SFB 823

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

Axel Bücher, Ivan Kojadinovic

Nr. 25/2013

Discussion Paper

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

Axel Bücher*

Université catholique de Louvain & Ruhr-Universität Bochum

Ivan Kojadinovic†

Université de Pau et des Pays de l'Adour

June 19, 2013

Abstract

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 (2009) frequently appears to lead to inference procedures with the best finite-sample properties. Bücher and Ruppert (2013) recently proposed an extension of this technique to strictly stationary strongly mixing observations by adapting the *dependent* multiplier bootstrap of Bühlmann (1993, 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 (2013) along two directions. First, the resampling scheme is now genuinely sequential, thereby allowing to transpose to the strongly mixing setting all 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 (block length) 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 weak convergence of the sequential empirical copula process is obtained under many serial dependence conditions, and 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.

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

*Université catholique de Louvain, Institut de statistique, Voie du Roman Pays 20, 1348 Louvain-la-Neuve, Belgium. E-mail: axel.buecher@rub.de

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

1 Introduction

Let $\boldsymbol{X}$ be a d-dimensional random vector with continuous marginal cumulative distribution functions (c.d.f.s) $F_1, \ldots, F_d$. From the work of Sklar (1959), the c.d.f. F of $\boldsymbol{X}$ can be written in a unique way as

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

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 $\boldsymbol{X}$. 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 et al., 2005), econometric modeling (Patton, 2012), environmental modeling (Salvadori et al., 2007), to name a very few.

Assume that C and $F_1, \ldots, F_d$ are unknown and let $\boldsymbol{X}_1, \ldots, \boldsymbol{X}_n$ be drawn from a strictly stationary sequence of continuous d-dimensional random vectors with c.d.f. F. For any $i \in \{1, \ldots, n\}$ and $j \in \{1, \ldots, d\}$, denote by $R_{ij}^{1:n}$ the rank of X_{ij} among $X_{1j}, \ldots, X_{nj}$ and let $\hat{U}_{ij}^{1:n} = R_{ij}^{1:n}/n$. The random vectors $\hat{\boldsymbol{U}}_i^{1:n} = (\hat{U}_{i1}^{1:n}, \ldots, \hat{U}_{id}^{1:n})$, $i \in \{1, \ldots, n\}$, are often referred to as *pseudo-observations* from the copula C, and a natural nonparametric estimator of C is the *empirical copula* of $\boldsymbol{X}_1, \ldots, \boldsymbol{X}_n$ (Rüschendorf, 1976; Deheuvels, 1981), frequently defined as the empirical c.d.f. computed from the pseudo-observations, i.e.,

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

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 (2005), Rémillard and Scaillet (2009), and Genest et al. (2009), 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 $(s,t) \in \Delta = \{(s,t) \in [0,1]^2 : s \leq t\}$ and $\boldsymbol{u} \in [0,1]^d$, by

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

where, for any $y \geq 0$, $\lfloor y \rfloor$ 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 $(X_{\lfloor ns \rfloor+1}, Y_{\lfloor ns \rfloor+1}), \ldots, (X_{\lfloor nt \rfloor}, Y_{\lfloor nt \rfloor})$ as

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

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

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, 1976; Gänssler and Stute, 1987;

Tsukahara, 2005; van der Vaart and Wellner, 2007; Segers, 2012; Bücher and Volgushev, 2013) while the quantity $\mathbb{C}_n(0,\cdot,\cdot,\cdot)$ is a sequential version of the latter investigated in Kojadinovic and Rohmer (2012) in the case of i.i.d. observations with nonparametric tests for change-point detection in mind. Notice that the process $\mathbb{C}_n(0,\cdot,\cdot,\cdot)$ does not coincide with the sequential process initially studied by Rüschendorf (1976) and defined by

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

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 (2013) 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). 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 et al. (2004) to the multiplier technique of Rémillard and Scaillet (2009). Their finite-sample properties were compared in Bücher and Dette (2010) who concluded that the multiplier bootstrap of Rémillard and Scaillet (2009) has, overall, the best finite-sample behavior. In the case of strongly mixing observations, Bücher and Ruppert (2013) recently proposed a similar resampling scheme by adapting the *dependent* multiplier bootstrap of Bühlmann (1993, Section 3.3) to the process $\mathbb{C}_n(0,1,\cdot,\cdot)$. Their empirical investigations indicate that the latter outperforms in finite samples a block bootstrap for $\mathbb{C}_n(0,1,\cdot,\cdot)$ based on the work of Künsch (1989) and Bühlmann (1994). Note that the dependent multiplier bootstrap of Bühlmann (1993, Section 3.3) was recently independently rediscovered by Shao (2010) in the context of the *smooth function model* but not in the empirical process setting. For the sample mean as statistic of interest, the author connected this resampling technique to the *tapered block bootstrap* of Paparoditis and Politis (2001).

The main aim of this work is to provide an extended version of the multiplier resampling scheme of Bücher and Ruppert (2013) adapted to the two-sided sequential process $\mathbb{C}_n$ defined in (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 analogue to that of the block length in the block bootstrap. As a practical consequence, the resulting dependent multiplier for $\mathbb{C}_n$ can be used in a fully automatic way and all 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, including the nonparametric tests for change-point detection investigated in Kojadinovic and Rohmer (2012).

There are two important by-products of this work that can be of independent interest. First, the weak convergence of the two-sided sequential empirical copula process $\mathbb{C}_n$ is established under many serial dependence scenarios as a consequence of the weak convergence of the multivariate sequential empirical process. The weak convergence of $\mathbb{C}_n$

under strong mixing is a mere corollary of this very general result. Second, the validity of a sequential version of the dependent multiplier bootstrap for empirical processes of Bühlmann (1993, Section 3.3) (which has also been considered in Bücher and Ruppert, 2013, 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.

The paper is organized as follows. In the second section, the weak convergence of the two-sided sequential empirical copula process $\mathbb{C}_n$ is established for many serial dependence conditions. The third section presents a sequential extension of the seminal work of Bühlmann (1993, Section 3.3). Based on this generalization, a dependent multiplier bootstrap for $\mathbb{C}_n$ is derived in the next section. In the fifth section, 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 (2004), among others. The sixth section discusses two ways to generate dependent multiplier sequences which are central to this resampling technique. 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 (2000), 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 Weak convergence of the sequential empirical copula process under serial dependence

Let $\boldsymbol{U}_1, \dots, \boldsymbol{U}_n$ be the unobservable sample obtained from $\boldsymbol{X}_1, \dots, \boldsymbol{X}_n$ by the probability integral transforms $U_{ij} = F_j(X_{ij})$, $i \in \{1, \dots, n\}$, $j \in \{1, \dots, d\}$. It follows that $\boldsymbol{U}_1, \dots, \boldsymbol{U}_n$ is a marginally uniform d-dimensional sample from the unknown c.d.f. C. The corresponding sequential empirical process is then defined as

$$\tilde{\mathbb{B}}_n(s, \boldsymbol{u}) = \frac{1}{\sqrt{n}} \sum_{i=1}^{\lfloor ns \rfloor} \{\mathbf{1}(\boldsymbol{U}_i \leq \boldsymbol{u}) - C(\boldsymbol{u})\}, \qquad (s, \boldsymbol{u}) \in [0,1]^{d+1}. \tag{2}$$

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

We shall derive the weak limit of the two-sided empirical process $\mathbb{C}_n$ defined in (1) under the following condition. In the case of i.i.d. observations, it is an immediate consequence of Theorem 2.12.1 in van der Vaart and Wellner (2000).

Condition 2.1. *The sample $\boldsymbol{U}_1, \dots, \boldsymbol{U}_n$ is drawn from a strictly stationary sequence $(\boldsymbol{U}_i)_{i \in \mathbb{Z}}$ 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*

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

We shall also consider the following smoothness condition on C proposed by Segers (2012). 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 and have continuous sample paths.

Condition 2.2. *For any $j \in \{1,\dots,d\}$, the partial derivatives $\dot{C}_j = \partial C/\partial u_j$ exist and are continuous on $\{\boldsymbol{u} \in [0,1]^d : u_j \in (0,1)\}$.*

As we continue, for any $j \in \{1,\dots,d\}$, we define $\dot{C}_j$ to be zero on the set $\{\boldsymbol{u} \in [0,1]^d : u_j \in \{0,1\}\}$ (see also Segers, 2012; Bücher and Volgushev, 2013). It then follows that, under Condition 2.2, $\dot{C}_j$ is defined on the whole of $[0,1]^d$. Also, for any $j \in \{1,\dots,d\}$ and any $\boldsymbol{u} \in [0,1]^d$, $\boldsymbol{u}^{(j)}$ is the vector of $[0,1]^d$ defined by $u_i^{(j)} = u_j$ if $i = j$ and 1 otherwise.

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

Theorem 2.3 (Weak convergence of the sequential empirical copula process)**.** *Assume that the sample $\boldsymbol{U}_1,\dots,\boldsymbol{U}_n$ satisfies Condition 2.1 and that C satisfies Condition 2.2. Then,*

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

where

$$\tilde{\mathbb{C}}_n(s,t,\boldsymbol{u}) = \{\tilde{\mathbb{B}}_n(t,\boldsymbol{u}) - \tilde{\mathbb{B}}_n(s,\boldsymbol{u})\} - \sum_{j=1}^d \dot{C}_j(\boldsymbol{u})\{\tilde{\mathbb{B}}_n(t,\boldsymbol{u}^{(j)}) - \tilde{\mathbb{B}}_n(s,\boldsymbol{u}^{(j)})\}. \tag{3}$$

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

$$\mathbb{C}_C(s,t,\boldsymbol{u}) = \{\mathbb{B}_C(t,\boldsymbol{u}) - \mathbb{B}_C(s,\boldsymbol{u})\} - \sum_{j=1}^d \dot{C}_j(\boldsymbol{u})\{\mathbb{B}_C(t,\boldsymbol{u}^{(j)}) - \mathbb{B}_C(s,\boldsymbol{u}^{(j)})\}. \tag{4}$$

The weak convergence of $\mathbb{C}_n$ under strong mixing immediately follows from the previous theorem. For a sequence of d-dimensional random vectors $(\boldsymbol{Y}_i)_{i\in\mathbb{Z}}$, the σ-field generated by $(Y_i)_{a\le i\le b}$, $a,b \in \mathbb{Z}\cup\{-\infty,+\infty\}$, is denoted by $\mathcal{F}_a^b$. The strong mixing coefficients corresponding to the sequence $(\boldsymbol{Y}_i)_{i\in\mathbb{Z}}$ are then defined by $\alpha_0 = 1/2$ and

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

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

The following corollary is an immediate consequence of the strong approximation result of Dhompongsa (1984), which implies that, if $\boldsymbol{U}_1,\dots,\boldsymbol{U}_n$ is drawn from a strictly stationary sequence $(\boldsymbol{U}_i)_{i\in\mathbb{Z}}$ whose strong mixing coefficients satisfy $\alpha_r = O(r^{-a})$, $a > 2 + d$, then $\tilde{\mathbb{B}}_n \rightsquigarrow \mathbb{B}_C$ in $\ell^\infty([0,1]^{d+1})$, that is, $\boldsymbol{U}_1,\dots,\boldsymbol{U}_n$ satisfies Condition 2.1.

Corollary 2.4. *Assume that $\boldsymbol{X}_1, \dots, \boldsymbol{X}_n$ is drawn from a strictly stationary sequence $(\boldsymbol{X}_i)_{i\in\mathbb{Z}}$ whose strong mixing coefficients satisfy $\alpha_r = O(r^{-a})$, $a > 2 + d$. Then, provided C satisfies Condition 2.2,*

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

where $\tilde{\mathbb{C}}_n$ is defined in (3).

The conditions of the above corollary are for instance satisfied (with much to spare) when $\boldsymbol{X}_1, \dots, \boldsymbol{X}_n$ is drawn from a stationary ARMA process (see e.g. Mokkadem, 1988).

3 A multiplier central limit theorem under strong mixing

The multiplier bootstrap of Rémillard and Scaillet (2009) 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, 2008, Theorem 10.1 and Corollary 10.3). A sequential version of the previous result can be proved (see Holmes et al., 2013, Theorem 1) by using the method of proof adopted in van der Vaart and Wellner (2000, Theorem 2.12.1). While investigating the block bootstrap for empirical processes constructed from strongly mixing observations, Bühlmann (1993, Section 3.3) obtained what resembles to a conditional version of the multiplier central limit theorem. The main idea of his approach 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:

(A1) 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 $\mathrm{E}(|\xi_{0,n}|^\nu) < \infty$ for all $\nu \geq 1$, and is independent of the available sample $\boldsymbol{X}_1, \dots, \boldsymbol{X}_n$.

(A2) There exists a sequence $\ell_n \to \infty$ of strictly positive constants such that $\ell_n = o(n)$ and the sequence $(\xi_{i,n})_{i\in\mathbb{Z}}$ is ℓ_n-dependent, i.e., $\xi_{i,n}$ is independent of $\xi_{i+h,n}$ for all $h > \ell_n$ and $i \in \mathbb{N}$.

(A3) 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 $h \in \mathbb{Z}$.

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 $(s, \boldsymbol{u}) \in [0,1]^{d+1}$, let

$$\tilde{\mathbb{B}}_n^{(m)}(s,\boldsymbol{u}) = \frac{1}{\sqrt{n}} \sum_{i=1}^{\lfloor ns \rfloor} \xi_{i,n}^{(m)} \{\mathbf{1}(\boldsymbol{U}_i \leq \boldsymbol{u}) - C(\boldsymbol{u})\}. \tag{5}$$

From the previous display, we see the bandwidth sequence ℓ_n defined in Assumption (A2) plays an analogue role 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 6.

The following result, inspired by Bühlmann (1993, Section 3.3), could be regarded as an extension of the multiplier central limit theorem to the sequential and strongly mixing case for empirical processes indexed by lower-left orthants. Its proof is given in Appendix B.

Theorem 3.1 (Dependent multiplier central limit theorem)**.** *Assume that* $\ell_n = O(n^{1/2-\varepsilon})$ *for some* $0 < \varepsilon < 1/2$ *and that* $\boldsymbol{U}_1, \dots, \boldsymbol{U}_n$ *is drawn from a strictly stationary sequence* $(\boldsymbol{U}_i)_{i\in\mathbb{Z}}$ *whose strong mixing coefficients satisfy* $\alpha_r = O(r^{-a})$*,* $a > 3 + 3d/2$*. 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), and* $\mathbb{B}_C^{(1)}, \dots, \mathbb{B}_C^{(M)}$ *are independent copies of* $\mathbb{B}_C$*.*

Before commenting on 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 (1993), and may be of interest for applications of empirical processes outside the scope of copulas. Recall that $\boldsymbol{X}_1, \dots, \boldsymbol{X}_n$ 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 $F_1, \dots, F_d$. Then, let

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

be the usual sequential empirical process and, for any $m \in \{1, \dots, M\}$, let

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

where $\overline{\mathbb{R}} = [-\infty, \infty]$ and F_n is the empirical c.d.f. computed from $\boldsymbol{X}_1, \dots, \boldsymbol{X}_n$. The following corollary is then a consequence of the fact that $\mathbb{Z}_n(s, \boldsymbol{x}) = \tilde{\mathbb{B}}_n\{s, F_1(x_1), \dots, F_d(x_d)\}$ for all $(s, \boldsymbol{x}) \in [0,1] \times \overline{\mathbb{R}}^d$ and that, under the conditions of Theorem 3.1, for all $m \in \{1, \dots, M\}$,

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

a proof of which follows from the proof of Lemma C.2.

Corollary 3.2 (Dependent multiplier bootstrap for $\mathbb{Z}_n$)**.** *Assume that* $\ell_n = O(n^{1/2-\varepsilon})$ *for some* $0 < \varepsilon < 1/2$ *and that* $\boldsymbol{X}_1, \dots, \boldsymbol{X}_n$ *is drawn from a strictly stationary sequence* $(\boldsymbol{X}_i)_{i\in\mathbb{Z}}$ *of continuous d-dimensional random vectors whose strong mixing coefficients satisfy* $\alpha_r = O(r^{-a})$*,* $a > 3 + 3d/2$*. Then,*

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

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* $\mathbb{Z}_F^{(1)},\dots,\mathbb{Z}_F^{(M)}$ *are independent copies of* $\mathbb{Z}_F$*.*

Remark. 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, 1998, 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, 1998, page 329). An appropriate unconditional weak convergence result of the form of the one established in Corollary 3.2 already includes the repetition of the random resampling mechanism and can be used directly to deduce "consistency" of a bootstrap procedure. It also allows for instance to show that related statistical tests hold their levels and are consistent, as $n\to\infty$ followed by $M\to\infty$. An additional advantage is that the usual workhorses for weak convergence in function spaces such as the functional delta method and the (extended) continuous mapping theorem can be directly applied, whereas the application of the conditional versions of the latter results appear to be more restrictive (Kosorok, 2008, Section 10.1.4).

From a practical perspective, Corollary 3.2 allows for instance to transpose to the strongly mixing setting the goodness-of-fit and nonparametric change-point tests considered in Kojadinovic and Yan (2012) and Holmes et al. (2013), respectively.

We end this section by a few comments on the assumptions of Theorem 3.1 and Corollary 3.2:

- The requirement that $\ell_n = O(n^{1/2-\varepsilon})$ for some $0<\varepsilon<1/2$ is used for proving the finite-dimensional convergence involved in Theorem 3.1, while the condition $\alpha_r = O(r^{-a})$, $a > 3+3d/2$, is needed for the proof of the asymptotic equicontinuity.
- Theorem 3.2 of Bühlmann (1993) can be regarded as a non sequential conditional analogue of Corollary 3.2 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 3.1 which remains comparable with that of Dhompongsa (1984) for the sequential empirical process.
- The condition on the strong mixing coefficients in Theorem 3.1 and Corollary 3.2 is clearly satisfied if $\boldsymbol{X}_1,\dots,\boldsymbol{X}_n$ 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 (2013) suggest that a simpler scheme with i.i.d. multipliers (based for instance on Theorem 1 of Holmes et al., 2013) might lead to better finite-sample performance.

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

By analogy with the approach adopted in Rémillard and Scaillet (2009) (see also Segers, 2012), the multiplier central limit theorem obtained in the previous section can be used to obtain a dependent multiplier bootstrap for $\mathbb{C}_n$ under strong mixing. The main result of this section can be regarded as an extension of Theorem 3 in Bücher and Ruppert (2013), where a similar but conditional weak convergence result has been established in the non-sequential setting and under stricter conditions on the mixing rate and the multipliers.

