigma_C(\mathbf{u}, \mathbf{v})$ could be estimated by

$$\check{\sigma}_n(\mathbf{u}, \mathbf{v}) = \sum_{k=-L}^{L} \kappa_{F,0.5}(k/L) \check{\gamma}_n(k, \mathbf{u}, \mathbf{v}).$$

As $\mathbf{U}_1, \ldots, \mathbf{U}_n$ is unobservable, it is natural to consider the sample of pseudo-observations $\hat{\mathbf{U}}_1^{1:n}, \ldots, \hat{\mathbf{U}}_n^{1:n}$ instead, and to replace $\check{\gamma}_n(k, \mathbf{u}, \mathbf{v})$ by

$$\hat{\gamma}_n(k, \mathbf{u}, \mathbf{v}) = \begin{cases} n^{-1} \displaystyle\sum_{i=1}^{n-k} \big\{\mathbf{1}\big(\hat{\mathbf{U}}_i^{1:n} \leq \mathbf{u}\big) - C_{1:n}(\mathbf{u})\big\}\big\{\mathbf{1}\big(\hat{\mathbf{U}}_{i+k}^{1:n} \leq \mathbf{v}\big) - C_{1:n}(\mathbf{v})\big\}, & k \geq 0, \\ n^{-1} \displaystyle\sum_{i=1-k}^{n} \big\{\mathbf{1}\big(\hat{\mathbf{U}}_i^{1:n} \leq \mathbf{u}\big) - C_{1:n}(\mathbf{u})\big\}\big\{\mathbf{1}\big(\hat{\mathbf{U}}_{i+k}^{1:n} \leq \mathbf{v}\big) - C_{1:n}(\mathbf{v})\big\}, & k \leq 0, \end{cases}$$

which gives the computable estimators

$$\begin{aligned}\hat{\sigma}_n(\mathbf{u},\mathbf{v}) &= \sum_{k=-L}^{L} \kappa_{F,0.5}(k/L)\hat{\gamma}_n(k,\mathbf{u},\mathbf{v}) \quad \text{and} \\ \hat{K}_n(\mathbf{u},\mathbf{v}) &= \sum_{k=-L}^{L} \kappa_{F,0.5}(k/L)k^2\hat{\gamma}_n(k,\mathbf{u},\mathbf{v})\end{aligned} \tag{5.6}$$

of $\sigma_C(\mathbf{u},\mathbf{v})$ and $\sum_{k\in\mathbb{Z}} k^2\gamma(\mathbf{u},\mathbf{v})$, respectively.

To estimate $\bar{\Gamma}^2$ and $\bar{\Delta}$ defined in (5.3), we then propose to use a grid $\{\mathbf{u}_i\}_{i\in\{1,\ldots,g\}}$ of g points uniformly spaced over $(0,1)^d$, and to compute

$$\hat{\bar{\Gamma}}_n^2 = \frac{\{\varphi''(0)\}^2}{4}\frac{1}{g^2}\sum_{i,j=1}^{g}\{\hat{K}_n(\mathbf{u}_i,\mathbf{u}_j)\}^2$$

and

$$\hat{\bar{\Delta}}_n = \left\{\int_{-1}^{1}\varphi(x)^2\,\mathrm{d}x\right\}\left(\left\{\frac{1}{g}\sum_{i=1}^{g}\hat{\sigma}_n(\mathbf{u}_i,\mathbf{u}_i)\right\}^2 + \frac{1}{g^2}\sum_{i,j=1}^{g}\{\hat{\sigma}_n(\mathbf{u}_i,\mathbf{u}_j)\}^2\right),$$

respectively. Plugging these into (5.4), we obtain an estimator of ℓ_n^{opt} which shall be denoted as $\hat{\ell}_n^{\mathrm{opt}}$ as we continue.

The above estimator depends on the choice of the integer L appearing in (5.6). To estimate L, we suggest proceeding along the lines of Politis and White ([38], Section 3.2) (see also Paparoditis and Politis [33], page 1112). Let $\hat{\rho}_j(k)$, $j\in\{1,\ldots,d\}$, be the autocorrelation function at lag k estimated from the sample $X_{1j},\ldots,X_{nj}$. For any $j\in\{1,\ldots,d\}$, let L_j be the smallest integer after which $\hat{\rho}_j(k)$ appears negligible. Notice that the latter can be determined automatically by means of the algorithm described in detail in Politis and White ([38], Section 3.2). Our implementation is based on Matlab code by A.J. Patton (available on his web page) and its R version by J. Racine and C. Parmeter. Then, we merely suggest taking $L=2\psi(L_1,\ldots,L_d)$, where ψ is some aggregation function such as the median, the mean, the minimum or the maximum. The previous approach is clearly not the only possible multivariate extension of the procedure of Politis and White [38]. Nonetheless, the choice $\psi=$ median was found to give meaningful results in our Monte Carlo experiments partially reported in Section 6.

5.2. Generation of dependent multiplier sequences

The practical use of the results stated in Sections 2 and 4 requires the generation of dependent multiplier random variables satisfying assumptions (M1), (M2) and (M3). We describe two ways of constructing such dependent sequences. The first one generalizes the moving average approach proposed by Bühlmann ([15], Section 6.2) (see also Bücher and Ruppert [11]) and produces multipliers that satisfy assumption (M3) only asymptotically. The second one was suggested by

Shao [44] and is based on the calculation of the square root of the covariance matrix implicitly defined in assumption (M3).

5.2.1. *The moving average approach*

Let κ be some positive bounded real function symmetric around zero such that $\kappa(x) > 0$ for all $|x| < 1$. Let b_n be a sequence of integers such that $b_n \to \infty$, $b_n = \mathrm{o}(n)$ and $b_n \geq 1$ for all $n \in \mathbb{N}$. Let $Z_1, \dots, Z_{n+2b_n-2}$ be i.i.d. random variables independent of the available sample $\mathbf{X}_1, \dots, \mathbf{X}_n$ such that $\mathrm{E}(Z_1) = 0$, $\mathrm{E}(Z_1^2) = 1$ and $\mathrm{E}(|Z_1|^\nu) < \infty$ for all $\nu \geq 1$. Then, let $\ell_n = 2b_n - 1$ and, for any $j \in \{1, \dots, \ell_n\}$, let $w_{j,n} = \kappa\{(j - b_n)/b_n\}$ and $\tilde{w}_{j,n} = w_{j,n}(\sum_{j'=1}^{\ell_n} w_{j',n}^2)^{-1/2}$. Finally, for all $i \in \{1, \dots, n\}$, let

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

Clearly, $\xi_{1,n}, \dots, \xi_{n,n}$ are identically distributed with $\mathrm{E}(\xi_{1,n}) = 0$, $\mathrm{E}(\xi_{1,n}^2) = 1$ and it can be verified that $\sup_{n \geq 1} \mathrm{E}(|\xi_{1,n}|^\nu) < \infty$ for all $\nu \geq 1$. Furthermore, $\xi_{1,n}, \dots, \xi_{n,n}$ are $(\ell_n - 1)$-dependent and, for any $i \in \{1, \dots, n\}$ and $r \in \{0, \dots, (\ell_n - 1) \wedge n\}$,

$$\begin{aligned}
\mathrm{Cov}(\xi_{i,n}\xi_{i+r,n}) &= \sum_{j=1}^{\ell_n} \sum_{j'=1}^{\ell_n} \tilde{w}_{j,n} \tilde{w}_{j',n} \mathrm{E}(Z_{j+i-1} Z_{j'+i+r-1}) = \sum_{j=r+1}^{\ell_n} \tilde{w}_{j,n} \tilde{w}_{j-r,n} \\
&= \left(\sum_{j=1}^{\ell_n} w_{j,n}^2 \right)^{-1} \sum_{j=r+1}^{\ell_n} \kappa\bigl\{(j - b_n)/b_n\bigr\} \kappa\bigl\{(j - r - b_n)/b_n\bigr\}.
\end{aligned}$$

For practical reasons, only a sequence of size n has been generated. From the previous developments, we immediately have that the infinite size version of $\xi_{1,n}, \dots, \xi_{n,n}$ satisfies assumptions (M1) and (M2) (as $(\ell_n - 1)$-dependence clearly implies ℓ_n-dependence). Let us now verify that it satisfies assumption (M3) asymptotically.

Assume additionally that $\kappa(x) = 0$ for all $|x| > 1$, and, for any $f, g : \mathbb{Z} \to \mathbb{R}$, let $f * g$ denote the discrete convolution of f and g, that is, $f * g(r) = \sum_{j=-\infty}^{\infty} f(j) g(r - j)$, $r \in \mathbb{Z}$. Then, let $\kappa_{b_n}(j) = \kappa(j/b_n)$, $j \in \mathbb{Z}$, and notice that the previous covariance can be written as

$$\mathrm{Cov}(\xi_{i,n}\xi_{i+r,n}) = \frac{\sum_{j=-\infty}^{\infty} \kappa_{b_n}(j - b_n) \kappa_{b_n}(j - r - b_n)}{\kappa_{b_n} * \kappa_{b_n}(0)} + \mathrm{o}(1) = \frac{\kappa_{b_n} * \kappa_{b_n}(r)}{\kappa_{b_n} * \kappa_{b_n}(0)} + \mathrm{o}(1)$$

for all $i \in \{1, \dots, n\}$ and $r \in \{0, \dots, n - i\}$, where the $\mathrm{o}(1)$ term comes from the fact that $\kappa(1)$ is not necessarily equal to 0.

Assume furthermore that there exists $\lambda > 0$ such that $|\kappa(x) - \kappa(y)| \leq \lambda |x - y|$ for all $x, y \in [-1, 1]$ and let r_n be a positive sequence such that $r_n/b_n \to \gamma \in [0, 1]$. We shall now check that $b_n^{-1} \kappa_{b_n} * \kappa_{b_n}(r_n) \to \kappa \star \kappa(\gamma)$, where $\star$ denotes the convolution operator between real functions. We have

$$\frac{1}{b_n} \kappa_{b_n} * \kappa_{b_n}(r_n) = \frac{1}{b_n} \sum_{j=-b_n}^{b_n} \kappa(j/b_n) \kappa\bigl\{(r_n - j)/b_n\bigr\}.$$

On the one hand,

$$\left| \frac{1}{b_n} \sum_{j=-b_n}^{b_n} \kappa(j/b_n)\kappa\{(r_n - j)/b_n\} - \frac{1}{b_n} \sum_{j=-b_n}^{b_n} \kappa(j/b_n)\kappa(\gamma - j/b_n) \right|$$
$$\leq \lambda |r_n/b_n - \gamma| \frac{2b_n + 1}{b_n} \sup_{x \in \mathbb{R}} \kappa(x) \to 0,$$

and, and on the other hand,

$$\frac{1}{b_n} \sum_{j=-b_n}^{b_n} \kappa(j/b_n)\kappa(\gamma - j/b_n) \to \int_{-1}^{1} \kappa(x)\kappa(\gamma - x)\,\mathrm{d}x = \kappa \star \kappa(\gamma).$$

It follows that

$$\frac{\kappa_{b_n} * \kappa_{b_n}(r_n)}{\kappa_{b_n} * \kappa_{b_n}(0)} \to \frac{\kappa \star \kappa(\gamma)}{\kappa \star \kappa(0)}.$$

Now, let

$$\varphi(x) = \frac{\kappa \star \kappa(2x)}{\kappa \star \kappa(0)}, \qquad x \in \mathbb{R}, \tag{5.7}$$

where the factor 2 ensures that $\varphi(x) = 0$ for all $|x| > 1$. Then, for large n, $\mathrm{Cov}(\xi_{i,n}\xi_{j,n}) \approx \varphi\{(i-j)/\ell_n\}$, for any $i, j \in \{1, \ldots, n\}$. Hence, the infinite size version of $\xi_{1,n}, \ldots, \xi_{n,n}$ satisfies assumption (M3) asymptotically.

In our numerical experiments, we considered several popular *kernels* for the function κ (see, e.g., Andrews [2]), defined, for any $x \in \mathbb{R}$, as

$$\begin{aligned} &\text{Truncated:} && \kappa_T(x) = \mathbf{1}\big(|x| \leq 1\big), \\ &\text{Bartlett:} && \kappa_B(x) = \big(1 - |x|\big) \vee 0, \\ &\text{Parzen:} && \kappa_P(x) = \big(1 - 6x^2 + 6|x|^3\big)\mathbf{1}\big(|x| \leq 1/2\big) + 2\big(1 - |x|\big)^3 \mathbf{1}\big(1/2 < |x| \leq 1\big), \end{aligned}$$

as well as the flat top kernel already defined in (5.5). The above kernels satisfy all the assumptions on the function κ mentioned previously. Their graphs are represented in Figure 1. The flat top (or trapezoidal) kernel, parametrized by $c \in [0, 1]$, was used in Paparoditis and Politis [33] in the context of the tapered block bootstrap for the mean. These authors found that, within the class of trapezoidal kernels symmetric around 0.5 and with support (0, 1), $\kappa_{F,0.14}$, rescaled and shifted to have support (0, 1), minimizes the asymptotic mean squared error of the bootstrapping procedure. The latter kernel was also used in Shao [44] who connected the tapered block bootstrap with the dependent multiplier bootstrap for the mean.

5.2.2. *The covariance matrix approach*

Let ℓ_n be a sequence of strictly positive constants such that $\ell_n \to \infty$ and $\ell_n = \mathrm{o}(n)$. Let φ be a function satisfying assumption (M3) such that, additionally, $\int_{-\infty}^{\infty} \varphi(u) e^{-\mathrm{i}ux}\,\mathrm{d}u \geq 0$ for all

$x \in \mathbb{R}$, and let $\mathbf{\Sigma}_n$ be the $n \times n$ (covariance) matrix whose elements are defined by $\varphi\{(i-j)/\ell_n\}$, $i, j \in \{1, \ldots, n\}$. The integral condition on φ ensures that $\mathbf{\Sigma}_n$ is positive definite which in turn ensures the existence of $\mathbf{\Sigma}_n^{1/2}$. From a practical perspective, $\mathbf{\Sigma}_n^{1/2}$ can be computed either by diagonalization, singular value decomposition or Cholesky factorization of $\mathbf{\Sigma}_n$. We use the first approach. Then, let $Z_1, \ldots, Z_n$ be i.i.d. standard normal random variables independent of the available sample $\mathbf{X}_1, \ldots, \mathbf{X}_n$. A dependent multiplier sequence $\xi_{1,n}, \ldots, \xi_{n,n}$ can then be simply obtained as

$$[\xi_{1,n}, \ldots, \xi_{n,n}]^\top = \mathbf{\Sigma}_n^{1/2}[Z_1, \ldots, Z_n]^\top.$$

If $\varphi(1) > 0$, then the above construction generates ℓ_n-dependent multipliers, while if $\varphi(1) = 0$, the generated sequence is $(\ell_n - 1)$-dependent. Clearly, the infinite size version of $\xi_{1,n}, \ldots, \xi_{n,n}$ satisfies assumptions (M1), (M2) and (M3).

From a practical perspective, for the function φ, we considered the Bartlett and Parzen kernels κ_B and κ_P, as well as $\kappa_{U,6}$ and $\kappa_{U,8}$, where $\kappa_{U,p}$ is the density function of the sum of p independent uniforms centered at 0, normalized so that it equals 1 at 0, and rescaled to have support $(-1, 1)$. The functions $\kappa_{U,6}$ and $\kappa_{U,8}$ are represented in Figure 1. Notice that $\kappa_T = \kappa_{U,1}$, $\kappa_B = \kappa_{U,2}$ and $\kappa_P = \kappa_{U,4}$. This also implies that $\kappa_{U,8}$ is a rescaled and normalized version of the convolution of κ_P with itself, that is, $\kappa_{U,8}(x) = \kappa_P \star \kappa_P(2x)/\kappa_P \star \kappa_P(0)$ for all $x \in \mathbb{R}$. A numerically stable and efficient way of computing $\kappa_{U,p}$ consists of using *divided differences* (see, e.g., Agarwal, Dalpatadu and Singh [1]). Finally, note that the truncated and flat top kernels cannot be used as they do not satisfy the integral condition ensuring that Σ_n is positive definite.

