SFB 823

Weak convergence of the empirical copula process with respect to weighted metrics

Betina Berghaus, Axel Bücher,
Stanislav Volgushev

Nr. 38/2014

Discussion Paper

Weak convergence of the empirical copula process with respect to weighted metrics

Betina Berghaus, Axel Bücher and Stanislav Volgushev*

November 21, 2014

Abstract

The empirical copula process plays a central role in the asymptotic analysis of many statistical procedures which are based on copulas or ranks. Among other applications, results regarding its weak convergence can be used to develop asymptotic theory for estimators of dependence measures or copula densities, they allow to derive tests for stochastic independence or specific copula structures, or they may serve as a fundamental tool for the analysis of multivariate rank statistics. In the present paper, we establish weak convergence of the empirical copula process (for observations that are allowed to be serially dependent) with respect to weighted supremum distances. The usefulness of our results is illustrated by applications to general bivariate rank statistics and to estimation procedures for the Pickands dependence function arising in multivariate extreme-value theory.

Keywords and Phrases: Empirical copula process; weighted weak convergence; strongly mixing; bivariate rank statistics; Pickands dependence function.

AMS Subject Classification: 62G30, 60F17.

1 Introduction

The theory of weak convergence of empirical processes can be regarded as one of the most powerful tools in mathematical statistics. Through the continuous mapping theorem or the functional delta method, it greatly facilitates the development of asymptotic theory in a vast variety of situations (Van der Vaart and Wellner, 1996).

*Ruhr-Universität Bochum, Fakultät für Mathematik, Universitätsstr. 150, 44780 Bochum, Germany. E-mail: betina.berghaus@rub.de, axel.buecher@rub.de, stanislav.volgushev@rub.de. This work has been supported by the Collaborative Research Center "Statistical modeling of nonlinear dynamic processes" (SFB 823) of the German Research Foundation (DFG) which is gratefully acknowledged.

For applying the continuous mapping theorem or the functional delta method, the course of action is often similar. Consider for instance the continuous mapping theorem: starting from some abstract weak convergence result, say $\mathbb{F}_n \rightsquigarrow \mathbb{F}$ in some metric space $(\mathcal{D}, d_{\mathcal{D}})$, one would like to deduce weak convergence of $\phi(\mathbb{F}_n) \rightsquigarrow \phi(\mathbb{F})$, where ϕ is some mapping defined on $(\mathcal{D}, d_{\mathcal{D}})$ with values in another metric space $(\mathcal{E}, d_{\mathcal{E}})$. This conclusion is possible provided ϕ is continuous at every point of a set which contains the limit $\mathbb{F}$, almost surely (Van der Vaart and Wellner, 1996).

The continuity of ϕ is linked to the strength of the metric $d_{\mathcal{D}}$ – a stronger metric will make more functions continuous. For example, let $\mathcal{D} = \ell^\infty([0,1])$ denote the space of bounded functions on $[0,1]$ and consider the real-valued functional $\phi(f) := \int_{(0,1)} f(x)/x \, dx$ (with ϕ defined on a suitable subspace of $\mathcal{D}$). In Section 3.2 below, this functional will turn out to be of great interest for the estimation of Pickands dependence function and it is also closely related to the classical Anderson-Darling statistic. Now, if we equip $\mathcal{D}$ with the supremum distance, as is typically done in empirical process theory, the map ϕ is not continuous because $1/x$ is not integrable. Continuity of ϕ can be ensured by considering a weighted distance, such as for instance $\sup_{x\in[0,1]} |f_1(x) - f_2(x)|/g(x)$ for a positive weight function g such that $g(x)/x$ is integrable. Similar phenomenas arise with the functional delta method, see Beutner and Zähle (2010). It thus is desirable to establish weak convergence results with the metric $d_{\mathcal{D}}$ taken as strong as possible. One class of metrics which is of particular interest in many statistical applications is given by weighted supremum distances.

For classical empirical processes, corresponding weak convergence results are well known. For example, the standard d-dimensional empirical process $\mathbb{F}_n(\boldsymbol{x}) = \sqrt{n}\{F_n(\boldsymbol{x}) - F(\boldsymbol{x})\}$ with F having standard uniform marginals, converges weakly with respect to the metric induced by the weighted norm

$$\|G\|_\omega = \sup_{\boldsymbol{u}\in[0,1]^d} \left| \frac{G(\boldsymbol{u})}{\{g(\boldsymbol{u})\}^\omega} \right|, \quad g(\boldsymbol{u}) = \big(\min_{j=1}^{d} u_j\big) \wedge \big(1 - \min_{j=1}^{d} u_j\big),$$

$\omega \in (0, 1/2)$. See, e.g., Shorack and Wellner (1986) for the one-dimensional i.i.d.-case, Shao and Yu (1996) for the one-dimensional time series case or Genest and Segers (2009) for the bivariate i.i.d.-case. For $d = 2$, the graph of the function g is depicted in Figure 1.

The present paper is motivated by the apparent lack of such results for the empirical copula process $\hat{\mathbb{C}}_n$. This process, an element of $D([0,1]^d)$ precisely defined in Section 2 below, plays a crucial role in the asymptotic analysis of statistical procedures which are based on copulas or ranks. Unweighted weak convergence of $\hat{\mathbb{C}}_n$ has been investigated by several authors under a variety of assumptions on the smoothness of the copula and on the temporal dependence of the underlying observations, see Gaenssler and Stute (1987); Fermanian et al. (2004); Segers (2012); Bücher and Volgushev (2013); Bücher

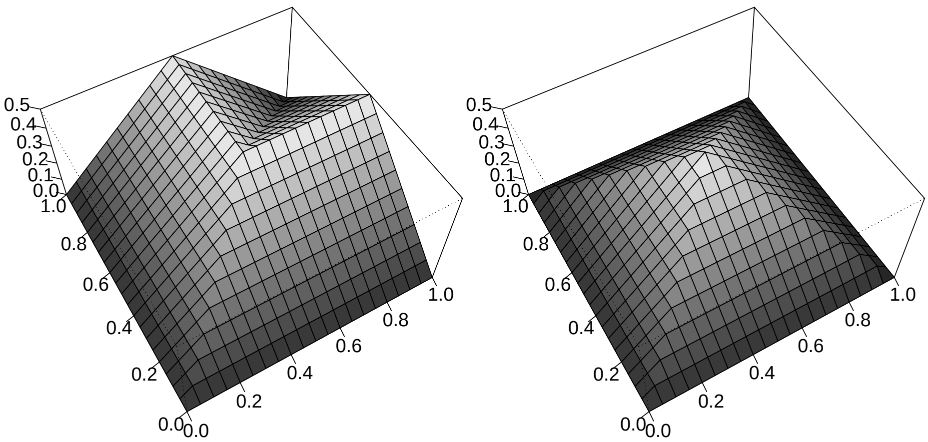

Figure 1: *Graphs of* $g(u,v) = \min\{u, v, 1 - \min(u,v)\}$ *(left picture) and of* $\tilde{g}(u,v) = \min\{u, v, (1-u), (1-v)\}$ *(right picture).*

et al. (2014), among others. However, results regarding its weighted weak convergence are almost non-existent. To the best of our knowledge, the only reference appears to be Rüschendorf (1976), where, however, weight functions are only allowed to approach zero at the lower boundary of the unit cube. The restrictiveness of this condition becomes particularly visible in dimension $d = 2$ where it is known that the limit of the empirical copula process is zero on the entire boundary of the unit square (Genest and Segers, 2010). This observation suggests that, for $d = 2$, it should be possible to maintain weak convergence of the empirical copula process when dividing by functions of the form $\{\tilde{g}(u,v)\}^\omega$ where

$$\tilde{g}(u,v) = u \wedge v \wedge (1-u) \wedge (1-v).$$

A picture of the graph of $\tilde{g}$ can be found in Figure 1, obviously, we have $\tilde{g} \leq g$. The main result of this paper confirms the last-mentioned conjecture. More precisely, we establish weighted weak convergence of the empirical copula process in general dimension $d \geq 2$ with weight functions that approach zero wherever the potential limit approaches zero. We also do not require the observations to be i.i.d. and allow for exponential alpha mixing.

Potential applications of the new weighted weak convergence results are extensive. As a direct corollary, one can derive the asymptotic behavior of Anderson-Darling type goodness-of-fit statistics for copulas. The derivation of the asymptotic behavior of rank-based estimators for the Pickands dependence functions (Genest and Segers, 2009) can be greatly simplified and, moreover, can be simply extended to time series observations. Through a suitable partial integration formula, the results can also be exploited to derive weak convergence of multivariate rank statistics as for instance of certain scalar measures of (serial) dependence. The latter two applications are worked out in detail in Section 3 of this paper.

The remaining part of this paper is organized as follows. In Section 2, the empirical copula process is introduced and the main result of the paper, its weighted weak convergence, is stated. In Section 3, the main result is illustratively exploited to derive the asymptotics of multivariate rank statistics and of common estimators for extreme-value copulas. All proofs are deferred to Section 4, with some auxiliary results postponed to Section 5. Finally, Appendix A in the supplementary material contains some general results on (locally) bounded variation and integration for two-variate functions which are needed for the proof of Theorem 3.3.

2 Weighted empirical copula processes

Let $\boldsymbol{X} = (X_1, \ldots, X_d)'$ be a d-dimensional random vector with joint cumulative distribution function (c.d.f.) F and continuous marginal c.d.f.s $F_1, \ldots, F_d$. The copula C of F, or, equivalently, the copula of $\boldsymbol{X}$, is defined as the c.d.f. of the random vector $\boldsymbol{U} = (U_1, \ldots, U_d)'$ that arises from marginal application of the probability integral transform, i.e., $U_j = F_j(X_j)$ for $j = 1, \ldots, d$. By construction, the marginal c.d.f.s of C are standard uniform on $[0, 1]$. By Sklar's Theorem, C is the unique function for which we have

$$F(x_1, \ldots, x_d) = C\{F_1(x_1), \ldots, F_d(x_d)\}$$

for all $\boldsymbol{x} = (x_1, \ldots, x_d) \in \mathbb{R}^d$.

Let $\boldsymbol{X}_i, i = 1, \ldots, n$ be an observed stretch of a strictly stationary time series such that $\boldsymbol{X}_i$ is equal in distribution to $\boldsymbol{X}$. Set $\boldsymbol{U}_i = (U_{i1}, \ldots, U_{id}) \sim C$ with $U_{ij} = F_j(X_{ij})$. Define (observable) pseudo observations $\hat{\boldsymbol{U}}_i = (\hat{U}_{i1}, \ldots, \hat{U}_{id})$ of C through $\hat{U}_{ij} = nF_{nj}(X_{ij})(/n + 1)$ for $i = 1, \ldots, n$ and $j = 1, \ldots, d$. The empirical copula $\hat{C}_n$ of the sample $\boldsymbol{X}_1, \ldots, \boldsymbol{X}_n$ is defined as the empirical distribution function of $\hat{\boldsymbol{U}}_1, \ldots, \hat{\boldsymbol{U}}_n$, i.e.,

$$\hat{C}_n(\boldsymbol{u}) = \frac{1}{n} \sum_{i=1}^{n} \mathbb{1}(\hat{\boldsymbol{U}}_i \le \boldsymbol{u}), \qquad \boldsymbol{u} \in [0, 1]^d.$$

The corresponding empirical copula process is defined as

$$\boldsymbol{u} \mapsto \hat{\mathbb{C}}_n(\boldsymbol{u}) = \sqrt{n}\{\hat{C}_n(\boldsymbol{u}) - C(\boldsymbol{u})\}.$$

For $\omega \ge 0$, define a weight function

$$g_\omega(\boldsymbol{u}) = \min\{\wedge_{j=1}^{d} u_j, \wedge_{j=1}^{d}[1 - (u_1 \wedge \cdots \wedge \widehat{u_j} \wedge \cdots \wedge u_d)]\}^\omega,$$

where the hat-notation $u_1 \wedge \cdots \wedge \widehat{u_j} \wedge \cdots \wedge u_d$ is used as a shorthand for $\min\{u_1, \ldots, u_{j-1}, u_{j+1}, \ldots, u_d\}$. For $d = 2$, the function is particularly nice and reduces to $g_\omega(u_1, u_2) = \min(u_1, u_2, 1 - u_1, 1 - u_2)^\omega$, see Figure 1. Note that for vectors $\boldsymbol{u} \in [0, 1]^d$ such that at least one coordinate is equal to 0 or such that $d-1$ coordinates are equal to 1, we have $g_\omega(\boldsymbol{u}) = 0$. As already mentioned in the introduction for the case $d = 2$, these vectors are exactly the points where the limit of the empirical copula process is equal to 0, almost surely, whence one might hope to obtain a weak convergence result for $\mathbb{C}_n/g_\omega$. To prove such a result, a smoothness condition on C has to be imposed.

Condition 2.1. For every $j \in \{1, \dots, d\}$, the first oder partial derivative $\dot{C}_j(\boldsymbol{u}) := \partial C(\boldsymbol{u})/\partial u_j$ exists and is continuous on $V_j = \{\boldsymbol{u} \in [0,1]^d : u_j \in (0,1)\}$. For every $j_2, j_2 \in \{1, \dots, d\}$, the second order partial derivative $\ddot{C}_{j_1 j_2}(\boldsymbol{u}) := \partial^2 C(\boldsymbol{u})/\partial u_{j_1} \partial u_{j_2}$ exists and is continuous on $V_{j_1} \cap V_{j_2}$. Moreover, there exists a constant $K > 0$ such that

$$|\ddot{C}_{j_1 j_2}(\boldsymbol{u})| \le K \min\left\{ \frac{1}{u_{j_1}(1-u_{j_1})}, \frac{1}{u_{j_2}(1-u_{j_2})} \right\}, \qquad \forall \boldsymbol{u} \in V_{j_1} \cap V_{j_2}.$$

For completeness, define $\dot{C}_j(\boldsymbol{u}) = \limsup_{h\to 0}\{C(\boldsymbol{u}+h\boldsymbol{e}_j)-C(\boldsymbol{u})\}/h$ wherever it does not exist. Note, that Condition 2.1 coincides with Condition 2.1 and Condition 4.1 in Segers (2012), who used it to prove Stute's representation of an almost sure remainder term (Stute, 1984). The condition is satisfied for many commonly occurring copulas (Segers, 2012).