The underlying idea is as follows: Theorem 3.1 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 $\boldsymbol{U}_1, \dots, \boldsymbol{U}_n$ is unobservable. Estimating C by the empirical copula $C_{1:n}$ and $\boldsymbol{U}_1, \dots, \boldsymbol{U}_n$ by the pseudo-observations $\hat{\boldsymbol{U}}_1^{1:n} \dots, \hat{\boldsymbol{U}}_n^{1:n}$, we obtain the following computable version of $\tilde{\mathbb{B}}_n^{(m)}$ defined, for any $(s, \boldsymbol{u}) \in [0,1]^{d+1}$, by

$$\hat{\mathbb{B}}_n^{(m)}(s, \boldsymbol{u}) = \frac{1}{\sqrt{n}} \sum_{i=1}^{\lfloor ns \rfloor} \xi_{i,n}^{(m)} \{\mathbf{1}(\hat{\boldsymbol{U}}_i^{1:n} \leq \boldsymbol{u}) - C_{1:n}(\boldsymbol{u})\}. \tag{6}$$

To define "almost" independent copies of $\mathbb{C}_n$ for large n in the spirit of Rémillard and Scaillet (2009), we additionally need to estimate the partial derivatives $\dot{C}_j$, $j \in \{1, \dots, d\}$ appearing in (4). As we continue, we consider estimators $\dot{C}_{j,n}$ of $\dot{C}_j$ satisfying the following condition put forward in Segers (2012):

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

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

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 $m \in \{1, \dots, M\}$ and $(s, t, \boldsymbol{u}) \in \Delta \times [0,1]^d$, let

$$\hat{\mathbb{C}}_n^{(m)}(s, t, \boldsymbol{u}) = \{\hat{\mathbb{B}}_n^{(m)}(t, \boldsymbol{u}) - \hat{\mathbb{B}}_n^{(m)}(s, \boldsymbol{u})\} - \sum_{j=1}^{d} \dot{C}_{j,n}(\boldsymbol{u}) \{\hat{\mathbb{B}}_n^{(m)}(t, \boldsymbol{u}^{(j)}) - \hat{\mathbb{B}}_n^{(m)}(s, \boldsymbol{u}^{(j)})\}. \tag{7}$$

The following proposition is proved in Appendix C by adapting the arguments of Segers (2012) to the current sequential and strongly mixing setting.

Proposition 4.2 (Dependent multiplier bootstrap for $\mathbb{C}_n$)**.** *Assume that $\ell_n = O(n^{1/2-\varepsilon})$ for some $0 < \varepsilon < 1/2$ and that $\boldsymbol{X}_1, \dots, \boldsymbol{X}_n$ is drawn from a strictly stationary sequence $(\boldsymbol{X}_i)_{i \in \mathbb{Z}}$ whose strong mixing coefficients satisfy $\alpha_r = O(r^{-a})$, $a > 3 + 3d/2$. Then, under Conditions 2.2 and 4.1,*

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

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 (4), and* $\mathbb{C}_C^{(1)}, \dots, \mathbb{C}_C^{(M)}$ *are independent copies of* $\mathbb{C}_C$*.*

Let us now briefly illustrate how Proposition 4.2 would be typically used in the context of change-point detection. As discussed in Kojadinovic and Rohmer (2012), a broad class of nonparametric tests for change-point detection particularly sensitive to changes in the copula can be derived from the process

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

The above definition is a mere transposition to the copula context of the "classical construction" adopted for instance in Csörgő and Horváth (1997, Section 2.6). Under the null hypothesis of no change in the distribution, the process $\mathbb{D}_n$ can be simply rewritten as

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

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

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

$m \in \{1, \dots, M\}$, 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 and the continuous mapping theorem 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. The latter result is the key step for establishing that tests based on $\mathbb{D}_n$ hold their level asymptotically. A more involved use of Proposition 4.2 for change-point testing will be presented in a companion paper.

We end this section by briefly mentioning three possible choices for the estimators of the partial derivatives. Rémillard and Scaillet (2009) proposed to estimate the partial derivatives $\dot{C}_j$, $j \in \{1, \dots, d\}$, by finite-differences as

$$\begin{aligned}\dot{C}_{j,n}(\boldsymbol{u}) = \frac{1}{2n^{-1/2}}\,&\{C_n(u_1, \dots, u_{j-1}, u_j + n^{-1/2}, u_{j+1}, \dots, u_d) \\ &\quad - C_n(u_1, \dots, u_{j-1}, u_j - n^{-1/2}, u_{j+1}, \dots, u_d)\}, \qquad \boldsymbol{u} \in [0,1]^d. \quad (8)\end{aligned}$$

A slightly different definition consisting of a "boundary correction" was proposed in Kojadinovic et al. (2011, page 706). Yet another definition is mentioned in Bücher and Ruppert (2013, page 212). Note that, for any $\delta \in (0, 1/2)$, all three definitions coincide on the set $\{\boldsymbol{u} \in [0,1]^d : u_j \in [\delta, 1-\delta]\}$ provided n is taken large enough. Now, under the assumptions of Corollary 2.4, 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 et al. (2011), 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 et al. (2011, page 706) and Bücher and Ruppert (2013, page 212), respectively.

5 Estimation of the bandwidth parameter ℓ_n

The bandwidth parameter ℓ_n defined in Assumption (A2) plays a role similar to that of the block length in the block bootstrap of Künsch (1989). 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 (2001) for the tapered block bootstrap using results from Künsch (1989). Related results are presented in Bühlmann (1993, Lemmas 3.12 and 3.13) and Shao (2010, 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 (2001) and Politis and White (2004) for other resampling schemes. Since the dependent multiplier bootstrap for $\mathbb{C}_n$ is based on the corresponding bootstrap 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 $\boldsymbol{X}_1, \dots, \boldsymbol{X}_n$, and, for any $\boldsymbol{u}, \boldsymbol{v} \in [0,1]^d$, let $\sigma_C(\boldsymbol{u}, \boldsymbol{v}) = \mathrm{Cov}\{\mathbb{B}_C(1, \boldsymbol{u}), \mathbb{B}_C(1, \boldsymbol{v})\}$. Now, fix $m \in \{1, \dots, M\}$ and, for any $\boldsymbol{u}, \boldsymbol{v} \in [0,1]^d$, let

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

where $\tilde{\mathbb{B}}_n^{(m)}$ is defined in (5). For the moment, although it is based on the unobservable sample $\boldsymbol{U}_1, \dots, \boldsymbol{U}_n$ and the unknown copula C, we shall regard $\tilde{\sigma}_n(\boldsymbol{u}, \boldsymbol{v})$ as an estimator of $\sigma_C(\boldsymbol{u}, \boldsymbol{v})$.

The following two results extend Lemmas 3.12 and 3.13 of Bühlmann (1993) and Proposition 2.1 of Shao (2010), and are proved in Appendix D.

Proposition 5.1. *Assume that $\ell_n = O(n^{1/2-\varepsilon})$ for some $0 < \varepsilon < 1/2$, that $\boldsymbol{U}_1, \dots, \boldsymbol{U}_n$ is drawn from a strictly stationary sequence $(\boldsymbol{U}_i)_{i \in \mathbb{Z}}$ whose strong mixing coefficients satisfy $\alpha_r = O(r^{-a})$, $a > 3$, and that φ defined in Assumption (A3) is additionally twice continuously differentiable on $[-1,1]$ with $\varphi''(0) \neq 0$. Then, for any $\boldsymbol{u}, \boldsymbol{v} \in [0,1]^d$,*

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

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

$$\Gamma(\boldsymbol{u}, \boldsymbol{v}) = \frac{\varphi''(0)}{2} \sum_{k=-\infty}^{\infty} k^2 \gamma(k, \boldsymbol{u}, \boldsymbol{v}) \quad \textit{with} \quad \gamma(k, \boldsymbol{u}, \boldsymbol{v}) = \mathrm{Cov}\{\mathbf{1}(\boldsymbol{U}_0 \leq \boldsymbol{u}), \mathbf{1}(\boldsymbol{U}_k \leq \boldsymbol{v})\}.$$

Proposition 5.2. *Assume that $\boldsymbol{U}_1, \dots, \boldsymbol{U}_n$ is drawn from a strictly stationary sequence $(\boldsymbol{U}_i)_{i \in \mathbb{Z}}$ whose strong mixing coefficients satisfy $\alpha_r = O(r^{-a})$, $a > 3$, and that there exists $\lambda > 0$ such that φ defined in Assumption (A3) additionally satisfies $|\varphi(x) - \varphi(y)| \leq \lambda |x - y|$ for all $x, y \in \mathbb{R}$. Then, for any $\boldsymbol{u}, \boldsymbol{v} \in [0,1]^d$,*

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

where

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

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

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

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

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

$$\mathrm{IMSE}_n = \int_{[0,1]^{2d}} \mathrm{MSE}\{\tilde{\sigma}_n(\boldsymbol{u}, \boldsymbol{v})\} \mathrm{d}\boldsymbol{u} \mathrm{d}\boldsymbol{v} \sim \frac{\bar{\Gamma}^2}{\ell_n^4} + \bar{\Delta} \frac{\ell_n}{n}, \tag{9}$$

where

$$\bar{\Gamma}^2 = \int_{[0,1]^{2d}} \{\Gamma(\boldsymbol{u}, \boldsymbol{v})\}^2 \mathrm{d}\boldsymbol{u} \mathrm{d}\boldsymbol{v} \qquad \text{and} \qquad \bar{\Delta} = \int_{[0,1]^{2d}} \Delta(\boldsymbol{u}, \boldsymbol{v}) \mathrm{d}\boldsymbol{u} \mathrm{d}\boldsymbol{v}. \tag{10}$$

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

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

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,

$$\ell_n^{opt} = \left(\frac{4\bar{\Gamma}^2}{\bar{\Delta}} \right)^{1/5} n^{1/5}. \tag{11}$$

Note that, in the context of the dependent multiplier bootstrap for $\mathbb{C}_n$, it might seem more meaningful to base the previous construction on the mean squared error of an estimator of the covariance $\mathrm{Cov}\{\mathbb{C}_C(1, \boldsymbol{u}), \mathbb{C}_C(1, \boldsymbol{v})\}$ instead of working with the estimator $\tilde{\sigma}_n(\boldsymbol{u}, \boldsymbol{v})$ of $\sigma_C(\boldsymbol{u}, \boldsymbol{v}) = \mathrm{Cov}\{\mathbb{B}_C(1, \boldsymbol{u}), \mathbb{B}_C(1, \boldsymbol{v})\}$. This would be in principle possible since, for any $\boldsymbol{u}, \boldsymbol{v} \in [0,1]^d$, starting from (4), $\mathrm{Cov}\{\mathbb{C}_C(1, \boldsymbol{u}), \mathbb{C}_C(1, \boldsymbol{v})\}$ can be expressed in terms of the function σ_C and the partial derivatives $\dot{C}_j$, $j \in \{1, \dots, d\}$, of C (see e.g. Genest and Segers, 2010, Equation (5) and the proof of Proposition 1 for the case $d = 2$). An estimator of $\mathrm{Cov}\{\mathbb{C}_C(1, \boldsymbol{u}), \mathbb{C}_C(1, \boldsymbol{v})\}$ follows immediately from the aforementioned expression by replacing σ_C by $\tilde{\sigma}_n$. By linearity of the expectation, we then immediately obtain the analogue of Proposition 5.1. The analogue of Proposition 5.2 requires more work as covariances of the form $\mathrm{Cov}\{\tilde{\sigma}_n(\boldsymbol{u},\boldsymbol{v}),\tilde{\sigma}_n(\boldsymbol{s},\boldsymbol{t})\}$ are needed. From a practical perspective, we feel however that the possible gain resulting from this theoretically more meaningful approach is not worth the additional computation brought in by the partial derivatives of C and the covariance terms $\mathrm{Cov}\{\tilde{\sigma}_n(\boldsymbol{u},\boldsymbol{v}),\tilde{\sigma}_n(\boldsymbol{s},\boldsymbol{t})\}$.

From (11), we see that, to estimate ℓ_n^{opt}, we need to estimate the infinite sums $K(\boldsymbol{u},\boldsymbol{v}) = \sum_{k\in\mathbb{Z}} k^2\gamma(k,\boldsymbol{u},\boldsymbol{v})$ and $\sigma_C(\boldsymbol{u},\boldsymbol{v}) = \sum_{k\in\mathbb{Z}} \gamma(k,\boldsymbol{u},\boldsymbol{v})$ for all $\boldsymbol{u},\boldsymbol{v}\in[0,1]^d$. Should $\boldsymbol{U}_1,\dots,\boldsymbol{U}_n$ be observable, this could done by adapting the procedures described in Paparoditis and Politis (2001, page 1111) or Politis and White (2004, Section 3) to the current empirical process setting. Let $L\geq 1$ be an integer to be determined from $\boldsymbol{X}_1,\dots,\boldsymbol{X}_n$ later and fix $\boldsymbol{u},\boldsymbol{v}\in[0,1]^d$. Proceeding in the spirit of Politis and Romano (1995) and Politis (2003), the quantity $K(\boldsymbol{u},\boldsymbol{v})$ could be estimated by $\check{K}_n(\boldsymbol{u},\boldsymbol{v}) = \sum_{k=-L}^{L}\kappa_{F,0.5}(k/L)k^2\check{\gamma}_n(k,\boldsymbol{u},\boldsymbol{v})$, where

$$\kappa_{F,c}(x) = [\{(1-|x|)/(1-c)\}\vee 0]\wedge 1, \quad c\in[0,1], \tag{12}$$

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

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

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

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

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

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

which gives the computable estimators

$$\hat{\sigma}_n(\boldsymbol{u},\boldsymbol{v}) = \sum_{k=-L}^{L}\kappa_{F,0.5}(k/L)\hat{\gamma}_n(k,\boldsymbol{u},\boldsymbol{v}) \quad\text{and}\quad \hat{K}_n(\boldsymbol{u},\boldsymbol{v}) = \sum_{k=-L}^{L}\kappa_{F,0.5}(k/L)k^2\hat{\gamma}_n(k,\boldsymbol{u},\boldsymbol{v}) \tag{13}$$

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

To estimate $\bar{\Gamma}^2$ and $\bar{\Delta}$ defined in (10), we then propose to use a grid $\{\boldsymbol{u}_i\}_{i\in\{1,\dots,g\}}$ of g points uniformly spaced over $(0,1)^d$, and to compute

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

and

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

respectively. Plugging these into (11), we obtain an estimator of ℓ_n^{opt} which shall be denoted as $\hat{\ell}_n^{opt}$ as we continue.

The above estimator depends on the choice of the integer L appearing in (13). To estimate L, we suggest proceeding along the lines of Politis and White (2004, Section 3.2) (see also Paparoditis and Politis, 2001, page 1112). Let $\hat{\rho}_j(k)$, $j \in \{1, \dots, d\}$, be the autocorrelation function at lag k estimated from the sample $X_{1j}, \dots, X_{nj}$. For any $j \in \{1, \dots, d\}$, 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 (2004, Section 3.2). Then, we merely suggest taking $L = 2\psi(L_1, \dots, L_d)$, 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 (2004). Nonetheless, the choice $\psi =$ median was found to give meaningful results in our Monte Carlo experiments partially reported in Section 7.

6 Generation of dependent multiplier sequences

The practical use of the results stated in Sections 3 and 4 requires the generation of dependent multiplier random variables satisfying Assumptions (A1), (A2) and (A3). We describe two ways of constructing such dependent sequences. The first one generalizes the moving average approach proposed by Bühlmann (1993, Section 6.2) (see also Bücher and Ruppert, 2013) and produces multipliers that satisfy Assumption (A3) only asymptotically. The second one extends the method used by Shao (2010) and is based on the calculation of the square root of the covariance matrix implicitly defined in Assumption (A3).

6.1 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 $b_n \to \infty$, $b_n = o(n)$ 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 $\boldsymbol{X}_1, \dots, \boldsymbol{X}_n$ 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 $\mathrm{E}(|\xi_{1,n}|^\nu) < \infty$ for all $\nu \geq 1$. Furthermore, $\xi_{1,n}, \dots, \xi_{n,n}$ are $(\ell_n - 1)$-dependent and, for any $i \in$

$\{1,\ldots,n\}$ and $r \in \{0,\ldots,(\ell_n-1)\wedge n\}$,

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

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 $\xi_{1,n},\ldots,\xi_{n,n}$ satisfies Assumption (A1) and Assumption (A2) (as (ℓ_n-1)-dependence clearly implies ℓ_n-dependence). Let us now verify that it satisfies Assumption (A3) 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

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

for all $i\in\{1,\ldots,n\}$ and $r\in\{0,\ldots,n-i\}$, where the $o(1)$ 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

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

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

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

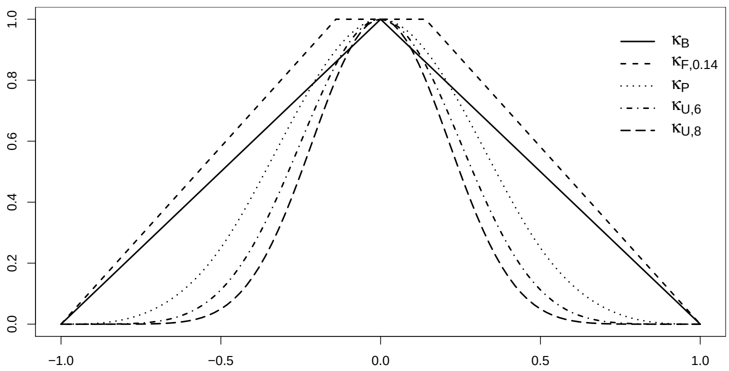

Figure 1: Graphs of the functions κ_B, $\kappa_{F,0.14}$ and κ_P, as well as $\kappa_{U,6}$ and $\kappa_{U,8}$ defined in Section 6.2.

Now, let

$$\varphi(x) = \frac{\kappa \star \kappa(2x)}{\kappa \star \kappa(0)}, \qquad x \in \mathbb{R}, \tag{14}$$

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 $i, j \in \{1, \dots, n\}$. Hence, the infinite size version of $\xi_{1,n}, \dots, \xi_{n,n}$ satisfies Assumption (A3) asymptotically.

In our numerical experiments, we considered several popular *kernels* for the function κ (see e.g. Andrews, 1991), defined, for any $x \in \mathbb{R}$, as

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

as well as the flat top kernel already defined in (12). 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 (2001) 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 (2010) who connected the tapered block bootstrap with the dependent multiplier bootstrap for the mean.

6.2 The covariance matrix approach

Let ℓ_n be a sequence of strictly positive constants such that $\ell_n \to \infty$ and $\ell_n = o(n)$. Let φ be a function satisfying Assumption (A3) such that, additionally, $\int_{-\infty}^{\infty} \varphi(u) e^{-iux} \mathrm{d}u \ge 0$ for all $x \in \mathbb{R}$, and let $\mathbf{\Sigma}_n$ be the $n \times n$ (covariance) matrix whose elements are defined by $\varphi\{(i-j)/\ell_n\}$, $i, j \in \{1, \dots, n\}$. 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 $Z_1, \dots, Z_n$ be i.i.d. random variables independent of the available sample $\boldsymbol{X}_1, \dots, \boldsymbol{X}_n$ such that $\mathrm{E}(Z_1) = 0$, $\mathrm{E}(Z_1^2) = 1$ and $\mathrm{E}(|Z_1|^\nu) < \infty$ for all $\nu \geq 1$. A dependent multiplier sequence $\xi_{1,n} \dots \xi_{n,n}$ can then be simply obtained as

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

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 $\xi_{1,n}, \dots, \xi_{n,n}$ satisfies Assumptions (A1), (A2) and (A3).