Remark 5.3. In the case of the moving average approach presented in Section 5.2.1, we have seen that κ determines φ asymptotically through (5.7). It follows that, for an initial standard normal i.i.d. sequence, the same value of ℓ_n and for large n, we could expect the dependent multiplier sequences generated by the moving average and the covariance matrix approaches, respectively, to give close results when κ in Section 5.2.1 and φ in Section 5.2.2 are related through (5.7). For instance, all other parameters being similar, using the Bartlett kernel for κ in Section 5.2.1 should produce similar results to using the Parzen kernel for φ in Section 5.2.2.

5.3. Estimation of the partial derivatives of the copula

For the estimators of the partial derivatives appearing in (4.2), we considered three possible definitions proposed in the literature. The first one is that of Rémillard and Scaillet [39] who suggested to estimate the partial derivatives $\dot{C}_j$, $j \in \{1, \ldots, d\}$, by finite-differences as

$$\begin{aligned}\dot{C}_{j,n}(\mathbf{u}) = \frac{1}{2n^{-1/2}}\big\{&C_n\big(u_1, \ldots, u_{j-1}, u_j + n^{-1/2}, u_{j+1}, \ldots, u_d\big)\\ &- C_n\big(u_1, \ldots, u_{j-1}, u_j - n^{-1/2}, u_{j+1}, \ldots, u_d\big)\big\}, \qquad \mathbf{u} \in [0,1]^d.\end{aligned} \tag{5.8}$$

A slightly different definition consisting of a "boundary correction" was proposed in Kojadinovic, Segers and Yan ([27], page 706). Yet another definition is mentioned in Bücher and Ruppert ([11], page 212). Note that, for any $\delta \in (0, 1/2)$, all three definitions coincide on the set $\{\mathbf{u} \in [0,1]^d : u_j \in [\delta, 1-\delta]\}$ provided n is taken large enough. Now, under the assumptions of Corollary 3.5, we have that $\mathbb{C}_n(0,1,\cdot) \rightsquigarrow \mathbb{C}_C(0,1,\cdot)$ in $\ell^\infty([0,1]^d)$. The latter weak convergence implies the first statement of Lemma 2 of Kojadinovic, Segers and Yan [27], which in turn implies that Condition 4.1 is satisfied for the above defined $\dot{C}_{j,n}$ as well as for the two slightly different definitions considered in Kojadinovic, Segers and Yan ([27], page 706) and Bücher and Ruppert ([11], page 212), respectively.

6. Monte Carlo experiments

To investigate the finite-sample performance of the proposed dependent multiplier bootstrap, we considered several statistics derived from the sequential empirical copula process $\mathbb{C}_n$ defined in (1.1). With applications to statistical tests in mind, we mostly focus in this section on Cramér–von-Mises and Kolomogorov–Smirnov statistics obtained from $\mathbb{C}_n(0,1,\cdot)$. Results for some simpler functionals can be found in Section G of the supplementary material (Bücher and Kojadinovic [9]).

Recall that M is a large integer, and let

$$\begin{aligned} S_n &= \int_{[0,1]^d} \bigl\{\mathbb{C}_n(0,1,\mathbf{u})\bigr\}^2 \,\mathrm{d}\mathbf{u} \quad \text{and} \\ S_n^{(m)} &= \int_{[0,1]^d} \bigl\{\hat{\mathbb{C}}_n^{(m)}(0,1,\mathbf{u})\bigr\}^2 \,\mathrm{d}\mathbf{u}, \qquad m \in \{1,\dots,M\}, \end{aligned} \tag{6.1}$$

where $\hat{\mathbb{C}}_n^{(m)}$ is defined in (4.2) with the partial derivative estimators defined as discussed later in this section. Under the conditions of Proposition 4.2 and from the continuous mapping theorem, we then immediately have that $(S_n, S_n^{(1)}, \dots, S_n^{(M)})$ converges weakly to $(S, S^{(1)}, \dots, S^{(M)})$, where $S = \int_{[0,1]^d} \{\mathbb{C}_C(0,1,\mathbf{u})\}^2 \,\mathrm{d}\mathbf{u}$ and $S^{(1)}, \dots, S^{(M)}$ are independent copies of S.

The first aim of our Monte Carlo experiments was to assess the quality of the estimation of the quantiles of S by the empirical quantiles of the sample $S_n^{(1)}, \dots, S_n^{(M)}$. Let $S_n^{(1:M)} \le \dots \le S_n^{(M:M)}$ denote the corresponding order statistics. An estimator of the quantile of S of order $p \in (0,1)$ is then simply $S_n^{(\lfloor pM \rfloor:M)}$. For each data generating scenario, the target theoretical quantiles of S of order p were accurately estimated empirically from 10^5 realizations of S_{1000} for $p \in \mathcal{P} = \{0.25, 0.5, 0.75, 0.9, 0.95, 0.99\}$. Then, for each data generating scenario, $N = 1000$ samples $\mathbf{X}_1, \dots, \mathbf{X}_n$ were generated and, for each sample, $S_n^{(\lfloor pM \rfloor:M)}$ was computed for each $p \in \mathcal{P}$ using the dependent multiplier bootstrap with $M = 2500$ yielding, for each $p \in \mathcal{P}$, N estimates of the quantile of S of order p. This allowed us to compute, for each data generating scenario and each $p \in \mathcal{P}$, the empirical bias and the empirical mean squared error (MSE) of the estimators of the quantiles of S of order p. Similar simulations were performed for the Kolmogorov–Smirnov statistic. Specifically, let

$$T_n = \sup_{\mathbf{u} \in [0,1]^d} \bigl|\mathbb{C}_n(0,1,\mathbf{u})\bigr| \quad \text{and} \quad T_n^{(m)} = \sup_{\mathbf{u} \in [0,1]^d} \bigl|\hat{\mathbb{C}}_n^{(m)}(0,1,\mathbf{u})\bigr|, \qquad m \in \{1,\dots,M\}. \tag{6.2}$$

The dimension d was fixed to two, and the integrals and the suprema in (6.1) and (6.2), respectively, were computed approximately using a fine grid on $(0,1)^2$ of 400 uniformly spaced points.

Four data generating models were considered. The first one is a simple AR1 model. Let $\mathbf{U}_i$, $i \in \{-100, \dots, 0, \dots, n\}$, be a bivariate i.i.d. sample from a copula C. Then, set $\boldsymbol{\epsilon}_i = (\Phi^{-1}(U_{i1}), \Phi^{-1}(U_{i2}))$, where Φ is the c.d.f. of the standard normal distribution, and $\mathbf{X}_{-100} = \boldsymbol{\epsilon}_{-100}$. Finally, for any $j \in \{1, 2\}$ and $i \in \{-99, \dots, 0, \dots, n\}$, compute recursively

$$X_{ij} = 0.5 X_{i-1,j} + \epsilon_{ij}. \tag{AR1}$$

The second and third data generating models are related to the nonlinear autoregressive (NAR) model used in Paparoditis and Politis ([33], Section 3.3), and to the exponential autoregressive (EXPAR) model considered in Auestad and Tjøstheim [3] and Paparoditis and Politis ([33], Section 3.3). The sample $\mathbf{X}_1, \dots, \mathbf{X}_n$ is generated as previously with (AR1) replaced by

$$X_{ij} = 0.6 \sin(X_{i-1,j}) + \epsilon_{ij} \tag{NAR}$$

and

$$X_{ij} = \{0.8 - 1.1 \exp(-50 X_{i-1,j}^2)\} X_{i-1,j} + 0.1 \epsilon_{ij}, \tag{EXPAR}$$

respectively. The fourth and last data generating model is the bivariate GARCH-like model considered in Bücher and Ruppert [11]. The sample of innovations is defined as for the models above. In addition, for any $j \in \{1, 2\}$, let $\sigma_{-100,j} = \sqrt{\omega_j / (1 - \alpha_j - \beta_j)}$ where ω_j, α_j and β_j are usual GARCH(1, 1) parameters whose values will be set below, and, for any $j \in \{1, 2\}$ and $i \in \{-99, \dots, 0, \dots, n\}$, compute recursively

$$\sigma_{ij}^2 = \omega_j + \beta_j \sigma_{i-1,j}^2 + \alpha_j \epsilon_{i-1,j}^2 \quad \text{and} \quad X_{ij} = \sigma_{ij} \epsilon_{ij}. \tag{GARCH}$$

Following Bücher and Ruppert [11], we take $(\omega_1, \beta_1, \alpha_1) = (0.012, 0.919, 0.072)$ and $(\omega_2, \beta_2, \alpha_2) = (0.037, 0.868, 0.115)$. The latter values were estimated by Jondeau, Poon and Rockinger [26] from SP500 and DAX daily logreturns, respectively.

The other factors of the experiments are as follows. Four different copulas were considered: Clayton copulas with parameter values 1 and 4, respectively, and Gumbel–Hougaard copulas with parameter value 1.5 and 3, respectively. The lower (resp., higher) parameter values correspond to a Kendall's tau of 1/3 (resp., 2/3), that is, to mild (resp., strong) dependence. Notice that the Clayton copula is lower-tail dependent while the Gumbel–Hougaard is upper-tail dependent (see, e.g., McNeil, Frey and Embrechts [31], Chapter 5). The values 100, 200 and 400 were considered for n.

We report the results of the experiments very partially (additional results are available in the supplementary material, see Bücher and Kojadinovic [9]) and when based on the estimators of the partial derivatives given in (5.8). Figure 2 displays the empirical MSE of the estimator $S_n^{(\lfloor pM \rfloor : M)}$ of the quantile of order $p = 0.95$ of S_n versus the bandwidth parameter ℓ_n for the different choices of κ/φ mentioned in Section 5.2. The top (resp., middle, bottom) line of graphs was obtained from datasets generated under the NAR (resp., EXPAR, GARCH) scenario with C being the Gumbel–Hougaard copula with parameter value 1.5. The line segments in the lower-right corners of the graphs correspond to the empirical MSEs of the estimator $S_n^{(\lfloor 0.95M \rfloor : M)}$ based on the estimated bandwidth $\hat{\ell}_n^{\mathrm{opt}}$ computed as explained in Section 5.1. The line styles of the segments correspond to the choice of φ. The results for the AR1 scenario being very similar

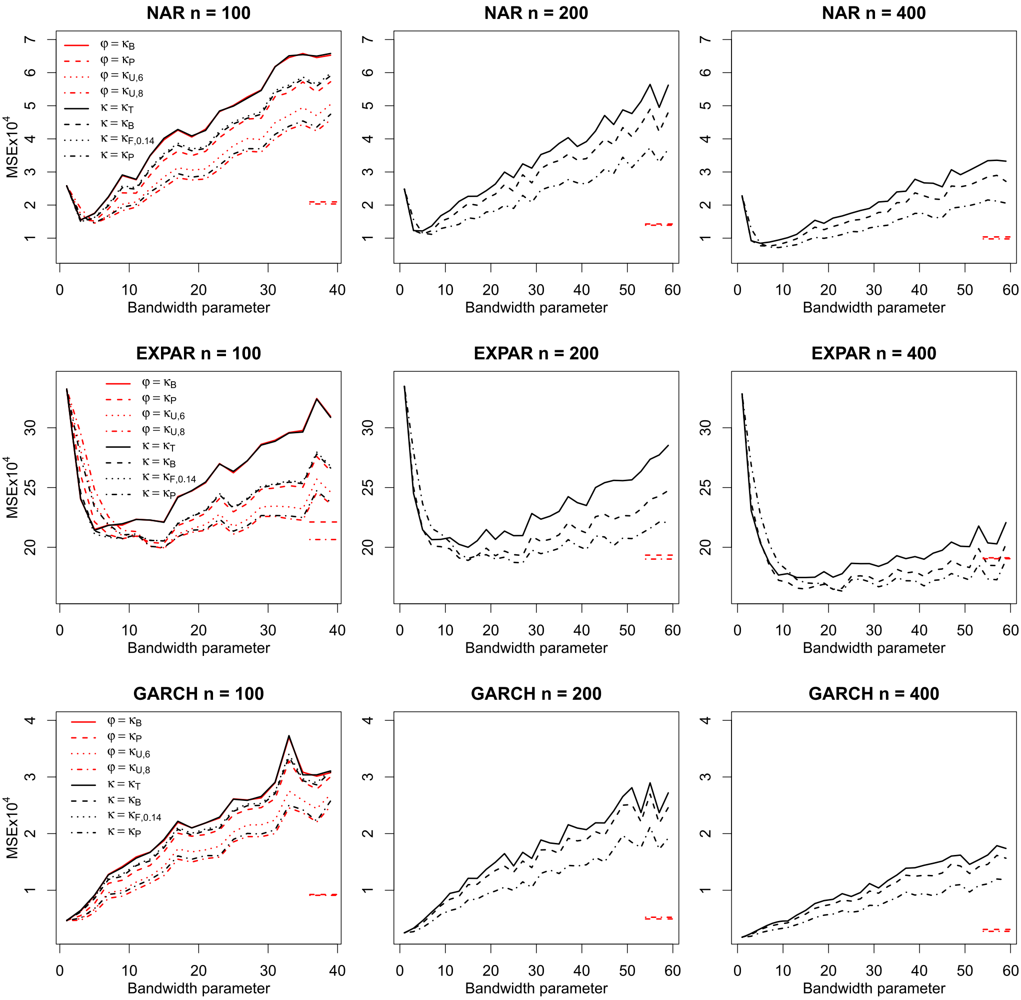

Figure 2. For various choices of the function κ/φ (see Section 5.2), empirical MSE $\times\, 10^4$ of the estimator $S_n^{(\lfloor 0.95M \rfloor : M)}$ with $M = 2500$ versus the bandwidth parameter ℓ_n under the NAR, EXPAR and GARCH data generating scenarios with C being the Gumbel–Hougaard copula with parameter 1.5. The line segments in the lower-right corners of the graphs correspond to the empirical MSEs of the estimator with estimated bandwidth parameter following the procedure described in Section 5.1. The line styles of the segments correspond to the choice of φ.

to those for the NAR scenario are not reported. Similar looking graphs were obtained for the other three copulas used in the simulations and when replacing the Cramér–von Mises statistics by the Kolmogorov–Smirnov statistics defined in (6.2). In a related manner, the shapes of the graphs were not too much affected by the value p of the quantile order: the empirical MSEs were smaller for $p < 0.95$ and higher for $p = 0.99$. Figures analogue to Figure 2 for other values of p and/or for the Kolmogorov–Smirnov statistic T_n can actually be found in Section G of the supplementary material (Bücher and Kojadinovic [9]).