For $-\infty \le a < b \le \infty$, let $\mathcal{F}_a^b$ denote the sigma-field generated by those $\boldsymbol{X}_i$ for which $i \in \{a, a+1, \dots, b\}$ and define, for $k \ge 1$,

$$\alpha^{[\boldsymbol{X}]}(k) = \sup\left\{ |\mathbb{P}(A \cap B) - \mathbb{P}(A)\mathbb{P}(B)| : A \in \mathcal{F}_{-\infty}^i, B \in \mathcal{F}_{i+k}^\infty, i \in \mathbb{Z} \right\}$$

as the alpha-mixing coefficient of the time series $(\boldsymbol{X}_i)_{i\in\mathbb{Z}}$. The sequence is called strongly mixing (or alpha-mixing) if $\alpha^{[\boldsymbol{X}]}(k) \to 0$ for $k \to \infty$. Finally,

$$\alpha_n(\boldsymbol{u}) = \sqrt{n}\{G_n(\boldsymbol{u}) - C(\boldsymbol{u})\}, \qquad G_n(\boldsymbol{u}) = n^{-1}\textstyle\sum_{i=1}^n \mathbb{1}(\boldsymbol{U}_i \le \boldsymbol{u}),$$

denotes the (unobservable) empirical process based on $\boldsymbol{U}_1, \dots, \boldsymbol{U}_n$.

Theorem 2.2. (Weighted weak convergence of the empirical copula process) *Suppose that $\boldsymbol{X}_1, \boldsymbol{X}_2, \dots$ is a stationary, alpha-mixing sequence with $\alpha^{[\boldsymbol{X}]}(k) = O(a^k)$, as $k \to \infty$, for some $a \in (0,1)$. If the marginals of the stationary distribution are continuous and if the corresponding copula C satisfies Condition 2.1, then, for any $c \in (0,1)$ and any $\omega \in (0, 1/2)$,*

$$\sup_{\boldsymbol{u} \in [\frac{c}{n}, 1-\frac{c}{n}]^d} \left| \frac{\hat{\mathbb{C}}_n(\boldsymbol{u})}{g_\omega(\boldsymbol{u})} - \frac{\bar{\mathbb{C}}_n(\boldsymbol{u})}{g_\omega(\boldsymbol{u})} \right| = o_P(1)$$

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

$$\bar{\mathbb{C}}_n(\boldsymbol{u}) := \alpha_n(\boldsymbol{u}) - \sum_{j=1}^d \dot{C}_j(\boldsymbol{u}) \alpha_n(\boldsymbol{u}^{(j)}),$$

with $\boldsymbol{u}^{(j)} = (1,\dots,1,u_j,1,\dots,1)$. *Moreover, we have* $\bar{\mathbb{C}}_n/\tilde{g}_\omega \rightsquigarrow \mathbb{C}_C/\tilde{g}_\omega$ *in* $(\ell^\infty([0,1]^d), \|\cdot\|_\infty)$, *where* $\tilde{g}_\omega(\boldsymbol{u}) = g_\omega(\boldsymbol{u}) + \mathbb{1}\{g_\omega(\boldsymbol{u}) = 0\}$, *where*

$$\mathbb{C}_C(\boldsymbol{u}) = \alpha_C(\boldsymbol{u}) - \sum_{j=1}^{d} \dot{C}_j(\boldsymbol{u})\alpha_C(\boldsymbol{u}^{(j)}),$$

and where α_C *denotes a tight, centered Gaussian process with covariance*

$$\mathrm{Cov}\{\alpha_C(\boldsymbol{u}), \alpha_C(\boldsymbol{v})\} = \sum\nolimits_{i\in\mathbb{Z}} \mathrm{Cov}\{\mathbb{1}(\boldsymbol{U}_0 \le u), \mathbb{1}(\boldsymbol{U}_i \le v)\}.$$

The proof of Theorem 2.2 is given in Section 4.1 below. In fact, we state a more general result which is based on conditions on the usual empirical process α_n. These conditions are subsequently shown to be valid for exponentially alpha-mixing time series.

3 Applications

Theorem 2.2 may be exploited in numerous ways. For instance, many of the most powerful goodness-of-fit tests for copulas are based on distances between the empirical copula and a parametric estimator for C (Genest et al., 2009). The results of Theorem 2.2 can be exploited to validate tests for a richer class of distances, as for weighted Kolomogorov-Smirnov or L^2-distances. Second, estimators for extreme-value copulas can often be expressed through improper integrals involving the empirical copula (see Genest and Segers, 2009, among others). Weighted weak convergence as in Theorem 2.2 facilitates the anlysis of their asymptotic behavior and allows to extend the available results to time series observations. Details regarding the CFG- and the Pickands estimator are worked out in Section 3.2 below.

Theorem 2.2 may also be used outside the genuine copula framework, for instance, for proving asymptotic normality of multivariate rank statistics. The power of that approach lies in the fact that proofs for time series are essentially the same as for i.i.d. data sets. In Section 3.1, we derive a general weak convergence result for bivariate rank statistics.

3.1 Bivariate rank statistics

Bivariate rank statistics constitute an important class of real-valued statistics that can be written as

$$R_n = \frac{1}{n}\sum_{i=1}^{n} J(\hat{U}_{i1}, \hat{U}_{i2})$$

for some function $J : (0,1)^2 \to \mathbb{R}$, called score function. R_n can also be expressed as a Lebesgue-Stieltjes integral with respect to $\hat{C}_n$, i.e.,

$$R_n = \int_{[\frac{1}{n+1},\frac{n}{n+1}]^2} J(u,v)\mathrm{d}\hat{C}_n(u,v),$$

which offers the way to derive the asymptotic behavior of R_n from the asymptotic behavior of the empirical copula. This idea has already been exploited in Fermanian et al. (2004): however, in their Theorem 6, J has to be a bounded function which is not the case for many interesting examples. Also, the uniform central limit theorems for multivariate rank statistics in van der Vaart and Wellner (2007) require rather strong smoothness assumptions on J (which imply boundedness of J).

Example 3.1. (Rank Autocorrelation Coefficients) Suppose $Y_1, \dots, Y_n$ are drawn from a stationary, univariate time series $(Y_i)_{i\in\mathbb{Z}}$. Rank autocorrelation coefficients of lag $k \in \mathbb{N}$ are statistics of the form

$$r_{n,k} = \frac{1}{n-k} \sum_{i=k+1}^{n} J_1\Big\{\tfrac{n}{n+1} F_n(Y_i)\Big\} J_2\Big\{\tfrac{n}{n+1} F_n(Y_{i-k})\Big\},$$

where J_1, J_2 are real-valued functions on $(0,1)$ and F_n denotes the empirical cdf of $Y_1, \dots, Y_n$. For example, the van der Waerden autocorrelation (Hallin and Puri, 1988) is given by

$$r_{n,k,vdW} = \frac{1}{n-k} \sum_{i=k+1}^{n} \Phi^{-1}\Big\{\tfrac{n}{n+1} F_n(Y_i)\Big\} \Phi^{-1}\Big\{\tfrac{n}{n+1} F_n(Y_{i-k})\Big\},$$

(with Φ and Φ^{-1} denoting the cdf of the standard normal distribution and its inverse, respectively) and the Wilcoxon autocorrelation (Hallin and Puri, 1988) is defined as

$$r_{n,k,W} = \frac{1}{n-k} \sum_{i=k+1}^{n} \Big\{\tfrac{n}{n+1} F_n(Y_i) - \frac{1}{2}\Big\} \log\Big\{\frac{\frac{n}{n+1} F_n(Y_{i-k})}{1 - \frac{n}{n+1} F_n(Y_{i-k})}\Big\}.$$

Obviously, the corresponding score functions are unbounded. Asymptotic normality for these and similar rank statistics has been shown for i.i.d. observations and for $ARMA$-processes (Hallin et al., 1985). To the best of our knowledge, no general tool to handle the asymptotic behavior of such statistics for dependent observations seems to be available. Theorem 3.3 below aims at partially filling that gap.

Example 3.2. (The pseudo-maximum likelihood estimator) As a common practice in bivariate copula modeling one assumes to observe a sample $\boldsymbol{X}_1, \dots, \boldsymbol{X}_n$ from a bivariate distribution whose copula belongs to a parametric copula family, parametrized by a finite-dimensional parameter $\theta \in \Theta \subset \mathbb{R}^p$. Except for the assumption of absolute continuity, the marginal distributions are often left unspecified in order to allow for maximal robustness with respect to potential miss-specification. In such a setting, the pseudo-maximum likelihood estimator (see Genest et al. (1995) for a theoretical investigation) provides the most common estimator for the parameter θ. If c_θ denotes the corresponding copula density, the estimator is defined as

$$\hat\theta_n = \arg\max_{\theta\in\Theta} \sum_{i=1}^n \log\{c_\theta(\hat U_{i1}, \hat U_{i2})\}.$$

Using standard arguments from maximum-likelihood theory and imposing suitable regularity conditions, the asymptotic distribution of $\sqrt{n}(\hat\theta_n - \theta_0)$ can be derived from the asymptotic behavior of

$$R_n = \frac{1}{n}\sum_{i=1}^n J_{\theta_0}(\hat U_{i1}, \hat U_{i2}), \tag{3.1}$$

where θ_0 denotes the unknown true parameter and where $J_\theta = (\partial \log c_\theta)/(\partial\theta)$ denote the score function. Typically, this function is unbounded, as for instance in case of the bivariate Gaussian copula model where θ is the correlation coefficient and the score function takes the form

$$J_\theta(u,v) = \frac{\theta(1-\theta^2) - \theta\{\Phi^{-1}(u)^2 + \Phi^{-1}(v)^2\} + (1+\theta^2)\Phi^{-1}(u)\Phi^{-1}(v)}{1+\theta^2}.$$

Still, the conditions of Theorem 3.3 below can be shown to be valid.

Finally, note that pseudo-maximum likelihood estimators also arise in Markovian copula models (Chen and Fan, 2006) where copulas are used to model the serial dependence of a stationary time series at lag one. Again, their asymptotic distribution may be derived from rank statistics as in (3.1).

The following theorem is the central result of this section. It establishes weak convergence of bivariate rank-statistics by exploiting weighted weak convergence of the empirical copula process. For the definition of the space of functions of locally bounded total variation in the sense of Hardy-Krause, $BVHK_{loc}((0,1)^2)$, and for Lebesgue-Stieltjes integrals with respect to such functions, we refer the reader to Definition A.8 in the supplementary material. The proof is given in Section 4.4.

Theorem 3.3. *Suppose the conditions of Theorem 2.2 are met. Moreover, suppose that $J \in BVHK_{loc}((0,1)^2)$ is right-continuous and that there exists $\omega \in (0,1/2)$ such that $|J(\boldsymbol{u})| \le \mathrm{const} \times g_\omega(\boldsymbol{u})^{-1}$ and such that*

$$\int_{(0,1)^2} g_\omega(\boldsymbol{u}) |\mathrm{d}J(\boldsymbol{u})| < \infty. \tag{3.2}$$

Moreover, for $\delta \to 0$, suppose that

$$\int_{(\delta,1-\delta]} |J(\mathrm{d}u, \delta)| = O(\delta^{-\omega}) \text{ and } \int_{(\delta,1-\delta]} |J(\mathrm{d}u, 1-\delta)| = O(\delta^{-\omega}), \tag{3.3}$$

$$\int_{(\delta,1-\delta]} |J(\delta, \mathrm{d}v)| = O(\delta^{-\omega}) \text{ and } \int_{(\delta,1-\delta]} |J(1-\delta, \mathrm{d}v)| = O(\delta^{-\omega}). \tag{3.4}$$

Then, as $n \to \infty$,

$$\sqrt{n}\{R_n - \mathbb{E}[J(\boldsymbol{U})]\} \rightsquigarrow \int_{(0,1)^2} \mathbb{C}_C(\boldsymbol{u})\mathrm{d}J(\boldsymbol{u}).$$

The weak limit is normally distributed with mean 0 *and variance*

$$\sigma^2 = \int_{(0,1)^2}\int_{(0,1)^2} \mathbb{E}[\mathbb{C}_C(\boldsymbol{u})\mathbb{C}_C(\boldsymbol{v})]\mathrm{d}J(\boldsymbol{u})\mathrm{d}J(\boldsymbol{v}).$$

Remark 3.4. (i) Provided the second order partial derivative $\ddot{J}_{12}(u,v) := \partial^2 J(u,v)/\partial u \partial v$ exists, then the conditions (3.2)–(3.4) are equivalent to $\int_{(0,1)^2} g_\omega(u,v)|\ddot{J}_{12}(u,v)|\mathrm{d}(u,v) < \infty$ and, as $\delta \to 0$,

$$\int_\delta^{1-\delta} |\dot{J}_1(u,\delta)|\mathrm{d}u = O(\delta^{-\omega}) \text{ and } \int_\delta^{1-\delta} |\dot{J}_1(u,1-\delta)|\mathrm{d}u = O(\delta^{-\omega}),$$
$$\int_\delta^{1-\delta} |\dot{J}_2(\delta,v)|\mathrm{d}v = O(\delta^{-\omega}) \text{ and } \int_\delta^{1-\delta} |\dot{J}_2(1-\delta,v)|\mathrm{d}v = O(\delta^{-\omega}),$$

where $\dot{J}_1(u,v) := \partial J(u,v)/\partial u$, $\dot{J}_2(u,v) := \partial J(u,v)/\partial v$.

(ii) A careful check of the proof of Theorem 3.3 shows that the theorem actually remains valid under the more general conditions of Theorem 4.5 below, with $\omega \in (0,1/2)$ replaced by $\omega \in (0, \frac{\theta_1}{2(1-\theta_1)} \wedge \frac{\theta_2}{2(1-\theta_2)} \wedge (\theta_3 - 1/2))$.

As a simple application of Theorem 3.3 let us return to the autocorrelation coefficients from Example 3.1. It can easily be shown that both $J_{vdW}(u,v) = \Phi^{-1}(u)\Phi^{-1}(v)$ and $J_W(u,v) = (u-\frac{1}{2})\log(\frac{v}{1-v})$ satisfy the conditions of Theorem 3.3. To prove this for J_{vdW} use that $|\Phi^{-1}(u)| \le \{u(1-u)\}^{-\varepsilon}$ for any $\varepsilon > 0$ and that $\frac{1}{\phi\{\Phi^{-1}(u)\}} \le \{u(1-u)\}^{-1}$, with ϕ denoting the density of the standard normal distribution. Therefore, both coefficients are asymptotically normally distributed for any stationary, exponentially alpha-mixing time series provided that the copula of (Y_t, Y_{t-k}) satisfies Condition 2.1. This broadens results from Hallin et al. (1985), which may be further extended along the lines of Remark 3.4(ii) by a more thorough investigation of Conditions 4.1–4.3. Details are omitted for the sake of brevity.