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, i.e., $\kappa_{U,8}(x) = \kappa_P \star \kappa_P(2x)/\kappa_P \star \kappa_P(0)$ 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 et al., 2002). 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.

6.3 Relationship between the two multiplier generation methods

In the case of the moving average approach presented in Section 6.1, we have seen that κ determines φ asymptotically through (14). It follows that, for the same type of initial 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 6.1 and φ in Section 6.2 are related through (14). For instance, all other parameters being similar, using the Bartlett kernel for κ in Section 6.1 should produce similar results to using the Parzen kernel for φ in Section 6.2.

7 Monte Carlo experiments

Recall that M is a large integer. For $m \in \{1, \dots, M\}$, let

$$S_n = \int_{[0,1]^d} \{\mathbb{C}_n(0, 1, \boldsymbol{u})\}^2 \mathrm{d}\boldsymbol{u} \quad \text{and} \quad S_n^{(m)} = \int_{[0,1]^d} \{\hat{\mathbb{C}}_n^{(m)}(0, 1, \boldsymbol{u})\}^2 \mathrm{d}\boldsymbol{u}, \tag{15}$$

where $\mathbb{C}_n$ is defined in (1) and where $\hat{\mathbb{C}}_n^{(m)}$ is defined in (7) with the partial derivative estimators defined as in (8). 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 $S = \int_{[0,1]^d} \{\mathbb{C}_C(0, 1, \boldsymbol{u})\}^2 \mathrm{d}\boldsymbol{u}$ 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 $S_n^{(1:M)} \leq \cdots \leq S_n^{(M:M)}$ 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, for $p \in \mathcal{P} = \{0.25, 0.5, 0.75, 0.9, 0.95, 0.99\}$, were accurately estimated empirically from 10^5 realizations of S_{1000}. Then, for each data generating scenario, $N = 1000$ samples $\boldsymbol{X}_1, \dots, \boldsymbol{X}_n$ 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 statistics

$$T_n = \sup_{\boldsymbol{u} \in [0,1]^d} |\mathbb{C}_n(0,1,\boldsymbol{u})| \quad \text{and} \quad T_n^{(m)} = \sup_{\boldsymbol{u} \in [0,1]^d} |\hat{\mathbb{C}}_n^{(m)}(0,1,\boldsymbol{u})|, \quad m \in \{1, \dots, M\}. \tag{16}$$

The dimension d was fixed to two, and the suprema and the integrals in (15) and (16), 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 $\boldsymbol{U}_i$, $i \in \{-100, \dots, 0, \dots, n\}$, 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 $\boldsymbol{X}_{-100} = \boldsymbol{\epsilon}_{-100}$. 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 (2001, Section 3.3), and to the exponential autoregressive (EXPAR) model considered in Auestad and Tjøstheim (1990) and Paparoditis and Politis (2001, Section 3.3). The sample $\boldsymbol{X}_1, \dots, \boldsymbol{X}_n$ 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 (2013). 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 (2013), 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 et al. (2007) from SP500 and DAX daily logreturns, respectively.

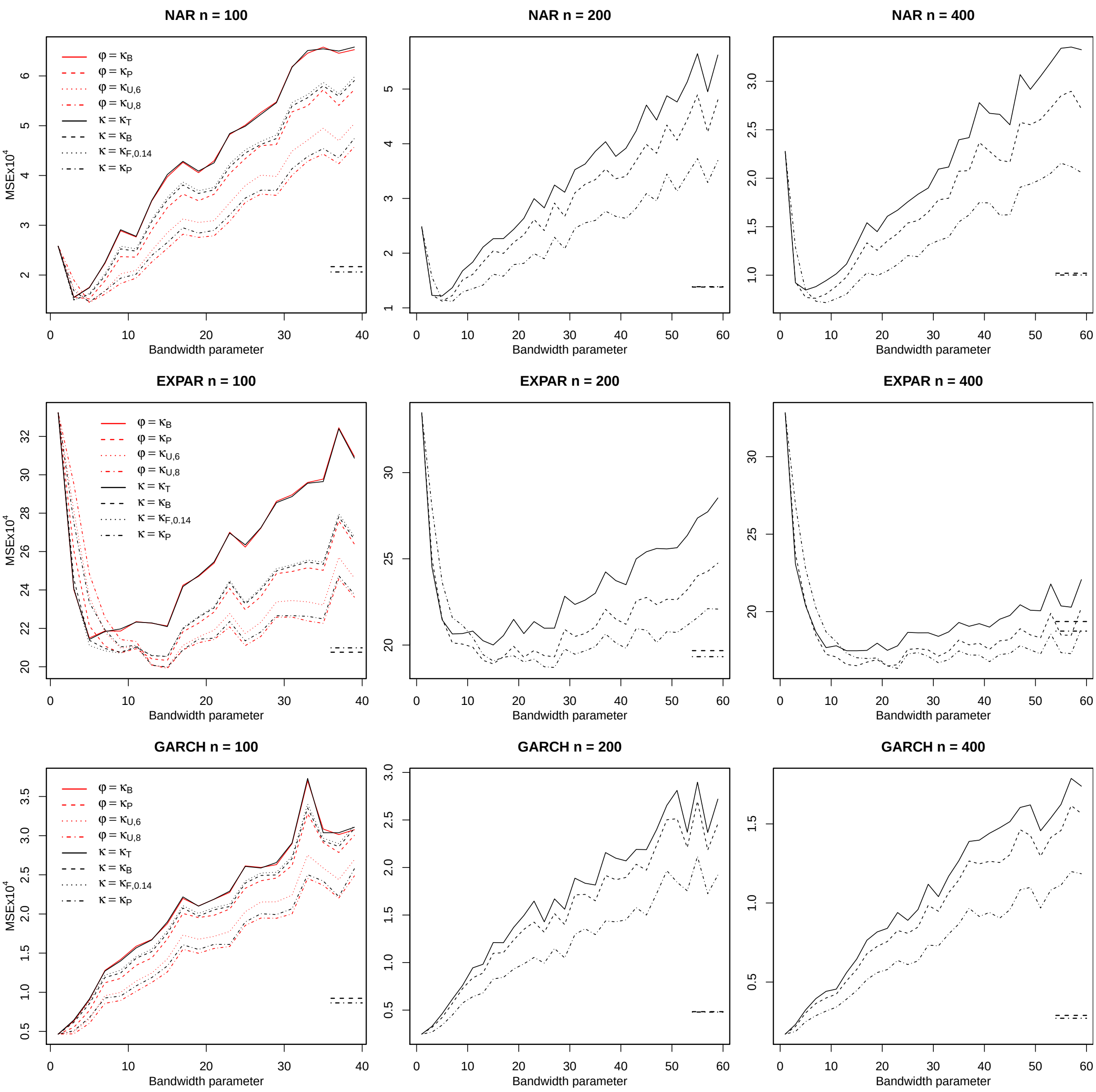

Figure 2: For various choices of the function κ/φ (see Section 6), empirical MSE$\times 10^4$ of the estimator $S_n^{(\lfloor 0.95M \rfloor : M)}$ with $M = 2500$ versus the bandwidth parameter ℓ_n under the NAR, EXPAR and GARCH data generating scenarios with C being the Gumbel–Hougaard copula with parameter 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. The line styles of the segments correspond to the choice of φ.

The other factors of the experiments are as follows. Four different copulas were considered: Clayton copulas with parameter value 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 et al., 2005, Chapter 5). The values 100, 200 and 400 were considered for n.

We report the results of the experiments very partially. 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 6. 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 results for the AR1 scenario being very similar to those for the NAR scenario are not reported. Very 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 (16). 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 $p = 0.99$.

The black (resp. red) curves in Figure 2 were obtained for dependent multiplier sequences generated using the moving average (resp. covariance matrix) approach described in Section 6.1 (resp. Section 6.2). 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 Section 6.3, 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 (2009) provides the best results. 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 will lead to the lowest variance. Another observation is that, unlike what was expected by Shao (2010, 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 $\kappa = \kappa_P$.

In view of the small differences between the moving average and covariance matrix approaches for generating dependent multipliers, 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 the estimation of ℓ_n using the results of Section 5, let us mention two observations of practical interest. Firstly, the distribution of the initial i.i.d. sequence necessary for generating dependent multipliers in Sections 6.1 and 6.2 did not seem to affect the results much. Indeed, we considered i.i.d. Rademacher, Gamma and normal sequences, centered and scaled to have variance one, and did not notice any substantial changes in terms of empirical MSE, all other parameters remaining fixed. The results reported above are those obtained with an initial normal i.i.d. sequence. Secondly, work-

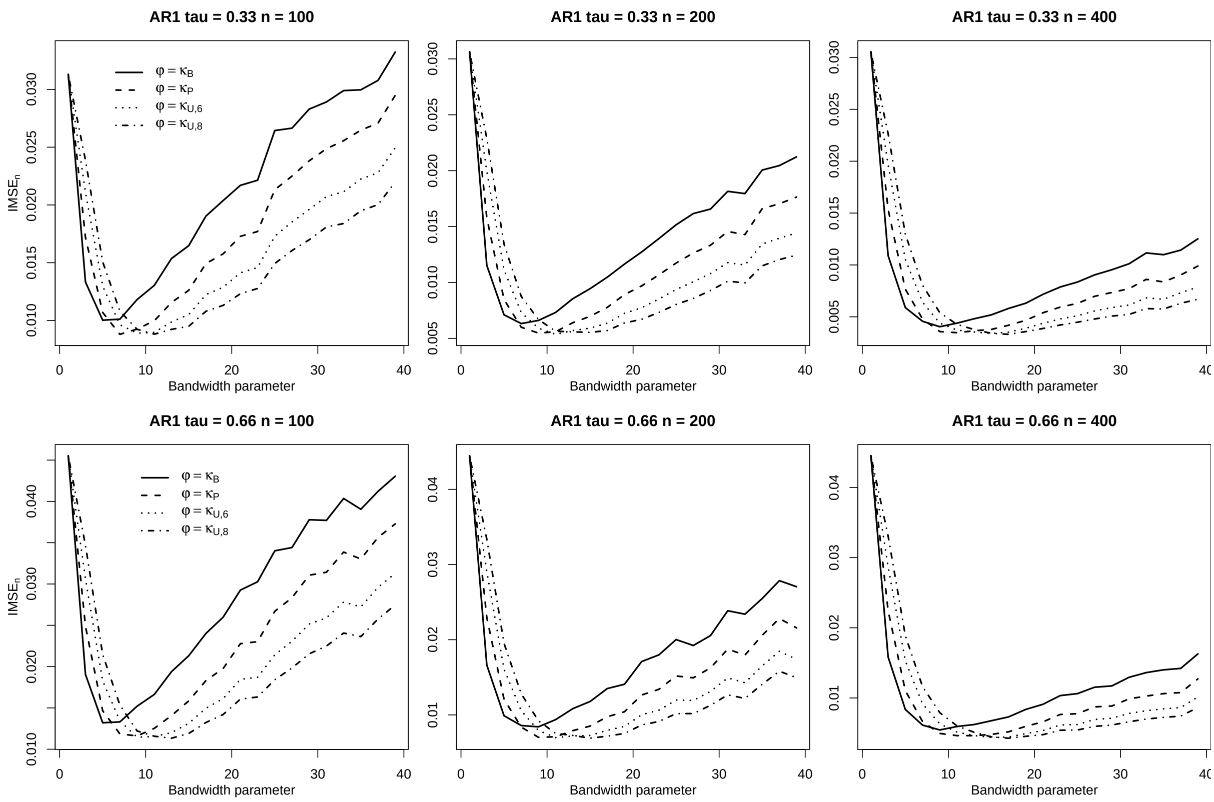

Figure 3: For the previously considered choices of the function φ, IMSE_n defined in (9), computed approximately using a grid of 25 uniformly spaced points on $(0,1)^2$ and 1000 samples versus the bandwidth parameter ℓ_n under the AR1 data generating scenario with C being the Gumbel–Hougaard copula with parameter 1.5 (top row) and parameter 3 (bottom row), respectively.

ing with the same random seed, we replicated the experiments described above using the two alternative definitions of the partial derivative estimators mentioned at the end of Section 4. To our surprise, the best results, overall, were obtained with the proposal of Rémillard and Scaillet (2009) given in (8).

We end this section by an empirical investigation of the estimator $\hat{\ell}_n^{opt}$ of ℓ_n^{opt} (see (11) and Section 5). We first report an experiment based on the AR1 model which will serve as a benchmark for judging about the performance of $\hat{\ell}_n^{opt}$. The setting is the following: a grid $\{\boldsymbol{u}_i\}_{i\in\{1,\dots,g\}}$ of $g = 25$ points uniformly spaced over $(0,1)^2$ was created, and $\sigma_C(\boldsymbol{u}_i,\boldsymbol{u}_j)$ 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 $\boldsymbol{X}_1,\dots,\boldsymbol{X}_n$ generated from the AR1 model, the marginally standard uniform sample $\boldsymbol{U}_1,\dots,\boldsymbol{U}_n$ was formed using the fact that the marginal c.d.f.s of the $\boldsymbol{X}_i$ 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); for any $i,j \in \{1,\dots,g\}$, $\sigma_C(\boldsymbol{u}_i,\boldsymbol{u}_j)$ was finally accurately estimated as the sample covariance of 10^5 independent realizations of $(\tilde{\mathbb{B}}_n(1,\boldsymbol{u}_i),\tilde{\mathbb{B}}_n(1,\boldsymbol{u}_j))$.
Next, for $n \in \{100,200,400\}$ and $\ell_n \in \{1,3,\dots,39\}$, IMSE_n defined in (9) was approxi-

Table 1: Mean and standard deviation of 1000 estimates of ℓ_n^{opt}, defined in (11), computed as explained in Section 5 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	9.07	4.34	12.60	7.03
	200	10.60	4.33	14.50	5.09
	400	13.00	3.72	17.62	5.15
3.0	100	9.09	5.25	12.17	5.29
	200	10.59	3.89	14.69	5.81
	400	12.85	3.91	17.73	5.44

mated as follows: 1000 samples $\boldsymbol{X}_1, \dots, \boldsymbol{X}_n$ 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 (6); computing sample covariances, this allowed us to obtain 1000 bootstrap estimates of $\sigma_C(\boldsymbol{u}_i, \boldsymbol{u}_j)$ 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 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 $n \in \{100, 200, 400\}$, $\varphi \in \{\kappa_P, \kappa_{U,8}\}$ 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 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.

For the final experiment, we reverted to the setting of the estimation of the quantiles of the statistic S_n defined in (15). This time, with $M = 2500$, we focused on the estimator $S_n^{(\lfloor 0.95M \rfloor : M)}$ based on the estimated bandwidth $\hat{\ell}_n^{opt}$ computed as explained in Section 5. For the NAR, EXPAR and GARCH data generating scenarios and $\varphi \in \{\kappa_P, \kappa_{U,8}\}$, its MSE was estimated from 1000 samples of size $n \in \{100, 200, 400\}$. The results are represented by line segments in the lower-right corners of the graphs of Figure 2. The line styles of the segments correspond to the choice of φ. As it can be seen, the empirical MSEs of the estimator with estimated bandwidth decrease with n and are, overall, reasonably close to the lowest observed MSE. The choice $\varphi = \kappa_{U,8}$ appears in general to lead to a slightly lower MSE, although the improvement might not be of practical interest.

Acknowledgements. 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.

A Proof of Theorem 2.3

The proof of Theorem 2.3 is based on the extended continuous mapping theorem (van der Vaart and Wellner, 2000, 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 and such that we may conclude that, as a process indexed by $s, t, \boldsymbol{u}$, $\mathbb{C}_n(s,t,\boldsymbol{u}) \approx g_n\{\tilde{\mathbb{B}}_n(t,\boldsymbol{u}) - \tilde{\mathbb{B}}_n(s,\boldsymbol{u})\}$ converges weakly to $g\{\tilde{\mathbb{B}}(t,\boldsymbol{u}) - \tilde{\mathbb{B}}(s,\boldsymbol{u})\} = \mathbb{C}(s,t,\boldsymbol{u})$.

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,

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

let

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

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

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

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

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

where, for any $u \in [0,1]$, $\boldsymbol{u}_{\{j\}}$ 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

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

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

$$\Phi_n(H^\star)(s,t,\boldsymbol{u}) = H^\star\{s,t,I_n(H_1^\star)(s,t,u_1),\dots,I_n(H_d^\star)(s,t,u_d)\}. \tag{17}$$

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

Also, let

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

let $\mathcal{D}_n^\star = \{\alpha^\star \in \mathcal{D}^\star : C_n^\star + n^{-1/2}\alpha^\star \in \mathcal{E}_{n,d}^\star\}$, and let $\mathcal{D}_0^\star = \mathcal{D}^\star \cap \mathcal{C}(\Delta \times [0,1]^d)$. Finally, for any $\alpha_n^\star \in \mathcal{D}_n^\star$ and any $(s,t,\boldsymbol{u}) \in \Delta \times [0,1]^d$, let

$$g_n(\alpha_n^\star)(s,t,\boldsymbol{u}) = \sqrt{n}\left\{\Phi_n(C_n^\star + n^{-1/2}\alpha_n^\star)(s,t,\boldsymbol{u}) - \Phi_n(C_n^\star)(s,t,\boldsymbol{u})\right\}, \tag{18}$$

and, for any $\alpha^\star \in \mathcal{D}_0^\star$ and any $(s,t,\boldsymbol{u}) \in \Delta \times [0,1]^d$, let

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

The following Lemma is the main ingredient for the proof of Theorem 2.3. Its proof is given subsequent to the proof of Theorem 2.3.