The black (resp., red) curves in the first column of panels of Figure 2 were obtained for dependent multiplier sequences generated from initial standard normal i.i.d. sequences using the moving average (resp., covariance matrix) approach described in Section 5.2.1 (resp., Section 5.2.2). The functions κ_T, κ_B, $\kappa_{F,0.14}$ and κ_P were considered for κ in the case of the moving average approach, while the function φ in the covariance matrix approach was successively taken equal to κ_B, κ_P, $\kappa_{U,6}$ and $\kappa_{U,8}$. Looking at the graphs for $n = 100$, we see that, when the functions κ and φ are chosen to match in the sense of Remark 5.3, the resulting empirical MSEs are very close. For that reason, to facilitate reading of the plots, only the curves obtained with the moving average approach and $\kappa \in \{\kappa_T, \kappa_B, \kappa_P\}$ are plotted when $n \in \{200, 400\}$. As it can be seen, for the NAR and EXPAR scenarios, the empirical MSEs tend to decrease first with ℓ_n, reach a minimum, and increase again. It is not the case for the GARCH setting for which it seems that $\ell_n = 1$ always leads to the smallest MSE. In other words, the use of the dependent multiplier bootstrap does not seem necessary in that context as the usual i.i.d. multiplier of Rémillard and Scaillet [39] provides the best results. This might be due to the fact that in this setting the contributions of the lagged covariances to the long-run variance of the empirical process are very small. Looking again at the graphs for the NAR and EXPAR settings, we see that the smallest MSEs are reached by choosing $\kappa = \kappa_P/\varphi = \kappa_{U,8}$, which is in accordance with Proposition 5.2 which states that, asymptotically, kernels with the smallest integral lead to the lowest variance. Another observation is that, unlike what was expected by Shao ([44], Remark 2.1) in the case of the mean as statistic of interest, the choice $\kappa = \kappa_{F,0.14}$ did not lead to better results than the choice $\kappa = \kappa_P$. Finally, let us comment on the empirical MSEs of the estimator $S_n^{(\lfloor 0.95M \rfloor:M)}$ based on the estimated bandwidth $\hat{\ell}_n^{\mathrm{opt}}$ computed as explained in Section 5.1. As it can be seen from the line segments in the lower-right corners of the graphs, the achieved empirical MSEs decrease with n and are, overall, reasonably close to the lowest observed MSE. Considering all the available results (see Section G of the supplementary material Bücher and Kojadinovic [9], for additional figures), the choice $\varphi = \kappa_{U,8}$ appears to lead to a slightly lower MSE, overall, when $n = 100$. For $n \in \{200, 400\}$, the choices $\varphi = \kappa_P$ and $\varphi = \kappa_{U,8}$ do not seem to lead to differences of practical interest.

In view of the small differences between the moving average and covariance matrix approaches for generating dependent multipliers (black versus red curves in the first column of graphs of Figure 2), we suggest to use the former which is faster and more stable numerically as it does not require the computation of the square root of a large covariance matrix.

Before discussing further the estimation of ℓ_n using the results of Section 5.1, let us mention an observation of practical interest. Working with the same random seed, we replicated the experiments described above using the two alternative definitions of the partial derivative estimators mentioned below (5.8). To our surprise, the best results, overall, were obtained with the proposal of Rémillard and Scaillet [39] given in (5.8), although the differences seem too small to be of practical interest.

We end this section with a more direct empirical investigation of the estimator $\hat{\ell}_n^{\mathrm{opt}}$ of ℓ_n^{opt} (see (5.4) and Section 5.1). We report an experiment based on the AR1 model which will serve as a benchmark for judging about the performance of $\hat{\ell}_n^{\mathrm{opt}}$. The setting is the following: a grid $\{\mathbf{u}_i\}_{i \in \{1,\dots,g\}}$ of $g = 25$ points uniformly spaced over $(0,1)^2$ was created, and $\sigma_C(\mathbf{u}_i, \mathbf{u}_j)$ was accurately estimated for all $i, j \in \{1, \dots, g\}$ from 10^5 samples of size 1000 generated under the AR1 model described previously. The latter estimation was carried out as follows: given a sample $\mathbf{X}_1, \dots, \mathbf{X}_n$ generated from the AR1 model, the marginally standard uniform sample $\mathbf{U}_1, \dots, \mathbf{U}_n$ was formed using the fact that the marginal c.d.f.s of the $\mathbf{X}_i$ are centered normal with variance $1/(1-0.5^2)$ in this case; this enabled us to compute $\tilde{\mathbb{B}}_n(1, \cdot)$ at the grid points, where $\tilde{\mathbb{B}}_n$ is defined in (2.1); for any $i, j \in \{1, \dots, g\}$, $\sigma_C(\mathbf{u}_i, \mathbf{u}_j)$ was finally accurately estimated as the sample covariance of 10^5 independent realizations of $(\tilde{\mathbb{B}}_n(1, \mathbf{u}_i), \tilde{\mathbb{B}}_n(1, \mathbf{u}_j))$.

Next, for $n \in \{100, 200, 400\}$ and $\ell_n \in \{1, 3, \dots, 39\}$, IMSE_n defined in (5.2) was approximated as follows: 1000 samples $\mathbf{X}_1, \dots, \mathbf{X}_n$ were generated under the AR1 model, and, for each sample, the processes $\hat{\mathbb{B}}_n^{(1)}(1, \cdot), \dots, \hat{\mathbb{B}}_n^{(M)}(1, \cdot)$ with $M = 1000$ were evaluated at the grid points, with $\hat{\mathbb{B}}_n^{(m)}$ defined in (4.1); computing sample covariances, this allowed us to obtain 1000 bootstrap estimates of $\sigma_C(\mathbf{u}_i, \mathbf{u}_j)$ for all $i, j \in \{1, \dots, g\}$, from which we approximated IMSE_n. The results are represented in the graphs of Figure 3 for the previously considered choices of the function φ. The top (resp., bottom) row of graphs was obtained when C in the AR1 data generating scenario is the Gumbel–Hougaard copula with parameter 1.5 (resp., 3).

The procedure described in Section 5.1 was finally used to obtain 1000 estimates of ℓ_n^{opt} under the AR1 model based on the Gumbel–Hougaard copula with parameter θ, for $n \in$

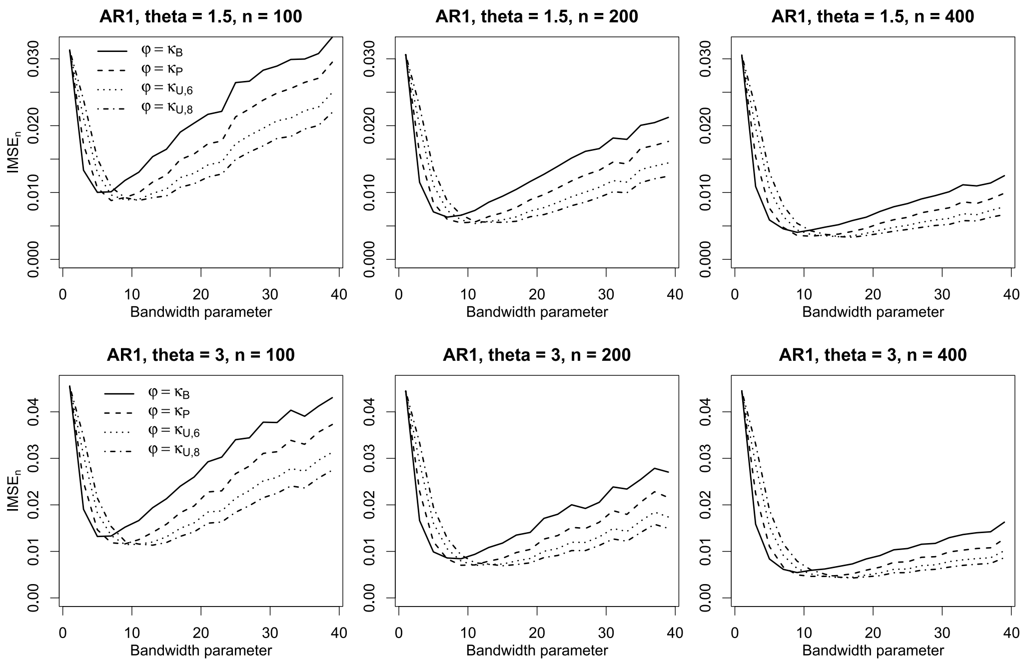

Figure 3. For several choices of the function φ, IMSE_n defined in (5.2), computed approximately using a grid of 25 uniformly spaced points on $(0, 1)^2$ and 1000 samples versus the bandwidth parameter ℓ_n under the AR1 data generating scenario with C being the Gumbel–Hougaard copula with parameter 1.5 (top row) and parameter 3 (bottom row).

Table 1. Mean and standard deviation of 1000 estimates of ℓ_n^{opt}, defined in (5.4), computed as explained in Section 5.1 from 1000 samples generated from the AR1 model in which C is the Gumbel–Hougaard copula with parameter θ. The computations were carried out for the choices $\varphi = \kappa_P$ and $\varphi = \kappa_{U,8}$

		$\varphi = \kappa_P$		$\varphi = \kappa_{U,8}$	
θ	n	Mean	Std.	Mean	Std.
1.5	100	8.93	3.85	12.41	5.92
	200	10.67	4.05	14.74	5.15
	400	12.81	3.94	17.73	4.99
3.0	100	9.11	5.18	12.75	8.13
	200	10.64	4.08	14.69	5.74
	400	12.77	3.94	17.66	5.31

$\{100, 200, 400\}$, $\varphi \in \{\kappa_P, \kappa_{U,8}\}$ and $\theta \in \{1.5, 3\}$. The mean and standard deviation of the estimates are reported in Table 1. A comparison with Figure 3 reveals that the procedure described in Section 5.1 for estimating ℓ_n^{opt} gives surprisingly good results on average for the experiment at hand. Another observation is that the estimates do not seem much affected by the value of θ, that is, the strength of the dependence.

Appendix A: Proof of Theorem 2.1

The proof of Theorem 2.1 is based on three lemmas. The first lemma establishes weak convergence of the finite-dimensional distributions, while the second and third lemmas concern asymptotic tightness.

The following result can be proved using a well-known blocking technique (see, e.g., Dehling and Philipp [20], page 31). Its proof is given in the supplementary material (Bücher and Kojadinovic [9]).

Lemma A.1 (Finite-dimensional convergence). *Assume that $\ell_n = \mathrm{O}(n^{1/2-\varepsilon})$ for some $0 < \varepsilon < 1/2$ and that $(\mathbf{U}_i)_{i\in\mathbb{Z}}$ is a strictly stationary sequence whose strong mixing coefficients satisfy $\alpha_r = \mathrm{O}(r^{-a})$, $a > 2$. Then, the finite-dimensional distributions of $(\mathbb{B}_n, \tilde{\mathbb{B}}_n^{(1)}, \dots, \tilde{\mathbb{B}}_n^{(M)})$ converge weakly to those of $(\mathbb{B}_C, \mathbb{B}_C^{(1)}, \dots, \mathbb{B}_C^{(M)})$.*

Regarding the tightness, let us first extend $\tilde{\mathbb{B}}_n^{(m)}$, $m \in \{1, \dots, M\}$, to blocks in $[0,1]^{d+1}$ in the spirit of Bickel and Wichura [6]. For any $(s,t] \subset [0,1]$ and $A = (u_1, v_1] \times \cdots \times (u_d, v_d] \subset [0,1]^d$, we define $\tilde{\mathbb{B}}_n^{(m)}((s,t] \times A)$ to be

$$\tilde{\mathbb{B}}_n^{(m)}\bigl((s,t] \times A\bigr) = \frac{1}{\sqrt{n}} \sum_{i=\lfloor ns\rfloor+1}^{\lfloor nt\rfloor} \xi_{i,n}^{(m)} \bigl[\mathbf{1}(\mathbf{U}_i \in A) - \nu(A)\bigr],$$

where

$$\begin{aligned}\nu(A) &= \mathrm{P}(\mathbf{U}_1 \in A) \\ &= \sum_{(\epsilon_1,\dots,\epsilon_d)\in\{0,1\}^d} (-1)^{\sum_{i=1}^d \epsilon_i} C\big\{(1-\epsilon_1)v_1+\epsilon_1 u_1,\dots,(1-\epsilon_d)v_d+\epsilon_d u_d\big\}.\end{aligned}$$

In the next two lemmas, the sequences $(\xi_{i,n}^{(m)})_{i\in\mathbb{Z}}$ are only assumed to satisfy (M1) with $\mathrm{E}[\{\xi_{0,n}^{(m)}\}^2] > 0$ not necessarily equal to one.

Lemma A.2 (Moment inequality). *Assume that $(\mathbf{U}_i)_{i\in\mathbb{Z}}$ is a strictly stationary sequence whose strong mixing coefficients satisfy $\alpha_r = \mathrm{O}(r^{-a})$, $a > 6$. Then, for any $m \in \{1,\dots,M\}$, $q \in (2a/(a-3), 4)$, $(s,t] \subset [0,1]$ and $A = (u_1,v_1]\times\cdots\times(u_d,v_d] \subset [0,1]^d$, we have*

$$\mathrm{E}\big[\big\{\tilde{\mathbb{B}}_n^{(m)}\big((s,t]\times A\big)\big\}^4\big] \le \kappa\big[\lambda_n(s,t)^2\big\{\nu(A)\big\}^{4/q} + n^{-1}\lambda_n(s,t)\big\{\nu(A)\big\}^{2/q}\big],$$

where $\kappa > 0$ is a constant.

Proof. The proof is similar to that of Lemma 3.22 in Dehling and Philipp [20]. Fix $m \in \{1,\dots,M\}$. For any $i \in \mathbb{Z}$, let $Y_i = \mathbf{1}(\mathbf{U}_i \in A) - \nu(A)$. Then,

$$\begin{aligned}&\mathrm{E}\big[\big\{\tilde{\mathbb{B}}_n^{(m)}\big((s,t]\times A\big)\big\}^4\big] \\ &\quad= \frac{1}{n^2}\sum_{i_1,i_2,i_3,i_4=\lfloor ns\rfloor+1}^{\lfloor nt\rfloor} \mathrm{E}\big[\xi_{i_1,n}^{(m)}\xi_{i_2,n}^{(m)}\xi_{i_3,n}^{(m)}\xi_{i_4,n}^{(m)}\big]\mathrm{E}[Y_{i_1}Y_{i_2}Y_{i_3}Y_{i_4}] \qquad (A.1)\\ &\quad\le \frac{4!\lambda_n(s,t)}{n}\sum_{\substack{0\le i,j,k\le\lfloor nt\rfloor-\lfloor ns\rfloor-1\\ i+j+k\le\lfloor nt\rfloor-\lfloor ns\rfloor-1}} \big|\mathrm{E}\big[\xi_{0,n}^{(m)}\xi_{i,n}^{(m)}\xi_{i+j,n}^{(m)}\xi_{i+j+k,n}^{(m)}\big]\mathrm{E}[Y_0Y_iY_{i+j}Y_{i+j+k}]\big|.\end{aligned}$$

On one hand, $|\mathrm{E}[\xi_{0,n}^{(m)}\xi_{i,n}^{(m)}\xi_{i+j,n}^{(m)}\xi_{i+j+k,n}^{(m)}]| \le \mathrm{E}[\{\xi_{0,n}^{(m)}\}^4]$. On the other hand, by Lemma 3.11 of Dehling and Philipp [20], for any $q \in (2a/(a-3), 4)$ and $p \in (2, a/3)$ such that $1/p + 2/q = 1$, we have

$$\begin{aligned}&\mathrm{E}\big[Y_0(Y_iY_{i+j}Y_{i+j+k})\big] \le 10\alpha_i^{1/p}\|Y_0\|_q\|Y_iY_{i+j}Y_{i+j+k}\|_q \le 10\alpha_i^{1/p}\|Y_0\|_q^2,\\ &\mathrm{E}\big[(Y_0Y_iY_{i+j})Y_{i+j+k}\big] \le 10\alpha_k^{1/p}\|Y_0\|_q^2\end{aligned}$$

and

$$\begin{aligned}\big|\mathrm{E}\big[(Y_0Y_i)(Y_{i+j}Y_{i+j+k})\big]\big| &\le \big|\mathrm{E}[Y_0Y_i]\mathrm{E}[Y_{i+j}Y_{i+j+k}]\big| + 10\alpha_j^{1/p}\|Y_0Y_i\|_q\|Y_{i+j}Y_{i+j+k}\|_q\\ &\le 100\alpha_i^{1/p}\alpha_k^{1/p}\|Y_0\|_q^4 + 10\alpha_j^{1/p}\|Y_0\|_q^2.\end{aligned}$$