3.2 Nonparametric estimation of Pickands dependence function

Theorem 2.2 can be used to extend recent results for the estimation of Pickands dependence functions. Recall that C is a multivariate extreme-value copula if and only if C has a representation of the form

$$C(\boldsymbol{u}) = \exp\left\{\Big(\sum_{j=1}^d \log u_j\Big) A\Big(\frac{\log u_1}{\sum_{j=1}^d \log u_j}, \dots, \frac{\log u_{d-1}}{\sum_{j=1}^d \log u_j}\Big)\right\}, \ \boldsymbol{u} \in (0,1)^d,$$

for some function $A : \Delta_{d-1} \to [1/d, 1]$, where Δ_{d-1} denotes the unit simplex $\Delta_{d-1} = \{\boldsymbol{w} = (w_1, \dots, w_{d-1}) \in [0,1]^{d-1} : \sum_{j=1}^{d-1} w_j \le 1\}$. In that case, A is necessarily convex and satisfies the relationship

$$\max(w_1, \dots, w_d) \le A(w_1, \dots, w_{d-1}) \le 1 \quad (w_d = 1 - \textstyle\sum_{j=1}^{d-1} w_j),$$

for all $\boldsymbol{w} \in \Delta_{d-1}$. By reference to Pickands (1981), A is called Pickands dependence function. Nonparametric estimation methods for A in the i.i.d. case and under the additional assumption that the marginal distributions are known have been considered in Pickands (1981); Deheuvels (1991); Capéraà et al. (1997); Jiménez et al. (2001), among others. In the more realistic case of unknown marginal distribution, rank-based estimators have for instance been investigated in Genest and Segers (2009); Bücher et al. (2011); Gudendorf and Segers (2012); Berghaus et al. (2013), among others. For illustrative purposes, we restrict attention to the rank-based versions of the Pickands estimator in Gudendorf and Segers (2012) in the following, even though the results easily carry over to, for instance, the CFG-estimator. The Pickands-estimator is defined as

$$\hat{A}_n^P(\boldsymbol{w}) = \left[\frac{1}{n} \sum_{i=1}^{n} \min \left\{ \frac{-\log(\hat{U}_{i1})}{w_1}, \dots, \frac{-\log(\hat{U}_{id})}{w_d} \right\} \right]^{-1}$$

and it follows by simple algebra (see Lemma 1 in Gudendorf and Segers, 2012) that $\mathbb{A}_n^P := \sqrt{n}(\hat{A}_n^P - A) = -A^2 \mathbb{B}_n^P / (1 + n^{1/2} \mathbb{B}_n^P)$, where

$$\mathbb{B}_n^P(\boldsymbol{w}) = \int_0^1 \hat{\mathbb{C}}_n(u^{w_1}, \dots, u^{w_d}) \, \frac{\mathrm{d}u}{u}.$$

Note that $\int_0^1 u^{-1} \, \mathrm{d}u$ does not converge, which hinders a direct application of the continuous mapping theorem to deduce weak convergence of $\mathbb{B}_n^P$ (and hence of $\mathbb{A}_n^P$) in $\ell^\infty(\Delta_{d-1})$ just on the basis of (unweighted) weak convergence of $\hat{\mathbb{C}}_n$. Deeper results are necessary and in fact, Genest and Segers (2009) and Gudendorf and Segers (2012) deduce weak convergence of $\mathbb{B}_n^P$ by using Stute's representation for the empirical copula process based on i.i.d. observations (see Stute, 1984; Tsukahara, 2005) and by exploiting a weighted weak convergence result for α_n.

With Theorem 2.2, we can give a much simpler proof. Write

$$\mathbb{B}_n^P(\boldsymbol{w}) = \int_0^1 \frac{\hat{\mathbb{C}}_n(u^{w_1}, \dots, u^{w_d})}{\min(u^{w_1}, \dots, u^{w_d})^\omega} \, \frac{\min(u^{w_1}, \dots, u^{w_d})^\omega}{u} \, \mathrm{d}u.$$

Then, since $\int_0^1 \min(u^{w_1}, \dots, u^{w_d})^\omega \, \frac{\mathrm{d}u}{u} \le \int_0^1 u^{\omega/d - 1} \, \mathrm{d}u$ exists for any $\omega > 0$, weak convergence of $\mathbb{B}_n^P$ is a direct consequence of the continuous mapping theorem and Theorem 2.2. Note that this method of proof is not restricted to the i.i.d. case.

4 Proofs

4.1 Proof of Theorem 2.2

Theorem 2.2 will be proved by an application of a more general result on the empirical copula process. For its formulation, we need a couple of additional conditions which, subsequently, will be shown to be satisfied for exponentially alpha-mixing time series.

Condition 4.1. There exists some $\theta_1 \in (0, 1/2]$ such that, for all $\mu \in (0, \theta_1)$ and all sequences $\delta_n \to 0$, we have

$$M_n(\delta_n, \mu) := \sup_{|\boldsymbol{u}-\boldsymbol{v}| \le \delta_n} \frac{|\alpha_n(\boldsymbol{u}) - \alpha_n(\boldsymbol{v})|}{|\boldsymbol{u}-\boldsymbol{v}|^{\mu} \vee n^{-\mu}} = o_P(1).$$

Condition 4.1 can for instance be verified in the i.i.d. case with $\theta_1 = 1/2$, exploiting a bound for the multivariate oscillation modulus derived in Proposition A.1 in Segers (2012).

Condition 4.2. The empirical process α_n converges weakly in $\ell^{\infty}([0,1]^d)$ to some limit process α_C which has continuous sample paths, almost surely.

For i.i.d. samples, the latter condition is satisfies with α_C being a C-Brownian bridge, i.e., a centered Gaussian process with continuous sample paths, a.s., and with $\mathrm{Cov}\{\alpha_C(\boldsymbol{u}), \alpha_C(\boldsymbol{v})\} = C(\boldsymbol{u} \wedge \boldsymbol{v}) - C(\boldsymbol{u})C(\boldsymbol{v})$.

Condition 4.3. There exist $\theta_2 \in (0, 1/2]$ and $\theta_3 \in (1/2, 1]$ such that, for any $\omega \in (0, \theta_2)$, any $\lambda \in (0, \theta_3)$ and all $j = 1, \ldots, d$, we have

$$\sup_{u_j \in (0,1)} \left| \frac{\alpha_{nj}(u_j)}{u_j^{\omega}(1-u_j)^{\omega}} \right| = O_P(1), \qquad \sup_{u_j \in (1/n^{\lambda}, 1-1/n^{\lambda})} \left| \frac{\beta_{nj}(u_j)}{u_j^{\omega}(1-u_j)^{\omega}} \right| = O_P(1),$$

where $\alpha_{nj}(u_j) = \sqrt{n}\{G_{nj}(u_j) - u_j\}$ and $\beta_{nj}(u_j) = \sqrt{n}\{G_{nj}^{-}(u_j) - u_j\}$.

Here, $G_{nj}(u_j) = n^{-1} \sum_{i=1}^{n} \mathbb{1}(U_{ij} \le u_j)$ and, for a distribution function H on the reals, H^- denotes the (left-continuous) generalized inverse function of H defined as

$$H^-(u) := \inf\{x \in \mathbb{R} : H(x) \ge u\}, \quad 0 < u \le 1,$$

and $H^-(0) = \sup\{x \in \mathbb{R} : H(x) = 0\}$. In the i.i.d. case, Condition 4.3 is a mere consequence of results in Csörgő et al. (1986), with $\theta_2 = 1/2$, $\theta_3 = 1$.

The following proposition shows that the (probabilistic) Conditions 4.1, 4.2 and 4.3 are satisfied for sequences that are exponentially alpha-mixing.

Proposition 4.4. *Suppose that $\boldsymbol{X}_1, \boldsymbol{X}_2, \ldots$ is a stationary, alpha-mixing sequence with $\alpha^{[\boldsymbol{X}]}(k) = O(a^k)$, as $k \to \infty$, for some $a \in (0,1)$. Then, Conditions 4.1, 4.2 and 4.3 are satisfied with $\theta_1 = \theta_2 = 1/2$ and $\theta_3 = 1$.*

Here, Condition 4.3 is a mere consequence of results in Shao and Yu (1996) and Csörgő and Yu (1996), whereas Condition 4.2 has been shown in Rio (2000). For the proof of Condition 4.1, we can rely on results from Kley et al. (2014). The precise arguments are given in Section 4.2 below.

The following theorem can be regarded as a generalization of Theorem 2.2: weighted weak convergence of the empirical copula process takes place provided the abstract Conditions 4.1, 4.2 and 4.3 are met. The proof is given in Section 4.3 below.

Theorem 4.5. **(Weighted weak convergence of empirical copula processes)** *Suppose Conditions 2.1, 4.1 and 4.3 are met. Then, for any $c \in (0,1)$ and any $\omega \in (0, \frac{\theta_1}{2(1-\theta_1)} \wedge \frac{\theta_2}{2(1-\theta_2)} \wedge (\theta_3 - 1/2))$,*

$$\sup_{\boldsymbol{u} \in [\frac{c}{n}, 1-\frac{c}{n}]^d} \left| \frac{\hat{\mathbb{C}}_n(\boldsymbol{u})}{g_\omega(\boldsymbol{u})} - \frac{\bar{\mathbb{C}}_n(\boldsymbol{u})}{g_\omega(\boldsymbol{u})} \right| = o_P(1).$$

If additionally Condition 4.2 is met, then $\bar{\mathbb{C}}_n/\tilde{g}_\omega \rightsquigarrow \mathbb{C}_C/\tilde{g}_\omega$ in $(\ell^\infty([0,1]^d), \|\cdot\|_\infty)$.

Proof of Theorem 2.2. The theorem is a mere consequence of Proposition 4.4 and Theorem 4.5. □

4.2 Proof of Proposition 4.4

For an r-dimensional random vector $(Y_1, \dots Y_r)'$, define the rth order joint cumulant by

$$\operatorname{cum}(Y_1, \dots Y_r) = \sum_{\{\nu_1, \dots, \nu_p\}} (-1)^{p-1}(p-1)! \mathbb{E}\Big(\prod_{j \in \nu_1} Y_j\Big) \times \dots \times \mathbb{E}\Big(\prod_{j \in \nu_p} Y_j\Big),$$

where the summation extends over all partitions $\{\nu_1, \dots, \nu_p\}$, $p \in \{1, \dots, r\}$, of $\{1, \dots, r\}$. The following lemma will be one of the main tools for establishing Condition 4.1 under exponentially alpha-mixing.

Lemma 4.6. *If $Y_1, Y_2, \dots$ is a strictly stationary sequence of random variables with $|Y_i| \le K < \infty$ and if there exist constants $\rho \in (0,1)$ and $K' < \infty$ such that for any $p \in \mathbb{N}$ and arbitrary $i_1, \dots, i_p \in \mathbb{Z}$*

$$|\operatorname{cum}(Y_{i_1}, \dots, Y_{i_p})| \le K' \rho^{\max_{k,\ell} |i_k - i_\ell|},$$

then, there exist constants $C_1, C_2 < \infty$ only depending on K, K' and $|\nu_r|$ such that

$$\left| \operatorname{cum}\left(\sum_{i=1}^n Y_i, j \in \nu_r \right) \right| \le C_1 (n+1) \varepsilon (|\log \varepsilon| + 1)^{C_2},$$

where $\varepsilon = \mathbb{E}[|Y_i|]$.

Proof. The proof is almost identical to the proof of Lemma 7.4 in Kley et al. (2014) and is therefore omitted. □

Proof of Proposition 4.4. The weak convergence result in Condition 4.2 has been shown in Theorem 7.3 in Rio (2000).

Regarding Condition 4.3, note that exponentially alpha-mixing implies that $\alpha^{[\boldsymbol{X}]}(k) = O(k^{-b-\delta})$ for any $b > 1 + \sqrt{2}$ and any $\delta > 0$. Therefore, by Theorem 3.1 in Shao and Yu (1996),

$$\sup_{u \in [0,1]} \left| \frac{\sqrt{n}\{G_{nj}(u) - u\}}{\{u(1-u)\}^{(1-1/b)/2}} \right| = O_P(1)$$

Since $(1-1/b)/2$ converges to $1/2$ for $b \to \infty$, we indeed have the first display in Condition 4.3 with $\theta_2 = 1/2$. Regarding the second display, Csörgő and Yu (1996) have shown that

$$\sup_{u \in [\delta_n, 1-\delta_n]} \left| \frac{\sqrt{n}\{G_{nj}^{-}(u) - u\}}{\{u(1-u)\}^{(1-1/b)/2}} \right| = O_P(1),$$

for $\delta_n = n^{-b/(1+b)} \to n^{-1}$ as $b \to \infty$, which implies that we may choose $\theta_3 = 1$.

Finally, consider Condition 4.1. It follows from a simple multivariate extension of Proposition 3.1 in Kley et al. (2014) that, in our case of an exponentially alpha-mixing sequence $(\boldsymbol{X}_i)_{i \in \mathbb{Z}}$, there exist constants $\rho \in (0,1)$ and $K < \infty$ such that, for any $p \in \mathbb{N}$ and any arbitrary hyper-rectangles $A_1, \dots, A_p \subset \mathbb{R}^d$ and arbitrary $i_1, \dots, i_p \in \mathbb{Z}$,

$$|\operatorname{cum}(\mathbb{1}\{\boldsymbol{X}_{i_1} \in A_1\}, \dots, \mathbb{1}\{\boldsymbol{X}_{i_p} \in A_p\})| \le K \rho^{\max_{k,\ell} |i_k - i_\ell|}. \tag{4.1}$$

The latter display will be the main tool to establish Condition 4.1. First, decompose

$$M_n(\delta_n, \mu) = \sup_{|\boldsymbol{u}-\boldsymbol{v}| \le \delta_n} \frac{|\alpha_n(\boldsymbol{u}) - \alpha_n(\boldsymbol{v})|}{|\boldsymbol{u}-\boldsymbol{v}|^{\mu} \vee n^{-\mu}} = \max\{S_{n1}, S_{n2}\}$$

where

$$S_{n1} = \sup_{n^{-1} \le |\boldsymbol{u}-\boldsymbol{v}| \le \delta_n} \frac{|\alpha_n(\boldsymbol{u}) - \alpha_n(\boldsymbol{v})|}{|\boldsymbol{u}-\boldsymbol{v}|^{\mu}}, \quad S_{n2} = \sup_{|\boldsymbol{u}-\boldsymbol{v}| \le n^{-1}} n^{\mu} |\alpha_n(\boldsymbol{u}) - \alpha_n(\boldsymbol{u})|.$$

It suffices to show that $S_{n1} = o_P(1)$ and $S_{n2} = o_P(1)$ as $n \to \infty$.