Lemma A.1. *Suppose that C satisfies Condition 2.2, and 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, $g_n(\alpha_n^\star) \to g(\alpha^\star) \in \ell^\infty(\Delta \times [0,1]^d)$.*

Proof of Theorem 2.3. Under Condition 2.1, we have that $\tilde{\mathbb{B}}_n \rightsquigarrow \mathbb{B}_C$ in $\ell^\infty([0,1]^{d+1})$. Now, for any $(s,t,\boldsymbol{u}) \in \Delta \times [0,1]^d$, define $\tilde{\mathbb{B}}_n^\Delta(s,t,\boldsymbol{u}) = \tilde{\mathbb{B}}_n(t,\boldsymbol{u}) - \tilde{\mathbb{B}}_n(s,\boldsymbol{u})$, $\mathbb{B}_C^\Delta(s,t,\boldsymbol{u}) = \mathbb{B}_C(t,\boldsymbol{u}) - \mathbb{B}_C(s,\boldsymbol{u})$, and

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

Notice that $\tilde{\mathbb{B}}_n^\Delta = \sqrt{n}(\tilde{H}_n^\star - C_n^\star)$ and that, by the continuous mapping theorem, $\tilde{\mathbb{B}}_n^\Delta \rightsquigarrow \mathbb{B}_C^\Delta$ in $\ell^\infty(\Delta \times [0,1]^d)$. Clearly, $\tilde{\mathbb{B}}_n^\Delta$, as a function of ω, takes its values in $\mathcal{D}_n^\star$ and $\mathbb{B}_C^\Delta$ is Borel measurable and separable by Condition 2.1, and, as a function of ω, takes its values in $\mathcal{D}_0^\star$. Now, consider the map h_n from $\mathcal{D}_n^\star$ to $\{\ell^\infty(\Delta \times [0,1]^d)\}^2$, defined, for any $\alpha_n^\star \in \mathcal{D}_n^\star$ and any $(s,t,\boldsymbol{u}) \in \Delta \times [0,1]^d$, by

$$h_n(\alpha_n^\star)(s,t,\boldsymbol{u}) = \left(g_n(\alpha_n^\star)(s,t,\boldsymbol{u}), g(\alpha_n^\star)(s,t,\boldsymbol{u})\right).$$

Using Lemma A.1 and the fact that g is linear and bounded, we have from the extended continuous mapping theorem (van der Vaart and Wellner, 2000, 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 $(s,t,\boldsymbol{u}) \in \Delta \times [0,1]^d$,

$$h(\alpha^\star)(s,t,\boldsymbol{u}) = \left(g(\alpha^\star)(s,t,\boldsymbol{u}), g(\alpha^\star)(s,t,\boldsymbol{u})\right).$$

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). To complete the proof, it remains to show that

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

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

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

Now, let $\tilde{H}_{\lfloor ns\rfloor+1:\lfloor nt\rfloor}$ be the empirical c.d.f. computed from the sample $\boldsymbol{U}_{\lfloor ns\rfloor+1},\dots,\boldsymbol{U}_{\lfloor nt\rfloor}$, 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 $F^{-1}(u)=\inf\{v\in[0,1]:F(v)\geq u\}$. Then, let

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

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

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

Hence, we obtain that

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

Proceeding for instance as in Lemma 1 of Kojadinovic and Rohmer (2012), it can be verified that

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

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

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

Lemma A.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 $j\in\{1,\dots,d\}$,*

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

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

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

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 $u \in [0,1]$, $I_n(U_n^\star + n^{-1/2}\alpha_{n,j}^\star)(s,t,u) = F_{s,t,n}^{-1}(u)$, where $F_{s,t,n} = \lambda_n(s,t)^{-1}(U_n^\star + n^{-1/2}\alpha_{n,j}^\star)(s,t,\cdot) \in \mathcal{E}$. It follows that, for any $\lfloor ns \rfloor < \lfloor nt \rfloor$ and $u \in (0,1]$, $\xi_n(s,t,u) = I_n(U_n^\star + n^{-1/2}\alpha_{n,j}^\star)(s,t,u) > 0$, 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 $u \in (0,1]$,

$$(U_n^\star + n^{-1/2}\alpha_{n,j}^\star)\{s,t,\xi_n(s,t,u) - \varepsilon_n(s,t,u)\} \leq \lambda_n(s,t)u \leq (U_n^\star + n^{-1/2}\alpha_{n,j}^\star)\{s,t,\xi_n(s,t,u)\},$$

that is

$$\begin{aligned} &-n^{-1/2}\alpha_{n,j}^\star\{s,t,\xi_n(s,t,u)\} \leq \lambda_n(s,t)\{\xi_n(s,t,u) - u\} \\ &\qquad \leq \lambda_n(s,t)\varepsilon_n(s,t,u) - n^{-1/2}\alpha_{n,j}^\star\{s,t,\xi_n(s,t,u) - \varepsilon_n(s,t,u)\}, \end{aligned} \tag{19}$$

which in turn implies that

$$\sup_{\lfloor ns \rfloor < \lfloor nt \rfloor, u \in (0,1]} |\lambda_n(s,t)\{\xi_n(s,t,u) - u\}| \to 0 \tag{20}$$

since, by uniform convergence of $\alpha_n^\star$ to $\alpha^\star$ and the fact that $\alpha^\star \in \mathcal{D}_0^\star$, $\sup_{(s,t,u) \in \Delta \times [0,1]} |\alpha_{n,j}^\star(s,t,u)|$ is bounded. From (19), exploiting the fact that $\varepsilon_n(s,t,u) \leq n^{-1}$, we then obtain that

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

where

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

and

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

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

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

and

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

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 $\sup_{t-s<\delta, u \in [0,1]} |\alpha_j^\star(s,t,u)| \leq \varepsilon$. We have $B_{n,2} = \max\{B_{n,3}, B_{n,4}\}$, where

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

and

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

Now, it is easy to verify that $t-s \leq \lambda_n(s,t)+1/n$, so that, for n sufficiently large, $t-s\geq\delta$ implies that $\lambda_n(s,t)\geq\delta/2$. Then, from (20) and the fact that $\xi_n(\cdot,\cdot,0)=0$, we immediately have that, for n sufficiently large,

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

Hence, we can write

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

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 A.1. Starting from the definitions of g_n and Φ_n given in (18) and (17), respectively, we have the decomposition

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

where

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

and

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

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

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

where

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

for sufficiently large n, where

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

and

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

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

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

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

$$a_{n,j} = \sup_{t-s\geq\delta,u\in[0,1]} |I_n(U_n^\star + n^{-1/2}\alpha_{n,j}^\star)(s,t,u) - u| \to 0, \tag{21}$$

and the fact that

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

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, $\sup_{(s,t,\boldsymbol{u})\in\Delta\times[0,1]^d} |A_{n,1}(s,t,\boldsymbol{u}) - \alpha^\star(s,t,\boldsymbol{u})| \leq 4\varepsilon$, and therefore that $\sup_{(s,t,\boldsymbol{u})\in\Delta\times[0,1]^d} |A_{n,1}(s,t,\boldsymbol{u}) - \alpha^\star(s,t,\boldsymbol{u})| \to 0$.

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 $\boldsymbol{u} \in [0,1]^d$, $j \in \{1,\dots,d\}$ 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 $\bar{\boldsymbol{u}}(r) = (\bar u_1(r),\dots,\bar u_d(r))$. Now, fix $\boldsymbol{u} \in (0,1)^d$ and let f be the function defined by

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

Obviously, we have that $0 < \bar u_j(r) < 1$ for all $r \in (0,1)$ and $j \in \{1,\dots,d\}$. Therefore, the function f is continuous on $[0,1]$, and, by Condition 2.2, 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

$$A_{n,2}(s,t,\boldsymbol{u}) = \sum_{j=1}^d \dot C_j\{\bar{\boldsymbol{u}}(r^*)\}\lambda_n(s,t)\sqrt{n}\{I_n(U_n^\star + n^{-1/2}\alpha_{n,j}^\star)(s,t,u_j) - u_j\}. \tag{22}$$

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 $\boldsymbol{u}$ is on the boundary of $[0,1]^d$. When $u_j = 0$ for some $j \in \{1,\dots,d\}$, $I_n(U_n^\star + n^{-1/2}\alpha_{n,j}^\star)(\cdot,\cdot,u_j) = 0$, which implies that $\bar u_j(r) = 0$ for all $r \in [0,1]$. It then immediately follows that the left side of (22) 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 (22) are actually also zero because, for any $k \in \{1,\dots,d\}$, $k \neq j$, $\dot C_k(\boldsymbol{v}) = 0$ for all $\boldsymbol{v} \in [0,1]^d$ such that $v_k = 0$. Hence, (22) remains true whenever $u_j = 0$ for some $j \in \{1,\dots,d\}$.

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 (22) is zero so that neither the left nor the right side of (22) depend on u_j anymore. It follows that, when some components of $\boldsymbol{u}$ 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

$$A_{n,2}(s,t,\boldsymbol{u}) = \sum_{j=1}^{d} \dot{C}_j(\boldsymbol{u})\lambda_n(s,t)\sqrt{n}\{I_n(U_n^\star + n^{-1/2}\alpha_{n,j}^\star)(s,t,u_j) - u_j\} + r_n(s,t,\boldsymbol{u}), \quad (23)$$

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

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

By Lemma A.2 and from the fact that $0 \leq \dot{C}_j \leq 1$ for all $j \in \{1,\dots,d\}$, the dominating term in decomposition (23) converges to

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

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

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

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

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

where

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

and

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

From the fact that $0 \leq \dot{C}_1 \leq 1$ and Lemma A.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 $|\alpha_1^\star(s,t,u)| \leq \varepsilon$ for all $t - s < \delta$ and all $u \in [0,1]$. Then, $B_{n,5} \leq B_{n,6} + B_{n,7}$, where

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

and

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

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

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

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

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

and

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

From (21), we obtain that

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

Since $\dot{C}_1$ is uniformly continuous on $[\kappa/2, 1-\kappa/2] \times [0,1]^{d-1}$ according to Condition 2.2, 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, $B_{n,9} \leq \varepsilon$. It follows that, for n sufficiently large, $B_{n,5} \leq 5\varepsilon$, which implies that $\sup_{(s,t,\boldsymbol{u})\in\Delta\times[0,1]^d} |r_{n,1}(s,t,\boldsymbol{u})| \to 0$. One can proceed similarly for $r_{n,j}$, $j \in \{2,\dots,d\}$. Hence, $\sup_{s\leq t, \boldsymbol{u}\in[0,1]^d} |r_n(s,t,\boldsymbol{u})| \to 0$. ■

B Proof of Theorem 3.1

The proof of Theorem 3.1 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.

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

Proof. Fix $m \in \{1,\dots,M\}$. For the sake of brevity, we shall only show that the finite-dimensional distributions of $(\tilde{\mathbb{B}}_n, \tilde{\mathbb{B}}_n^{(m)})$ converge weakly to those of $(\mathbb{B}_C, \mathbb{B}_C^{(m)})$, the proof of the stated result being a more notationally complex version of the proof of the latter result.

Let $q \in \mathbb{N}$, $q > 1$, be arbitrary, and let $(\boldsymbol{u}_1, s_1, \boldsymbol{v}_1, t_1), \dots, (\boldsymbol{u}_q, s_q, \boldsymbol{v}_q, t_q) \in [0,1]^{2(d+1)}$. The result is proved if we show that

$$\left(\tilde{\mathbb{B}}_n(s_1,\boldsymbol{u}_1), \tilde{\mathbb{B}}_n^{(m)}(t_1,\boldsymbol{v}_1), \dots, \tilde{\mathbb{B}}_n(s_q,\boldsymbol{u}_q), \tilde{\mathbb{B}}_n^{(m)}(t_q,\boldsymbol{v}_q)\right) \\ \rightsquigarrow \left(\mathbb{B}_C(s_1,\boldsymbol{u}_1), \mathbb{B}_C^{(m)}(t_1,\boldsymbol{v}_1), \dots, \mathbb{B}_C(s_q,\boldsymbol{u}_q), \mathbb{B}_C^{(m)}(t_q,\boldsymbol{v}_q)\right).$$

Let $c_1, d_1, \ldots, c_q, d_q \in \mathbb{R}$ be arbitrary. By the Cramér–Wold device, it then suffices to show that

$$Z_n = \sum_{l=1}^{q} c_l \tilde{\mathbb{B}}_n(s_l, \boldsymbol{u}_l) + \sum_{l=1}^{q} d_l \tilde{\mathbb{B}}_n^{(m)}(t_l, \boldsymbol{v}_l) \rightsquigarrow Z = \sum_{l=1}^{q} c_l \mathbb{B}_C(s_l, \boldsymbol{u}_l) + \sum_{l=1}^{q} d_l \mathbb{B}_C^{(m)}(t_l, \boldsymbol{v}_l).$$

Now, for any $i \in \{1, \ldots, n\}$, let

$$Z_{i,n} = \sum_{l=1}^{q} c_l \{\mathbf{1}(\boldsymbol{U}_i \leq \boldsymbol{u}_l) - C(\boldsymbol{u}_l)\} \mathbf{1}(i \leq \lfloor n s_l \rfloor)$$

and

$$Z_{i,n}^{(m)} = \xi_{i,n}^{(m)} \sum_{l=1}^{q} d_l \{\mathbf{1}(\boldsymbol{U}_i \leq \boldsymbol{v}_l) - C(\boldsymbol{v}_l)\} \mathbf{1}(i \leq \lfloor n t_l \rfloor).$$

Hence, $Z_n = n^{-1/2} \sum_{i=1}^{n} (Z_{i,n} + Z_{i,n}^{(m)})$. To prove the convergence in distribution of Z_n to Z, we employ a blocking technique (see e.g. Dehling and Philipp, 2002, page 31). Each block is composed of a big subblock followed by a small subblock. Let $0 < \eta_b < \eta_s < \varepsilon$ such that $\eta_s < 1/2 - 1/a$. The length of the small subblocks is $s_n = \lfloor n^{1/2-\eta_s} \rfloor$ and the length of the big subblocks is $b_n = \lfloor n^{1/2-\eta_b} \rfloor$ so that the length of a block is $b_n + s_n$. The total number of blocks is $k_n = \lfloor n/(b_n + s_n) \rfloor$, and we can write $n = k_n(b_n + s_n) + \{n - k_n(b_n + s_n)\}$. Note that $s_n \sim n^{1/2-\eta_s}$, $b_n \sim n^{1/2-\eta_b}$ and $k_n \sim n^{1/2+\eta_b}$ and that both b_n and s_n dominate ℓ_n. As we continue, n is taken sufficiently large so that $b_n > s_n > \ell_n$. Notice also that the condition $\eta_s < 1/2 - 1/a$ implies that $n s_n^{-a} \to 0$. Now, for any $j \in \{1, \ldots, k_n\}$, let

$$B_{j,n} = \sum_{i=(j-1)(b_n+s_n)+1}^{(j-1)(b_n+s_n)+b_n} (Z_{i,n} + Z_{i,n}^{(m)}) \qquad \text{and} \qquad S_{j,n} = \sum_{i=(j-1)(b_n+s_n)+b_n+1}^{j(b_n+s_n)} (Z_{i,n} + Z_{i,n}^{(m)})$$

be the sums of the $(Z_{i,n} + Z_{i,n}^{(m)})$ in the jth big subblock and the jth small subblock, respectively. Then,

$$Z_n = n^{-1/2} \sum_{j=1}^{k_n} B_{j,n} + n^{-1/2} \sum_{j=1}^{k_n} S_{j,n} + n^{-1/2} R_n,$$

where $R_n = \sum_{i=k_n(b_n+s_n)+1}^{n} (Z_{i,n} + Z_{i,n}^{(m)})$ is the sum of the $(Z_{i,n} + Z_{i,n}^{(m)})$ after the last small subblock. It follows that

$$\begin{aligned} \mathrm{Var}(Z_n) = \mathrm{Var}\left(n^{-1/2} \sum_{j=1}^{k_n} B_{j,n} \right) + 2n^{-1} \sum_{j,j'=1}^{k_n} \mathrm{E}(B_{j,n} S_{j',n}) + 2n^{-1} \sum_{j=1}^{k_n} \mathrm{E}(B_{j,n} R_n) \\ + n^{-1} \sum_{j,j'=1}^{k_n} \mathrm{E}(S_{j,n} S_{j',n}) + 2n^{-1} \sum_{j=1}^{k_n} \mathrm{E}(S_{j,n} R_n) + \mathrm{E}(n^{-1} R_n^2). \quad (24) \end{aligned}$$

We shall now show that all the terms on the right except the first one tend to zero. Notice that the convergence of the fourth and sixth term to zero will imply that $|Z_n -$

$n^{-1/2}\sum_{j=1}^{k_n} B_{j,n}| = |n^{-1/2}\sum_{j=1}^{k_n} S_{j,n} + n^{-1/2}R_n| \xrightarrow{\mathrm{P}} 0$. We start with the second one. For any $i \in \mathbb{Z}$ and any $\boldsymbol{u}, \boldsymbol{v} \in [0,1]^d$, let $\gamma(i, \boldsymbol{u}, \boldsymbol{v}) = \mathrm{Cov}\{\mathbf{1}(\boldsymbol{U}_0 \leq \boldsymbol{u}), \mathbf{1}(\boldsymbol{U}_i \leq \boldsymbol{v})\}$. We have

$$\mathrm{E}(B_{j,n}S_{j',n}) = \sum_{i=(j-1)(b_n+s_n)+1}^{j(b_n+s_n)+b_n} \sum_{i'=(j'-1)(b_n+s_n)+b_n+1}^{j'(b_n+s_n)} \{\mathrm{E}(Z_{i,n}Z_{i',n}) + \mathrm{E}(Z_{i,n}^{(m)}Z_{i',n}^{(m)})\}.$$

Now,

$$|\mathrm{E}(Z_{i,n}Z_{i',n})| \leq \sum_{l,l'=1}^{q} |c_l c_{l'}||\gamma(i'-i, \boldsymbol{u}_l, \boldsymbol{u}_{l'})| \leq 4\alpha_{|i'-i|} \sum_{l,l'=1}^{q} |c_l c_{l'}|,$$

where the last inequality is a consequence of Lemma 3.9 in Dehling and Philipp (2002), which implies that

$$\sup_{\boldsymbol{u},\boldsymbol{v}\in[0,1]^d} |\gamma(i, \boldsymbol{u}, \boldsymbol{v})| \leq 4\alpha_{|i|}, \qquad i \in \mathbb{Z}. \tag{25}$$

Similarly,

$$|\mathrm{E}(Z_{i,n}^{(m)}Z_{i',n}^{(m)})| \leq \mathbf{1}(|i'-i| \leq \ell_n)4\alpha_{|i'-i|} \sum_{l,l'=1}^{q} |d_l d_{l'}|$$

since, by Cauchy-Schwarz's inequality, $\mathrm{E}(\xi_{i,n}^{(m)}\xi_{i',n}^{(m)}) \leq \mathrm{E}\{(\xi_{0,n}^{(m)})^2\} = 1$. It follows that

$$|\mathrm{E}(B_{j,n}S_{j,n})| \leq \text{const} \times \sum_{i=1}^{b_n} \sum_{i'=b_n+1}^{b_n+s_n} \alpha_{|i-i'|} \leq \text{const} \times \sum_{i=1}^{b_n+s_n-1} i\alpha_i < \infty$$

since $\sum_{i=1}^{\infty} i\alpha_i < \infty$. Similarly, we obtain that $|\mathrm{E}(B_{j,n}S_{j-1,n})| < \infty$. For $j' \geq j+1$ or $j > j'+1$, $\mathrm{E}(B_{j,n}S_{j',n}) = O(b_n s_n \alpha_{b_n}) = O(b_n s_n b_n^{-a})$. Hence,

$$2n^{-1}\sum_{j,j'=1}^{k_n} \mathrm{E}(B_{j,n}S_{j',n}) = O(n^{-1}k_n) + O(n^{-1}k_n^2 b_n s_n b_n^{-a}) = O(b_n^{-1}) + O(n s_n b_n^{-a-1}).$$

Since $n s_n b_n^{-a-1} < n s_n^{-a}$, the previous term converges to zero. In a similar way, for the third summand in (24), we have

$$\begin{aligned} 2n^{-1}\sum_{j=1}^{k_n} \mathrm{E}(B_{j,n}R_n) &= 2n^{-1}\sum_{j=1}^{k_n-1} \mathrm{E}(B_{j,n}R_n) + 2n^{-1}\mathrm{E}(B_{k_n,n}R_n) \\ &= O(n^{-1}k_n b_n(n - k_n(b_n+s_n))\alpha_{b_n}) + O(n^{-1}b_n(n - k_n(b_n+s_n))) \\ &= O(b_n^{-a+1}) + O(n^{-1}b_n^2) \to 0 \end{aligned}$$

using the fact that $n - k_n(b_n+s_n) < b_n + s_n$. The case of the fifth summand is similar. Regarding the fourth summand in (24), we have

$$\mathrm{E}(S_{j,n}S_{j',n}) = \sum_{i=(j-1)(b_n+s_n)+b_n+1}^{j(b_n+s_n)} \sum_{i'=(j'-1)(b_n+s_n)+b_n+1}^{j'(b_n+s_n)} \{\mathrm{E}(Z_{i,n}Z_{i',n}) + \mathrm{E}(Z_{i,n}^{(m)}Z_{i',n}^{(m)})\},$$

which implies that

$$\mathrm{E}(S_{j,n}^2) \leq \text{const} \times \sum_{i,i'=b_n+1}^{b_n+s_n} \alpha_{|i'-i|} \leq \text{const} \times 2 \sum_{i=0}^{s_n-1} (s_n - i)\alpha_i \leq \text{const} \times s_n \sum_{i=0}^{\infty} \alpha_i = O(s_n)$$

and that, for $j \neq j'$, $\mathrm{E}(S_{j,n}S_{j',n}) = O(s_n^2 \alpha_{b_n}) = O(s_n^2 b_n^{-a})$. Hence,

$$n^{-1} \sum_{j,j'=1}^{k_n} \mathrm{E}(S_{j,n}S_{j',n}) = O(n^{-1}k_n s_n) + O(n^{-1}k_n^2 s_n^2 b_n^{-a}) = O(b_n^{-1}s_n) + O(n s_n^2 b_n^{-a-2})$$

which converges to 0 since $b_n^{-1}s_n \to 0$ and $n s_n^2 b_n^{-a-2} < n s_n^{-a}$. Finally, for the sixth summand in (24), we have

$$\mathrm{E}(n^{-1}R_n^2) = O(n^{-1}(n - k_n(b_n + s_n))^2) = O(n^{-1}(b_n + s_n)^2) = O(n^{-1}b_n^2)$$

since $n - k_n(b_n + s_n) < b_n + s_n$.