Proceeding as in Lemma 3.22 of Dehling and Philipp [20], we split the sum on the right of (A.1) into three sums according to which of the indices i, j, k is the largest. Combining this decomposition with the three previous inequalities, we obtain

$$\begin{aligned}
&\mathrm{E}\big[\big|\tilde{\mathbb{B}}_n^{(m)}\big((s,t]\times A\big)\big|^4\big] \\
&\quad\le \frac{24\mathrm{E}[\{\xi_{0,n}^{(m)}\}^4]\lambda_n(s,t)}{n} \\
&\qquad\times\left\{100\|Y_0\|_q^4 \sum_{j=0}^{\lfloor nt\rfloor-\lfloor ns\rfloor-1}\sum_{i,k\le j}\alpha_i^{1/p}\alpha_k^{1/p} + 30\|Y_0\|_q^2 \sum_{i=0}^{\lfloor nt\rfloor-\lfloor ns\rfloor-1}\sum_{j,k\le i}\alpha_i^{1/p}\right\}.
\end{aligned}$$

Observing that $\sum_{i=1}^{\infty}\alpha_i^{1/p}<\infty$ and $\sum_{i=1}^{\infty} i^2\alpha_i^{1/p}<\infty$ (note that $p<a/3$ by construction), we can bound the expression on the right of the previous inequality by

$$\kappa\big\{\lambda_n(s,t)^2\|Y_0\|_q^4 + n^{-1}\lambda_n(s,t)\|Y_0\|_q^2\big\},$$

where $\kappa>0$ is a constant depending on the mixing coefficients and $\mathrm{E}[\{\xi_{0,n}^{(m)}\}^4]$. Finally, since $q>2$ by construction, the assertion follows from the fact that $\mathrm{E}[|Y_0|^q]\le \mathrm{E}[Y_0^2]=\nu(A)-\nu(A)^2\le\nu(A)$. □

Let us introduce additional notation. For any $\delta\ge 0$, $T\subset[0,1]^{d+1}$ and $f\in\ell^\infty([0,1]^{d+1})$, let

$$w_\delta(f,T)=\sup_{\substack{x,y\in T\\ \|x-y\|_1\le\delta}}\big|f(x)-f(y)\big|,$$

where $\|\cdot\|_1$ denotes the 1-norm.

Lemma A.3 (Asymptotic equicontinuity). *Assume that $(\mathbf{U}_i)_{i\in\mathbb{Z}}$ is a strictly stationary sequence whose strong mixing coefficients satisfy $\alpha_r=\mathrm{O}(r^{-a})$, $a>3+3d/2$. Then, for any $m\in\{1,\dots,M\}$, $\tilde{\mathbb{B}}_n^{(m)}$ is asymptotically uniformly $\|\cdot\|_1$-equicontinuous in probability, that is, for any $\varepsilon>0$,*

$$\lim_{\delta\downarrow 0}\limsup_{n\to\infty}\mathrm{P}\big\{w_\delta\big(\tilde{\mathbb{B}}_n^{(m)},[0,1]^{d+1}\big)>\varepsilon\big\}=0.$$

Proof. Fix $m\in\{1,\dots,M\}$. Let $K>0$ be a constant and let us first assume that, for any $n\ge 1$ and $i\in\{1,\dots,n\}$, $\xi_{i,n}^{(m)}\ge -K$. Then, let $Z_{i,n}^{(m)}=\xi_{i,n}^{(m)}+K\ge 0$. Furthermore, let $\gamma\in(0,1/2]$ be a real parameter to be chosen later, and define

$$I_n=\{i/n: i=0,\dots,n\},\qquad I_{n,\gamma}=\big\{i/\big\lfloor n^{1/2+\gamma}\big\rfloor : i=0,\dots,\big\lfloor n^{1/2+\gamma}\big\rfloor\big\},$$

and $T_n=I_n\times I_{n,\gamma}^d$. Also, for any $s\in[0,1]$, let $\underline{s}=\lfloor sn\rfloor/n$ and $\bar{s}=\lceil sn\rceil/n$; clearly, $\underline{s},\bar{s}\in I_n$ and are such that $\underline{s}\le s\le\bar{s}$ and $\bar{s}-\underline{s}\le 1/n$. Similarly, for any $u\in[0,1]$, let $\underline{u}_\gamma,\bar{u}_\gamma\in I_{n,\gamma}$ such that

$\underline{u}_\gamma \leq u \leq \bar{u}_\gamma$ and $\bar{u}_\gamma - \underline{u}_\gamma \leq 1/\lfloor n^{1/2+\gamma} \rfloor$. Then, for any $\mathbf{u} \in [0,1]^d$, we define $\underline{\mathbf{u}}_\gamma \in I^d_{n,\gamma}$ (resp., $\bar{\mathbf{u}}_\gamma \in I^d_{n,\gamma}$) as $\underline{\mathbf{u}}_\gamma = (\underline{u}_{1,\gamma}, \dots, \underline{u}_{d,\gamma})$ (resp., $\bar{\mathbf{u}}_\gamma = (\bar{u}_{1,\gamma}, \dots, \bar{u}_{d,\gamma})$).

Now, for any $(s, \mathbf{u}) \in [0,1]^{d+1}$,

$$\begin{aligned}\tilde{\mathbb{B}}_n^{(m)}(s, \mathbf{u}) - \tilde{\mathbb{B}}_n^{(m)}(s, \underline{\mathbf{u}}_\gamma) \leq{}& \frac{1}{\sqrt{n}} \sum_{i=1}^{\lfloor ns \rfloor} Z_{i,n}^{(m)} \big\{ \mathbf{1}(\mathbf{U}_i \leq \bar{\mathbf{u}}_\gamma) - \mathbf{1}(\mathbf{U}_i \leq \underline{\mathbf{u}}_\gamma) \big\} \\ & + \sqrt{n} K \big\{ C(\bar{\mathbf{u}}_\gamma) - C(\underline{\mathbf{u}}_\gamma) \big\}.\end{aligned}$$

Thus,

$$\begin{aligned}\tilde{\mathbb{B}}_n^{(m)}(s, \mathbf{u}) - \tilde{\mathbb{B}}_n^{(m)}(s, \underline{\mathbf{u}}_\gamma) \leq{}& \tilde{\mathbb{B}}_n^{(m)}(s, \bar{\mathbf{u}}_\gamma) - \tilde{\mathbb{B}}_n^{(m)}(s, \underline{\mathbf{u}}_\gamma) + K \big\{ \tilde{\mathbb{B}}_n(s, \bar{\mathbf{u}}_\gamma) - \tilde{\mathbb{B}}_n(s, \underline{\mathbf{u}}_\gamma) \big\} \\ & + \left(\sqrt{n} K + \frac{1}{\sqrt{n}} \sum_{i=1}^{\lfloor ns \rfloor} Z_{i,n}^{(m)} \right) \big\{ C(\bar{\mathbf{u}}_\gamma) - C(\underline{\mathbf{u}}_\gamma) \big\},\end{aligned}$$

and therefore

$$\begin{aligned}\tilde{\mathbb{B}}_n^{(m)}(s, \mathbf{u}) - \tilde{\mathbb{B}}_n^{(m)}(s, \underline{\mathbf{u}}_\gamma) \leq{}& \big| \tilde{\mathbb{B}}_n^{(m)}(s, \bar{\mathbf{u}}_\gamma) - \tilde{\mathbb{B}}_n^{(m)}(s, \underline{\mathbf{u}}_\gamma) \big| + K \big| \tilde{\mathbb{B}}_n(s, \bar{\mathbf{u}}_\gamma) - \tilde{\mathbb{B}}_n(s, \underline{\mathbf{u}}_\gamma) \big| \\ & + d \big(n^\gamma - 1 \big)^{-1} \Big(K + \max_{1 \leq i \leq n} \big| Z_{i,n}^{(m)} \big| \Big),\end{aligned}$$

using the fact that C satisfies the Lipschitz condition

$$\big| C(\mathbf{u}) - C(\mathbf{v}) \big| \leq \| \mathbf{u} - \mathbf{v} \|_1 \qquad \forall \mathbf{u}, \mathbf{v} \in [0,1]^d, \tag{A.2}$$

and that $n^{1/2} (\lfloor n^{1/2+\gamma} \rfloor)^{-1} \leq (n^\gamma - 1)^{-1}$ for all $n \geq 1$. Similarly, for any $(s, \mathbf{u}) \in [0,1]^{d+1}$,

$$\begin{aligned}&\tilde{\mathbb{B}}_n^{(m)}(s, \underline{\mathbf{u}}_\gamma) - \tilde{\mathbb{B}}_n^{(m)}(s, \mathbf{u}) \\ &\quad \leq \frac{1}{\sqrt{n}} \sum_{i=1}^{\lfloor ns \rfloor} Z_{i,n}^{(m)} \big\{ C(\bar{\mathbf{u}}_\gamma) - C(\underline{\mathbf{u}}_\gamma) \big\} + \frac{K}{\sqrt{n}} \sum_{i=1}^{\lfloor ns \rfloor} \big\{ \mathbf{1}(\mathbf{U}_i \leq \bar{\mathbf{u}}_\gamma) - \mathbf{1}(\mathbf{U}_i \leq \underline{\mathbf{u}}_\gamma) \big\} \\ &\quad \leq d \big(n^\gamma - 1 \big)^{-1} \Big(K + \max_{1 \leq i \leq n} \big| Z_{i,n}^{(m)} \big| \Big) + K \big| \tilde{\mathbb{B}}_n(s, \bar{\mathbf{u}}_\gamma) - \tilde{\mathbb{B}}_n(s, \underline{\mathbf{u}}_\gamma) \big|.\end{aligned}$$

Hence, for any $(s, \mathbf{u}) \in [0,1]^{d+1}$, we have that

$$\begin{aligned}&\big| \tilde{\mathbb{B}}_n^{(m)}(s, \mathbf{u}) - \tilde{\mathbb{B}}_n^{(m)}(s, \underline{\mathbf{u}}_\gamma) \big| \\ &\quad \leq \big| \tilde{\mathbb{B}}_n^{(m)}(s, \bar{\mathbf{u}}_\gamma) - \tilde{\mathbb{B}}_n^{(m)}(s, \underline{\mathbf{u}}_\gamma) \big| + K \big| \tilde{\mathbb{B}}_n(s, \bar{\mathbf{u}}_\gamma) - \tilde{\mathbb{B}}_n(s, \underline{\mathbf{u}}_\gamma) \big| \\ &\qquad + d \big(n^\gamma - 1 \big)^{-1} \Big(K + \max_{1 \leq i \leq n} \big| Z_{i,n}^{(m)} \big| \Big).\end{aligned} \tag{A.3}$$

Then, noticing that, for any $s \in [0,1]$, $\tilde{\mathbb{B}}_n^{(m)}(s,\cdot) = \tilde{\mathbb{B}}_n^{(m)}(\underline{s},\cdot)$, and applying (A.3) to the first and the third summand on the right-hand side of the decomposition

$$\tilde{\mathbb{B}}_n^{(m)}(s,\mathbf{u}) - \tilde{\mathbb{B}}_n^{(m)}(t,\mathbf{v}) = \{\tilde{\mathbb{B}}_n^{(m)}(\underline{s},\mathbf{u}) - \tilde{\mathbb{B}}_n^{(m)}(\underline{s},\underline{\mathbf{u}}_\gamma)\} + \{\tilde{\mathbb{B}}_n^{(m)}(\underline{s},\underline{\mathbf{u}}_\gamma) - \tilde{\mathbb{B}}_n^{(m)}(\underline{t},\underline{\mathbf{v}}_\gamma)\}$$
$$+ \{\tilde{\mathbb{B}}_n^{(m)}(\underline{t},\underline{\mathbf{v}}_\gamma) - \tilde{\mathbb{B}}_n^{(m)}(\underline{t},\mathbf{v})\},$$

we obtain that, for any $\delta > 0$,

$$\begin{aligned} w_\delta(\tilde{\mathbb{B}}_n^{(m)}, [0,1]^{d+1}) &\le 3w_{\delta+(d+1)/\lfloor n^{1/2+\gamma}\rfloor}(\tilde{\mathbb{B}}_n^{(m)}, T_n) + 2K\,w_{\delta+d/\lfloor n^{1/2+\gamma}\rfloor}(\tilde{\mathbb{B}}_n, [0,1]^{d+1}) \\ &\quad + 2d(n^\gamma - 1)^{-1}\Bigl(K + \max_{1\le i\le n}|Z_{i,n}^{(m)}|\Bigr) \\ &\le 3w_{2\delta}(\tilde{\mathbb{B}}_n^{(m)}, T_n) + 2K\,w_{2\delta}(\tilde{\mathbb{B}}_n, [0,1]^{d+1}) \\ &\quad + 2d(n^\gamma - 1)^{-1}\Bigl(K + \max_{1\le i\le n}|Z_{i,n}^{(m)}|\Bigr), \end{aligned}$$

for sufficiently large n. Now, from the previous inequality, for any $\varepsilon > 0$,

$$\begin{aligned} \mathrm{P}\{w_\delta(\tilde{\mathbb{B}}_n^{(m)}, [0,1]^{d+1}) > \varepsilon\} &\le \mathrm{P}\{3w_{2\delta}(\tilde{\mathbb{B}}_n^{(m)}, T_n) > \varepsilon/3\} \\ &\quad + \mathrm{P}\{2K\,w_{2\delta}(\tilde{\mathbb{B}}_n, [0,1]^{d+1}) > \varepsilon/3\} \\ &\quad + \mathrm{P}\Bigl\{2d(n^\gamma - 1)^{-1}\Bigl(K + \max_{1\le i\le n}|Z_{i,n}^{(m)}|\Bigr) > \varepsilon/3\Bigr\}. \end{aligned}$$

Since $a > 1$, we have from Bücher ([7], Lemma 2) that $\tilde{\mathbb{B}}_n$ is asymptotically uniformly $\|\cdot\|_1$-equicontinuous in probability. This implies that the second term on the right of the previous display converges to 0 as $n \to \infty$ followed by $\delta \downarrow 0$. The third term converges to zero because $n^{-\gamma}\max_{1\le i\le n}|Z_{i,n}^{(m)}| \xrightarrow{\mathrm{P}} 0$. Indeed, for any $\eta > 0$ and $\nu > 1/\gamma \ge 2$, by Markov's inequality and (M1),

$$\mathrm{P}\Bigl(n^{-\gamma}\max_{1\le i\le n}|Z_{i,n}^{(m)}| > \eta\Bigr) \le n\mathrm{P}(|Z_{1,n}^{(m)}| \ge \eta n^\gamma) \le \eta^{-\nu}n^{1-\gamma\nu}\sup_{n\ge 1}\mathrm{E}(|Z_{1,n}^{(m)}|^\nu) \to 0.$$

Thus, it remains to show that, for any $\varepsilon > 0$, $\lim_{\delta\downarrow 0}\limsup_{n\to\infty}\mathrm{P}\{w_\delta(\tilde{\mathbb{B}}_n^{(m)}, T_n) > \varepsilon\} = 0$, or equivalently (see, e.g., van der Vaart and Wellner [48], Problem 2.1.5) that, for any positive sequence $\delta_n \downarrow 0$, $\lim_{n\to\infty}\mathrm{P}\{w_{\delta_n}(\tilde{\mathbb{B}}_n^{(m)}, T_n) > \varepsilon\} = 0$. To do so, we shall use Lemma A.2 together with Lemma 2 of Balacheff and Dupont [4] (see also Bickel and Wichura [6], Theorem 3 and the remarks on page 1665).