First consider S_{n2}. We will show that, for any $\ell \in \mathbb{N}$ and any $\beta \in (0,1)$, there exist constants K_1 and K_2 only depending on d, ℓ, β and the constants in (4.1) such that

$$\mathbb{P}\Big(\sup_{|\boldsymbol{u}-\boldsymbol{v}| \le n^{-1}} |\alpha_n(\boldsymbol{u}) - \alpha_n(\boldsymbol{v})| > \varepsilon \Big) \le 3\mathbb{1}(n^{-1/2} > K_1 \varepsilon) + K_2 \varepsilon^{-2\ell} n^{1-\beta\ell}. \tag{4.2}$$

Indeed, $S_{n2} = o_P(1)$ follows by setting $\varepsilon = n^{-\mu}\varepsilon'$, by choosing $\beta > 2\mu$ and by finally choosing ℓ sufficiently large.

In order to prove (4.2), we begin by bounding the left-hand side of that display by

$$\mathbb{P}\Big(\sup_{|\boldsymbol{u}-\boldsymbol{v}|\le n^{-1}} \Big| \frac{1}{\sqrt{n}} \sum_{i=1}^n \mathbb{1}(\boldsymbol{U}_i \le \boldsymbol{u}) - \mathbb{1}(\boldsymbol{U}_i \le \boldsymbol{v}) \Big| > \frac{\varepsilon}{2} \Big) + \mathbb{P}\Big(\sup_{|\boldsymbol{u}-\boldsymbol{v}|\le n^{-1}} \sqrt{n} |C(\boldsymbol{u}) - C(\boldsymbol{v})| > \frac{\varepsilon}{2} \Big),$$

where the second probability is smaller than $\mathbb{1}(n^{-1/2} > \frac{\varepsilon}{2})$ by Lipschitz-continuity of C. Furthermore, we have

$$\begin{aligned}
&\sup_{|\boldsymbol{u}-\boldsymbol{v}|<n^{-1}} \Big| \frac{1}{\sqrt{n}} \sum_{i=1}^n \mathbb{1}(\boldsymbol{U}_i \le \boldsymbol{u}) - \mathbb{1}(\boldsymbol{U}_i \le \boldsymbol{v}) \Big| \\
&\quad \le \sum_{j=1}^d \sup_{0\le v_j-u_j\le n^{-1}} \frac{1}{\sqrt{n}} \sum_{i=1}^n \mathbb{1}(U_{ij} \le v_j) - \mathbb{1}(U_{ij} \le u_j) \\
&\quad = \sum_{j=1}^d \sup_{0\le v_j-u_j\le n^{-1}} \sqrt{n}\{G_{nj}(v_j) - G_{nj}(u_j)\} \\
&\quad \le \sum_{j=1}^d \sup_{|v_j-u_j|\le n^{-1}} \sqrt{n}|G_{nj}(v_j) - G_{nj}(u_j) - (v_j - u_j)| + \frac{d}{\sqrt{n}} \\
&\quad = \sum_{j=1}^d \sup_{|v_j-u_j|\le n^{-1}} |\alpha_{nj}(v_j) - \alpha_{nj}(u_j)| + \frac{d}{\sqrt{n}}.
\end{aligned}$$

We now proceed similar as in the proof of Lemma 8.6 in Kley et al. (2014) to bound the sum on the right-hand side. Set $M_n = \{0, \frac{1}{n}, \frac{2}{n}, \dots, 1\}$. Monotonicity of G_{nj} yields

$$\begin{aligned}
&\sup_{|u_j-v_j|\le n^{-1}} \sqrt{n}|G_{nj}(u_j) - G_{nj}(v_j) - (u_j - v_j)| \\
&\quad \le \max_{u_j,v_j\in M_n:|u_j-v_j|\le 2/n} \sqrt{n}|G_{nj}(u_j) - G_{nj}(v_j) - (u_j - v_j)| + 2/\sqrt{n}.
\end{aligned}$$

Therefore, we get

$$\begin{aligned}
&\mathbb{P}\Big(\sup_{|\boldsymbol{u}-\boldsymbol{v}|\le n^{-1}} \Big| \frac{1}{\sqrt{n}} \sum_{i=1}^n \mathbb{1}(\boldsymbol{U}_i \le \boldsymbol{u}) - \mathbb{1}(\boldsymbol{U}_i \le \boldsymbol{v}) \Big| > \frac{\varepsilon}{2} \Big) \\
&\quad \le \mathbb{P}\Big(\sum_{j=1}^d \max_{u_j,v_j\in M_n:|u_j-v_j|\le 2/n} |\alpha_{nj}(v_j) - \alpha_{nj}(u_j)| > \frac{\varepsilon}{6} \Big) \\
&\qquad + \mathbb{1}\Big(n^{-1/2} > \frac{\varepsilon}{6d}\Big) + \mathbb{1}\Big(n^{-1/2} > \frac{\varepsilon}{12}\Big).
\end{aligned}$$

Now, note that the set $\{(u,v) \in M_n^2 : |u-v| \le 2/n\}$ contains $O(n)$ elements. Since $\mathbb{E}(\max_{i=1,\dots,m} |Y_i|^p) \le m \times \max_{i=1,\dots,m} \mathbb{E}(|Y_i|^p)$ for any random variables $Y_1, \dots, Y_m$, we can conclude that

$$\begin{aligned}
&\mathbb{P}\Big(\sum_{j=1}^d \max_{u_j,v_j \in M_n : |u_j - v_j| \le 2/n} |\alpha_{nj}(v_j) - \alpha_{nj}(u_j)| > \frac{\varepsilon}{8}\Big) \\
&\quad \le \sum_{j=1}^d \mathbb{P}\Big(\max_{u_j,v_j \in M_n : |u_j - v_j| \le 2/n} |\alpha_{nj}(v_j) - \alpha_{nj}(u_j)| > \frac{\varepsilon}{8d}\Big) \\
&\quad \le \sum_{j=1}^d (8d)^{2\ell} \varepsilon^{-2\ell} \mathbb{E}\Big[\max_{u_j,v_j \in M_n : |u_j - v_j| \le 2/n} |\alpha_{nj}(v_j) - \alpha_{nj}(u_j)|^{2\ell}\Big] \\
&\quad \le \text{const} \times \varepsilon^{-2\ell} \sum_{j=1}^d n \sup_{|u_j - v_j| \le 2/n} \mathbb{E}\Big[|\alpha_{nj}(v_j) - \alpha_{nj}(u_j)|^{2\ell}\Big].
\end{aligned}$$

The assertion in (4.2) now follows from an inequality in the proof of Lemma 8.6 in Kley et al. (2014). These authors showed that, if (4.1) is satisfied, then, for any $\ell \in \mathbb{N}$, there exist constants c_1 and c_2 which only depend on ℓ and the constants in (4.1) such that

$$\sup_{u_j,v_j \in [0,1] : |u_j - v_j| \le \delta} \mathbb{E}|\alpha_{nj}(u_j) - \alpha_{nj}(v_j)|^{2\ell} \le c_2[\{\delta(1 + |\log \delta|^{c_1})\} \vee n^{-1}]^{\ell}.$$

Set $\delta = 2/n$ and exploit that $\log n \le n^{(1-\beta)/c_1}$ for $\beta \in (0,1)$ to get rid of the logarithmic term on the right-hand side to finally arrive at (4.2).

It remains to be shown that $S_{n1} = o_P(1)$. We have

$$\begin{aligned}
&\mathbb{P}\Big(\sup_{n^{-1} \le |\boldsymbol{u} - \boldsymbol{v}| \le \delta_n} \frac{|\alpha_n(\boldsymbol{u}) - \alpha_n(\boldsymbol{v})|}{|\boldsymbol{u} - \boldsymbol{v}|^{\mu}} > \varepsilon\Big) \\
&\quad \le \mathbb{P}\Big(\max_{k : n^{-1} < 2^{-k}\delta_n} \sup_{2^{-(k+1)}\delta_n \le |\boldsymbol{u} - \boldsymbol{v}| \le 2^{-k}\delta_n} \frac{|\alpha_n(\boldsymbol{u}) - \alpha_n(\boldsymbol{v})|}{|\boldsymbol{u} - \boldsymbol{v}|^{\mu}} > \varepsilon\Big) \\
&\quad \le \sum_{k : n^{-1} < 2^{-k}\delta_n} \mathbb{P}\Big(\sup_{2^{-(k+1)}\delta_n \le |\boldsymbol{u} - \boldsymbol{v}| \le 2^{-k}\delta_n} |\alpha_n(\boldsymbol{u}) - \alpha_n(\boldsymbol{v})| > \varepsilon(2^{-k}\delta_n)^{\mu} 2^{-\mu}\Big) \\
&\quad \le \sum_{k : n^{-1} < 2^{-k}\delta_n} \mathbb{P}\Big(\sup_{|\boldsymbol{u} - \boldsymbol{v}| \le 2^{-k}\delta_n} |\alpha_n(\boldsymbol{u}) - \alpha_n(\boldsymbol{v})| > \varepsilon(2^{-k}\delta_n)^{\mu} 2^{-\mu}\Big).
\end{aligned}$$

Therefore, we only have to show, that

$$\sum_{k : n^{-1} < 2^{-k}\delta_n} \mathbb{P}\Big(\sup_{|\boldsymbol{u} - \boldsymbol{v}| \le 2^{-k}\delta_n} |\alpha_n(\boldsymbol{u}) - \alpha_n(\boldsymbol{v})| > \varepsilon(2^{-k}\delta_n)^{\mu} 2^{-\mu}\Big) = o(1). \tag{4.3}$$

We will show later that, for any $L \in \mathbb{N}$ and for any $\gamma \in (0, 1/2)$, there exists a constant $K = K(\gamma, L) > 0$ such that, for all $\boldsymbol{u}, \boldsymbol{v} \in [0,1]^d$ with

$|\boldsymbol{u}-\boldsymbol{v}| \geq n^{-1}$,

$$\|\alpha_n(\boldsymbol{u}) - \alpha_n(\boldsymbol{v})\|_{2L} \leq K|\boldsymbol{u}-\boldsymbol{v}|^\gamma =: K\mathrm{d}(\boldsymbol{u},\boldsymbol{v}), \tag{4.4}$$

where $\|X\|_p = \mathbb{E}[|X|^p]^{1/p}$. Note that the packing number $D(\varepsilon, \mathrm{d})$ of the metric space $([0,1]^d, \mathrm{d})$ satisfies $D(\varepsilon, \mathrm{d}) \leq \text{const} \times \varepsilon^{-d/\gamma}$. Then, using the notation $\Psi(x) = x^{2L}$, $\Psi^{-1}(x) = x^{1/(2L)}$, $\delta = (2^{-k}\delta_n)^\gamma$ and $\bar{\eta} = 2n^{-\gamma}$, Lemma 7.1 in Kley et al. (2014) yields the existence of a random variable S_1 such that

$$\begin{aligned}
&\mathbb{P}\Big(\sup_{|\boldsymbol{u}-\boldsymbol{v}|\leq 2^{-k}\delta_n} |\alpha_n(\boldsymbol{u}) - \alpha_n(\boldsymbol{v})| > \varepsilon(2^{-k}\delta_n)^\mu 2^{-\mu}\Big) \\
&\qquad \leq \mathbb{P}\Big(S_1 > (2^{-k}\delta_n)^\mu 2^{-\mu-1}\varepsilon\Big) \\
&\qquad\qquad + \mathbb{P}\Big(2 \sup_{|\boldsymbol{u}-\boldsymbol{v}|\leq n^{-1}, \boldsymbol{u}\in\tilde{T}} |\alpha_n(\boldsymbol{u}) - \alpha_n(\boldsymbol{v})| > (2^{-k}\delta_n)^\mu 2^{-\mu-1}\varepsilon\Big),
\end{aligned}$$

where $\tilde{T}$ denotes a finite set of cardinality $O(n^d)$ and where, for any $\eta > \bar{\eta}$,

$$\begin{aligned}
&\mathbb{P}(|S_1| > (2^{-k}\delta_n)^\mu 2^{-\mu-1}\varepsilon) \\
&\qquad\qquad \leq \text{const} \times \left[\frac{\int_0^\eta \varepsilon^{-\frac{d}{2\gamma L}}\, \mathrm{d}\varepsilon + \{(2^{-k}\delta_n)^\gamma + 4n^{-\gamma}\}\eta^{-\frac{d}{\gamma L}}}{(2^{-k}\delta_n)^\mu 2^{-\mu-1}\varepsilon}\right]^{2L}.
\end{aligned}$$

Set $\eta = 2(2^{-k}\delta_n)^{\gamma/(1+\frac{d}{\gamma 2L})}$, choose γ and L such that $d < 2\gamma L$ and note that $4n^{-\gamma} \leq 4(2^{-k}\delta_n)^\gamma$. Then

$$\mathbb{P}(|S_1| > (2^{-k}\delta_n)^\mu 2^{-\mu-1}\varepsilon) \leq \text{const} \times \Big\{(2^{-k}\delta_n)^{-\mu+\gamma\frac{1-\frac{d}{2\gamma L}}{1+\frac{d}{2\gamma L}}}\Big\}^{2L},$$

where the constant may depend on $\varepsilon, \gamma, \mu, d, L$. Therefore, choosing L and γ sufficiently large, we obtain that

$$\mathbb{P}(|S_1| > (2^{-k}\delta_n)^\mu 2^{-\mu-1}\varepsilon) \leq \text{const} \times (2^{-k}\delta_n)^\kappa$$

for some $\kappa > 0$.