In order to prove that Z_n converges in distribution to Z, it suffices therefore to prove that $n^{-1/2}\sum_{j=1}^{k_n} B_{j,n}$ converges in distribution to Z. Let $\psi_{j,n}(t) = \exp(itn^{-1/2}B_{j,n})$, $t \in \mathbb{R}$, $j \in \{1, \ldots, k_n\}$, and observe that the characteristic function of $n^{-1/2}\sum_{j=1}^{k_n} B_{j,n}$ can be written as $t \mapsto \mathrm{E}\left\{\prod_{j=1}^{k_n} \psi_{j,n}(t)\right\}$. Also, for two σ-fields $\mathcal{F}_1$ and $\mathcal{F}_2$, let

$$\alpha(\mathcal{F}_1, \mathcal{F}_2) = \sup_{A \in \mathcal{F}_1, B \in \mathcal{F}_2} |\mathrm{P}(A \cap B) - \mathrm{P}(A)\mathrm{P}(B)|.$$

Now, for any $t \in \mathbb{R}$, we can write

$$\begin{aligned}
\left|\mathrm{E}\left\{\prod_{j=1}^{k_n} \psi_{j,n}(t)\right\} - \prod_{j=1}^{k_n} \mathrm{E}\{\psi_{j,n}(t)\}\right| &\leq \left|\mathrm{E}\left\{\prod_{j=1}^{k_n} \psi_{j,n}(t)\right\} - \mathrm{E}\{\psi_{1,n}(t)\}\mathrm{E}\left\{\prod_{j=2}^{k_n} \psi_{j,n}(t)\right\}\right| \\
&+ |\mathrm{E}\{\psi_{1,n}(t)\}| \left|\mathrm{E}\left\{\prod_{j=2}^{k_n} \psi_{j,n}(t)\right\} - \mathrm{E}\{\psi_{2,n}(t)\}\mathrm{E}\left\{\prod_{j=3}^{k_n} \psi_{j,n}(t)\right\}\right| + \ldots \\
&\cdots + \left|\prod_{j=1}^{k_n-2} \mathrm{E}\{\psi_{j,n}(t)\}\right| \left|\mathrm{E}\left\{\prod_{j=k_n-1}^{k_n} \psi_{j,n}(t)\right\} - \prod_{j=k_n-1}^{k_n} \mathrm{E}\{\psi_{j,n}(t)\}\right|.
\end{aligned}$$

Using the fact that the modulus of a characteristic function is smaller than one and applying $k_n - 1$ times Lemma 3.9 of Dehling and Philipp (2002), we obtain

$$\left|\mathrm{E}\left\{\prod_{j=1}^{k_n} \psi_{j,n}(t)\right\} - \prod_{j=1}^{k_n} \mathrm{E}\{\psi_{j,n}(t)\}\right| \leq 2\pi k_n \max_{1 \leq i \leq k_n - 1} \alpha\left[\sigma\{\psi_{i,n}(t)\}, \sigma\left\{\prod_{j=i+1}^{k_n} \psi_{j,n}(t)\right\}\right].$$

Since the big subblocks are s_n observations apart, the right-hand side of the previous inequality is smaller than $2\pi k_n \alpha_{s_n} = O(k_n s_n^{-a})$ which tends to zero as $k_n s_n^{-a} \leq n s_n^{-a} \to 0$. Hence, for any $t \in \mathbb{R}$,

$$\left|\mathrm{E}\left\{\prod_{j=1}^{k_n} \psi_{j,n}(t)\right\} - \prod_{j=1}^{k_n} \mathrm{E}\{\psi_{j,n}(t)\}\right| \to 0.$$

In other words, the characteristic function of $n^{-1/2}\sum_{j=1}^{k_n} B_{j,n}$ is asymptotically equivalent to the characteristic function of $n^{-1/2}\sum_{j=1}^{k_n} B'_{j,n}$, where $B'_{1,n},\dots,B'_{k_n,n}$ are independent and $B'_{j,n}$ and $B_{j,n}$ have the same distribution for all $j\in\{1,\dots,k_n\}$. To conclude that $n^{-1/2}\sum_{j=1}^{k_n} B_{j,n}$ converges in distribution to Z, it suffices therefore to show that $n^{-1/2}\sum_{j=1}^{k_n} B'_{j,n}$ converges in distribution to Z. This will be accomplished using the Lindeberg–Feller central limit theorem for triangular arrays. Hence, let us first show that $\mathrm{Var}\left(n^{-1/2}\sum_{j=1}^{k_n} B'_{j,n}\right)\to\mathrm{Var}(Z)$.

We have

$$\mathrm{Var}(Z)=\sum_{l,l'=1}^{q} c_l c_{l'}(s_l\wedge s_{l'})\sum_{i\in\mathbb{Z}}\gamma(i,\boldsymbol{u}_l,\boldsymbol{u}_{l'})+\sum_{l,l'=1}^{q} d_l d_{l'}(t_l\wedge t_{l'})\sum_{i\in\mathbb{Z}}\gamma(i,\boldsymbol{v}_l,\boldsymbol{v}_{l'}).$$

Note that, for any $\boldsymbol{u},\boldsymbol{v}\in[0,1]^d$, $\sum_{i\in\mathbb{Z}}\gamma(i,\boldsymbol{u},\boldsymbol{v})=\mathrm{Cov}\{\mathbb{B}_C(1,\boldsymbol{u}),\mathbb{B}_C(1,\boldsymbol{v})\}<\infty$ since, from the fact that $\gamma(i,\boldsymbol{v},\boldsymbol{u})=\gamma(-i,\boldsymbol{u},\boldsymbol{v})$ and (25),

$$\sum_{i\in\mathbb{Z}}|\gamma(i,\boldsymbol{u},\boldsymbol{v})|=\sum_{i=1}^{\infty}|\gamma(i,\boldsymbol{v},\boldsymbol{u})|+|\gamma(0,\boldsymbol{u},\boldsymbol{v})|+\sum_{i=1}^{\infty}|\gamma(i,\boldsymbol{u},\boldsymbol{v})|\le 1+8\sum_{i=1}^{\infty}\alpha_i<\infty. \quad (26)$$

Now, we shall first show that

$$\mathrm{Var}\left(n^{-1/2}\sum_{j=1}^{k_n} B'_{j,n}\right)=\mathrm{Var}(Z_n)+o(1)$$

and then that $\mathrm{Var}(Z_n)\to\mathrm{Var}(Z)$. We have

$$\begin{aligned}\mathrm{Var}\left(n^{-1/2}\sum_{j=1}^{k_n} B'_{j,n}\right)&=n^{-1}\sum_{j=1}^{k_n}\mathrm{Var}\left(B'_{j,n}\right)=n^{-1}\sum_{j=1}^{k_n}\mathrm{Var}\left(B_{j,n}\right)\\&=\mathrm{Var}\left(n^{-1/2}\sum_{j=1}^{k_n} B_{j,n}\right)-n^{-1}\sum_{\substack{j,j'=1\\ j\neq j'}}^{k_n}\mathrm{E}\left(B_{j,n}B_{j',n}\right).\end{aligned}$$

From (24), we know that $\mathrm{Var}(n^{-1/2}\sum_{j=1}^{k_n} B_{j,n})=\mathrm{Var}(Z_n)+o(1)$. Hence, it remains to show that the double sum in the last displayed formula converges to 0. Proceeding as for the summands on the right of (24), we have that, for $j\neq j'$, $\mathrm{E}(B_{j,n}B_{j',n})=O(b_n^2\alpha_{s_n})=O(b_n^2 s_n^{-a})$. Hence,

$$n^{-1}\sum_{\substack{j,j'=1\\ j\neq j'}}^{k_n}\mathrm{E}\left(B_{j,n}B_{j',n}\right)=O(n^{-1}k_n^2 b_n^2 s_n^{-a})=O(ns_n^{-a})\to 0.$$

Thus, it remains to show that $\mathrm{Var}(Z_n)\to\mathrm{Var}(Z)$. Now,

$$\mathrm{Var}(Z_n)=n^{-1}\sum_{i,i'=1}^{n}\{\mathrm{E}(Z_{i,n}Z_{i',n})+\mathrm{E}(Z_{i,n}^{(m)}Z_{i',n}^{(m)})\}.$$

It follows that

$$\operatorname{Var}(Z_n) = \sum_{l,l'=1}^{q} c_l c_{l'} n^{-1} \sum_{i=1}^{\lfloor ns_l \rfloor} \sum_{i'=1}^{\lfloor ns_{l'} \rfloor} \gamma(i'-i, \boldsymbol{u}_l, \boldsymbol{u}_{l'}) \\ + \sum_{l,l'=1}^{q} d_l d_{l'} n^{-1} \sum_{i=1}^{\lfloor nt_l \rfloor} \sum_{i'=1}^{\lfloor nt_{l'} \rfloor} \varphi\{(i'-i)/\ell_n\} \gamma(i'-i, \boldsymbol{v}_l, \boldsymbol{v}_{l'}), \quad (27)$$

where φ is the function appearing in Assumption (A3). Let us first deal with the second term on the right. Let $l, l' \in \{1, \dots, q\}$ be arbitrary and suppose without loss of generality that $t_l \leq t_{l'}$. Then,

$$n^{-1} \sum_{i=1}^{\lfloor nt_l \rfloor} \sum_{i'=1}^{\lfloor nt_{l'} \rfloor} v\{(i'-i)/\ell_n\} \gamma(i'-i, \boldsymbol{v}_l, \boldsymbol{v}_{l'}) = n^{-1} \sum_{i=1}^{\lfloor nt_l \rfloor} \sum_{i'=1}^{\lfloor nt_l \rfloor} v\{(i'-i)/\ell_n\} \gamma(i'-i, \boldsymbol{v}_l, \boldsymbol{v}_{l'}) \\ + n^{-1} \sum_{i=1}^{\lfloor nt_l \rfloor} \sum_{i'=\lfloor nt_l \rfloor + 1}^{\lfloor nt_{l'} \rfloor} v\{(i'-i)/\ell_n\} \gamma(i'-i, \boldsymbol{v}_l, \boldsymbol{v}_{l'}). \quad (28)$$

The first sum on the right-hand side is equal to

$$n^{-1} \sum_{i=-\lfloor nt_l \rfloor}^{\lfloor nt_l \rfloor} \{\lfloor nt_l \rfloor - |i|\} \varphi(i/\ell_n) \gamma(i, \boldsymbol{v}_l, \boldsymbol{v}_{l'}) = \sum_{i=-\lfloor nt_l \rfloor}^{\lfloor nt_l \rfloor} \{\lambda_n(0, t_l) - |i|/n\} \varphi(i/\ell_n) \gamma(i, \boldsymbol{v}_l, \boldsymbol{v}_{l'})$$

and converges to $t_l \sum_{i \in \mathbb{Z}} \gamma(i, \boldsymbol{v}_l, \boldsymbol{v}_{l'})$ by Assumption (A3), (26) and dominated convergence. The second sum on the right-hand side of (28) is bounded in absolute value by $4n^{-1} \sum_{i=1}^{\lfloor nt_{l'} \rfloor - 1} i\alpha_i \to 0$. Hence, the second term on the right of (27) converges to $\sum_{l,l'=1}^{q} d_l d_{l'} (t_l \wedge t_{l'}) \sum_{i \in \mathbb{Z}} \gamma(i, \boldsymbol{v}_l, \boldsymbol{v}_{l'})$. Similarly, the first term on the right of (27) converges to $\sum_{l,l'=1}^{q} c_l c_{l'} (s_l \wedge s_{l'}) \sum_{i \in \mathbb{Z}} \gamma(i, \boldsymbol{u}_l, \boldsymbol{u}_{l'})$. Thus, $\operatorname{Var}(Z_n) \to \operatorname{Var}(Z)$.

To be able to conclude that $n^{-1/2} \sum_{j=1}^{k_n} B'_{j,n}$ converges in distribution to Z, it remains to prove the Lindeberg condition of the Lindeberg-Feller theorem, i.e., that, for every $\delta > 0$,

$$n^{-1} \sum_{j=1}^{k_n} \mathrm{E}\{(B'_{j,n})^2 \mathbf{1}(|B'_{j,n}| > n^{1/2}\delta)\} = n^{-1} \sum_{j=1}^{k_n} \mathrm{E}\{B_{j,n}^2 \mathbf{1}(|B_{j,n}| > n^{1/2}\delta)\} \to 0.$$

Let $\delta > 0$ be arbitrary. Using Hölder's inequality with $p = 1 + \nu/2$, where $\nu > 0$ is to be chosen later on, and Markov's inequality, we have

$$
\begin{aligned}
&n^{-1}\sum_{j=1}^{k_n}\mathrm{E}\{B_{j,n}^2\mathbf{1}(|B_{j,n}|>n^{1/2}\delta)\}\\
&\qquad\le n^{-1}\sum_{j=1}^{k_n}\{\mathrm{E}(|B_{j,n}|^{2+\nu})\}^{2/(2+\nu)}\{\mathrm{P}(|B_{j,n}|>n^{1/2}\delta)\}^{\nu/(2+\nu)}\\
&\qquad\le n^{-1}\sum_{j=1}^{k_n}\{\mathrm{E}(|B_{j,n}|^{2+\nu})\}^{2/(2+\nu)}\{\mathrm{P}(|B_{j,n}|^{2+\nu}>n^{(2+\nu)/2}\delta^{2+\nu})\}^{\nu/(2+\nu)}\\
&\qquad\le n^{-1}\sum_{j=1}^{k_n}\{\mathrm{E}(|B_{j,n}|^{2+\nu})\}^{2/(2+\nu)}\{\mathrm{E}(|B_{j,n}|^{2+\nu})\}^{\nu/(2+\nu)}(n^{1/2}\delta)^{-\nu}\\
&\qquad\le n^{-1}\sum_{j=1}^{k_n}\mathrm{E}(|B_{j,n}|^{2+\nu})n^{-\nu/2}\delta^{-\nu}.
\end{aligned}
$$

Now, from Minkowski's inequality,

$$
\{\mathrm{E}(|B_{j,n}|^{2+\nu})\}^{1/(2+\nu)}\le\sum_{i=(j-1)(b_n+s_n)+1}^{(j-1)(b_n+s_n)+b_n}\left[\{\mathrm{E}(|Z_{i,n}|^{2+\nu})\}^{1/(2+\nu)}+\{\mathrm{E}(|Z_{i,n}^{(m)}|^{2+\nu})\}^{1/(2+\nu)}\right]=O(b_n)
$$

since, for any $\nu>0$,

$$
\max_{1\le i\le n}\mathrm{E}(|Z_{i,n}|^{2+\nu})\le\mathrm{E}\left(\left[\sum_{l=1}^{q}|c_l||\mathbf{1}(\boldsymbol{U}_0\le\boldsymbol{u}_l)-C(\boldsymbol{u}_l)|\right]^{2+\nu}\right)<\infty,
$$

and

$$
\max_{1\le i\le n}\mathrm{E}(|Z_{i,n}^{(m)}|^{2+\nu})\le\mathrm{E}\left(\left[|\xi_{1,n}^{(m)}|\sum_{l=1}^{q}|d_l||\mathbf{1}(\boldsymbol{U}_0\le\boldsymbol{v}_l)-C(\boldsymbol{v}_l)|\right]^{2+\nu}\right)<\infty.
$$

It follows that

$$
n^{-1}\sum_{j=1}^{k_n}\mathrm{E}\{B_{j,n}'^2\mathbf{1}(|B_{j,n}'|>n^{1/2}\delta)\}=O(n^{-1}k_nb_n^{2+\nu}n^{-\nu/2})=O(b_n^{1+\nu}n^{-\nu/2})=O(n^{1/2-\eta_b(1+\nu)}),
$$

which converges to zero for $\nu>1/(2\eta_b)-1$. ■

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 (1971). 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

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

where

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

Lemma B.2 (Moment inequality). *Assume that $(\boldsymbol{U}_i)_{i\in\mathbb{Z}}$ is a strictly stationary sequence whose strong mixing coefficients satisfy $\alpha_r = O(r^{-a})$, $a > 6$. 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*

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

where $\kappa > 0$ is a constant.