Recall that ν is the measure on $[0,1]^d$ corresponding to the c.d.f. C, and let μ be a measure on $[0,1]^{d+1}$ defined by $\mu = 2\lambda \otimes \nu$, where λ denotes the one-dimensional Lebesgue measure. Next, for some real $q \in (2a/(a-3), 6a/(2a-3)) \subset (2,4)$, let $\beta = 2 - 2/q - 3/a \in (1, 4/q)$. Furthermore, consider a non-empty set $(s,t] \times A = (s,t] \times (u_1,v_1] \times \cdots \times (u_d,v_d]$ of $[0,1]^{d+1}$ whose boundary points are all distinct and lie in T_n. Then, starting from Lemma A.2, for any

$q \in (2a/(a-3), 6a/(2a-3)) \subset (2,4)$,

$$\begin{aligned}
&\mathrm{E}\big[\big\{\tilde{\mathbb{B}}_n^{(m)}\big((s,t]\times A\big)\big\}^4\big] \\
&\quad\le \kappa\big[\lambda_n(s,t)^2\big\{\nu(A)\big\}^{4/q} + n^{-1}\lambda_n(s,t)\big\{\nu(A)\big\}^{2/q}\big] \\
&\quad\le \kappa\big[\big\{\lambda_n(s,t)\nu(A)\big\}^{4/q} + n^{-1}\big\{\lambda_n(s,t)\nu(A)\big\}^{2/q}\big] \\
&\quad\le \kappa\mu\big((s,t]\times A\big)^{\beta}\big\{\mu\big((s,t]\times A\big)^{4/q-\beta} + n^{-1}\mu\big((s,t]\times A\big)^{2/q-\beta}\big\} \\
&\quad\le \kappa\mu\big((s,t]\times A\big)^{\beta}\big\{2^{4/q-\beta} + n^{-1}n^{-(1+d/2+d\gamma)(2/q-\beta)}\big\} \\
&\quad= \kappa\mu\big((s,t]\times A\big)^{\beta}\big\{2^{4/q-\beta} + n^{(\beta-2/q)(1+d/2+d\gamma)-1}\big\}.
\end{aligned}$$

Note that $\inf_{q>2a/(a-3)}(\beta - 2/q) = 3/a$. Hence, because $3/a < 2/(2+d)$ from the assumption on the mixing rate, it is possible to choose $q \in (2a/(a-3), 6a/(2a-3))$ and $\gamma > 0$ (the parameter involved in the grid $I_{n,\gamma}^d$) small enough such that $\beta - 2/q < 2/(2+d+2d\gamma)$. For the aforementioned parameter choices, $(\beta - 2/q)(1 + d/2 + d\gamma) - 1 < 0$, which implies that $n^{(\beta-2/q)(1+d/2+d\gamma)-1} \le 1$ for all $n \ge 1$.

With some abuse of notation consisting of incorporating the constant $\{\kappa(2^{4/q-\beta}+1)\}^{1/\beta}$ into the measure, we obtain

$$\mathrm{E}\big[\big\{\tilde{\mathbb{B}}_n^{(m)}\big((s,t]\times A\big)\big\}^4\big] \le \mu\big((s,t]\times A\big)^{\beta},$$

which, by Markov's inequality, implies that, for any $\varepsilon > 0$,

$$\mathrm{P}\big\{\big|\tilde{\mathbb{B}}_n^{(m)}\big((s,t]\times A\big)\big| \ge \varepsilon\big\} \le \varepsilon^{-4}\mu\big((s,t]\times A\big)^{\beta}.$$

Now, let $\tilde{\mu}_n$ denote a finite measure on T_n defined from its values on the singletons $\{(s,\mathbf{u})\}$ of T_n as

$$\tilde{\mu}_n\big(\big\{(s,\mathbf{u})\big\}\big) = \begin{cases} 0, & \text{if } s\wedge u_1\wedge\cdots\wedge u_d = 0, \\ \mu\big(\big(s',s\big]\times\big(u_1',u_1\big]\times\cdots\times\big(u_d',u_d\big]\big), & \text{otherwise,}\end{cases}$$

where $s' = \max\{t \in I_n : t < s\}$ and $u_j' = \max\{u \in I_{n,\gamma} : u < u_j\}$ for all $j \in \{1,\ldots,d\}$. By additivity of $\tilde{\mu}_n$, the previous estimation reads

$$\mathrm{P}\big\{\big|\tilde{\mathbb{B}}_n^{(m)}\big((s,t]\times A\big)\big| \ge \varepsilon\big\} \le \varepsilon^{-4}\tilde{\mu}_n\big[\big\{(s,t]\times A\big\}\cap T_n\big]^{\beta}.$$

We shall now conclude by an application of Lemma 2 of Balacheff and Dupont [4]. Consider a positive sequence $\delta_n \downarrow 0$, and let $\delta_n' \downarrow 0$ such that, for any $n \in \mathbb{N}$, $\delta_n' \in \{1/i : i \in \mathbb{N}\}$ and $\delta_n' \ge \max\{\delta_n, 1/\lfloor n^{1/2+\gamma}\rfloor\}$. Applying Lemma 2 of Balacheff and Dupont [4] (note that $1/\lfloor n^{1/2+\gamma}\rfloor = \max\{1/n, 1/\lfloor n^{1/2+\gamma}\rfloor\}$ is denoted by τ in the lemma) and using the fact that $\|\cdot\|_2 \le \|\cdot\|_1$, we obtain that, for any $\varepsilon > 0$, there exists a constant $\lambda > 0$ depending on ε, β and d, such that

$$\begin{aligned}
&\mathrm{P}\big\{w_{\delta_n}\big(\tilde{\mathbb{B}}_n^{(m)}, T_n\big) > \varepsilon\big\} \\
&\quad\le \mathrm{P}\big\{w_{\delta_n'}\big(\tilde{\mathbb{B}}_n^{(m)}, T_n\big) > \varepsilon\big\}
\end{aligned}$$

$$
\begin{aligned}
&\le \lambda \tilde{\mu}_n(T_n) \\
&\quad \times \Bigg[\max\Bigg\{ \sup_{\substack{s,t \in I_n \\ |s-t| \le 3\delta_n'}} \big|\tilde{\mu}_n\big(\{0,\dots,s\} \times I_{n,\gamma}^d\big) - \tilde{\mu}_n\big(\{0,\dots,t\} \times I_{n,\gamma}^d\big)\big|, \\
&\qquad \sup_{\substack{u,v \in I_{n,\gamma} \\ |u-v| \le 3\delta_n'}} \big|\tilde{\mu}_n\big(I_n \times \{0,\dots,u\} \times I_{n,\gamma}^{d-1}\big) - \tilde{\mu}_n\big(I_n \times \{0,\dots,v\} \times I_{n,\gamma}^{d-1}\big)\big|, \\
&\qquad \dots, \\
&\qquad \sup_{\substack{u,v \in I_{n,\gamma} \\ |u-v| \le 3\delta_n'}} \big|\tilde{\mu}_n\big(I_n \times I_{n,\gamma}^{d-1} \times \{0,\dots,u\}\big) - \tilde{\mu}_n\big(I_n \times I_{n,\gamma}^{d-1} \times \{0,\dots,v\}\big)\big| \Bigg\}\Bigg]^{\beta-1},
\end{aligned}
$$

which implies that,

$$
\begin{aligned}
&\mathrm{P}\big\{w_{\delta_n}\big(\tilde{\mathbb{B}}_n^{(m)}, T_n\big) > \varepsilon\big\} \\
&\quad \le \lambda \mu\big([0,1]^{d+1}\big) \\
&\qquad \times \Bigg[\max\Bigg\{ \sup_{\substack{s,t \in [0,1] \\ |s-t| \le 3\delta_n'}} \big|\mu\big([0,s] \times [0,1]^d\big) - \mu\big([0,t] \times [0,1]^d\big)\big|, \\
&\qquad\quad \sup_{\substack{u,v \in [0,1] \\ |u-v| \le 3\delta_n'}} \big|\mu\big([0,1] \times [0,u] \times [0,1]^{d-1}\big) - \mu\big([0,1] \times [0,v] \times [0,1]^{d-1}\big)\big|, \\
&\qquad\quad \dots, \\
&\qquad\qquad\qquad \sup_{\substack{u,v \in [0,1] \\ |u-v| \le 3\delta_n'}} \big|\mu\big([0,1]^d \times [0,u]\big) - \mu\big([0,1]^d \times [0,v]\big)\big| \Bigg\}\Bigg]^{\beta-1},
\end{aligned}
$$

which converges to 0 by uniform continuity of the functions $s \mapsto \mu([0,s] \times [0,1]^d)$, $u \mapsto \mu([0,1] \times [0,u] \times [0,1]^{d-1}), \dots, u \mapsto \mu([0,1]^d \times [0,u])$ on $[0,1]$. This concludes the proof for the case $\xi_{i,n}^{(m)} \ge -K$.

Let us now consider the general case. Let $Z_{i,n}^+ = \max(\xi_{i,n}^{(m)}, 0)$, $Z_{i,n}^- = \max(-\xi_{i,n}^{(m)}, 0)$, $K^+ = \mathrm{E}(Z_{0,n}^+)$ and $K^- = \mathrm{E}(Z_{0,n}^-)$. Furthermore, define $\xi_{i,n}^{(m),+} = Z_{i,n}^+ - K^+$ and $\xi_{i,n}^{(m),-} = Z_{i,n}^- - K^-$. Then, using the fact that $K^+ - K^- = 0$, we can write

$$
\xi_{i,n}^{(m)} = Z_{i,n}^+ - Z_{i,n}^- = Z_{i,n}^+ - K^+ - \big(Z_{i,n}^- - K^-\big) = \xi_{i,n}^{(m),+} - \xi_{i,n}^{(m),-}.
$$

Setting

$$
\tilde{\mathbb{B}}_n^{(m),\pm}(s,\mathbf{u}) = n^{-1/2} \sum_{i=1}^{\lfloor ns \rfloor} \xi_{i,n}^{(m),\pm} \big\{\mathbf{1}(\mathbf{U}_i \le \mathbf{u}) - C(\mathbf{u})\big\}, \qquad (s,\mathbf{u}) \in [0,1]^{d+1},
$$

we obtain that $\tilde{\mathbb{B}}_n^{(m)} = \tilde{\mathbb{B}}_n^{(m),+} - \tilde{\mathbb{B}}_n^{(m),-}$. The case treated above immediately yields asymptotic equicontinuity of $\tilde{\mathbb{B}}_n^{(m),+}$ and of $\tilde{\mathbb{B}}_n^{(m),-}$, which implies asymptotic equicontinuity of $\tilde{\mathbb{B}}_n^{(m)}$. □

Proof of Theorem 2.1. Weak convergence of the finite-dimensional distributions is established in Lemma A.1. Asymptotic tightness of $\tilde{\mathbb{B}}_n$ is a consequence of the weak convergence of $\tilde{\mathbb{B}}_n$ to $\mathbb{B}_C$ in $\ell^\infty([0,1]^d)$, which follows from Theorem 1 in Bücher [7]. From Lemma A.3, we have that, for any $m \in \{1, \dots, M\}$, $\tilde{\mathbb{B}}_n^{(m)}$ is asymptotically uniformly $\|\cdot\|_1$-equicontinuous in probability. Together with the fact that $[0,1]^{d+1}$ is totally bounded for $\|\cdot\|_1$ and Lemma A.1, we have, for instance, from Theorem 2.1 in Kosorok [29], that, for any $m \in \{1, \dots, M\}$, $\tilde{\mathbb{B}}_n^{(m)} \rightsquigarrow \mathbb{B}_C^{(m)}$ in $\ell^\infty([0,1]^d)$, which implies asymptotic tightness of $\tilde{\mathbb{B}}_n^{(m)}$. The proof is complete as marginal asymptotic tightness implies joint asymptotic tightness. □

Appendix B: Proof of Theorem 3.4

The proof of Theorem 3.4 is based on the extended continuous mapping theorem (van der Vaart and Wellner [48], Theorem 1.11.1). The intuition of the proof is as follows: the aim is to construct suitable maps g_n and g such that g_n continuously converges to g (i.e., $g_n(\alpha_n)$ converges uniformly to $g(\alpha)$ for all sequences α_n converging uniformly to α) and such that we may conclude that, as a process indexed by $s, t, \mathbf{u}$, $\mathbb{C}_n(s,t,\mathbf{u}) \approx g_n\{\tilde{\mathbb{B}}_n(t,\mathbf{u}) - \tilde{\mathbb{B}}_n(s,\mathbf{u})\}$ converges weakly to $g\{\tilde{\mathbb{B}}(t,\mathbf{u}) - \tilde{\mathbb{B}}(s,\mathbf{u})\} = \mathbb{C}(s,t,\mathbf{u})$.

In the following, all the convergences are with respect to $n \to \infty$. Let $\mathcal{E}$ be the set of c.d.f.s on $[0,1]$ with no mass at 0, that is,

$$\mathcal{E} = \bigl\{F : [0,1] \to [0,1] : F \text{ is right-continuous and nondecreasing with } F(0) = 0 \text{ and } F(1) = 1\bigr\},$$

let

$$\mathcal{E}_n^\star = \bigl\{F^\star : \Delta \times [0,1] \to [0,1] : u \mapsto \lambda_n(s,t)^{-1} F^\star(s,t,u) \in \mathcal{E} \text{ if } \lfloor ns \rfloor < \lfloor nt \rfloor \text{ and } F^\star(s,t,\cdot) = 0 \text{ if } \lfloor ns \rfloor = \lfloor nt \rfloor\bigr\},$$

where $\lambda_n(s,t) = (\lfloor nt \rfloor - \lfloor ns \rfloor)/n$, and let I_n be the sequence of maps defined, for any $F^\star \in \mathcal{E}_n^\star$ and any $(s,t,u) \in \Delta \times [0,1]$, by

$$I_n\bigl(F^\star\bigr)(s,t,u) = \inf\bigl\{v \in [0,1] : F^\star(s,t,v) \ge \lambda_n(s,t)u\bigr\}.$$

Furthermore, given a function $H^\star \in \ell^\infty(\Delta \times [0,1]^d)$, for any $j \in \{1, \dots, d\}$, we define

$$H_j^\star(s,t,u) = H^\star(s,t,\mathbf{u}_{\{j\}}), \qquad (s,t,u) \in \Delta \times [0,1],$$

where, for any $u \in [0,1]$, $\mathbf{u}_{\{j\}}$ is the vector of $[0,1]^d$ whose components are all equal to 1 except the jth one which is equal to u. Then, let

$$\mathcal{E}_{n,d}^\star = \bigl\{H^\star : \Delta \times [0,1]^d \to [0,1] : H_j^\star \in \mathcal{E}_n^\star \text{ for all } j \in \{1, \dots, d\}\bigr\}$$

and let Φ_n be the map from $\mathcal{E}^\star_{n,d}$ to $\ell^\infty(\Delta \times [0,1]^d)$ defined, for any $H^\star \in \mathcal{E}^\star_{n,d}$ and $(s,t,\mathbf{u}) \in \Delta \times [0,1]^d$, by

$$\Phi_n\big(H^\star\big)(s,t,\mathbf{u}) = H^\star\big\{s,t,I_n\big(H^\star_1\big)(s,t,u_1),\dots,I_n\big(H^\star_d\big)(s,t,u_d)\big\}. \tag{B.1}$$

Let additionally $U^\star_n \in \mathcal{E}^\star_n$ be defined as $U^\star_n(s,t,u) = \lambda_n(s,t)u$ for all $(s,t,u) \in \Delta \times [0,1]$, and let $C^\star_n(s,t,\mathbf{u}) = \lambda_n(s,t)C(\mathbf{u})$ for all $(s,t,\mathbf{u}) \in \Delta \times [0,1]^d$. Clearly, we have that $C^\star_{n,1} = \cdots = C^\star_{n,d} = U^\star_n$. Moreover, $\Phi_n(C^\star_n) = C^\star_n$.