Furthermore, (4.2) and the fact that $2^{-k}\delta_n \geq n^{-1}$ implies that

$$\begin{aligned}
&\mathbb{P}\Big(2 \sup_{|\boldsymbol{u}-\boldsymbol{v}|\leq n^{-1}, \boldsymbol{u}\in\tilde{T}} |\alpha_n(\boldsymbol{u}) - \alpha_n(\boldsymbol{v})| > (2^{-k}\delta_n)^\mu 2^{-\mu-1}\varepsilon\Big) \\
&\qquad\qquad\qquad\qquad \leq \text{const} \times n^{-\bar{\beta}} + 3\mathbb{1}(n^{\mu-1/2} > \text{const})
\end{aligned}$$

for some $\bar{\beta} > 0$, by choosing $\beta \in (2\mu, 1)$ and ℓ sufficiently large. Therefore,

$$\begin{aligned}
&\sum_{k: n^{-1} < 2^{-k}\delta_n} \mathbb{P}\Big(\sup_{|\boldsymbol{u}-\boldsymbol{v}|\leq 2^{-k}\delta_n} |\alpha_n(\boldsymbol{u}) - \alpha_n(\boldsymbol{v})| > \varepsilon(2^{-k}\delta_n)^\mu 2^{-\mu}\Big) \\
&\qquad \leq \text{const}\left\{\log(n)\{n^{-\bar{\beta}} + 3\mathbb{1}(n^{\mu-1/2} > \text{const})\} + \delta_n^\kappa \sum_{k=0}^\infty 2^{-k\kappa}\right\} = o(1),
\end{aligned}$$

where the logarithmic term is due to the fact that there are at most $O(\log n)$ summands such that $(2^{-k}\delta_n) > n^{-1}$. The last display is exactly (4.3).

Finally, it remains to be shown that (4.4) is satisfied. For $i = 1, \dots, n$, let $A_i(\boldsymbol{u}, \boldsymbol{v}) = \mathbb{1}(\boldsymbol{U}_i \le \boldsymbol{u}) - \mathbb{1}(\boldsymbol{U}_i \le \boldsymbol{v}) - \{C(\boldsymbol{u}) - C(\boldsymbol{v})\}$. Then, by Theorem 2.3.2 in Brillinger (1975),

$$
\begin{aligned}
\mathbb{E}\{\alpha_n(\boldsymbol{u}) - \alpha_n(\boldsymbol{v})\}^{2L} &= n^{-L}\mathbb{E}\Big\{\sum_{i=1}^n A_i(\boldsymbol{u}, \boldsymbol{v})\Big\}^{2L} \\
&= n^{-L} \operatorname{cum}\Big(\prod_{j=1}^{2L}\sum_{i=1}^n A_i(\boldsymbol{u}, \boldsymbol{v})\Big) \\
&= n^{-L} \sum_{\nu_1,\dots,\nu_R} \prod_{r=1}^R \operatorname{cum}\Big(\sum_{i=1}^n A_i(\boldsymbol{u}, \boldsymbol{v}), j \in \nu_r\Big),
\end{aligned}
$$

where the sum runs over all partitions of the set $\{1, \dots, 2L\}$ and where $\operatorname{cum}(Y_j, j \in \nu)$ denotes the joint cumulant of all random variables Y_j with $j \in \nu$. Note that, for ν_r with $|\nu_r| = 1$, we have $\operatorname{cum}\big(\sum_{i=1}^n A_i(\boldsymbol{u}, \boldsymbol{v}), j \in \nu_r\big) = \mathbb{E}\sum_{i=1}^n A_i(\boldsymbol{u}, \boldsymbol{v}) = 0$, whence it is sufficient to consider $R \le L$. In that case, an application of Lemma 4.6 implies that there exist constants $0 < C, C' < \infty$ such that

$$
\begin{aligned}
\operatorname{cum}\Big(\sum_{i=1}^n A_i(\boldsymbol{u}, \boldsymbol{v}), j \in \nu_r\Big) &\le C(n+1)|\boldsymbol{u} - \boldsymbol{v}|(1 + \big|\log|\boldsymbol{u} - \boldsymbol{v}|\big|)^{C'} \\
&\le \bar{K}(n+1)|\boldsymbol{u} - \boldsymbol{v}|^{2\gamma}.
\end{aligned}
$$

Hence, for any $\boldsymbol{u}, \boldsymbol{v} \in [0,1]^d$ such that $|\boldsymbol{u} - \boldsymbol{v}| > n^{-1}$,

$$
\begin{aligned}
\mathbb{E}\{\alpha_n(\boldsymbol{u}) - \alpha_n(\boldsymbol{v})\}^{2L} &\le \text{const} \times \sum_{\nu_1,\dots,\nu_R, R \le L} (n+1)^{R-L}|\boldsymbol{u} - \boldsymbol{v}|^{2R\gamma} \\
&\le \text{const} \times |\boldsymbol{u} - \boldsymbol{v}|^{2L\gamma},
\end{aligned}
$$

which is exactly (4.4). □

4.3 Proof of Theorem 4.5

Throughout the proof, we will use the following additional notations. Set

$$
C_n(\boldsymbol{u}) = G_n\{\boldsymbol{G}_n^-(\boldsymbol{u})\}, \qquad \boldsymbol{G}_n^-(\boldsymbol{u}) = \big(G_{n1}^-(u_1), \dots, G_{nd}^-(u_d)\big)
$$

and define a version of the empirical copula process based on C_n by

$$
\boldsymbol{u} \mapsto \mathbb{C}_n(\boldsymbol{u}) = \sqrt{n}\{C_n(\boldsymbol{u}) - C(\boldsymbol{u})\}.
$$

Moreover, for $0 < a < b < 1/2$, define

$$
N(a, b) = \{\boldsymbol{u} \in [0,1]^d | a < g_1(\boldsymbol{u}) \le b\}.
$$

Note that $[0,1]^d = \{\boldsymbol{u} : g_1(\boldsymbol{u}) = 0\} \cup N(0,a) \cup N(a,1/2)$. The set $N(a,1/2)$ consists of those vectors such that all of their coordinates are larger than a and such that at most $d-2$ coordinates are larger than or equal to $1-a$. In particular, for $d=2$, we have $N(a,1/2) = (a,1-a)^2$.

The proof of Theorem 4.5 will be based on the following sequence of Lemmas. All convergences are with respect to $n \to \infty$.

Lemma 4.7. *Under the conditions of Theorem 4.5,*

$$\sup_{\boldsymbol{u} \in N(cn^{-1},1/2)} \Big| \frac{\hat{\mathbb{C}}_n(\boldsymbol{u})}{g_\omega(\boldsymbol{u})} - \frac{\mathbb{C}_n(\boldsymbol{u})}{g_\omega(\boldsymbol{u})} \Big| = o_P(1).$$

Lemma 4.8. *Under the conditions of Theorem 4.5,*

$$\sup_{\boldsymbol{u} \in N(n^{-1/2},1/2)} \Big| \frac{\mathbb{C}_n(\boldsymbol{u})}{g_\omega(\boldsymbol{u})} - \frac{\bar{\mathbb{C}}_n(\boldsymbol{u})}{g_\omega(\boldsymbol{u})} \Big| = o_P(1).$$

Lemma 4.9. *Under the conditions of Theorem 4.5, for any $\delta_n \downarrow 0$ such that $\delta_n \geq cn^{-1}$,*

$$\sup_{\boldsymbol{u} \in N(cn^{-1},\delta_n)} \Big| \frac{\mathbb{C}_n(\boldsymbol{u})}{g_\omega(\boldsymbol{u})} \Big| = o_P(1).$$

Lemma 4.10. *Under the conditions of Theorem 4.5, for any $\delta_n \downarrow 0$,*

$$\sup_{\boldsymbol{u} \in N(0,\delta_n)} \Big| \frac{\bar{\mathbb{C}}_n(\boldsymbol{u})}{g_\omega(\boldsymbol{u})} \Big| = o_P(1).$$

Lemma 4.11. *Under the conditions of Theorem 4.5, for any $\delta_n \downarrow 0$*

$$\sup_{\boldsymbol{u},\boldsymbol{u}' \in [\frac{c}{n},1-\frac{c}{n}]^d : |\boldsymbol{u}-\boldsymbol{u}'| \leq \delta_n} \Big| \frac{\mathbb{C}_n(\boldsymbol{u})}{g_\omega(\boldsymbol{u})} - \frac{\mathbb{C}_n(\boldsymbol{u}')}{g_\omega(\boldsymbol{u}')} \Big| = o_P(1) \tag{4.5}$$

and

$$\sup_{\boldsymbol{u},\boldsymbol{u}' \in [0,1]^d : |\boldsymbol{u}-\boldsymbol{u}'| \leq \delta_n} \Big| \frac{\bar{\mathbb{C}}_n(\boldsymbol{u})}{\tilde{g}_\omega(\boldsymbol{u})} - \frac{\bar{\mathbb{C}}_n(\boldsymbol{u}')}{\tilde{g}_\omega(\boldsymbol{u}')} \Big| = o_P(1) \tag{4.6}$$

Proof of Theorem 4.5. Set $\delta_n = dn^{-1/2}$. Given $\boldsymbol{u} \in [\frac{c}{n}, 1-\frac{c}{n}]^d$, choose $\boldsymbol{u}' \in [\frac{1}{\sqrt{n}}, 1-\frac{1}{\sqrt{n}}]^d$ such that $|\boldsymbol{u}-\boldsymbol{u}'| \leq \delta_n$. Since

$$\Big| \frac{\hat{\mathbb{C}}_n(\boldsymbol{u})}{g_\omega(\boldsymbol{u})} - \frac{\bar{\mathbb{C}}_n(\boldsymbol{u})}{g_\omega(\boldsymbol{u})} \Big| \leq \Big| \frac{\hat{\mathbb{C}}_n(\boldsymbol{u})}{g_\omega(\boldsymbol{u})} - \frac{\mathbb{C}_n(\boldsymbol{u})}{g_\omega(\boldsymbol{u})} \Big| + \Big| \frac{\mathbb{C}_n(\boldsymbol{u})}{g_\omega(\boldsymbol{u})} - \frac{\mathbb{C}_n(\boldsymbol{u}')}{g_\omega(\boldsymbol{u}')} \Big|$$
$$+ \Big| \frac{\mathbb{C}_n(\boldsymbol{u}')}{g_\omega(\boldsymbol{u}')} - \frac{\bar{\mathbb{C}}_n(\boldsymbol{u}')}{g_\omega(\boldsymbol{u}')} \Big| + \Big| \frac{\bar{\mathbb{C}}_n(\boldsymbol{u}')}{g_\omega(\boldsymbol{u}')} - \frac{\bar{\mathbb{C}}_n(\boldsymbol{u})}{g_\omega(\boldsymbol{u})} \Big|$$

the first assertion of the theorem follows from Lemma 4.7, 4.8 and 4.11.

Next, let us show that $\bar{\mathbb{C}}_n/\tilde{g}_\omega \rightsquigarrow \mathbb{C}_C/\tilde{g}_\omega$ in $(\ell^\infty([0,1]^d), \|\cdot\|_\infty)$. From Problem 2.1.5 in Van der Vaart and Wellner (1996) and Lemma 4.11 we obtain that $\bar{\mathbb{C}}_n/\tilde{g}_\omega$ is asymptotically equicontinuous. Furthermore, Condition 4.2 yields that the finite dimensional distributions of $\bar{\mathbb{C}}_n/\tilde{g}_\omega$ converge weakly to the finite dimensional distributions of $\mathbb{C}_C/\tilde{g}_\omega$. Note that $\mathbb{C}_C/\tilde{g}_\omega(\boldsymbol{u}) = \bar{\mathbb{C}}_n/\tilde{g}_\omega(\boldsymbol{u}) = 0$ for any $\boldsymbol{u}$ with at least one entry equal to 0 or with $d-1$ entries equal to 1. □

Proof of Lemma 4.7. It suffices to show that, there exists $\mu \in (\omega, \theta_1)$ such that

$$\sup_{\boldsymbol{u}\in[0,1]^d} |\hat{C}_n(\boldsymbol{u}) - C_n(\boldsymbol{u})| = o_P(n^{-1/2-\mu}).$$

Note that $F_{nj}(X_{ij}) = G_{nj}(U_j)$, whence

$$\begin{aligned}
&\sup_{\boldsymbol{u}\in[0,1]^d} |\hat{C}_n(\boldsymbol{u}) - C_n(\boldsymbol{u})| \\
&\quad\le \sup_{\boldsymbol{u}\in[0,1]^d} \left| \frac{1}{n}\sum_{i=1}^n \mathbb{1}\{\boldsymbol{G}_n(\boldsymbol{U}_i) \le \tfrac{n+1}{n}\boldsymbol{u}\} - \mathbb{1}\{\boldsymbol{G}_n(\boldsymbol{U}_i) \le \boldsymbol{u}\} \right| \\
&\qquad + \sup_{\boldsymbol{u}\in[0,1]^d} \left| \frac{1}{n}\sum_{i=1}^n \mathbb{1}\{\boldsymbol{G}_n(\boldsymbol{U}_i) \le \boldsymbol{u}\} - \mathbb{1}\{\boldsymbol{U}_i \le \boldsymbol{G}_n^-(\boldsymbol{u})\} \right| \\
&\quad\le \sum_{j=1}^d \Bigg[\sup_{u\in[0,1]} \frac{1}{n}\sum_{i=1}^n \mathbb{1}\{u < G_{nj}(U_{ij}) \le \tfrac{n+1}{n}u\} \\
&\qquad + \sup_{u\in[0,1]} \frac{1}{n}\sum_{i=1}^n \Big| \mathbb{1}\{G_{nj}(U_{ij}) \le u\} - \mathbb{1}\{U_{ij} \le G_{nj}^-(u)\} \Big| \Bigg]
\end{aligned}$$

From the definition of the empirical distribution function and the generalized inverse function we have that, for any fixed u, both $\sum_{i=1}^n \mathbb{1}\{u < G_{nj}(U_{ij}) \le \frac{n+1}{n}u\}$ and $\sum_{i=1}^n |\mathbb{1}\{G_{nj}(U_{ij}) \le u\} - \mathbb{1}\{U_{ij} \le F_{nj}^-(u)\}|$ are bounded by the maximal number of U_{ij} which are equal. Note that this maximal number is equal to $n \times \sup_{u\in[0,1]} |G_{nj}(u) - G_{nj}(u-)|$. Provided there are no ties among $U_{1j}, \dots, U_{nj}$, for any $j = 1, \dots, d$ (which, for instance, occurs in the i.i.d. case), this expression is equal to 1 and the Lemma is proven. In the general case, we have

$$\begin{aligned}
\sup_{u\in[0,1]} |G_{nj}(u) - G_{nj}(u-)| &\le \sup_{\substack{u,v\in[0,1]\\|u-v|\le 1/n}} |G_{nj}(u) - G_{nj}(v)| \\
&\le \sup_{\substack{u,v\in[0,1]\\|u-v|\le 1/n}} |G_{nj}(u) - G_{nj}(v) - (u-v)| + \frac{1}{n} \\
&\le \frac{1}{\sqrt{n}} \sup_{\substack{\boldsymbol{u},\boldsymbol{v}\in[0,1]^d\\|\boldsymbol{u}-\boldsymbol{v}|\le 1/n}} |\alpha_n(\boldsymbol{u}) - \alpha_n(\boldsymbol{v})| + \frac{1}{n} \qquad (4.7)
\end{aligned}$$

Then, the assertion follows from Condition 4.1. □

Proof of Lemma 4.8. First of all, we write

$$\mathbb{C}_n(\boldsymbol{u}) - \bar{\mathbb{C}}_n(\boldsymbol{u}) = (B_{n1} + B_{n2} + B_{n3})(\boldsymbol{u})$$

where

$$B_{n1}(\boldsymbol{u}) = \alpha_n\{\boldsymbol{G}_n^-(\boldsymbol{u})\} - \alpha_n(\boldsymbol{u})$$
$$B_{n2}(\boldsymbol{u}) = \sqrt{n}\left[C\{\boldsymbol{G}_n^-(\boldsymbol{u})\} - C(\boldsymbol{u})\right] - \sum_{j=1}^d \dot{C}_j(\boldsymbol{u})\beta_{nj}(u_j)$$
$$B_{n3}(\boldsymbol{u}) = \sum_{j=1}^d \dot{C}_j(\boldsymbol{u})\left\{\beta_{nj}(u_j) + \alpha_{nj}(u_j)\right\}.$$

For $p = 1, 2, 3$, set $A_{np}(\boldsymbol{u}) = B_{np}(\boldsymbol{u})/g_\omega(\boldsymbol{u})$. The Lemma is proved if we show uniform negligibility of each term individually.