Proof. The proof is similar to that of Lemma 3.22 in Dehling and Philipp (2002). Fix $m \in \{1,\dots,M\}$. For any $i \in \mathbb{Z}$, let $Y_i = \mathbf{1}(\boldsymbol{U}_i \in A) - \nu(A)$. Then,

$$\begin{aligned}\mathrm{E}\left[\{\tilde{\mathbb{B}}_n^{(m)}((s,t]\times A)\}^4\right] &= \frac{1}{n^2}\sum_{i_1,i_2,i_3,i_4=\lfloor ns\rfloor+1}^{\lfloor nt\rfloor} \mathrm{E}[\xi_{i_1,n}\xi_{i_2,n}\xi_{i_3,n}\xi_{i_4,n}]\mathrm{E}[Y_{i_1}Y_{i_2}Y_{i_3}Y_{i_4}],\\ &\le \frac{4!\lambda_n(s,t)}{n}\sum_{\substack{0\le i,j,k\le\lfloor nt\rfloor-\lfloor ns\rfloor-1\\ i+j+k\le\lfloor nt\rfloor-\lfloor ns\rfloor-1}} |\mathrm{E}[\xi_{0,n}\xi_{i,n}\xi_{i+j,n}\xi_{i+j+k,n}]\mathrm{E}[Y_0Y_iY_{i+j}Y_{i+j+k}]|.\end{aligned} \tag{29}$$

On one hand, $|\mathrm{E}[\xi_{0,n}\xi_{i,n}\xi_{i+j,n}\xi_{i+j+k,n}]| \le \mathrm{E}[\xi_{0,n}^4]\,\mathbf{1}(i\le\ell_n, k\le\ell_n)$. On the other hand, by Lemma 3.11 of Dehling and Philipp (2002), for any $q \in (2a/(a-3), 4)$ and $p \in (2, a/3)$ such that $1/p + 2/q = 1$, we have

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

and

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

Proceeding as in Lemma 3.22 of Dehling and Philipp (2002), we split the sum on the right of (29) 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

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

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

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

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

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

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

where $\|\cdot\|_1$ is the Euclidean L_1-norm.

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

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

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

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

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 $\underline{s} \leq s \leq \bar{s}$ and $\bar{s} - \underline{s} \leq 1/n$. 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 $\boldsymbol{u} \in [0,1]^d$, we define $\underline{\boldsymbol{u}}_\gamma \in I_{n,\gamma}^d$ (resp. $\bar{\boldsymbol{u}}_\gamma \in I_{n,\gamma}^d$) as $\underline{\boldsymbol{u}}_\gamma = (\underline{u}_{1,\gamma},\dots,\underline{u}_{d,\gamma})$ (resp. $\bar{\boldsymbol{u}}_\gamma = (\bar{u}_{1,\gamma},\dots,\bar{u}_{d,\gamma})$).

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

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

that is,

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

and therefore

$$\tilde{\mathbb{B}}_n^{(m)}(s,\boldsymbol{u}) - \tilde{\mathbb{B}}_n^{(m)}(s,\underline{\boldsymbol{u}}_\gamma) \leq \left|\tilde{\mathbb{B}}_n^{(m)}(s,\bar{\boldsymbol{u}}_\gamma) - \tilde{\mathbb{B}}_n^{(m)}(s,\underline{\boldsymbol{u}}_\gamma)\right| \\ + K\left|\tilde{\mathbb{B}}_n(s,\bar{\boldsymbol{u}}_\gamma) - \tilde{\mathbb{B}}_n(s,\underline{\boldsymbol{u}}_\gamma)\right| + d(n^\gamma - 1)^{-1}(K + \max_{1\leq i\leq n}|Z_{i,n}^{(m)}|),$$

using the fact that C satisfies the Lipschitz condition

$$|C(\boldsymbol{u}) - C(\boldsymbol{v})| \leq \|\boldsymbol{u}-\boldsymbol{v}\|_1, \qquad \forall\, \boldsymbol{u},\boldsymbol{v} \in [0,1]^d, \tag{30}$$

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 $(s,\boldsymbol{u}) \in [0,1]^{d+1}$,

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

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

$$\left|\tilde{\mathbb{B}}_n^{(m)}(s,\boldsymbol{u}) - \tilde{\mathbb{B}}_n^{(m)}(s,\underline{\boldsymbol{u}}_\gamma)\right| \leq \left|\tilde{\mathbb{B}}_n^{(m)}(s,\bar{\boldsymbol{u}}_\gamma) - \tilde{\mathbb{B}}_n^{(m)}(s,\underline{\boldsymbol{u}}_\gamma)\right| \\ + K\left|\tilde{\mathbb{B}}_n(s,\bar{\boldsymbol{u}}_\gamma) - \tilde{\mathbb{B}}_n(s,\underline{\boldsymbol{u}}_\gamma)\right| + d(n^\gamma - 1)^{-1}(K + \max_{1\leq i\leq n}|Z_{i,n}^{(m)}|). \tag{31}$$

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 (31) to the first and the third summand on the right-hand side of the decomposition

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

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

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

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

$$\mathrm{P}\{w_\delta(\tilde{\mathbb{B}}_n^{(m)},[0,1]^{d+1} > \varepsilon\} \leq \mathrm{P}\{3w_{2\delta}(\tilde{\mathbb{B}}_n^{(m)},T_n) > \varepsilon/3\} \\ + \mathrm{P}\{2Kw_{2\delta}(\tilde{\mathbb{B}}_n,[0,1]^{d+1}) > \varepsilon/3\} + \mathrm{P}\{2d(n^\gamma-1)^{-1}(K + \max_{1\leq i\leq n}|Z_{i,n}^{(m)}|) > \varepsilon/3\}.$$

Since $a > 2+d$, we have from Dhompongsa (1984) that $\tilde{\mathbb{B}}_n \rightsquigarrow \mathbb{B}_C$ in $\ell^\infty([0,1]^{d+1})$. Because the limiting process $\mathbb{B}_C$ is Gaussian, $\tilde{\mathbb{B}}_n$ is asymptotically uniformly ρ-equicontinuous in probability, where ρ is the *standard deviation semimetric* (see van der Vaart and Wellner, 2000, page 41). Proceeding as in van der Vaart and Wellner (2000, page 226) and using the fact that C satisfies the Lipschitz condition (30), it can be verified 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 $n^{-\gamma} \max_{1\leq i\leq n} |Z_{i,n}^{(m)}| \xrightarrow{\mathrm{P}} 0$. Indeed, since $\mathrm{E}(|Z_{1,n}^{(m)}|^{1/\gamma+1}) = \int_0^\infty (1/\gamma+1)x^{1/\gamma}\mathrm{P}(|Z_{1,n}^{(m)}| > x)\mathrm{d}x < \infty$, we immediately obtain that $x^{1/\gamma}\mathrm{P}(|Z_{1,n}^{(m)}| > x) \to 0$ as $x \to +\infty$. Then, for any $\eta > 0$,

$$\mathrm{P}(n^{-\gamma} \max_{1\leq i\leq n} |Z_{i,n}^{(m)}| > \eta) \leq n\mathrm{P}(|Z_{1,n}^{(m)}| > \eta n^\gamma) \to 0.$$

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, 2000, 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 B.2 together with Lemma 2 of Balacheff and Dupont (1980) (see also Bickel and Wichura, 1971, 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 $\mu = \lambda^{(1)} \otimes \nu + \lambda^{(d)}$, where $\lambda^{(d)}$ denotes the d-dimensional Lebesgue measure. Now, it is easy to verify that $q \in (2a/(a-3), 4)$ if and only if $1 - 2/q \in (3/a, 1/2)$. Consequently, it is possible to choose $q \in (2a/(a-3), 4)$ such that $1-2/q = 3/a+\epsilon$ for some small $\epsilon \in (0, 4/q-1)$. Next, let $\beta = 1+\epsilon/2$. Clearly, $\beta \in (1, 4/q)$. 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 B.2,

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

For the above choice of β, we have $\beta - 2/q = 3/a + \epsilon/2$. Because $3/a < 2/(2+d)$ from the assumption on the mixing rate, it is possible to choose $\epsilon \in (0, 4/q-1)$ and $\gamma > 0$ (the parameter involved in the grid $I_{n,\gamma}^d$) small enough such that $3/a+\epsilon/2 < 2/(2+d+2d\gamma)$. For the aforementioned parameter choices, $(\beta-2/q)(1+d/2+d\gamma)-1 < 0$, which implies that $n^{(\beta-2/q)(1+d/2+d\gamma)-1} \leq 1$ for all $n \geq 1$.

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

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

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

$$\mathrm{P}\left\{|\tilde{\mathbb{B}}_n^{(m)}((s,t]\times A)| \geq \varepsilon\right\} \leq \varepsilon^{-4}\mu((s,t]\times A)^\beta.$$

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

$$\tilde{\mu}_n(\{(s, \boldsymbol{u})\}) = \begin{cases} 0 & \text{if } s \wedge u_1 \wedge \cdots \wedge u_d = 0, \\ \mu((s', s] \times (u'_1, u_1] \times \cdots \times (u'_d, u_d]) & \text{otherwise,} \end{cases}$$

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

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

We shall now conclude by an application of Lemma 2 of Balacheff and Dupont (1980). 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 $\delta'_n \geq \max\{\delta_n, 1/\lfloor n^{1/2+\gamma} \rfloor\}$. Applying Lemma 2 of Balacheff and Dupont (1980) (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 $\|\cdot\|_2 \leq \|\cdot\|_1$, we obtain that, for any $\varepsilon > 0$, there exists a constant $\lambda > 0$ depending on ε, β and d, such that

$$\begin{aligned} \mathrm{P}\{w_{\delta_n}(\tilde{\mathbb{B}}_n^{(m)}, T_n) > \varepsilon\} \leq & \mathrm{P}\{w_{\delta'_n}(\tilde{\mathbb{B}}_n^{(m)}, T_n) > \varepsilon\} \leq \lambda \tilde{\mu}_n(T_n) \\ & \times \Big[\max\{ \sup_{\substack{s,t \in I_n \\ |s-t| \leq 3\delta'_n}} |\tilde{\mu}_n(\{0, ..., s\} \times I_{n,\gamma}^d) - \tilde{\mu}_n(\{0, ..., t\} \times I_{n,\gamma}^d)|, \\ & \qquad \sup_{\substack{u,v \in I_{n,\gamma} \\ |u-v| \leq 3\delta'_n}} |\tilde{\mu}_n(I_n \times \{0, \dots, u\} \times I_{n,\gamma}^{d-1}) - \tilde{\mu}_n(I_n \times \{0, \dots, v\} \times I_{n,\gamma}^{d-1})|, \\ & \qquad \dots, \\ & \qquad \sup_{\substack{u,v \in I_{n,\gamma} \\ |u-v| \leq 3\delta'_n}} |\tilde{\mu}_n(I_n \times I_{n,\gamma}^{d-1} \times \{0, \dots, u\}) - \tilde{\mu}_n(I_n \times I_{n,\gamma}^{d-1} \times \{0, \dots, v\})|\} \Big]^{\beta-1}, \end{aligned}$$

that is,

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

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

Let us now consider the general case. Let $Z_{i,n}^+ = \max(\xi_{i,n}, 0)$, $Z_{i,n}^- = \max(-\xi_{i,n}, 0)$, $K^+ = \mathrm{E}(Z_{0,n}^+)$ 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

$$\xi^{(m)}_{i,n} = Z^+_{i,n} - Z^-_{i,n} = Z^+_{i,n} - K^+ - (Z^-_{i,n} - K^-) = \xi^{(m),+}_{i,n} - \xi^{(m),-}_{i,n}.$$

Setting

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

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

Proof of Theorem 3.1. Weak convergence of the finite-dimensional distributions is established in Lemma B.1. Asymptotic tightness of $\tilde{\mathbb{B}}_n$ is a consequence of the the weak convergence of $\tilde{\mathbb{B}}_n$ to $\mathbb{B}_C$ in $\ell^\infty([0,1]^d)$, which follows from Dhompongsa (1984) as $a > 2 + d$. From Lemma B.3, we have that, for any $m \in \{1, \dots, M\}$, $\tilde{\mathbb{B}}^{(m)}_n$ 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 B.1, we have, for instance from Theorem 2.1 in Kosorok (2008), that, for any $m \in \{1, \dots, M\}$, $\tilde{\mathbb{B}}^{(m)}_n \rightsquigarrow \mathbb{B}^{(m)}_C$ in $\ell^\infty([0,1]^d)$, which implies asymptotic tightness of $\tilde{\mathbb{B}}^{(m)}_n$. The proof is complete as marginal asymptotic tightness implies joint asymptotic tightness. ∎

C Proof of Proposition 4.2

Let us first introduce some additional notation similar to those used in Appendix A. Let $\tilde{H}_n$ denote the empirical c.d.f. of the unobservable sample $\boldsymbol{U}_1, \dots, \boldsymbol{U}_n$ and let $\tilde{H}_{n,1}, \dots, \tilde{H}_{n,d}$ denote the corresponding marginal c.d.f.s. Also, for any $j \in \{1, \dots, d\}$, let

$$\tilde{H}^{-1}_{n,j}(u) = \inf\{v \in [0,1] : \tilde{H}_{n,j}(v) \geq u\}, \qquad u \in [0,1],$$

and let

$$\tilde{H}^{-1}_n(\boldsymbol{u}) = \left(\tilde{H}^{-1}_{n,1}(u_1), \dots, \tilde{H}^{-1}_{n,d}(u_d)\right), \qquad \boldsymbol{u} \in [0,1]^d.$$

Finally, for any $m \in \{1, \dots, M\}$, let

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

The proof of Proposition 4.2 is based on two lemmas given after the proof.

Proof of Propositon 4.2. For any $m \in \{1, \dots, M\}$ and $(s, t, \boldsymbol{u}) \in \Delta \times [0,1]^d$, let

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

From Theorem 3.1 and the continuous mapping theorem, we then immediately obtain that

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

in $\{\ell^\infty(\Delta \times [0,1]^d)\}^{M+1}$, where $\tilde{\mathbb{C}}_n$ is defined in (3). Proceeding for instance as in Segers (2012, proof of Proposition 3.2) or as in Kojadinovic and Rohmer (2012, proof of Lemma 6), we obtain that, for any $m \in \{1, \dots, M\}$,

$$\sup_{(s,\boldsymbol{u}) \in [0,1]^{d+1}} \left[\left| \dot{C}_{j,n}(\boldsymbol{u}) - \dot{C}_j(\boldsymbol{u}) \right| \left| \tilde{\mathbb{B}}_n^{(m)}(s, \boldsymbol{u}) \right| \right] \xrightarrow{\mathrm{P}} 0$$

using Condition 4.1 and the fact that $\tilde{\mathbb{B}}_n^{(m)} \rightsquigarrow \mathbb{B}_C^{(m)}$ in $\ell^\infty([0,1]^{d+1})$. The desired result is finally a mere consequence of Corollary 2.4, Condition 4.1, and Lemmas C.1 and C.2 below. ■

Lemma C.1. *For any* $m \in \{1, \dots, M\}$,

$$\sup_{(s,\boldsymbol{u}) \in [0,1]^{d+1}} \left| \hat{\mathbb{B}}_n^{(m)}(s, \boldsymbol{u}) - \check{\mathbb{B}}_n^{(m)}\{\tilde{H}_n^{-1}(\boldsymbol{u})\} \right| \xrightarrow{\mathrm{P}} 0$$

where $\hat{\mathbb{B}}_n^{(m)}$ *is defined in* (6).

Proof. The proof is similar to that of Lemma 1 of Kojadinovic and Rohmer (2012). We here additionally use the fact that $\mathrm{E}(|\xi_{0,n}^{(m)}|^3) = \int_0^\infty 3x^2 \mathrm{P}(|\xi_{0,n}^{(m)}| > x)\mathrm{d}x < \infty$, which implies that $n^{-1/2} \max_{1 \le i \le n} |\xi_{i,n}^{(m)}| \xrightarrow{\mathrm{P}} 0$. ■

Lemma C.2. *Assume that* $\ell_n = O(n^{1/2-\varepsilon})$ *for some* $0 < \varepsilon < 1/2$ *and that* $(\boldsymbol{U}_i)_{i \in \mathbb{Z}}$ *is a strictly stationary sequence whose strong mixing coefficients satisfy* $\alpha_r = O(r^{-a})$, $a > 3 + 3d/2$. *Then, for any* $m \in \{1, \dots, M\}$,

$$\sup_{(s,\boldsymbol{u}) \in [0,1]^{d+1}} \left| \tilde{\mathbb{B}}_n^{(m)}(s, \boldsymbol{u}) - \check{\mathbb{B}}_n^{(m)}\{\tilde{H}_n^{-1}(\boldsymbol{u})\} \right| \xrightarrow{\mathrm{P}} 0.$$

where $\tilde{\mathbb{B}}_n^{(m)}$ *is defined in* (5).

Proof. Let us first show that

$$J_n = \sup_{(s,\boldsymbol{u}) \in [0,1]^{d+1}} \left| \tilde{\mathbb{B}}_n^{(m)}(s, \boldsymbol{u}) - \check{\mathbb{B}}_n^{(m)}(s, \boldsymbol{u}) \right| \xrightarrow{\mathrm{P}} 0.$$

Writing $Z_k = \sum_{i=1}^k \xi_{i,n}^{(m)}$ for $k \in \{1, \dots, n\}$, we have $J_n \le K_n \times L_n$, where

$$K_n = n^{-1} \max_{1 \le k \le n} |Z_k| \quad \text{and} \quad L_n = n^{1/2} \sup_{\boldsymbol{u} \in [0,1]^d} \left| \tilde{H}_n(\boldsymbol{u}) - C(\boldsymbol{u}) \right|.$$

Let $\nu > 1$. Then,

$$\mathrm{E}\left\{ \left(n^{-1} \max_{1 \le k \le n} |Z_k| \right)^\nu \right\} = n^{-\nu} \mathrm{E}\left(\max_{1 \le k \le n} |Z_k|^\nu \right) \le n^{-\nu} \mathrm{E}\left(\sum_{k=1}^n |Z_k|^\nu \right) \le n^{1-\nu} \max_{1 \le k \le n} \mathrm{E}\left(|Z_k|^\nu\right).$$

For $\nu = 4$, using the fact that the sequence $(\xi_{i,n}^{(m)})_{i\in\mathbb{Z}}$ is strictly stationary and ℓ_n-dependent, we obtain

$$\begin{aligned}
\mathrm{E}\left(|Z_k|^4\right) &= \sum_{i_1,i_2,i_3,i_4=1}^{k} \mathrm{E}\left(\xi_{i_1,n}^{(m)}\xi_{i_2,n}^{(m)}\xi_{i_3,n}^{(m)}\xi_{i_4,n}^{(m)}\right) \le 4! \sum_{1\le i_1\le i_2\le i_3\le i_4\le n} \left|\mathrm{E}\left(\xi_{i_1,n}^{(m)}\xi_{i_2,n}^{(m)}\xi_{i_3,n}^{(m)}\xi_{4,n}^{(m)}\right)\right| \\
&\le 4!\, n \sum_{1\le i_2\le i_3\le i_4\le n} \left|\mathrm{E}\left(\xi_{1,n}^{(m)}\xi_{i_2,n}^{(m)}\xi_{i_3,n}^{(m)}\xi_{i4,n}^{(m)}\right)\right| \le 4!\, n \sum_{i_2=1}^{\ell_n}\sum_{i_3=i_2}^{n}\sum_{i_4=i_3}^{\ell_n} \left|\mathrm{E}\left(\xi_{1,n}^{(m)}\xi_{i_2,n}^{(m)}\xi_{i_3,n}^{(m)}\xi_{i_4,n}^{(m)}\right)\right|.
\end{aligned}$$

From Cauchy-Swcharz's inequality, $|\mathrm{E}(\xi_{1,n}^{(m)}\xi_{i_2,n}^{(m)}\xi_{i_3,n}^{(m)}\xi_{i_4,n}^{(m)})| \le \mathrm{E}\{(\xi_{1,n}^{(m)})^4\}$, and therefore $\max_{1\le k\le n} \mathrm{E}\left(|Z_k|^4\right) = O(n^2\ell_n^2)$. It follows that

$$\mathrm{E}\left\{\left(n^{-1}\max_{1\le k\le n}|Z_k|\right)^4\right\} = O(n^{-2\varepsilon}) \to 0,$$

which implies that $K_n \xrightarrow{\mathrm{P}} 0$. As $a > 2+d$, from Dhompongsa (1984), $L_n = O_{\mathrm{P}}(1)$, which implies that $J_n \xrightarrow{\mathrm{P}} 0$.