Also, let

$$\begin{aligned}\mathcal{D}^\star = \big\{&\alpha^\star \in \ell^\infty\big(\Delta \times [0,1]^d\big) : \alpha^\star(s,t,\cdot) = 0 \text{ if } s = t, \text{ and} \\ &\alpha^\star(s,t,\mathbf{u}) = 0 \text{ if } s < t \text{ and if one of the components of } \mathbf{u} \text{ is } 0 \text{ or } \mathbf{u} = (1,\dots,1)\big\},\end{aligned}$$

let $\mathcal{D}^\star_n = \{\alpha^\star \in \mathcal{D}^\star : C^\star_n + n^{-1/2}\alpha^\star \in \mathcal{E}^\star_{n,d}\}$, and let $\mathcal{D}^\star_0 = \mathcal{D}^\star \cap \mathcal{C}(\Delta \times [0,1]^d)$. Finally, for any $\alpha^\star_n \in \mathcal{D}^\star_n$ and any $(s,t,\mathbf{u}) \in \Delta \times [0,1]^d$, let

$$g_n\big(\alpha^\star_n\big)(s,t,\mathbf{u}) = \sqrt{n}\big\{\Phi_n\big(C^\star_n + n^{-1/2}\alpha^\star_n\big)(s,t,\mathbf{u}) - \Phi_n\big(C^\star_n\big)(s,t,\mathbf{u})\big\}, \tag{B.2}$$

and, for any $\alpha^\star \in \mathcal{D}^\star_0$ and any $(s,t,\mathbf{u}) \in \Delta \times [0,1]^d$, let

$$g\big(\alpha^\star\big)(s,t,\mathbf{u}) = \alpha^\star(s,t,\mathbf{u}) - \sum_{j=1}^d \dot{C}_j(\mathbf{u})\alpha^\star\big(s,t,\mathbf{u}^{(j)}\big).$$

The following lemma is the main ingredient for the proof of Theorem 3.4. Its proof is given subsequent to the proof of Theorem 3.4.

Lemma B.1. *Suppose that C satisfies Condition 3.2, and let $\alpha^\star_n \to \alpha^\star$ with $\alpha^\star_n \in \mathcal{D}^\star_n$ for every n and $\alpha^\star \in \mathcal{D}^\star_0$. Then, $g_n(\alpha^\star_n) \to g(\alpha^\star) \in \ell^\infty(\Delta \times [0,1]^d)$.*

Proof of Theorem 3.4. Under Condition 3.1, we have that $\tilde{\mathbb{B}}_n \rightsquigarrow \mathbb{B}_C$ in $\ell^\infty([0,1]^{d+1})$. Now, for any $(s,t,\mathbf{u}) \in \Delta \times [0,1]^d$, define $\tilde{\mathbb{B}}^\Delta_n(s,t,\mathbf{u}) = \tilde{\mathbb{B}}_n(t,\mathbf{u}) - \tilde{\mathbb{B}}_n(s,\mathbf{u})$, $\mathbb{B}^\Delta_C(s,t,\mathbf{u}) = \mathbb{B}_C(t,\mathbf{u}) - \mathbb{B}_C(s,\mathbf{u})$, and

$$\tilde{H}^\star_n(s,t,\mathbf{u}) = \frac{1}{n}\sum_{i=\lfloor ns\rfloor+1}^{\lfloor nt\rfloor} \mathbf{1}(\mathbf{U}_i \le \mathbf{u}).$$

Notice that $\tilde{\mathbb{B}}^\Delta_n = \sqrt{n}(\tilde{H}^\star_n - C^\star_n)$ and that, by the continuous mapping theorem, $\tilde{\mathbb{B}}^\Delta_n \rightsquigarrow \mathbb{B}^\Delta_C$ in $\ell^\infty(\Delta \times [0,1]^d)$. Clearly, $\tilde{\mathbb{B}}^\Delta_n$, as a function of ω, takes its values in $\mathcal{D}^\star_n$ and $\mathbb{B}^\Delta_C$ is Borel measurable and separable by Condition 3.1, and, as a function of ω, takes its values in $\mathcal{D}^\star_0$. Now, consider the map h_n from $\mathcal{D}^\star_n$ to $\{\ell^\infty(\Delta \times [0,1]^d)\}^2$, defined, for any $\alpha^\star_n \in \mathcal{D}^\star_n$ and any $(s,t,\mathbf{u}) \in \Delta \times [0,1]^d$, by

$$h_n\big(\alpha^\star_n\big)(s,t,\mathbf{u}) = \big(g_n\big(\alpha^\star_n\big)(s,t,\mathbf{u}), g\big(\alpha^\star_n\big)(s,t,\mathbf{u})\big).$$

Using Lemma B.1 and the fact that g is linear and bounded, we have from the extended continuous mapping theorem (van der Vaart and Wellner [48], Theorem 1.11.1) that $h_n(\tilde{\mathbb{B}}_n^{\Delta}) \rightsquigarrow h(\mathbb{B}_C^{\Delta})$ in $\{\ell^{\infty}(\Delta \times [0,1]^d)\}^2$, where, for any $\alpha^{\star} \in \mathcal{D}_0^{\star}$ and any $(s,t,\mathbf{u}) \in \Delta \times [0,1]^d$,

$$h(\alpha^{\star})(s,t,\mathbf{u}) = \big(g(\alpha^{\star})(s,t,\mathbf{u}), g(\alpha^{\star})(s,t,\mathbf{u})\big).$$

An application of the continuous mapping theorem immediately yields that $g_n(\tilde{\mathbb{B}}_n^{\Delta}) - \tilde{\mathbb{C}}_n = g_n(\tilde{\mathbb{B}}_n^{\Delta}) - g(\tilde{\mathbb{B}}_n^{\Delta}) \rightsquigarrow 0$ in $\ell^{\infty}(\Delta \times [0,1]^d)$, where $\tilde{\mathbb{C}}_n$ is defined in (3.1). To complete the proof, it remains to show that

$$A_n = \sup_{(s,t,\mathbf{u}) \in \Delta \times [0,1]^d} \big| g_n(\tilde{\mathbb{B}}_n^{\Delta})(s,t,\mathbf{u}) - \mathbb{C}_n(s,t,\mathbf{u}) \big| = o_{\mathrm{P}}(1).$$

Note that it suffices to restrict the supremum over all pairs $(s,t) \in \Delta$ such that $\lfloor ns \rfloor < \lfloor nt \rfloor$. From the definition of g_n, we have that

$$\begin{aligned}
& g_n(\tilde{\mathbb{B}}_n^{\Delta})(s,t,\mathbf{u}) \\
&\quad = \sqrt{n}\big\{\Phi_n(\tilde{H}_n^{\star})(s,t,\mathbf{u}) - \Phi_n(C_n^{\star})(s,t,\mathbf{u})\big\} \\
&\quad = \frac{1}{\sqrt{n}} \sum_{i=\lfloor ns \rfloor+1}^{\lfloor nt \rfloor} \big[\mathbf{1}\big\{U_{i1} \le I_n(\tilde{H}_{n,1}^{\star})(s,t,u_1), \dots, U_{id} \le I_n(\tilde{H}_{n,d}^{\star})(s,t,u_d)\big\} - C(\mathbf{u})\big].
\end{aligned}$$

Now, let $\tilde{H}_{\lfloor ns \rfloor+1:\lfloor nt \rfloor}$ be the empirical c.d.f. computed from the sample $\mathbf{U}_{\lfloor ns \rfloor+1}, \dots, \mathbf{U}_{\lfloor nt \rfloor}$, and let $\tilde{H}_{\lfloor ns \rfloor+1:\lfloor nt \rfloor,1}, \dots, \tilde{H}_{\lfloor ns \rfloor+1:\lfloor nt \rfloor,d}$ be the corresponding marginal c.d.f.s. Given $F \in \mathcal{E}$, let F^{-1} be its generalized inverse defined by $F^{-1}(u) = \inf\{v \in [0,1] : F(v) \ge u\}$. Then, let

$$\tilde{\mathbf{H}}_{\lfloor ns \rfloor+1:\lfloor nt \rfloor}^{-1}(\mathbf{u}) = \big(\tilde{H}_{\lfloor ns \rfloor+1:\lfloor nt \rfloor,1}^{-1}(u_1), \dots, \tilde{H}_{\lfloor ns \rfloor+1:\lfloor nt \rfloor,d}^{-1}(u_d)\big), \qquad \mathbf{u} \in [0,1]^d.$$

Using the fact that, for any $j \in \{1,\dots,d\}$, $I_n(\tilde{H}_{n,j}^{\star})(s,t,u) = \tilde{H}_{\lfloor ns \rfloor+1:\lfloor nt \rfloor,j}^{-1}(u)$ for all $(s,t,u) \in \Delta \times [0,1]$ such that $\lfloor ns \rfloor < \lfloor nt \rfloor$, we obtain

$$\begin{aligned}
g_n(\tilde{\mathbb{B}}_n^{\Delta})(s,t,\mathbf{u}) &= \frac{1}{\sqrt{n}} \sum_{i=\lfloor ns \rfloor+1}^{\lfloor nt \rfloor} \big[\mathbf{1}\big\{\mathbf{U}_i \le \tilde{\mathbf{H}}_{\lfloor ns \rfloor+1:\lfloor nt \rfloor}^{-1}(\mathbf{u})\big\} - C(\mathbf{u})\big] \\
&= \sqrt{n}\lambda_n(s,t)\big[\tilde{H}_{\lfloor ns \rfloor+1:\lfloor nt \rfloor}\big\{\tilde{\mathbf{H}}_{\lfloor ns \rfloor+1:\lfloor nt \rfloor}^{-1}(\mathbf{u})\big\} - C(\mathbf{u})\big].
\end{aligned}$$

Hence, we obtain that

$$\begin{aligned}
A_n &= \sup_{(s,t,\mathbf{u}) \in \Delta \times [0,1]^d} \sqrt{n}\lambda_n(s,t) \big| C_{\lfloor ns \rfloor+1:\lfloor nt \rfloor}(\mathbf{u}) - \tilde{H}_{\lfloor ns \rfloor+1:\lfloor nt \rfloor}\big\{\tilde{\mathbf{H}}_{\lfloor ns \rfloor+1:\lfloor nt \rfloor}^{-1}(\mathbf{u})\big\} \big| \\
&= n^{-1/2} \max_{1 \le l < k \le n} \sup_{\mathbf{u} \in [0,1]^d} (k-l) \big| C_{l+1:k}(\mathbf{u}) - \tilde{H}_{l+1:k}\big\{\tilde{\mathbf{H}}_{l+1:k}^{-1}(\mathbf{u})\big\} \big|.
\end{aligned}$$

Under Condition 3.3, it can be verified, using properties of generalized inverses, that

$$\sup_{\mathbf{u}\in[0,1]^d} \left|C_{l+1:k}(\mathbf{u}) - \tilde{H}_{l+1:k}\left\{\tilde{\mathbf{H}}^{-1}_{l+1:k}(\mathbf{u})\right\}\right| \leq \frac{d}{k-l},$$

which implies that $A_n \to 0$ and completes the proof. □

It remains to prove Lemma B.1. For that purpose, another lemma is needed.

Lemma B.2. *Let $\alpha_n^\star \to \alpha^\star$ with $\alpha_n^\star \in \mathcal{D}_n^\star$ for every n and $\alpha^\star \in \mathcal{D}_0^\star$. Then, for any $j \in \{1,\ldots,d\}$,*

$$\sup_{(s,t,u)\in\Delta\times[0,1]} \left|\sqrt{n}\lambda_n(s,t)\left\{I_n\left(U_n^\star + n^{-1/2}\alpha^\star_{n,j}\right)(s,t,u) - u\right\} + \alpha_j^\star(s,t,u)\right| \to 0.$$

Proof. The assertion is trivial for $u = 0$ because $\alpha^\star \in \mathcal{D}_0^\star$ and $U_n^\star + n^{-1/2}\alpha^\star_{n,j} \in \mathcal{E}_n^\star$.

Clearly, for any $s \in [0,1]$, $ns \geq \lfloor ns\rfloor$, that is, $s \geq \lambda_n(0,s)$. Furthermore, under the constraint $s \leq t$, $\lfloor nt\rfloor = \lfloor ns\rfloor$ is equivalent to $0 \leq t - \lambda_n(0,s) < 1/n$, which can be written as $0 \leq t - s + s - \lambda_n(0,s) < 1/n$, which means that there exists $h_n \downarrow 0$ such that $t - s < h_n$. Then, we have

$$\begin{aligned}
&\sup_{\lfloor nt\rfloor=\lfloor ns\rfloor, u\in[0,1]} \left|\lambda_n(s,t)\sqrt{n}\left\{I_n\left(U_n^\star + n^{-1/2}\alpha^\star_{n,j}\right)(s,t,u) - u\right\} + \alpha_j^\star(s,t,u)\right| \\
&\quad\leq \sup_{t-s<h_n, u\in[0,1]} \left|\alpha_j^\star(s,t,u)\right| \to 0
\end{aligned}$$

by uniform continuity of $\alpha_j^\star$ on $\Delta \times [0,1]$.