Treatment of A_{n1}. Let Ω_n denote the event that $\sup_{\boldsymbol{u}\in[0,1]^d} |\boldsymbol{G}_n^-(\boldsymbol{u}) - \boldsymbol{u}| \le \delta_n = n^{-1/2+\kappa}$, with $\kappa > 0$ to be specified later on. Note that the probability of Ω_n converges to 1. Exploiting Condition 4.1 and the fact that $|g_\omega(\boldsymbol{u})|^{-1} \le n^{\omega/2}$ for $\boldsymbol{u} \in N(n^{-1/2}, 1/2)$ we obtain, for any $\mu \in (0, \theta_1)$,

$$\begin{aligned}\sup_{\boldsymbol{u}\in N(n^{-1/2},1/2)} |A_{n1}(\boldsymbol{u})| &\le n^{\omega/2} \sup_{\boldsymbol{u}\in[0,1]^d} \left|\alpha_n\{\boldsymbol{G}_n^-(\boldsymbol{u})\} - \alpha_n(\boldsymbol{u})\right| \\ &\le n^{\omega/2} M_n(\delta_n, \mu) \sup_{\boldsymbol{u}\in[0,1]^d} \left\{|\boldsymbol{G}_n^-(\boldsymbol{u}) - \boldsymbol{u}|^\mu \vee n^{-\mu}\right\} \mathbb{1}_{\Omega_n} \\ &\qquad + o_P(1) \\ &\le n^{\omega/2-\mu/2+\kappa\mu} o_P(1) + o_P(1)\end{aligned}$$

The right-hand side is $o_P(1)$ if we choose $\mu \in (\omega, \theta_1)$ sufficiently large and $\kappa > 0$ sufficiently small such that $\omega < \mu(1 - 2\kappa)$.

Treatment of A_{n2}. Fix $\boldsymbol{u} \in N(n^{-1/2}, 1/2)$. Let $S = S_{\boldsymbol{u}}$ denote the set of all $j \in \{1, \dots, d\}$ such that $u_j \in [n^{-1/2}, 1 - n^{-\gamma}]$, with $\gamma > 1/2$ to be specified later. Let $(\boldsymbol{G}_n^-(\boldsymbol{u}))_S$ denote the vector in $\mathbb{R}^d$ whose jth coordinate is equal to $G_{nj}^-(u_j)\mathbb{1}(j \in S) + u_j\mathbb{1}(j \notin S)$. Write $A_{n2}(\boldsymbol{u}) = D_{n1}(\boldsymbol{u}) + D_{n2}(\boldsymbol{u})$, where

$$D_{n1}(\boldsymbol{u}) = \left(\sqrt{n}\left[C\{\boldsymbol{G}_n^-(\boldsymbol{u})\} - C\{(\boldsymbol{G}_n^-(\boldsymbol{u}))_S\}\right] - \sum_{j\notin S} \dot{C}_j(\boldsymbol{u})\beta_{nj}(u_j)\right) g_\omega^{-1}(\boldsymbol{u}),$$
$$D_{n2}(\boldsymbol{u}) = \left(\sqrt{n}\left[C\{(\boldsymbol{G}_n^-(\boldsymbol{u}))_S\} - C(\boldsymbol{u})\right] - \sum_{j\in S} \dot{C}_j(\boldsymbol{u})\beta_{nj}(u_j)\right) g_\omega^{-1}(\boldsymbol{u}).$$

Since $\dot{C}_j \in [0, 1]$, we can bound

$$D_{n1}(\boldsymbol{u}) \le 2 \sum_{j\notin S} \left|\frac{\beta_{nj}(u_j)}{g_\omega(\boldsymbol{u})}\right| \le 2 \sum_{j=1}^d \sup_{u_j\in[1-n^{-\gamma},1]} \left|\frac{\beta_{nj}(u_j)}{n^{-\omega/2}}\right|.$$

The right-hand side is $o_P(1)$ by Lemma 5.3. Regarding D_{n2}, by Taylor's Theorem, $|D_{n2}(\boldsymbol{u})| = \frac{1}{2}\sum_{j_1,j_2\in S} D_{n2}^{j_1 j_2}(\boldsymbol{u})$, where

$$D_{n2}^{j_1 j_2}(\boldsymbol{u}) = n^{-1/2}\ddot{C}_{j_1 j_2}(\boldsymbol{\xi}_n)\beta_{nj_1}(u_{j_1})\beta_{nj_2}(u_{j_2})g_\omega(\boldsymbol{u})^{-1},$$

and where $\boldsymbol{\xi}_n = (\xi_{n1},\ldots,\xi_{nd})'$ is an intermediate point between $(\boldsymbol{G}_n^-(\boldsymbol{u}))_S$ and $\boldsymbol{u}$. By Condition 2.1, we have

$$|\ddot{C}_{j_1 j_2}(\boldsymbol{\xi}_n)| \le K\{\xi_{nj_1}(1-\xi_{nj_1})\}^{-1/2}\{\xi_{nj_2}(1-\xi_{nj_2})\}^{-1/2}.$$

Therefore, since $g_\omega(\boldsymbol{u})^{-1} \le n^{\omega/2}$,

$$\begin{aligned}|D_{n2}^{j_1 j_2}(\boldsymbol{u})| \le Kn^{-1/2+\omega/2} \sup_{\boldsymbol{u}\in[n^{-1/2},1-n^{-1/2}]^d} \Bigg| &\left\{\frac{u_{j_1}(1-u_{j_1})}{\xi_{nj_1}(1-\xi_{nj_1})}\right\}^{1/2} \\ \times \left\{\frac{u_{j_2}(1-u_{j_2})}{\xi_{nj_2}(1-\xi_{nj_2})}\right\}^{1/2} &\times \frac{|\beta_{nj_1}(u_{j_1})|}{\{u_{j_1}(1-u_{j_1})\}^\omega} \times \frac{|\beta_{nj_2}(u_{j_2})|}{\{u_{j_2}(1-u_{j_2})\}^\omega} \\ &\times \{u_{j_1}(1-u_{j_1})u_{j_2}(1-u_{j_2})\}^{\omega-1/2}\Bigg|.\end{aligned}$$

By an application of Lemma 5.2 and by Condition 4.3, the right-hand side is of order $O_P(n^{-1/2+\omega/2+\gamma(1-2\omega)}) = o_P(1)$, provided we choose $\gamma \in (1/2, \{1/2+\omega/(2-4\omega)\} \wedge \{1/(2(1-\theta_2))\} \wedge \theta_3)$. Since $\boldsymbol{u} \in N(n^{-1/2}, 1/2)$ was arbitrary, we can conclude that $\sup_{\boldsymbol{u}\in N(n^{-1/2},1/2)} |A_{n2}(\boldsymbol{u})| = o_P(1)$.

Treatment of A_{n3}. Since $|\dot{C}_j(\boldsymbol{u})| \le 1$ for any $\boldsymbol{u} \in [0,1]^d$, we have

$$\begin{aligned}\sup_{\boldsymbol{u}\in N(n^{-1/2},1/2)} |A_{n3}(\boldsymbol{u})| \le n^{\omega/2}\sum_{j=1}^d \sup_{u_j\in[0,1]} |\beta_{nj}(u_j) + \alpha_{nj}\{G_{nj}^-(u_j)\}| \\ + n^{\omega/2}\sum_{j=1}^d \sup_{u_j\in[0,1]} \left|\alpha_{nj}\{G_{nj}^-(u_j)\} - \alpha_{nj}(u_j)\right|.\end{aligned}$$

The second sum on the right-hand side is of order $o_P(1)$ as shown in the preceding treatment of the term A_{n1}. Negligibility of the first sum follows from Lemma 5.1, observing that $\alpha_{nj}\{G_{nj}^-(u_j)\} = \sqrt{n}[G_{nj}\{G_{nj}^-(u_j)\} - G_{nj}^-(u_j)]$ from the definition of α_n. □

Proof of Lemma 4.9. Note that, by a monotonicity argument, it suffices to treat sequences δ_n such that $\delta_n \gg n^{-1/2}$, i.e., $\delta_n\sqrt{n} \to \infty$. First of all, choose γ such that $1/2+\omega < \gamma < 1/\{2(1-\theta_2)\} \wedge \theta_3$. Set $M_{n\gamma} = N(n^{-\gamma}, \delta_n) \cap (n^{-\gamma}, 1-n^{-\gamma})^d$ and $M_{n\gamma}^c = N(n^{-\gamma}, \delta_n) \setminus (n^{-\gamma}, 1-n^{-\gamma})^d$, and note that $N(cn^{-1}, \delta_n) = N(cn^{-1}, n^{-\gamma}) \cup M_{n\gamma} \cup M_{n\gamma}^c$. Therefore,

$$\sup_{\boldsymbol{u}\in N(cn^{-1},\delta_n)} \left|\frac{\mathbb{C}_n(\boldsymbol{u})}{g_\omega(\boldsymbol{u})}\right| = R_n\{N(cn^{-1}, n^{-\gamma})\} \vee R_n(M_{n\gamma}) \vee R_n(M_{n\gamma}^c), \quad (4.8)$$

where, for $A \subset [0,1]^d$, $R_n(A) = \sup_{\boldsymbol{u}\in A} |\mathbb{C}_n(\boldsymbol{u})/g_\omega(\boldsymbol{u})|$. It suffices to show negligibility of each term on the right-hand side of (4.8).

Treatment of $R_n\{N(cn^{-1}, n^{-\gamma})\}$. We will distinguish the cases that either $g_\omega(\boldsymbol{u}) = u_1^\omega$ or $g_\omega(\boldsymbol{u}) = (1-u_1)^\omega$. The cases $g_\omega(\boldsymbol{u}) = u_j^\omega$ or $g_\omega(\boldsymbol{u}) = (1-u_j)^\omega$ for some $j > 1$ can be treated similarly.

Let us first consider $\boldsymbol{u}$ such that $g_\omega(\boldsymbol{u}) = u_1^\omega$. Obviously,

$$|C_n(\boldsymbol{u}) - C(\boldsymbol{u})| \le |C_n(\boldsymbol{u}) - C_n(0, u_2, \dots, u_d)| + |C(0, u_2, \dots, u_d) - C(\boldsymbol{u})|.$$

By Lipschitz-continuity of the copula function C, the second term on the right hand side can be bounded by $u_1 = g_1(\boldsymbol{u})$. For the first term, note that

$$\begin{aligned} |C_n(\boldsymbol{u}) - C_n(0, u_2, \dots, u_d)| &= \frac{1}{n}\sum_{i=1}^n \mathbb{1}\{\boldsymbol{U}_i \le \boldsymbol{G}_n^-(\boldsymbol{u})\} \\ &\le \frac{1}{n}\sum_{i=1}^n \mathbb{1}\{U_{i1} \le G_{n1}^-(u_1)\} = G_{n1}\{G_{n1}^-(u_1)\} \end{aligned} \tag{4.9}$$

By Lemma 5.1 the last expression is equal to $u_1 + o_P(n^{-1/2-\mu}) = g_1(\boldsymbol{u}) + o_P(n^{-1/2-\mu})$ for any $\mu \in (\omega, \theta_1)$, where the residual term is uniformly in $u_1 \in [0,1]$. Combined, this yields $|\mathbb{C}_n(\boldsymbol{u})| \le \sqrt{n} 2 g_1(\boldsymbol{u}) + o_P(n^{-\mu})$, and hence

$$\sup_{\boldsymbol{u}\in N(cn^{-1}, n^{-\gamma}), g_\omega(\boldsymbol{u}) = u_1^\omega} \left| \frac{\mathbb{C}_n(\boldsymbol{u})}{g_\omega(\boldsymbol{u})} \right| \le 2n^{1/2+\omega-\gamma} + o_P(n^{-\mu+\omega}) = o_P(1).$$

Now, consider the case $g_\omega(\boldsymbol{u}) = (1-u_1)^\omega$, i.e., $1 - u_1 = 1 - u_1 \wedge \dots \wedge \widehat{u_k} \wedge \dots \wedge u_d$ for some $k \in \{2, \dots, d\}$ and without loss of generality we may assume that $k = 2$. Then, in particular, $1 - u_1 \le 1 - u_2$ and $1 - u_1 \ge 1 - u_j$ for all $j \ge 3$. Now, decompose

$$|C_n(\boldsymbol{u}) - C(\boldsymbol{u})| \le |C_n(\boldsymbol{u}) - C_n(\boldsymbol{u}^{(2)})| + |C_n(\boldsymbol{u}^{(2)}) - C(\boldsymbol{u}^{(2)})| + |C(\boldsymbol{u}^{(2)}) - C(\boldsymbol{u})|.$$

Again by Lipschitz-continuity of the copula function, we have

$$|C(\boldsymbol{u}^{(2)}) - C(\boldsymbol{u})| \le \sum_{j\ne 2} |1 - u_j| \le (d-1)|1-u_1| = (d-1) g_1(\boldsymbol{u}).$$