It remains to show that

$$\sup_{(s,\boldsymbol{u})\in[0,1]^{d+1}} \left|\check{\mathbb{B}}_n^{(m)}(s,\boldsymbol{u}) - \check{\mathbb{B}}_n^{(m)}\{s,\tilde{H}_n^{-1}(\boldsymbol{u})\}\right| \xrightarrow{\mathrm{P}} 0.$$

The assumptions of Theorem 3.1 being satisfied, $\tilde{\mathbb{B}}_n^{(m)} \rightsquigarrow \mathbb{B}_C^{(m)}$ in $\ell^\infty([0,1]^{d+1})$. Since $J_n \xrightarrow{\mathrm{P}} 0$, we obtain that $\check{\mathbb{B}}_n^{(m)} \rightsquigarrow \mathbb{B}_C^{(m)}$ in $\ell^\infty([0,1]^{d+1})$. The previous result together with the Lipschitz continuity of C in (30) implies the asymptotic uniform $\|\cdot\|_1$-equicontinuity in probability of $\check{\mathbb{B}}_n^{(m)}$. The desired result finally follows from the fact that $\sup_{u\in[0,1]}|\tilde{H}_{n,j}^{-1}(u)-u| = \sup_{u\in[0,1]}|\tilde{H}_{n,j}(u)-u| \xrightarrow{\mathrm{P}} 0$ for all $j\in\{1,\dots,d\}$, the latter convergence being a consequence of the fact that $L_n = O_{\mathrm{P}}(1)$. ■

D Proofs of Propositions 5.1 and 5.2

Proof of Propositions 5.1. From (25) and the fact that $\alpha_r = O(r^{-a})$, $a > 3$, we have that

$$\sup_{\boldsymbol{u},\boldsymbol{v}\in[0,1]^d}\sum_{k\in\mathbb{Z}} k^2|\gamma(k,\boldsymbol{u},\boldsymbol{v})| \le 4\sum_{k\in\mathbb{Z}} k^2\alpha_{|k|} = 8\sum_{k=1}^{\infty} k^2\alpha_k < \infty.$$

Furthermore,

$$\begin{aligned}
\mathrm{E}\{\tilde{\sigma}_n(\boldsymbol{u},\boldsymbol{v})\} &= \frac{1}{n}\sum_{i,j=1}^{n}\varphi(|i-j|/\ell_n)\gamma(j-i,\boldsymbol{u},\boldsymbol{v}) = \frac{1}{n}\sum_{k=-n}^{n}(n-|k|)\varphi(k/\ell_n)\gamma(k,\boldsymbol{u},\boldsymbol{v}) \\
&= \sum_{k=-\ell_n}^{\ell_n}\varphi(k/\ell_n)\gamma(k,\boldsymbol{u},\boldsymbol{v}) + s_n(\boldsymbol{u},\boldsymbol{v}),
\end{aligned}$$

where $s_n(\boldsymbol{u},\boldsymbol{v}) = -\frac{1}{n}\sum_{k=-\ell_n}^{\ell_n} |k|\varphi(k/\ell_n)\gamma(k,\boldsymbol{u},\boldsymbol{v})$. Using the fact that φ is bounded, we get

$$\sup_{\boldsymbol{u},\boldsymbol{v}\in[0,1]^d} |s_n(\boldsymbol{u},\boldsymbol{v})| \le \frac{8}{n}\sum_{k=1}^{\infty} |k|\alpha_k = O(n^{-1}) = o(\ell_n^{-2}).$$

Now, from the second-order mean value theorem, for any $n \ge 1$ and $k \ne 0$, $|k| \le \ell_n$, there exists $\zeta_{k,n}$ strictly between 0 and k/ℓ_n such that

$$\begin{aligned}\varphi(k/\ell_n) &= \varphi(0) + \varphi'(0)k/\ell_n + \varphi''(\zeta_{k,n})k^2/(2\ell_n^2) = 1 + \varphi''(\zeta_{k,n})k^2/(2\ell_n^2)\\ &= 1 + \varphi''(0)k^2/(2\ell_n^2) + \{\varphi''(\zeta_{k,n}) - \varphi''(0)\}k^2/(2\ell_n^2).\end{aligned}$$

Hence,

$$\sum_{k=-\ell_n}^{\ell_n} \varphi(k/\ell_n)\gamma(k,\boldsymbol{u},\boldsymbol{v}) = \sum_{k=-\ell_n}^{\ell_n} \gamma(k,\boldsymbol{u},\boldsymbol{v}) + \frac{\varphi''(0)}{2\ell_n^2}\sum_{k=-\ell_n}^{\ell_n} k^2\gamma(k,\boldsymbol{u},\boldsymbol{v}) + t_n(\boldsymbol{u},\boldsymbol{v}),$$

where

$$t_n(\boldsymbol{u},\boldsymbol{v}) = \frac{1}{2\ell_n^2}\sum_{k=-\ell_n}^{\ell_n} \{\varphi''(\zeta_{k,n}) - \varphi''(0)\}k^2\gamma(k,\boldsymbol{u},\boldsymbol{v}).$$

Furthermore,

$$2\ell_n^2 \sup_{\boldsymbol{u},\boldsymbol{v}\in[0,1]^d} |t_n(\boldsymbol{u},\boldsymbol{v})| \le 8\sum_{k=1}^{\ell_n} |\varphi''(\zeta_{k,n}) - \varphi''(0)|k^2\alpha_k \to 0$$

by dominated convergence. Indeed, pointwise in k, $|\varphi''(\zeta_{k,n}) - \varphi''(0)|k^2\alpha_k \to 0$ and, for any $k \ge 1$, $|\varphi''(\zeta_{k,n}) - \varphi''(0)|k^2\alpha_k \le 2\sup_{x\in[-1,1]}|\varphi''(x)|k^2\alpha_k$ which converges absolutely. Therefore, $\sup_{\boldsymbol{u},\boldsymbol{v}\in[0,1]^d}|t_n(\boldsymbol{u},\boldsymbol{v})| = o(\ell_n^{-2})$. The result finally follows with $r_{n,1}(\boldsymbol{u},\boldsymbol{v}) = s_n(\boldsymbol{u},\boldsymbol{v}) + t_n(\boldsymbol{u},\boldsymbol{v}) + \sum_{|k|>\ell_n}\gamma(k,\boldsymbol{u},\boldsymbol{v})$ and the fact that

$$\ell_n^2 \sup_{\boldsymbol{u},\boldsymbol{v}\in[0,1]^d}\left|\sum_{|k|>\ell_n}\gamma(k,\boldsymbol{u},\boldsymbol{v})\right| \le 8\ell_n^2\sum_{k=\ell_n+1}^{\infty}\alpha_k \le 8\sum_{k=\ell_n+1}^{\infty}k^2\alpha_k \to 0.$$

■

Proof of Propositions 5.2. The proof is based on an appropriate version of Eq. (A.1) of Künsch (1989). For any $i \in \{1,\dots,n\}$ and $\boldsymbol{u} \in [0,1]^d$, let $Y_i(\boldsymbol{u}) = \mathbf{1}(\boldsymbol{U}_i \le \boldsymbol{u}) - C(\boldsymbol{u})$. Then, let us show that, for any $\boldsymbol{u},\boldsymbol{v},\boldsymbol{w},\boldsymbol{z} \in [0,1]^d$ and any $i,j,k,l \in \{0,\dots,n\}$ such that $1 \le i+j+k+l \le n$,

$$\begin{aligned}|\mathrm{E}\{Y_i(\boldsymbol{u})Y_{i+j}(\boldsymbol{v})Y_{i+j+k}(\boldsymbol{w})Y_{i+j+k+l}(\boldsymbol{z})\} &- \mathrm{E}\{Y_i(\boldsymbol{u})Y_{i+j}(\boldsymbol{v})\}\mathrm{E}\{Y_{i+j+k}(\boldsymbol{w})Y_{i+j+k+l}(\boldsymbol{z})\}\\ &- \mathrm{E}\{Y_i(\boldsymbol{u})Y_{i+j+k}(\boldsymbol{w})\}\mathrm{E}\{Y_{i+j}(\boldsymbol{v})Y_{i+j+k+l}(\boldsymbol{z})\}\\ &- \mathrm{E}\{Y_i(\boldsymbol{u})Y_{i+j+k+l}(\boldsymbol{z})\}\mathrm{E}\{Y_{i+j}(\boldsymbol{v})Y_{i+j+k}(\boldsymbol{w})\}| \le 16\alpha_{\max\{j,k,l\}}. \quad (32)\end{aligned}$$

Fix $\boldsymbol{u},\boldsymbol{v},\boldsymbol{w},\boldsymbol{z} \in [0,1]^d$ and $i,j,k,l \in \{0,\dots,n\}$ such that $1 \le i+j+k+l \le n$, and let A, B_1, B_2 and B_3 denote the four terms appearing in the absolute value on the left of (32). Assume first that $j = \max\{j,k,l\}$. Then, using Lemma 3.9 in Dehling and Philipp (2002) as well as the facts that $\mathrm{E}\{Y_1(\cdot)\} = 0$ and $|Y_1(\cdot)| \leq 1$, we obtain that $|A| \leq 4\alpha_j$, $|B_1| \leq 4\alpha_j$, $|B_2| \leq 4\alpha_{j+k} \leq 4\alpha_j$ and $|B_3| \leq 4\alpha_{j+k+l} \leq 4\alpha_j$, which, combined with the triangle inequality, implies (32). Assume now that $k = \max\{j,k,l\}$. Then, $|A - B_1| \leq 4\alpha_k$, $|B_2| \leq 4\alpha_{j+k} \leq \alpha_k$ and $|B_3| \leq 4\alpha_{j+k+l} \leq \alpha_k$, and (32) holds. Finally, suppose that $l = \max\{j,k,l\}$. Then, $|A| \leq 4\alpha_l$, $|B_1| \leq 4\alpha_l$, $|B_2| \leq 4\alpha_{k+l} \leq 4\alpha_l$ and $|B_3| \leq 4\alpha_{j+k+l} \leq 4\alpha_j$ and (32) follows again.

Fix $\boldsymbol{u},\boldsymbol{v} \in [0,1]^d$. Starting from the expression of $\tilde{\sigma}_n(\boldsymbol{u},\boldsymbol{v})$, we have that

$$\mathrm{Var}\{\tilde{\sigma}_n(\boldsymbol{u},\boldsymbol{v})\} = \frac{1}{n^2}\sum_{i,j,k,l=1}^{n} \varphi\{(i-j)/\ell_n\}\varphi\{(k-l)/\ell_n\}\mathrm{Cov}\{Y_i(\boldsymbol{u})Y_j(\boldsymbol{v}), Y_k(\boldsymbol{u})Y_l(\boldsymbol{v})\}.$$

Let us now write

$$\begin{aligned}\mathrm{Var}\{\tilde{\sigma}_n(\boldsymbol{u},\boldsymbol{v})\} = \frac{1}{n^2}\sum_{i,j,k,l=1}^{n} &\big[\varphi\{(i-j)/\ell_n\}\varphi\{(k-l)/\ell_n\}\\ &\times\{\gamma(k-i,\boldsymbol{u},\boldsymbol{u})\gamma(l-j,\boldsymbol{v},\boldsymbol{v}) + \gamma(l-i,\boldsymbol{u},\boldsymbol{v})\gamma(k-j,\boldsymbol{v},\boldsymbol{u})\}\big] + s_n(\boldsymbol{u},\boldsymbol{v}),\end{aligned}$$

and show that $\sup_{\boldsymbol{u},\boldsymbol{v}\in[0,1]^d} |s_n(\boldsymbol{u},\boldsymbol{v})| = o(\ell_n/n)$. Using the fact that φ is bounded by one, we have

$$\begin{aligned}\sup_{\boldsymbol{u},\boldsymbol{v}\in[0,1]^d} |s_n(\boldsymbol{u},\boldsymbol{v})| \leq \frac{1}{n^2}\sum_{i,j,k,l=1}^{n}\sup_{\boldsymbol{u},\boldsymbol{v}\in[0,1]^d}&\big|\mathrm{Cov}\{Y_i(\boldsymbol{u})Y_j(\boldsymbol{v}), Y_k(\boldsymbol{u})Y_l(\boldsymbol{v})\}\\ &-\gamma(k-i,\boldsymbol{u},\boldsymbol{u})\gamma(l-j,\boldsymbol{v},\boldsymbol{v}) - \gamma(l-i,\boldsymbol{u},\boldsymbol{v})\gamma(k-j,\boldsymbol{v},\boldsymbol{u})\big|,\end{aligned}$$

which implies that

$$\begin{aligned}\sup_{\boldsymbol{u},\boldsymbol{v}\in[0,1]^d} |s_n(\boldsymbol{u},\boldsymbol{v})| &\leq \frac{1}{n^2}\sum_{i,j,k,l=1}^{n}\sup_{\boldsymbol{u},\boldsymbol{v},\boldsymbol{w},\boldsymbol{z}\in[0,1]^d}\big|\mathrm{E}\{Y_i(\boldsymbol{u})Y_j(\boldsymbol{v})Y_k(\boldsymbol{w})Y_l(\boldsymbol{z})\}\\ &-\mathrm{E}\{Y_i(\boldsymbol{u})Y_j(\boldsymbol{v})\}\mathrm{E}\{Y_k(\boldsymbol{w})Y_l(\boldsymbol{z})\} - \mathrm{E}\{Y_i(\boldsymbol{u})Y_k(\boldsymbol{w})\}\mathrm{E}\{Y_j(\boldsymbol{v})Y_l(\boldsymbol{z})\}\\ &-\mathrm{E}\{Y_i(\boldsymbol{u})Y_l(\boldsymbol{z})\}\mathrm{E}\{Y_j(\boldsymbol{v})Y_k(\boldsymbol{w})\}\big|\\ &= \frac{4!}{n^2}\sum_{\substack{i,j,k,l=0\\1\leq i+j+k+l\leq n}}^{n}\sup_{\boldsymbol{u},\boldsymbol{v},\boldsymbol{w},\boldsymbol{z}\in[0,1]^d}\big|\mathrm{E}\{Y_i(\boldsymbol{u})Y_{i+j}(\boldsymbol{v})Y_{i+j+k}(\boldsymbol{w})Y_{i+j+k+l}(\boldsymbol{z})\}\\ &-\mathrm{E}\{Y_i(\boldsymbol{u})Y_{i+j}(\boldsymbol{v})\}\mathrm{E}\{Y_{i+j+k}(\boldsymbol{w})Y_{i+j+k+l}(\boldsymbol{z})\}\\ &-\mathrm{E}\{Y_i(\boldsymbol{u})Y_{i+j+k}(\boldsymbol{w})\}\mathrm{E}\{Y_{i+j}(\boldsymbol{v})Y_{i+j+k+l}(\boldsymbol{z})\}\\ &-\mathrm{E}\{Y_i(\boldsymbol{u})Y_{i+j+k+l}(\boldsymbol{z})\}\mathrm{E}\{Y_{i+j}(\boldsymbol{v})Y_{i+j+k}(\boldsymbol{w})\}\big|.\end{aligned}$$

Then, we obtain from (32),

$$\begin{aligned}\sup_{\boldsymbol{u},\boldsymbol{v}\in[0,1]^d} |s_n(\boldsymbol{u},\boldsymbol{v})| \leq \frac{4!}{n^2}\sum_{\substack{i,j,k,l=0\\1\leq i+j+k+l\leq n}}^{n} 16\alpha_{\max\{j,k,l\}} &\leq \frac{384}{n}\sum_{\substack{j,k,l=0\\1\leq j+k+l\leq n}}^{n}\alpha_{\max\{j,k,l\}}\\ &\leq 3\times\frac{384}{n}\sum_{j=1}^{n}\sum_{k,l=1}^{j}\alpha_j \leq \frac{1152}{n}\sum_{j=1}^{\infty} j^2\alpha_j = o(\ell_n/n).\end{aligned}$$

Hence, we have that

$$\operatorname{Var}\{\tilde{\sigma}_n(\boldsymbol{u},\boldsymbol{v})\} = \frac{1}{n^2}\sum_{i,j,k,l=1}^{n} \varphi\{(i-j)/\ell_n\}\varphi\{(k-l)/\ell_n\}\gamma(k-i,\boldsymbol{u},\boldsymbol{u})\gamma(l-j,\boldsymbol{v},\boldsymbol{v})$$
$$+\frac{1}{n^2}\sum_{i,j,k,l=1}^{n} \varphi\{(i-j)/\ell_n\}\varphi\{(k-l)/\ell_n\}\gamma(l-i,\boldsymbol{u},\boldsymbol{v})\gamma(k-j,\boldsymbol{v},\boldsymbol{u}) + s_n(\boldsymbol{u},\boldsymbol{v}). \quad (33)$$

Now, for any $\boldsymbol{u},\boldsymbol{v}\in[0,1]^d$, let

$$a_n(\boldsymbol{u},\boldsymbol{v}) = \frac{1}{n}\sum_{p,q,r=-\infty}^{\infty} \varphi(p/\ell_n)\varphi\{(q-r)/\ell_n\}\gamma(q,\boldsymbol{u},\boldsymbol{u})\gamma(r-p,\boldsymbol{v},\boldsymbol{v}).$$