Hence, it remains to consider the case $\lfloor ns\rfloor < \lfloor nt\rfloor$ and $u \in (0,1]$. Given $F \in \mathcal{E}$, let F^{-1} be its generalized inverse defined by $F^{-1}(u) = \inf\{v \in [0,1] : F(v) \geq u\}$. Then, notice that, for any $\lfloor ns\rfloor < \lfloor nt\rfloor$ and $u \in [0,1]$, $I_n(U_n^\star + n^{-1/2}\alpha^\star_{n,j})(s,t,u) = F^{-1}_{s,t,n}(u)$, where $F_{s,t,n} = \lambda_n(s,t)^{-1}(U_n^\star + n^{-1/2}\alpha^\star_{n,j})(s,t,\cdot) \in \mathcal{E}$. It follows that, for any $\lfloor ns\rfloor < \lfloor nt\rfloor$ and $u \in (0,1]$, $\xi_n(s,t,u) = I_n(U_n^\star + n^{-1/2}\alpha^\star_{n,j})(s,t,u) > 0$, and therefore that $\varepsilon_n(s,t,u) = n^{-1} \wedge \xi_n(s,t,u) > 0$. Also, for any $F \in \mathcal{E}$, it can be verified that $F\{F^{-1}(u) - \eta\} \leq u \leq F \circ F^{-1}(u)$ for all $u \in (0,1]$ and all $\eta > 0$ such that $F^{-1}(u) - \eta \geq 0$. Hence, for any $\lfloor ns\rfloor < \lfloor nt\rfloor$ and $u \in (0,1]$,

$$\begin{aligned}
\left(U_n^\star + n^{-1/2}\alpha^\star_{n,j}\right)\left\{s,t,\xi_n(s,t,u) - \varepsilon_n(s,t,u)\right\} &\leq \lambda_n(s,t)u \\
&\leq \left(U_n^\star + n^{-1/2}\alpha^\star_{n,j}\right)\left\{s,t,\xi_n(s,t,u)\right\},
\end{aligned}$$

that is

$$\begin{aligned}
&-n^{-1/2}\alpha^\star_{n,j}\left\{s,t,\xi_n(s,t,u)\right\} \\
&\quad\leq \lambda_n(s,t)\left\{\xi_n(s,t,u) - u\right\} \\
&\quad\leq \lambda_n(s,t)\varepsilon_n(s,t,u) - n^{-1/2}\alpha^\star_{n,j}\left\{s,t,\xi_n(s,t,u) - \varepsilon_n(s,t,u)\right\},
\end{aligned} \tag{B.3}$$

which in turn implies that

$$\sup_{\lfloor ns\rfloor<\lfloor nt\rfloor, u\in(0,1]} \big|\lambda_n(s,t)\big\{\xi_n(s,t,u)-u\big\}\big| \to 0 \tag{B.4}$$

since, by uniform convergence of $\alpha_n^\star$ to $\alpha^\star$ and the fact that $\alpha^\star \in \mathcal{D}_0^\star$, the quantity $\sup_{(s,t,u)\in\Delta\times[0,1]} |\alpha_{n,j}^\star(s,t,u)|$ is bounded. From (B.3), exploiting the fact that $\varepsilon_n(s,t,u)\le n^{-1}$, we then obtain that

$$\sup_{\lfloor ns\rfloor<\lfloor nt\rfloor, u\in(0,1]} \big|\sqrt{n}\lambda_n(s,t)\big\{\xi_n(s,t,u)-u\big\}+\alpha_j^\star(s,t,u)\big| \le A_n+B_n+n^{-1/2},$$

where

$$A_n=\sup_{\lfloor ns\rfloor<\lfloor nt\rfloor, u\in(0,1]} \big|\alpha_n^\star\big\{s,t,\xi_n(s,t,u)\big\}-\alpha_j^\star(s,t,u)\big|,$$

and

$$B_n=\sup_{\lfloor ns\rfloor<\lfloor nt\rfloor, u\in(0,1]} \big|\alpha_{n,j}^\star\big\{s,t,\xi_n(s,t,u)-\varepsilon_n(s,t,u)\big\}-\alpha_j^\star(s,t,u)\big|.$$

For B_n, we write $B_n\le B_{n,1}+B_{n,2}$, where

$$\begin{aligned} B_{n,1}&=\sup_{\substack{\lfloor ns\rfloor<\lfloor nt\rfloor\\ u\in(0,1]}} \big|\alpha_{n,j}^\star\big\{s,t,\xi_n(s,t,u)-\varepsilon_n(s,t,u)\big\}-\alpha_j^\star\big\{s,t,\xi_n(s,t,u)-\varepsilon_n(s,t,u)\big\}\big| \\ &\le \sup_{(s,t,u)\in\Delta\times[0,1]} \big|\alpha_{n,j}^\star(s,t,u)-\alpha_j^\star(s,t,u)\big| \to 0, \end{aligned}$$

and

$$B_{n,2}=\sup_{(s,t,u)\in\Delta\times[0,1]} \big|\alpha_j^\star\big\{s,t,\xi_n(s,t,u)-\varepsilon_n(s,t,u)\big\}-\alpha_j^\star(s,t,u)\big|.$$

It remains to show that $B_{n,2}\to 0$. Let $\varepsilon>0$. Since $\alpha^\star\in\mathcal{D}_0^\star$, there exists $\delta>0$ such that $\sup_{t-s<\delta,u\in[0,1]}|\alpha_j^\star(s,t,u)|\le\varepsilon$. We have $B_{n,2}=\max\{B_{n,3},B_{n,4}\}$, where

$$B_{n,3}=\sup_{t-s<\delta, u\in[0,1]} \big|\alpha_j^\star\big\{s,t,\xi_n(s,t,u)-\varepsilon_n(s,t,u)\big\}-\alpha_j^\star(s,t,u)\big| \le 2\varepsilon,$$

and

$$B_{n,4}=\sup_{t-s\ge\delta, u\in[0,1]} \big|\alpha_j^\star\big\{s,t,\xi_n(s,t,u)-\varepsilon_n(s,t,u)\big\}-\alpha_j^\star(s,t,u)\big|.$$

Now, it is easy to verify that $t-s\le\lambda_n(s,t)+1/n$, so that, for n sufficiently large, $t-s\ge\delta$ implies that $\lambda_n(s,t)\ge\delta/2$. Then, from (B.4) and the fact that $\xi_n(\cdot,\cdot,0)=0$, we immediately have that, for n sufficiently large,

$$a_n=\sup_{\substack{t-s\ge\delta\\ u\in[0,1]}} \big|\xi_n(s,t,u)-u\big| \le \sup_{\substack{t-s\ge\delta\\ u\in[0,1]}} \big|\lambda_n(s,t)\big\{\xi_n(s,t,u)-u\big\}\big| \times \sup_{t-s\ge\delta} \lambda_n(s,t)^{-1} \to 0.$$

Hence, we can write

$$B_{n,4} \leq \sup_{\substack{t-s\geq\delta,u,u'\in[0,1]\\|u'-u|\leq a_n+n^{-1}}} \left|\alpha_j^\star\left(s,t,u'\right) - \alpha_j^\star(s,t,u)\right| \to 0$$

since $\alpha_j^\star$ is uniformly continuous on $\Delta \times [0,1]$. Proceeding as for B_n, it can be verified that $A_n \to 0$, which completes the proof. □

Proof of Lemma B.1. Starting from the definitions of g_n and Φ_n given in (B.2) and (B.1), respectively, we have the decomposition

$$g_n\left(\alpha_n^\star\right)(s,t,\mathbf{u}) = A_{n,1}(s,t,\mathbf{u}) + A_{n,2}(s,t,\mathbf{u}),$$

where

$$A_{n,1}(s,t,\mathbf{u}) = \alpha_n^\star\left\{s,t,I_n\left(U_n^\star + n^{-1/2}\alpha_{n,1}^\star\right)(s,t,u_1),\ldots,I_n\left(U_n^\star + n^{-1/2}\alpha_{n,d}^\star\right)(s,t,u_d)\right\},$$

and

$$\begin{aligned}&A_{n,2}(s,t,\mathbf{u})\\&= \sqrt{n}\lambda_n(s,t)\left[C\left\{I_n\left(U_n^\star + n^{-1/2}\alpha_{n,1}^\star\right)(s,t,u_1),\ldots,I_n\left(U_n^\star + n^{-1/2}\alpha_{n,d}^\star\right)(s,t,u_d)\right\} - C(\mathbf{u})\right].\end{aligned}$$

We begin the proof by showing that $\sup_{(s,t,\mathbf{u})\in\Delta\times[0,1]^d} |A_{n,1}(s,t,\mathbf{u}) - \alpha^\star(s,t,\mathbf{u})| \to 0$. Let $\varepsilon > 0$. Using the fact that $\alpha^\star \in \mathcal{D}_0^\star$, there exists $\delta > 0$ such that $|\alpha^\star(s,t,\mathbf{u})| \leq \varepsilon$ for all $t-s<\delta$ and $\mathbf{u} \in [0,1]^d$. Then, we write

$$\sup_{(s,t,\mathbf{u})\in\Delta\times[0,1]^d} \left|A_{n,1}(s,t,\mathbf{u}) - \alpha^\star(s,t,\mathbf{u})\right| \leq B_{n,1} + B_{n,2} + B_{n,3},$$

where

$$\begin{aligned}B_{n,1} &= \sup_{(s,t,\mathbf{u})\in\Delta\times[0,1]^d} \Big|A_{n,1}(s,t,\mathbf{u}) - \alpha^\star\Big\{s,t,I_n\left(U_n^\star + n^{-1/2}\alpha_{n,1}^\star\right)(s,t,u_1),\ldots,\\&\qquad\qquad I_n\left(U_n^\star + n^{-1/2}\alpha_{n,d}^\star\right)(s,t,u_d)\Big\}\Big|\\&\leq \sup_{(s,t,\mathbf{u})\in\Delta\times[0,1]^d} \left|\alpha_n^\star(s,t,\mathbf{u}) - \alpha^\star(s,t,\mathbf{u})\right| \leq \varepsilon,\end{aligned}$$

for sufficiently large n, where

$$\begin{aligned}B_{n,2} = \sup_{\substack{t-s<\delta\\\mathbf{u}\in[0,1]^d}} \Big|&\alpha^\star\left\{s,t,I_n\left(U_n^\star + n^{-1/2}\alpha_{n,1}^\star\right)(s,t,u_1),\ldots,I_n\left(U_n^\star + n^{-1/2}\alpha_{n,d}^\star\right)(s,t,u_d)\right\}\\&- \alpha^\star(s,t,\mathbf{u})\Big|,\end{aligned}$$

and

$$B_{n,3} = \sup_{\substack{t-s\geq\delta \\ \mathbf{u}\in[0,1]^d}} \bigl|\alpha^\star\bigl\{s,t,I_n\bigl(U_n^\star + n^{-1/2}\alpha_{n,1}^\star\bigr)(s,t,u_1),\ldots,I_n\bigl(U_n^\star + n^{-1/2}\alpha_{n,d}^\star\bigr)(s,t,u_d)\bigr\} - \alpha^\star(s,t,\mathbf{u})\bigr|.$$

For $B_{n,2}$, using the triangle inequality, we have that

$$B_{n,2} \leq 2 \sup_{t-s<\delta,\mathbf{u}\in[0,1]^d} \bigl|\alpha^\star(s,t,\mathbf{u})\bigr| \leq 2\varepsilon.$$

For $B_{n,3}$, we use the fact that Lemma B.2 implies that, for any $j \in \{1,\ldots,d\}$,

$$a_{n,j} = \sup_{t-s\geq\delta,u\in[0,1]} \bigl|I_n\bigl(U_n^\star + n^{-1/2}\alpha_{n,j}^\star\bigr)(s,t,u) - u\bigr| \to 0, \tag{B.5}$$

and the fact that

$$B_{n,3} \leq \sup_{t-s\geq\delta,|u_1-v_1|\leq a_{n,1},\ldots,|u_d-v_d|\leq a_{n,d}} \bigl|\alpha^\star(s,t,\mathbf{u}) - \alpha^\star(s,t,\mathbf{v})\bigr|.$$

By uniform continuity of $\alpha^\star$, for sufficiently large n, we obtain that $B_{n,3} \leq \varepsilon$. Hence, we have shown that, for sufficiently large n, $\sup_{(s,t,\mathbf{u})\in\Delta\times[0,1]^d} |A_{n,1}(s,t,\mathbf{u}) - \alpha^\star(s,t,\mathbf{u})| \leq 4\varepsilon$, and therefore that $\sup_{(s,t,\mathbf{u})\in\Delta\times[0,1]^d} |A_{n,1}(s,t,\mathbf{u}) - \alpha^\star(s,t,\mathbf{u})| \to 0$.

Let us now deal with $A_{n,2}$. Fix $n \geq 1$ and $s < t$ such that $\lfloor ns\rfloor < \lfloor nt\rfloor$. For any $\mathbf{u} \in [0,1]^d$, $j \in \{1,\ldots,d\}$ and $r \in [0,1]$, let $\bar{u}_j(r) = u_j + r\{I_n(U_n^\star + n^{-1/2}\alpha_{n,j}^\star)(s,t,u_j) - u_j\}$ and define $\bar{\mathbf{u}}(r) = (\bar{u}_1(r),\ldots,\bar{u}_d(r))$. Now, fix $\mathbf{u} \in (0,1)^d$ and let f be the function defined by

$$f(r) = C_n^\star\bigl\{s,t,\bar{\mathbf{u}}(r)\bigr\} = \lambda_n(s,t)C\bigl\{\bar{\mathbf{u}}(r)\bigr\}.$$

Obviously, we have that $0 < \bar{u}_j(r) < 1$ for all $r \in (0,1)$ and $j \in \{1,\ldots,d\}$. Therefore, the function f is continuous on $[0,1]$, and, by Condition 3.2, is differentiable on $(0,1)$. Hence, by the mean value theorem, there exists $r^* \in (0,1)$ such that $f(1) - f(0) = f'(r^*)$, which implies that

$$A_{n,2}(s,t,\mathbf{u}) = \sum_{j=1}^d \dot{C}_j\bigl\{\bar{\mathbf{u}}\bigl(r^*\bigr)\bigr\}\lambda_n(s,t)\sqrt{n}\bigl\{I_n\bigl(U_n^\star + n^{-1/2}\alpha_{n,j}^\star\bigr)(s,t,u_j) - u_j\bigr\}. \tag{B.6}$$

The previous equality remains clearly valid when $\lfloor ns\rfloor = \lfloor nt\rfloor$. Let us now verify that it also holds when $\lfloor ns\rfloor < \lfloor nt\rfloor$ and $\mathbf{u}$ is on the boundary of $[0,1]^d$. When $u_j = 0$ for some $j \in \{1,\ldots,d\}$, $I_n(U_n^\star + n^{-1/2}\alpha_{n,j}^\star)(\cdot,\cdot,u_j) = 0$, which implies that $\bar{u}_j(r) = 0$ for all $r \in [0,1]$. It then immediately follows that the left-hand side of (B.6) is zero and that the jth term in the sum on the right is zero. The $d-1$ remaining terms in the sum on the right of (B.6) are actually also zero because, for any $k \in \{1,\ldots,d\}$, $k \neq j$, $\dot{C}_k(\mathbf{v}) = 0$ for all $\mathbf{v} \in [0,1]^d$ such that $v_k = 0$. Hence, (B.6) remains true whenever $u_j = 0$ for some $j \in \{1,\ldots,d\}$.

Let us now assume that $\lfloor ns \rfloor < \lfloor nt \rfloor$ and that $u_j = 1$ for some $j \in \{1, \dots, d\}$. Two cases can be distinguished according to whether $I_n(U_n^\star + n^{-1/2}\alpha_{n,j}^\star)(s,t,1) = 1$ or $I_n(U_n^\star + n^{-1/2}\alpha_{n,j}^\star)(s,t,1) < 1$. In the later case, $0 < \bar{u}_j(r) < 1$. In the former case, we obtain that $\bar{u}_j(r) = 1$ for all $r \in [0,1]$ and that the jth term in the sum on the right of (B.6) is zero so that neither the left nor the right-hand side of (B.6) depend on u_j anymore. It follows that, when some components of $\mathbf{u}$ are one, the previous equality can be recovered by an application of the mean value theorem similar to the one carried out above.