Furthermore, we have

$$\begin{aligned} |C_n(\boldsymbol{u}) - C_n(\boldsymbol{u}^{(2)})| \le &\ |C_n(\boldsymbol{u}) - C_n\{1, u_2, \dots, u_d\}| \\ &+ |C_n\{1, u_2, \dots, u_d\} - C_n\{1, u_2, 1, u_4, \dots, u_d\}| \\ &+ \dots + |C_n\{1, u_2, 1, 1, \dots, 1, u_d\} - C_n(\boldsymbol{u}^{(2)})| \end{aligned} \tag{4.10}$$

and thus, by similar arguments as in (4.9), $|C_n(\boldsymbol{u}) - C_n(\boldsymbol{u}^{(2)})| \le (d-1)g_1(\boldsymbol{u}) + o_P(n^{-1/2-\mu})$, uniformly in $\boldsymbol{u}$. Finally, from Lemma 5.1

$$|C_n(\boldsymbol{u}^{(2)}) - C(\boldsymbol{u}^{(2)})| = |G_{n2}\{G_{n2}^-(u_2)\} - u_2| = o_P(n^{-1/2-\mu}).$$

Altogether, we obtain

$$\sup_{\boldsymbol{u}\in N(cn^{-1},n^{-\gamma}), g_\omega(\boldsymbol{u})=(1-u_1)^\omega} \left|\frac{\mathbb{C}_n(\boldsymbol{u})}{g_\omega(\boldsymbol{u})}\right| \le 2(d-1)n^{1/2+\omega-\gamma} + o_P(n^{-\mu+\omega}) = o_P(1).$$

Treatment of $R_n(M_{n\gamma})$. Again, let us first treat the case where $g_\omega(\boldsymbol{u}) = u_1^\omega$. We can write $\mathbb{C}_n(\boldsymbol{u})/g_\omega(\boldsymbol{u}) = S_{1n}(\boldsymbol{u}) + S_{2n}(\boldsymbol{u}) + S_{3n}(\boldsymbol{u})$, where

$$\begin{aligned}
S_{1n}(\boldsymbol{u}) &= \sqrt{n}[G_n\{\boldsymbol{G}_n^-(\boldsymbol{u})\} - C\{\boldsymbol{G}_n^-(\boldsymbol{u})\}]/g_\omega(\boldsymbol{u}) \\
S_{2n}(\boldsymbol{u}) &= \sqrt{n}[C\{\boldsymbol{G}_n^-(\boldsymbol{u})\} - C\{G_{n1}^-(u_1), u_2, \dots, u_d\}]/g_\omega(\boldsymbol{u}) \\
S_{3n}(\boldsymbol{u}) &= \sqrt{n}[C\{G_{n1}^-(u_1), u_2, \dots, u_d\} - C(\boldsymbol{u})]/g_\omega(\boldsymbol{u}).
\end{aligned}$$

Lipschitz continuity of the copula C together with Condition 4.3 implies that $\sup_{\boldsymbol{u}\in M_{n\gamma}, g_\omega(\boldsymbol{u})=u_1^\omega} |S_{3n}(\boldsymbol{u})| = o_P(1)$.

Regarding S_{1n}, let Ω_n denote the event that $\sup_{u_1\in[0,\delta_n]} G_{n1}^-(u_1) \le 2\delta_n$. On Ω_n^c, we have $\sqrt{n}\delta_n < \sup_{u_1\in[0,\delta_n]} \sqrt{n}|G_{n1}^-(u_1) - u_1| = O_P(1)$, whence, by the assumption that $\sqrt{n}\delta_n \to \infty$, we get $\Pr(\Omega_n^c) \to 0$. Therefore, by Condition 4.1, for any $\mu \in (0, \theta_1)$, we have

$$\begin{aligned}
|S_{1n}(\boldsymbol{u})| &= \left|\frac{\alpha_n\{\boldsymbol{G}_n^-(\boldsymbol{u})\} - \alpha_n\{0, G_{n2}^-(u_2), \dots, G_{nd}^-(u_d)\}}{u_1^\omega}\right| \\
&\le M_n(2\delta_n, \mu)\left|\frac{\{G_{n1}^-(u_1)\}^\mu \vee n^{-\mu}}{u_1^\omega}\right| \mathbb{1}_{\Omega_n} + o_P(1) \\
&\le o_P(1)\left\{\frac{|G_{n1}^-(u_1) - u_1|^\mu}{u_1^\omega} + u_1^{\mu-\omega}\right\} \vee n^{-\mu+\gamma\omega} + o_P(1),
\end{aligned}$$

where we used subadditivity of the function $x \mapsto x^\mu$, $x \ge 0$. By Condition 4.3, we have

$$\begin{aligned}
\sup_{u_1\in[n^{-\gamma},\delta_n]} \frac{|G_{n1}^-(u_1) - u_1|^\mu}{u_1^\omega} &\le n^{-\mu/2} \sup_{u_1\in[n^{-\gamma},\delta_n]} |u_1^{\omega(\mu-1)}| O_P(1) \\
&= O_P(n^{-\mu/2-\gamma\omega(\mu-1)})
\end{aligned}$$

Exploit that $\gamma < 1$ and choose $\mu \in (\omega/(\omega + 1/2), \theta_1)$ to obtain that, as $n \to \infty$, $\sup_{\boldsymbol{u}\in M_{n\gamma}, g_\omega(\boldsymbol{u})=u_1^\omega} |S_{1n}(\boldsymbol{u})| = o_P(1)$.

Finally, we turn to S_{2n}. The mean value theorem allows to write

$$S_{2n}(\boldsymbol{u}) = \sum_{j=2}^d \frac{\dot{C}_j\{G_{n1}^-(u_1), \zeta_2, \dots, \zeta_d\}\sqrt{n}\{G_{nj}^-(u_j) - u_j\}}{g_\omega(\boldsymbol{u})} =: \sum_{j=2}^d S_{2nj}(\boldsymbol{u})$$

for some intermediate values ζ_j between u_j and $G_{nj}^-(u_j)$, for $j = 2, \ldots, d$. We may consider each summand individually; let us fix $j \in \{2, \ldots, d\}$ and distinguish two cases. First, suppose that $1 - u_j < u_1 = g_1(\boldsymbol{u})$. Then, with $\omega' \in (\omega, \theta_1)$,

$$|S_{2nj}(\boldsymbol{u})| \leq \frac{\sqrt{n}|G_{nj}^-(u_j) - u_j|}{(1-u_j)^{\omega'}}(1-u_j)^{\omega'-\omega} = o_P(1),$$

by Condition 4.3 and the fact that $n^{-\gamma} < (1-u_j) \leq \delta_n$. Now, suppose that $1 - u_j \geq u_1 = g_1(\boldsymbol{u}) > n^{-\gamma}$. Since $\dot{C}_j(0, u_2, \ldots, u_d) = 0$ for any $j = 2, \ldots, d$, another application of the mean value theorem allows to write

$$S_{2nj}(\boldsymbol{u}) = \frac{\ddot{C}_{j1}(\boldsymbol{\xi}_j) G_{n1}^-(u_1)\sqrt{n}\{G_{nj}^-(u_j) - u_j\}}{u_1^\omega},$$

where $\boldsymbol{\xi}_j = (\xi_{j1}, \zeta_2, \ldots, \zeta_d)$ satisfies $\xi_{j1} \in (0, G_{n1}^-(u_1))$. Now, fix $\omega' \in (0, \theta_2)$ such that $\omega' > (1 - \frac{1}{2\gamma}) \vee \omega$. By Condition 2.1 and Lemma 5.2, we have

$$\begin{aligned} |S_{2nj}(\boldsymbol{u})| &\leq \frac{G_{n1}^-(u_1)}{u_1^\omega}\left|\frac{\sqrt{n}\{G_{nj}^-(u_j) - u_j\}}{\{u_j(1-u_j)\}^{\omega'}}\right| \times K\frac{\{u_j(1-u_j)\}^{\omega'}}{\xi_{jj}(1-\xi_{jj})} \\ &\leq \left\{n^{-1/2}\frac{\sqrt{n}|G_{n1}^-(u_1) - u_1|}{u_1^\omega} + u_1^{1-\omega}\right\}\{u_j(1-u_j)\}^{\omega'-1}O_P(1) \end{aligned} \tag{4.11}$$

Observing that $u_j \geq u_1$ as a consequence of $g_\omega(\boldsymbol{u}) = u_1^\omega$ and that $1-u_j \geq u_1$ by assumption, we obtain

$$\{u_j(1-u_j)\}^{\omega'-1} \leq [\{u_j \wedge (1-u_j)\}/2]^{\omega'-1} \leq 2^{1-\omega'}u_1^{\omega'-1} \leq 2u_1^{\omega'-1},$$

where we used the fact that $u(1-u) \geq \{u \wedge (1-u)\}/2$ for all $u \in [0,1]$. Therefore, we can bound the right-hand side of (4.11) by

$$\left\{n^{-1/2}u_1^{\omega'-1}O_P(1) + u_1^{\omega'-\omega}\right\} \times O_P(1),$$

where all O_P-terms are uniform in $\{\boldsymbol{u} \in M_{n\gamma} : g_\omega(\boldsymbol{u}) = u_1^\omega\}$. Thus, by the choice of γ and ω', $\sup_{\boldsymbol{u} \in M_{n\gamma}, g_\omega(\boldsymbol{u}) = u_1^\omega} |S_{2n}(\boldsymbol{u})| = o_P(1)$.

For the treatment of $R_n(M_{n\gamma})$, it remains to consider the case $g_\omega(\boldsymbol{u}) = (1-u_1)^\omega$, i.e., $1 - u_1 = 1 - (u_1 \wedge \cdots \wedge \widehat{u_k} \wedge \cdots \wedge u_d)$ for some $k \in \{2, \ldots, d\}$. Again, without loss of generality, we may assume that $k = 2$, which implies that $1-u_1 \leq 1-u_2$ and $1-u_1 \geq 1-u_j$ for all $j \geq 3$. Note that, additionally, $1 - u_j > n^{-\gamma}$ for all $j = 1, \ldots, d$ since $\boldsymbol{u} \in M_{n\gamma}$. Now,

$$\frac{\mathbb{C}_n(\boldsymbol{u})}{g_\omega(\boldsymbol{u})} = \frac{\alpha_n\{\boldsymbol{G}_n^-(\boldsymbol{u})\} + \sqrt{n}[C\{\boldsymbol{G}_n^-(\boldsymbol{u})\} - C(\boldsymbol{u})]}{g_\omega(\boldsymbol{u})} = \sum_{p=1}^4 T_{pn}(\boldsymbol{u})$$

with

$$T_{1n}(\boldsymbol{u}) = \frac{\alpha_n\{\boldsymbol{G}_n^-(\boldsymbol{u})\} - \alpha_n\{1, G_{n2}^-(u_2), 1, \dots, 1\}}{g_\omega(\boldsymbol{u})}$$

$$T_{2n}(\boldsymbol{u}) = \frac{\alpha_n\{1, G_{n2}^-(u_2), 1, \dots, 1\} + \sqrt{n}\{G_{n2}^-(u_2) - u_2\}}{g_\omega(\boldsymbol{u})}$$

$$T_{3n}(\boldsymbol{u}) = \frac{\sqrt{n}[C\{\boldsymbol{G}_n^-(\boldsymbol{u})\} - C\{G_{n1}^-(u_1), u_2, G_{n3}^-(u_3), \dots, G_{nd}^-(u_d)\}]}{g_\omega(\boldsymbol{u})} - \frac{\sqrt{n}\{G_{n2}^-(u_2) - u_2\}}{g_\omega(\boldsymbol{u})}$$

$$T_{4n}(\boldsymbol{u}) = \frac{\sqrt{n}[C\{G_{n1}^-(u_1), u_2, G_{n3}^-(u_3), \dots, G_{nd}^-(u_d)\} - C(\boldsymbol{u})]}{g_\omega(\boldsymbol{u})}$$

Concerning T_{1n}, we can proceed similar as for S_{1n} above. Define the event Ω_n by $|\boldsymbol{G}_n^-(\boldsymbol{u}) - (1, G_{n2}^-(u_2), 1, \dots, 1)'| \le 2d\delta_n$ and note that $\mathbb{P}(\Omega_n^c) \to 0$. Then, by Condition 4.1 applied with $\mu \in (\omega/(\omega + \frac{1}{2}), \theta_1)$,

$$|T_{1n}(\boldsymbol{u})| \le M_n(2d\delta_n, \mu) \frac{|\boldsymbol{G}_n^-(\boldsymbol{u}) - (1, G_{n2}^-(u_2), 1, \dots, 1)'|^\mu \vee n^{-\mu}}{(1-u_1)^\omega} \mathbb{1}_{\Omega_n} + o_P(1).$$

Use the fact that $\gamma < 1$ and $1-u_1 \ge 1-u_j \ge n^{-\gamma}$ for $j \ge 3$ and subadditivity of $x \mapsto x^\mu$ to bound the right-hand side by

$$O_P(1) \sum_{j \ne 2} \frac{|G_{nj}^-(u_j) - u_j|^\mu + |1-u_j|^\mu}{(1-u_j)^\omega} + o_P(1)$$
$$\le O_P(1) \left\{ \sum_{j \ne 2} n^{-\mu/2 + \omega - \omega\mu} \left\{ \frac{\sqrt{n}|G_{nj}^-(u_j) - u_j|}{(1-u_j)^\omega} \right\}^\mu + \delta_n^{\mu - \omega} \right\} + o_P(1).$$

Therefore, by Condition 4.3 and by the choice of μ, $|T_{1n}(\boldsymbol{u})| = o_P(1)$ uniformly in $\{\boldsymbol{u} \in M_{n\gamma} : g_\omega(\boldsymbol{u}) = (1-u_1)^\omega\}$.

Regarding T_{2n}, by the definition of α_n and since $g_1(\boldsymbol{u}) = 1 - u_1 \ge n^{-1}$,

$$\sup_{\boldsymbol{u} \in M_{n\gamma}, g_1(\boldsymbol{u}) = 1-u_1} |T_{2n}(\boldsymbol{u})| \le n^\omega \sup_{u_2 \in [0,1]} \sqrt{n}|G_{n2}\{G_{n2}^-(u_2)\} - u_2|.$$

An application of Lemma 5.1 with $\mu \in (\omega, \theta_1)$ yields that the right-hand side is of order $o_P(n^{-\mu+\omega}) = o_P(1)$.