Our next goal is to show that the first term on the right of (33) is equal to $a_n(\boldsymbol{u},\boldsymbol{v})$ plus a term that is $o(\ell_n/n)$ uniformly in $\boldsymbol{u},\boldsymbol{v}\in[0,1]^d$. For any $\boldsymbol{u},\boldsymbol{v}\in[0,1]^d$, let

$$a_{i,n}(\boldsymbol{u},\boldsymbol{v}) = \frac{1}{n}\sum_{p=1-i}^{n-i}\sum_{q=1-i}^{n-i}\sum_{r=1-i}^{n-i} \varphi(p/\ell_n)\varphi\{(q-r)/\ell_n\}\gamma(q,\boldsymbol{u},\boldsymbol{u})\gamma(r-p,\boldsymbol{v},\boldsymbol{v}),$$

and notice that, by setting $j=i+p$, $k=i+q$ and $l=i+r$, the first term on the right of (33) can be written as $n^{-1}\sum_{i=1}^n a_{i,n}(\boldsymbol{u},\boldsymbol{v})$. Hence, it suffices to show that

$$\sup_{\boldsymbol{u},\boldsymbol{v}\in[0,1]^d} n^{-1}\sum_{i=1}^{n}|a_{i,n}(\boldsymbol{u},\boldsymbol{v})-a_n(\boldsymbol{u},\boldsymbol{v})| = o(\ell_n/n).$$

For any $i\in\{1,\dots,n\}$, let $A_{i,n}=\{1-i,2-i,\dots,n-i-1,n-i\}$, and let

$$\psi_n(p,q,r) = \varphi(p/\ell_n)\varphi\{(q-r)/\ell_n\}\alpha_{|q|}\alpha_{|r-p|}, \qquad p,q,r\in\mathbb{Z}.$$

Then, using (25),

$$\sup_{\boldsymbol{u},\boldsymbol{v}\in[0,1]^d}|a_{i,n}(\boldsymbol{u},\boldsymbol{v})-a_n(\boldsymbol{u},\boldsymbol{v})| \leq \text{const}\times\frac{1}{n}\sum_{p,q,r}\psi_n(p,q,r),$$

where the sum on the right is over all $p,q,r\in\mathbb{Z}$ such that at least one of the indices does not lie in $A_{i,n}$. Next, we can bound the right-hand side of the last display by the sum of three sums $B_{n,i,1}+B_{n,i,2}+B_{n,i,3}$, where the first one is over $p\in\mathbb{Z}\setminus A_{i,n}$ and $q,r\in\mathbb{Z}$, the second one over $q\in\mathbb{Z}\setminus A_{i,n}$ and $p,r\in\mathbb{Z}$ and the last one over $r\in\mathbb{Z}\setminus A_{i,n}$ and $p,q\in\mathbb{Z}$. We shall estimate each of them individually and have to show that $n^{-1}\sum_{i=1}^n B_{n,i,k}=o(\ell_n/n)$ for $k\in\{1,2,3\}$. Regarding $B_{n,i,1}$, we have

$$B_{n,i,1}\leq\frac{\text{const}}{n}\sum_{p=n-i+1}^{\infty}\sum_{q,r\in\mathbb{Z}}\psi_n(p,q,r)+\frac{\text{const}}{n}\sum_{p=-\infty}^{-i}\sum_{q,r\in\mathbb{Z}}\psi_n(p,q,r) =: B'_{n,i,1}+B''_{n,i,1}.$$

Now, $B'_{n,i,1}$ is equal to 0 if $n-i+1>\ell_n$, that is, if $i\leq n-\ell_n$, and is bounded by a constant multiple of $\{\ell_n-(n-i)\}/n\leq\ell_n/n$ for $i>n-\ell_n$. Similarly, $B''_{n,i,1}$ is equal to 0 if $i > \ell_n$ and is bounded by a constant multiple of $\{\ell_n - i + 1\}/n \le \ell_n/n$ for $i \le \ell_n$. This implies that

$$\frac{1}{n}\sum_{i=1}^{n} B_{n,i,1} = \frac{1}{n}\left\{\sum_{i=n-\ell_n+1}^{n} B'_{n,i,1} + \sum_{i=1}^{\ell_n} B''_{n,i,1}\right\} = O\{(\ell_n/n)^2\} = o(\ell_n/n).$$

Next, let us consider $B_{n,i,2}$, and, as previously, split the sum into two sums $B'_{n,i,2} + B''_{n,i,2}$, where the first considers the summands with $q \ge n - i + 1$ and the second those with $q \le -i$. Regarding $B'_{n,i,2}$, we distinguish two cases. On one hand, when $i \le n - 3\ell_n$,

$$\begin{aligned} B'_{n,i,2} &= \frac{\text{const}}{n}\sum_{q=n-i+1}^{\infty}\sum_{r=q-\ell_n}^{q+\ell_n}\sum_{p=-\ell_n}^{\ell_n}\psi_n(p,q,r) \le \frac{\text{const}}{n}\sum_{q=3\ell_n}^{\infty}\alpha_q\sum_{r=q-\ell_n}^{q+\ell_n}\sum_{p=-\ell_n}^{\ell_n}\alpha_{|r-p|} \\ &\le \frac{\text{const}}{n}\times\ell_n^2\times\alpha_{\ell_n}\times\sum_{q=3\ell_n}^{\infty}\alpha_q = O(n^{-1}\ell_n^{2-a})\times o(1) = o(\ell_n/n), \end{aligned}$$

where we used the fact that φ is bounded by one, the fact that $|r-p| \ge r-p \ge q-2\ell_n \ge \ell_n$ (which implies that $\alpha_{|r-p|} \le \alpha_{\ell_n}$) and the assumption on the mixing rate. On the other hand, when $i > n - 3\ell_n$,

$$B'_{n,i,2} \le \frac{\text{const}}{n}\sum_{q=1}^{\infty}\alpha_q\sum_{r=q-\ell_n}^{q+\ell_n}\sum_{p=-\infty}^{\infty}\alpha_{|r-p|} = O(\ell_n/n).$$

Therefore,

$$\frac{1}{n}\sum_{i=1}^{n} B'_{n,i,2} = \frac{n-3\ell_n}{n}o(\ell_n/n) + \frac{3\ell_n}{n}O(\ell_n/n) = o(\ell_n/n).$$

The quantity $B''_{n,i,2}$ can be treated similarly. It remains therefore to consider $B_{n,i,3}$ which we split again into two sums $B'_{n,i,3} + B''_{n,i,3}$, where the first considers the summands with $r \ge n - i + 1$ and the second those with $r \le -i$. On one hand, when $i \le n - 2\ell_n + 1$,

$$B'_{n,i,3} = \frac{\text{const}}{n}\sum_{r=n-i+1}^{\infty}\sum_{q=r-\ell_n}^{r+\ell_n}\sum_{p=-\ell_n}^{\ell_n}\psi_n(p,q,r) \le \frac{\text{const}}{n}\times\alpha_{\ell_n}\times\ell_n\times\sum_{r=n-i+1}^{\infty}\sum_{p=-\ell_n}^{\ell_n}\alpha_{r-p},$$

where we used the fact that $|q| \ge q \ge r - \ell_n \ge \ell_n$. The double sum on the right-hand side can be bounded by $(2\ell_n + 1)\sum_{s=\ell_n}^{\infty}\alpha_s = o(\ell_n)$, whence $B'_{n,i,3} = o(\ell_n/n)$ by similar arguments as before. On the other hand, when $i > n - 2\ell_n + 1$,

$$B'_{n,i,3} \le \frac{\text{const}}{n}\sum_{r=1}^{\infty}\sum_{q\in\mathbb{Z}}\alpha_{|q|}\sum_{p=-\ell_n}^{\ell_n}\alpha_{|r-p|} = O(\ell_n/n).$$

Hence,

$$\frac{1}{n}\sum_{i=1}^{n} B'_{n,i,3} = \frac{n-2\ell_n+1}{n}o(\ell_n/n) + \frac{2\ell_n-1}{n}O(\ell_n/n) = o(\ell_n/n),$$

The quantity $B''_{n,i,3}$ can be treated similarly, by considering the cases $i \ge 2\ell_n$ and $i < 2\ell_n$.

This completes the proof for the first term on the right (33). Proceeding similarly for the second term, we obtain that

$$\begin{aligned}\operatorname{Var}\{\tilde{\sigma}_n(\boldsymbol{u},\boldsymbol{v})\} &= \frac{1}{n}\sum_{p,q,r=-\infty}^{\infty} \varphi(p/\ell_n)\varphi\{(q-r)/\ell_n\}\gamma(q,\boldsymbol{u},\boldsymbol{u})\gamma(r-p,\boldsymbol{v},\boldsymbol{v}) \\ &\quad + \frac{1}{n}\sum_{p,q,r=-\infty}^{\infty} \varphi(p/\ell_n)\varphi\{(q-r)/\ell_n\}\gamma(q,\boldsymbol{u},\boldsymbol{v})\gamma(r-p,\boldsymbol{v},\boldsymbol{u}) + s_n'(\boldsymbol{u},\boldsymbol{v}),\end{aligned}$$

where $\sup_{\boldsymbol{u},\boldsymbol{v}\in[0,1]^d} s_n'(\boldsymbol{u},\boldsymbol{v}) = o(\ell_n/n)$. Setting $s = p - r$, we get

$$\begin{aligned}\operatorname{Var}\{\tilde{\sigma}_n(\boldsymbol{u},\boldsymbol{v})\} = \frac{1}{n}\sum_{s,q,r=-\infty}^{\infty} \varphi\{(s+r)/\ell_n\}\varphi\{(q-r)/\ell_n\}\,&\{\gamma(q,\boldsymbol{u},\boldsymbol{u})\gamma(s,\boldsymbol{v},\boldsymbol{v}) \\ &+\gamma(q,\boldsymbol{u},\boldsymbol{v})\gamma(s,\boldsymbol{u},\boldsymbol{v})\} + s_n'(\boldsymbol{u},\boldsymbol{v}).\end{aligned}$$

To finish the proof, it remains to show that $\sup_{\boldsymbol{u},\boldsymbol{v}\in[0,1]^d} |s_n''(\boldsymbol{u},\boldsymbol{v})| = o(\ell_n/n)$, where

$$\begin{aligned}s_n''(\boldsymbol{u},\boldsymbol{v}) = \frac{1}{n}\sum_{s,q,r=-\infty}^{\infty} \left[\varphi\{(s+r)/\ell_n\}\varphi\{(q-r)/\ell_n\} - \{\varphi(r/\ell_n)\}^2\right]&\{\gamma(q,\boldsymbol{u},\boldsymbol{u})\gamma(s,\boldsymbol{v},\boldsymbol{v}) \\ &+\gamma(q,\boldsymbol{u},\boldsymbol{v})\gamma(s,\boldsymbol{u},\boldsymbol{v})\}.\end{aligned}$$

Using (25), we have

$$\sup_{\boldsymbol{u},\boldsymbol{v}\in[0,1]^2} |s_n''(\boldsymbol{u},\boldsymbol{v})| \le \frac{\text{const}}{n}\sum_{s,q=-\infty}^{\infty}\alpha_{|q|}\alpha_{|s|}\sum_{r=-\infty}^{\infty}\left|\varphi\{(s+r)/\ell_n\}\varphi\{(q-r)/\ell_n\} - \{\varphi(r/\ell_n)\}^2\right|.$$

The sum over r on the right-hand side consists of at most $3\times(2\ell_n+1)$ non-zero summands (those r for which either $|s+r|, |q-r|$ or $|r|$ is smaller than or equal to ℓ_n). Now, for any s,q,r such that $|s+r|, |q-r|$ or $|r|$ is smaller than or equal to ℓ_n, we have, from the conditions on φ,

$$\begin{aligned}&\left|\varphi\{(s+r)/\ell_n\}\varphi\{(q-r)/\ell_n\} - \{\varphi(r/\ell_n)\}^2\right| \\ &\quad = \left|\varphi\{(s+r)/\ell_n\}\left[\varphi\{(q-r)/\ell_n\} - \varphi(r/\ell_n)\right] + \varphi(r/\ell_n)\left[\varphi\{(s+r)/\ell_n\} - \varphi(r/\ell_n)\right]\right| \\ &\quad \le \sup_{x\in[-1,1]}|\varphi(x)| \times \left[\left|\varphi\{(q-r)/\ell_n\} - \varphi(r/\ell_n)\right| + \left|\varphi\{(s+r)/\ell_n\} - \varphi(r/\ell_n)\right|\right] \\ &\quad \le \lambda\{|q|+|s|\}/\ell_n,\end{aligned}$$

and therefore, from the conditions on the strongly mixing coefficients,

$$\begin{aligned}&\sum_{s,q=-\infty}^{\infty}\alpha_{|q|}\alpha_{|s|}\sum_{r=-\infty}^{\infty}\left|\varphi\{(s+r)/\ell_n\}\varphi\{(q-r)/\ell_n\} - \{\varphi(r/\ell_n)\}^2\right| \\ &\qquad \le \sum_{s,q=-\infty}^{\infty}\alpha_{|q|}\alpha_{|s|}(6\ell_n+3)\lambda\{|q|+|s|\}/\ell_n \le \text{const}\sum_{s,q=-\infty}^{\infty}\alpha_{|q|}\alpha_{|s|}\{|q|+|s|\} < \infty.\end{aligned}$$

Hence, $\sup_{\boldsymbol{u},\boldsymbol{v}\in[0,1]^2}|s_n''(\boldsymbol{u},\boldsymbol{v})| = O(1/n) = o(\ell_n/n)$. The stated result finally follows from the fact that $\ell_n^{-1}\sum_{r=-\ell_n}^{\ell_n}\{\varphi(r/\ell_n)\}^2 = \int_{-1}^{1}\varphi(x)^2\mathrm{d}x + o(1)$. ■

References

G.G. Agarwal, R.J. Dalpatadu, and A.K. Singh. Linear functions of uniform order statistics and B-splines. *Commun. Stat., Theory Methods*, 31(2):181–192, 2002.

D.W.K. Andrews. Heteroskedasticity and autocorrelation consistent covariance matrix estimation. *Econometrica*, 59(3):817–858, 1991.

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

S. Balacheff and G. Dupont. Normalité asymptotique des processus empiriques tronqués et des processus de rang. *Lecture Notes in Mathematics*, 821:19–45, 1980.

P.J. Bickel and M.J. Wichura. Convergence criteria for multiparameter stochastic processes and some applications. *Annals of Mathematical Statistics*, 42(5):1656–1670, 1971.

A. Bücher and H. Dette. A note on bootstrap approximations for the empirical copula process. *Statistics and Probability Letters*, 80(23–24):1925–1932, 2010.

A. Bücher and M. Ruppert. Consistent testing for a constant copula under strong mixing based on the tapered block multiplier technique. *Journal of Multivariate Analysis*, 116: 208–229, 2013.

A. Bücher and S. Volgushev. Empirical and sequential empirical copula processes under serial dependence. *Journal of Multivariate Analysis*, 119:61–70, 2013.

P. Bühlmann. *The blockwise bootstrap in time series and empirical processes*. PhD thesis, ETH Zürich, 1993. Diss. ETH No. 10354.

P. Bühlmann. Blockwise bootstrapped empirical process for stationary sequences. *The Annals of Statistics*, 22(2):995–1012, 1994.

M. Csörgő and L. Horváth. *Limit theorems in change-point analysis*. Wiley Series in Probability and Statistics. John Wiley & Sons, Chichester, UK, 1997.

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

H. Dehling and W. Philipp. Empirical process techniques for dependent data. In H. Dehling, T. Mikosch, and M. Sorensen, editors, *Empirical process techniques for dependent data*, pages 1–113. Birkhäuser, Boston, 2002.

S. Dhompongsa. A note of the almost sure approximation of the empirical process of weakly dependent random vectors. *Yokohama Mathematical Journal*, 32:113–121, 1984.

J.-D. Fermanian, D. Radulovic, and M. Wegkamp. Weak convergence of empirical copula processes. *Bernoulli*, 10(5):847–860, 2004.

P. Gänssler and W. Stute. *Seminar on empirical processes*. DMV Seminar 9. Birkhäuser, Basel, 1987.

C. Genest and J. Segers. On the covariance of the asymptotic empirical copula process. *Journal of Multivariate Analysis*, 101:1837–1845, 2010.

C. Genest, B. Rémillard, and D. Beaudoin. Goodness-of-fit tests for copulas: A review and a power study. *Insurance: Mathematics and Economics*, 44:199–213, 2009.

M. Holmes, I. Kojadinovic, and J-F. Quessy. Nonparametric tests for change-point detection à la Gombay and Horváth. *Journal of Multivariate Analysis*, 115:16–32, 2013.

E. Jondeau, S.-H. Poon, and M. Rockinger. *Financial modeling under non-Gaussian distributions*. Springer, London, 2007.

I. Kojadinovic and T. Rohmer. Asymptotics and multiplier bootstrap of the sequential empirical copula process with applications to change-point detection. *arXiv:1206.2557*, 2012.

I. Kojadinovic and J. Yan. Goodness-of-fit testing based on a weighted bootstrap: A fast large-sample alternative to the parametric bootstrap. *The Canadian Journal of Statistics*, 40(3):480–500, 2012.

I. Kojadinovic, J. Segers, and J. Yan. Large-sample tests of extreme-value dependence for multivariate copulas. *The Canadian Journal of Statistics*, 39(4):703–720, 2011.

M.R. Kosorok. *Introduction to empirical processes and semiparametric inference*. Springer, New York, 2008.

H.R. Künsch. The jacknife and the bootstrap for general stationary observations. *Annals of Statistics*, 17(3):1217–1241, 1989.

A.J. McNeil, R. Frey, and P. Embrechts. *Quantitative risk management*. Princeton University Press, New Jersey, 2005.

A. Mokkadem. Mixing properties of arma processes. *Stochastic Processes Applications*, 29(2):309–315, 1988.

E. Paparoditis and D.N. Politis. Tapered block bootstrap. *Biometrika*, 88(4):1105–1119, 2001.

A. Patton. Copula methods for forecasting multivariate time series. In *Handbook of Economic Forecasting*, volume 2. Springer Verlag, 2012.

D.N. Politis. Adaptive bandwidth choice. *Journal of Nonparametric Statistics*, 15(4–5): 517–533, 2003.

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

D.P. Politis and J.P. Romano. Bias-corrected nonparametric spectral estimation. *Journal of Time Series Analysis*, 16:67–103, 1995.

B. Rémillard and O. Scaillet. Testing for equality between two copulas. *Journal of Multivariate Analysis*, 100(3):377–386, 2009.

L. Rüschendorf. Asymptotic distributions of multivariate rank order statistics. *Annals of Statistics*, 4:912–923, 1976.

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

J. Segers. Asymptotics of empirical copula processes under nonrestrictive smoothness assumptions. *Bernoulli*, 18:764–782, 2012.

X. Shao. The dependent wild bootstrap. *Journal of the American Statistical Association*, 105(489):218–235, 2010.

A. Sklar. Fonctions de répartition à n dimensions et leurs marges. *Publications de l'Institut de Statistique de l'Université de Paris*, 8:229–231, 1959.

H. Tsukahara. Semiparametric estimation in copula models. *The Canadian Journal of Statistics*, 33(3):357–375, 2005.

A.W. van der Vaart. *Asymptotic statistics.* Cambridge University Press, 1998.

A.W. van der Vaart and J.A. Wellner. Empirical processes indexed by estimated functions. In *Asymptotics: Particles, Processes and Inverse Problems*, pages 234–252. Institute of Mathematical Statistics, 2007.

A.W. van der Vaart and J.A. Wellner. *Weak convergence and empirical processes.* Springer, New York, 2000. Second edition.