Now, we write

$$A_{n,2}(s,t,\mathbf{u}) = \sum_{j=1}^d \dot{C}_j(\mathbf{u})\lambda_n(s,t)\sqrt{n}\bigl\{I_n\bigl(U_n^\star + n^{-1/2}\alpha_{n,j}^\star\bigr)(s,t,u_j) - u_j\bigr\} + r_n(s,t,\mathbf{u}), \tag{B.7}$$

where $r_n(s,t,\mathbf{u}) = \sum_{j=1}^d r_{n,j}(s,t,\mathbf{u})$ and, for any $j \in \{1,\dots,d\}$,

$$r_{n,j}(s,t,\mathbf{u}) = \bigl[\dot{C}_j\bigl\{\bar{\mathbf{u}}\bigl(r^*\bigr)\bigr\} - \dot{C}_j(\mathbf{u})\bigr]\lambda_n(s,t)\sqrt{n}\bigl\{I_n\bigl(U_n^\star + n^{-1/2}\alpha_{n,j}^\star\bigr)(s,t,u_j) - u_j\bigr\}.$$

By Lemma B.2 and from the fact that $0 \le \dot{C}_j \le 1$ for all $j \in \{1,\dots,d\}$, the dominating term in decomposition (B.7) converges to

$$-\sum_{j=1}^d \dot{C}_j(\mathbf{u})\alpha^\star\bigl(s,t,\mathbf{u}^{(j)}\bigr)$$

uniformly in $(s,t,\mathbf{u}) \in \Delta \times [0,1]^d$. It therefore remains to show that

$$\sup_{(s,t,\mathbf{u})\in\Delta\times[0,1]^d} \bigl|r_n(s,t,\mathbf{u})\bigr| \to 0.$$

Let us first show that $\sup_{(s,t,\mathbf{u})\in\Delta\times[0,1]^d} |r_{n,1}(s,t,\mathbf{u})| \to 0$. We have that

$$\sup_{(s,t,\mathbf{u})\in\Delta\times[0,1]^d} \bigl|r_{n,1}(s,t,\mathbf{u})\bigr| \le B_{n,4} + B_{n,5},$$

where

$$\begin{aligned} B_{n,4} = {} & \sup_{(s,t,\mathbf{u})\in\Delta\times[0,1]^d} \bigl|\dot{C}_1\bigl\{\bar{\mathbf{u}}\bigl(r^*\bigr)\bigr\} - \dot{C}_1(\mathbf{u})\bigr| \\ & \times \sup_{(s,t,\mathbf{u})\in\Delta\times[0,1]^d} \bigl|\lambda_n(s,t)\sqrt{n}\bigl\{I_n\bigl(U_n^\star + n^{-1/2}\alpha_{n,1}^\star\bigr)(s,t,u_1) - u_1\bigr\} + \alpha_1^\star(s,t,u_1)\bigr|, \end{aligned}$$

and

$$B_{n,5} = \sup_{(s,t,\mathbf{u})\in\Delta\times[0,1]^d} \bigl|\bigl[\dot{C}_1\bigl\{\bar{\mathbf{u}}\bigl(r^*\bigr)\bigr\} - \dot{C}_1(\mathbf{u})\bigr]\alpha_1^\star(s,t,u_1)\bigr|.$$

From the fact that $0 \le \dot{C}_1 \le 1$ and Lemma B.2, we immediately obtain that $B_{n,4} \to 0$. It remains to show that $B_{n,5} \to 0$. To this end, let $\varepsilon > 0$. Since $\alpha^\star \in \mathcal{D}_0^\star$, there exists $\delta > 0$ such that $|\alpha_1^\star(s,t,u)| \le \varepsilon$ for all $t - s < \delta$ and all $u \in [0,1]$. Then, $B_{n,5} \le B_{n,6} + B_{n,7}$, where

$$B_{n,6} = \sup_{(s,t,\mathbf{u}) \in \Delta \times [0,1]^d} \left| \dot{C}_1 \{ \bar{\mathbf{u}}(r^*) \} - \dot{C}_1(\mathbf{u}) \right| \times \sup_{t-s<\delta, u \in [0,1]} \left| \alpha_1^\star(s,t,u) \right| \le 2\varepsilon,$$

and

$$B_{n,7} = \sup_{t-s \ge \delta, \mathbf{u} \in [0,1]^d} \left| \left[\dot{C}_1 \{ \bar{\mathbf{u}}(r^*) \} - \dot{C}_1(\mathbf{u}) \right] \alpha_1^\star(s,t,u_1) \right|.$$

For $B_{n,7}$, we use the fact that, since $\alpha^\star \in \mathcal{D}_0^\star$, there exists $0 < \kappa < 1/2$ such that

$$\sup_{t-s \ge \delta, u \in [0,\kappa) \cup (1-\kappa,1]} \left| \alpha_1^\star(s,t,u) \right| \le \varepsilon.$$

Then, we write $B_{n,7} \le B_{n,8} + B_{n,9}$, where

$$B_{n,8} = \sup_{(s,t,\mathbf{u}) \in \Delta \times [0,1]^d} \left| \dot{C}_1 \{ \bar{\mathbf{u}}(r^*) \} - \dot{C}_1(\mathbf{u}) \right| \times \sup_{\substack{t-s \ge \delta, \mathbf{u} \in [0,1]^d \\ u_1 \in [0,\kappa) \cup (1-\kappa,1]}} \left| \alpha_1^\star(s,t,u_1) \right| \le 2\varepsilon,$$

and

$$B_{n,9} = \sup_{t-s \ge \delta, \mathbf{u} \in [0,1]^d, u_1 \in [\kappa, 1-\kappa]} \left| \dot{C}_1 \{ \bar{\mathbf{u}}(r^*) \} - \dot{C}_1(\mathbf{u}) \right| \times \sup_{(s,t,u) \in \Delta \times [0,1]} \left| \alpha_1^\star(s,t,u) \right|.$$

From (B.5), we obtain that

$$B_{n,9} \le \sup_{\substack{\mathbf{u},\mathbf{v} \in [0,1]^d, u_1, v_1 \in [\kappa/2, 1-\kappa/2] \\ |u_1 - v_1| \le a_{n,1}, \ldots, |u_d - v_d| \le a_{n,d}}} \left| \dot{C}_1(\mathbf{u}) - \dot{C}_1(\mathbf{v}) \right| \times \sup_{(s,t,u) \in \Delta \times [0,1]} \left| \alpha_1^\star(s,t,u) \right|.$$

Since $\dot{C}_1$ is uniformly continuous on $[\kappa/2, 1-\kappa/2] \times [0,1]^{d-1}$ according to Condition 3.2, and since $\sup_{(s,t,u) \in \Delta \times [0,1]} |\alpha_1^\star(s,t,u)|$ is bounded, we have that $B_{n,9} \to 0$, which implies that, for n sufficiently large, $B_{n,9} \le \varepsilon$. It follows that, for n sufficiently large, $B_{n,5} \le 5\varepsilon$, which implies that $\sup_{(s,t,\mathbf{u}) \in \Delta \times [0,1]^d} |r_{n,1}(s,t,\mathbf{u})| \to 0$. One can proceed similarly for $r_{n,j}$, $j \in \{2, \ldots, d\}$. Hence, $\sup_{s \le t, \mathbf{u} \in [0,1]^d} |r_n(s,t,\mathbf{u})| \to 0$. $\square$

Acknowledgements

The authors would like to thank two anonymous referees and an Associate Editor for their constructive comments on an earlier version of this manuscript, as well as Johan Segers for fruitful discussions as usual.

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.

Supplementary Material

Supplement to "A dependent multiplier bootstrap for the sequential empirical copula process under strong mixing" (DOI: 10.3150/14-BEJ682SUPP; .pdf). Additional proofs and simulation results can be found in (Bücher and Kojadinovic [9]).

References

[1] Agarwal, G.G., Dalpatadu, R.J. and Singh, A.K. (2002). Linear functions of uniform order statistics and B-splines. *Comm. Statist. Theory Methods* **31** 181–192. MR1894822

[2] Andrews, D.W.K. (1991). Heteroskedasticity and autocorrelation consistent covariance matrix estimation. *Econometrica* **59** 817–858. MR1106513

[3] Auestad, B. and Tjøstheim, D. (1990). Identification of nonlinear time series: First order characterization and order determination. *Biometrika* **77** 669–687. MR1086681

[4] Balacheff, S. and Dupont, G. (1980). Normalité asymptotique des processus empiriques tronqués et des processus de rang (cas multidimensionnel mélangeant). In *Nonparametric Asymptotic Statistics* (*Proc. Conf., Rouen*, 1979) (*French*). *Lecture Notes in Math.* **821** 19–45. Berlin: Springer. MR0604019

[5] Beare, B.K. (2010). Copulas and temporal dependence. *Econometrica* **78** 395–410. MR2642867

[6] Bickel, P.J. and Wichura, M.J. (1971). Convergence criteria for multiparameter stochastic processes and some applications. *Ann. Math. Statist.* **42** 1656–1670. MR0383482

[7] Bücher, A. (2014). A note on weak convergence of the sequential multivariate empirical process under strong mixing. *J. Theoret. Probab.* DOI:10.1007/s10959-013-0529-5.

[8] Bücher, A. and Dette, H. (2010). A note on bootstrap approximations for the empirical copula process. *Statist. Probab. Lett.* **80** 1925–1932. MR2734261

[9] Bücher, A. and Kojadinovic, I. (2014). Supplement to "A dependent multiplier bootstrap for the sequential empirical copula process under strong mixing." DOI:10.3150/14-BEJ682SUPP.

[10] Bücher, A., Kojadinovic, I., Rohmer, T. and Segers, J. (2014). Detecting changes in cross-sectional dependence in multivariate time series. *J. Multivariate Anal.* **132** 111–128. MR3266263

[11] Bücher, A. and Ruppert, M. (2013). Consistent testing for a constant copula under strong mixing based on the tapered block multiplier technique. *J. Multivariate Anal.* **116** 208–229. MR3049901

[12] Bücher, A. and Segers, J. (2014). Extreme value copula estimation based on block maxima of a multivariate stationary time series. *Extremes* **17** 495–528. MR3252823

[13] Bücher, A. and Volgushev, S. (2013). Empirical and sequential empirical copula processes under serial dependence. *J. Multivariate Anal.* **119** 61–70. MR3061415

[14] Bühlmann, P. (1994). Blockwise bootstrapped empirical process for stationary sequences. *Ann. Statist.* **22** 995–1012. MR1292553

[15] Bühlmann, P.L. (1993). The blockwise bootstrap in time series and empirical processes. Ph.D. thesis, ETH Zürich. MR2714738

[16] Chen, X. and Fan, Y. (1999). Consistent hypothesis testing in semiparametric and nonparametric models for econometric time series. *J. Econometrics* **91** 373–401. MR1703951

[17] Csörgő, M. and Horváth, L. (1997). *Limit Theorems in Change-Point Analysis*. *Wiley Series in Probability and Statistics*. Chichester: Wiley. MR2743035

[18] Darsow, W.F., Nguyen, B. and Olsen, E.T. (1992). Copulas and Markov processes. *Illinois J. Math.* **36** 600–642. MR1215798

[19] Deheuvels, P. (1981). A non parametric test for independence. *Publications de l'Institut de Statistique de l'Université de Paris* **26** 29–50.

[20] Dehling, H. and Philipp, W. (2002). Empirical process techniques for dependent data. In *Empirical Process Techniques for Dependent Data* (H. Dehling, T. Mikosch and M. Sorensen, eds.) 3–113. Boston, MA: Birkhäuser. MR1958777

[21] Fermanian, J.-D., Radulović, D. and Wegkamp, M. (2004). Weak convergence of empirical copula processes. *Bernoulli* **10** 847–860. MR2093613

[22] Gaenssler, P. and Stute, W. (1987). *Seminar on Empirical Processes*. *DMV Seminar* **9**. Basel: Birkhäuser. MR0902803

[23] Genest, C., Rémillard, B. and Beaudoin, D. (2009). Goodness-of-fit tests for copulas: A review and a power study. *Insurance Math*. *Econom*. **44** 199–213. MR2517885

[24] Holmes, M., Kojadinovic, I. and Quessy, J.-F. (2013). Nonparametric tests for change-point detection à la Gombay and Horváth. *J*. *Multivariate Anal*. **115** 16–32. MR3004542

[25] Inoue, A. (2001). Testing for distributional change in time series. *Econometric Theory* **17** 156–187. MR1863569

[26] Jondeau, E., Poon, S.-H. and Rockinger, M. (2007). *Financial Modeling Under Non-Gaussian Distributions*. *Springer Finance*. London: Springer. MR2260802

[27] Kojadinovic, I., Segers, J. and Yan, J. (2011). Large-sample tests of extreme-value dependence for multivariate copulas. *Canad*. *J*. *Statist*. **39** 703–720. MR2860835

[28] Kojadinovic, I. and Yan, J. (2012). Goodness-of-fit testing based on a weighted bootstrap: A fast large-sample alternative to the parametric bootstrap. *Canad*. *J*. *Statist*. **40** 480–500. MR2968394

[29] Kosorok, M.R. (2008). *Introduction to Empirical Processes and Semiparametric Inference*. *Springer Series in Statistics*. New York: Springer. MR2724368

[30] Künsch, H.R. (1989). The jackknife and the bootstrap for general stationary observations. *Ann*. *Statist*. **17** 1217–1241. MR1015147

[31] McNeil, A.J., Frey, R. and Embrechts, P. (2005). *Quantitative Risk Management*. *Princeton Series in Finance*. Princeton, NJ: Princeton Univ. Press. MR2175089

[32] Mokkadem, A. (1988). Mixing properties of ARMA processes. *Stochastic Process*. *Appl*. **29** 309–315. MR0958507

[33] Paparoditis, E. and Politis, D.N. (2001). Tapered block bootstrap. *Biometrika* **88** 1105–1119. MR1872222

[34] Patton, A., Politis, D.N. and White, H. (2009). Correction to "Automatic block-length selection for the dependent bootstrap" by D. Politis and H. White [MR2041534]. *Econometric Rev*. **28** 372–375. MR2532278

[35] Patton, A.J. (2012). Copula methods for forecasting multivariate time series. In *Handbook of Economic Forecasting*, *Vol*. 2. Amsterdam: North-Holland.

[36] Politis, D.N. (2003). Adaptive bandwidth choice. *J*. *Nonparametr. Stat*. **15** 517–533. MR2017485

[37] Politis, D.N. and Romano, J.P. (1995). Bias-corrected nonparametric spectral estimation. *J*. *Time Series Anal*. **16** 67–103. MR1323618

[38] Politis, D.N. and White, H. (2004). Automatic block-length selection for the dependent bootstrap. *Econometric Rev*. **23** 53–70. MR2041534

[39] Rémillard, B. and Scaillet, O. (2009). Testing for equality between two copulas. *J*. *Multivariate Anal*. **100** 377–386. MR2483426

[40] Rüschendorf, L. (1976). Asymptotic distributions of multivariate rank order statistics. *Ann*. *Statist*. **4** 912–923. MR0420794

[41] Salvadori, G., De Michele, C., Kottegoda, N.T. and Rosso, R. (2007). *Extremes in Nature*: *An Approach Using Copulas*. *Water Science and Technology Library* **56**. Berlin: Springer.

[42] Scaillet, O. (2005). A Kolmogorov–Smirnov type test for positive quadrant dependence. *Canad*. *J*. *Statist*. **33** 415–427. MR2193983

[43] Segers, J. (2012). Asymptotics of empirical copula processes under non-restrictive smoothness assumptions. *Bernoulli* **18** 764–782. MR2948900

[44] Shao, X. (2010). The dependent wild bootstrap. *J. Amer. Statist. Assoc.* **105** 218–235. MR2656050

[45] Sklar, M. (1959). Fonctions de répartition à n dimensions et leurs marges. *Publ. Inst. Stat. Univ. Paris* **8** 229–231. MR0125600

[46] Tsukahara, H. (2005). Semiparametric estimation in copula models. *Canad. J. Statist.* **33** 357–375. MR2193980

[47] van der Vaart, A.W. (1998). *Asymptotic Statistics*. *Cambridge Series in Statistical and Probabilistic Mathematics* **3**. Cambridge: Cambridge Univ. Press. MR1652247

[48] van der Vaart, A.W. and Wellner, J.A. (2000). *Weak Convergence and Empirical Processes*, 2nd ed. New York: Springer.

[49] van der Vaart, A.W. and Wellner, J.A. (2007). Empirical processes indexed by estimated functions. In *Asymptotics*: *Particles*, *Processes and Inverse Problems*. *Institute of Mathematical Statistics Lecture Notes—Monograph Series* **55** 234–252. Beachwood, OH: IMS. MR2459942

Received June 2013 and revised October 2014