Regarding T_{3n}, choose $\omega' \in (\omega \vee (1 - \frac{1}{2\gamma}), \theta_2)$. By the mean-value theorem, we can write

$$T_{3n}(\boldsymbol{u}) = \frac{\sqrt{n}[\dot{C}_2\{G_{n1}^-(u_1), \zeta_2, G_{n3}^-(u_3), \dots, G_{nd}^-(u_d)\} - 1]\{G_{n2}^-(u_2) - u_2\}}{g_\omega(\boldsymbol{u})}$$

for some intermediate value ζ_2 between $G_{n2}^-(u_2)$ and u_2. Due to the fact that $\dot{C}_2\{1, \zeta_2, 1, \ldots, 1\} = 1$, a second application of the mean value theorem allows to write the right-hand side of the last display as

$$T_{3n}(\boldsymbol{u}) = \sum_{j \neq 2} \frac{\sqrt{n} \ddot{C}_{2j}(\boldsymbol{\xi})\{G_{n2}^-(u_2) - u_2\}\{G_{nj}^-(u_j) - 1\}}{g_\omega(\boldsymbol{u})}$$

for some $\boldsymbol{\xi}$ lying between $\boldsymbol{G}_n^-(\boldsymbol{u})$ and $\boldsymbol{u}$. Hence, by Condition 2.1, Condition 4.3 and Lemma 5.2, we can bound T_{3n} as follows:

$$\begin{aligned} |T_{3n}(\boldsymbol{u})|| &\leq \frac{\sqrt{n}|G_{n2}^-(u_2) - u_2|}{\{u_2(1-u_2)\}^{\omega'}} \frac{\{u_2(1-u_2)\}^{\omega'}}{(1-u_1)^\omega} \frac{O_P(1)}{u_2(1-u_2)} \sum_{j \neq 2} |G_{nj}^-(u_j) - 1| \\ &= O_P(1)\{u_2(1-u_2)\}^{\omega'-1} \sum_{j \neq 2} \left\{ \frac{|G_{nj}^-(u_j) - u_j|}{(1-u_j)^\omega} + (1-u_j)^{1-\omega} \right\}. \end{aligned}$$

Since $1 - u_2 \geq 1 - u_1$ and $u_2 \geq 1 - u_1$, the right-hand side is of order $O_P\{n^{-1/2}(1-u_1)^{\omega'-1} + (1-u_1)^{\omega'-\omega}\} = o_P(1)$ uniformly in $\boldsymbol{u} \in M_{n\gamma}$ such that $g_\omega(\boldsymbol{u}) = (1-u_1)^\omega$, by the choice of ω'.

Finally, regarding T_{4n}, Lipschitz-continuity of the copula function and Condition 4.3 immediately imply that for any $\omega' \in (\omega, \theta_2)$

$$|T_{4n}(\boldsymbol{u})| \leq \sum_{j \neq 2} (1-u_1)^{-\omega}(1-u_j)^{\omega'} \frac{\sqrt{n}|G_{nj}^-(u_j) - u_j|}{(1-u_j)^{\omega'}} = O_P((1-u_1)^{\omega'-\omega}),$$

which is of order $o_P(1)$ uniformly in $\{\boldsymbol{u} \in M_{n\gamma} : g_\omega(\boldsymbol{u}) = (1-u_1)^\omega\}$.

Treatment of $R_n(M_{n\gamma}^c)$. First note that, from the definition of $N(n^{-\gamma}, \delta_n)$, for every $\boldsymbol{u} \in M_{n\gamma}^c$ there are at most $d-2$ components larger than or equal to $1 - n^{-\gamma}$. For that reason, we can write

$$M_{n\gamma}^c = \bigcup_{\boldsymbol{\ell} = (\ell_1, \ldots, \ell_d) \in \{0,1\}^d; |\boldsymbol{\ell}| \geq 2} S_{\ell_1} \times \cdots \times S_{\ell_d},$$

where $|\boldsymbol{\ell}| = \sum_{j=1}^d \ell_j$, $S_0 = [1 - n^{-\gamma}, 1]$ and $S_1 = (n^{-\gamma}, 1 - n^{-\gamma})$. In order to show negligibility of $R_n(M_{n\gamma}^c)$, it suffices to fix a vector ℓ with $|\ell| \geq 2$ and to show uniform negligibility of $\mathbb{C}_n / g_\omega$ over $\boldsymbol{u} \in S_{\boldsymbol{\ell}} := S_{\ell_1} \times \cdots \times S_{\ell_d}$.

For $\boldsymbol{u} \in [0,1]^d$, let $\boldsymbol{u}^{(\boldsymbol{\ell})}$ denote the vector whose jth component (with $j = 1, ..., d$) is equal to $\mathbb{1}(\ell_j = 0) + u_j \mathbb{1}(\ell_j = 1)$. Then,

$$\begin{aligned} \sup_{\boldsymbol{u} \in S_{\boldsymbol{\ell}}} \left| \frac{\mathbb{C}_n(\boldsymbol{u})}{g_\omega(\boldsymbol{u})} \right| \leq \sup_{\boldsymbol{u} \in S_{\boldsymbol{\ell}}} & \frac{\sqrt{n}|G_n\{\boldsymbol{G}_n^-(\boldsymbol{u})\} - G_n\{\boldsymbol{G}_n^-(\boldsymbol{u})^{(\boldsymbol{\ell})}\}|}{n^{-\omega\gamma}} \\ &+ \sup_{\boldsymbol{u} \in S_{\boldsymbol{\ell}}} \frac{\sqrt{n}|G_n\{\boldsymbol{G}_n^-(\boldsymbol{u})^{(\boldsymbol{\ell})}\} - C(\boldsymbol{u}^{(\boldsymbol{\ell})})|}{g_\omega(\boldsymbol{u})} \\ &+ \sup_{\boldsymbol{u} \in S_{\boldsymbol{\ell}}} \frac{\sqrt{n}|C(\boldsymbol{u}^{(\boldsymbol{\ell})}) - C(\boldsymbol{u})|}{n^{-\omega\gamma}} \\ &=: I_{n1} + I_{n2} + I_{n3}. \end{aligned}$$

For I_{n3}, by Lipschitz-continuity of C and by the choice of γ,

$$I_{n3} \le n^{1/2+\omega\gamma}\sqrt{d}|\boldsymbol{u} - \boldsymbol{u}^{(\boldsymbol{\ell})}| = O(n^{1/2+\omega\gamma-\gamma}) = o(1).$$

For the treatment of I_{n1}, we can proceed similar as in (4.10) to obtain that $|G_n\{\boldsymbol{G}_n^-(\boldsymbol{u})\} - G_n\{\boldsymbol{G}_n^-(\boldsymbol{u})^{(\boldsymbol{\ell})}\}| \le (d-2)n^{-\gamma} + o_P(n^{-1/2-\mu})$ for any $\mu \in (\omega, \theta_1)$. This yields $I_{n1} = o_P(n^{1/2+\omega\gamma-\gamma} + n^{\omega\gamma-\mu}) = o_P(1)$.

Finally, regarding I_{n2}, note that $g_\omega(\boldsymbol{u}) = g_\omega(\boldsymbol{u}^{(\boldsymbol{\ell})})$. Therefore,

$$I_{n2} = \sup_{\boldsymbol{u}\in S_\ell} \frac{|\mathbb{C}_n(\boldsymbol{u}^{(\boldsymbol{\ell})})|}{g_\omega(\boldsymbol{u}^{(\boldsymbol{\ell})})}$$

All coordinates of vectors in $S_{\boldsymbol{\ell}}$ which are not equal to 1 lie in $(n^{-\gamma}, 1-n^{-\gamma})$. Therefore, I_{n2} can be treated similar as $R_n(M_{n\gamma})$. □

Proof of Lemma 4.10. Again, by a monotonicity argument, it suffices to treat sequences δ_n such that $\delta_n \gg n^{-1/2}$, i.e., $\delta_n\sqrt{n} \to \infty$. Analogously to the proof of Lemma 4.9 we can decompose the supremum. For $1/\{2(1-\omega)\} < \gamma < 1$ we can write

$$\sup_{\boldsymbol{u}\in N(0,\delta_n)} \Big|\frac{\bar{\mathbb{C}}_n(\boldsymbol{u})}{g_\omega(\boldsymbol{u})}\Big| = \sup_{\boldsymbol{u}\in N(0,n^{-\gamma})} \Big|\frac{\bar{\mathbb{C}}_n(\boldsymbol{u})}{g_\omega(\boldsymbol{u})}\Big| \vee \sup_{\boldsymbol{u}\in N(n^{-\gamma},\delta_n)} \Big|\frac{\bar{\mathbb{C}}_n(\boldsymbol{u})}{g_\omega(\boldsymbol{u})}\Big| =: \bar{R}_{1n\gamma} \vee \bar{R}_{2n\gamma}.$$

Therefore, we only have to show that $\bar{R}_{1n\gamma} = o_P(1)$ and $\bar{R}_{2n\gamma} = o_P(1)$. For both of these terms, we will distinguish the cases that either $g_\omega(\boldsymbol{u}) = u_1^\omega$ or $g_\omega(\boldsymbol{u}) = (1-u_1)^\omega$. The cases $g_\omega(\boldsymbol{u}) = u_j^\omega$ or $g_\omega(\boldsymbol{u}) = (1-u_j)^\omega$ for $j > 1$ can be treated similarly.

Treatment of $\bar{R}_{1n\gamma}$. First of all note that by definition

$$\bar{\mathbb{C}}_n(u_1, \ldots, u_{j-1}, 0, u_{j+1}, \ldots, u_d) = \bar{\mathbb{C}}_n(\boldsymbol{u}^{(j)}) = 0 \tag{4.12}$$

and that $\alpha_n(u^{(j)}) = \sqrt{n}\{G_{nj})(u_j) - u_j\}$ for any $j = 1, \ldots, d$. Let us first consider the supremum over those $\boldsymbol{u} \in N(0, n^{-\gamma})$ that additionally satisfy $g_\omega(\boldsymbol{u}) = u_1^\omega$. Choose $\omega' \in (\omega, \theta_2)$, then

$$\begin{aligned}
\Big|\frac{\bar{\mathbb{C}}_n(\boldsymbol{u})}{g_\omega(\boldsymbol{u})}\Big| &= \Big|\frac{\bar{\mathbb{C}}_n(\boldsymbol{u}) - \bar{\mathbb{C}}_n(0, u_2, \ldots, u_d)}{u_1^\omega}\Big| \qquad (4.13)\\
&\le \sqrt{n}\Big|\frac{G_{n1}(u_1)}{u_1^\omega}\Big| + \sqrt{n}u_1^{1-\omega} + \sum_{j=1}^d \Big|\frac{\dot{C}_j(\boldsymbol{u})\sqrt{n}\{G_{nj}(u_j) - u_j\}}{u_1^\omega}\Big| \\
&\le \Big|\frac{\sqrt{n}\{G_{n1}(u_1) - u_1\}}{u_1^{\omega'}}\Big| u_1^{\omega'-\omega} + 2\sqrt{n}u_1^{1-\omega} \\
&\qquad + \sum_{j=1}^d \Big|\frac{\dot{C}_j(\boldsymbol{u})\sqrt{n}\{G_{nj}(u_j) - u_j\}}{u_1^\omega}\Big|
\end{aligned}$$

By Condition 4.3, the first summand on the right-hand side is of order $n^{-\gamma(\omega'-\omega)}O_P(1) = o_P(1)$, by the choice of ω'. The second summand can be bounded by $2n^{1/2-\gamma(1-\omega)} = o(1)$, by the choice of γ. Thus, it remains to be shown that, for any $j = 1, \ldots, d$,

$$\bar{S}_{nj}(\boldsymbol{u}) = \left| \frac{\dot{C}_j(\boldsymbol{u})\sqrt{n}\{G_{nj}(u_j) - u_j\}}{u_1^{\omega}} \right| = o_P(1) \tag{4.14}$$

uniformly in $\{\boldsymbol{u} \in N(0, n^{-\gamma}) : g_\omega(\boldsymbol{u}) = u_1^\omega\}$. For later reference, we even show uniform negligibility on $\boldsymbol{u} \in N(0, \delta_n)$ such that $g_\omega(\boldsymbol{u}) = u_1^\omega$. $\bar{S}_{n1}$ can be bounded by the first term on the right-hand side of (4.13), and, therefore, is $O_P(\delta_n^{\omega'-\omega})$. Now, fix $j \in \{2, \ldots, d\}$. Uniformly in $\boldsymbol{u} \in N(0, \delta_n)$ such that $1 - u_j \le u_1$, we have

$$\bar{S}_{nj}(\boldsymbol{u}) \le \sup_{u_j \in (0,1)} \left| \frac{\sqrt{n}\{G_{nj}(u_j) - u_j\}}{(1-u_j)^{\omega'}} \right| \times n^{-\gamma(\omega'-\omega)} = o_P(1)$$

from Condition 4.3 and since $|\dot{C}_j(\boldsymbol{u})| \le 1$ for any $\boldsymbol{u} \in [0,1]^d$. Note that $\bar{S}_{nj}(\boldsymbol{u}) = 0$, if $u_j = 1$. For the remaining case, i.e., for $\boldsymbol{u}$ such that $1 - u_j > u_1$, we may use the fact that $\dot{C}_j(0, u_2, \ldots, u_d) = 0$. A suitable application of the mean value theorem together with Condition 2.1 implies that, for any $\omega' \in (\omega, \theta_2)$,

$$\begin{aligned}
\bar{S}_{nj}(\boldsymbol{u}) &= |\ddot{C}_{1j}(\xi, u_2, \ldots, u_d) u_1 \sqrt{n}\{G_{nj}(u_j) - u_j\}| / u_1^\omega \\
&\le u_1^{1-\omega}\{u_j(1-u_j)\}^{\omega'-1} K \frac{|\sqrt{n}\{G_{nj}(u_j) - u_j\}|}{\{u_j(1-u_j)\}^{\omega'}} \\
&\le u_1^{\omega'-\omega} O_P(1) = o_P(\delta_n^{\omega'-\omega}) = o_P(1).
\end{aligned}$$

In order to finalize the treatment of $\bar{R}_{1n\gamma}$, let us now consider the case that $g_\omega(\boldsymbol{u}) = (1-u_1)^\omega$, i.e., $1 - u_1 = 1 - u_1 \wedge \cdots \wedge \widehat{u_k} \wedge \cdots \wedge u_d$ for some $k \in \{2, \ldots, d\}$. Without loss of generality, we may assume that $k = 2$, which implies that $1 - u_1 \le 1 - u_2$ and $1 - u_1 \ge 1 - u_j$ for all $j \ge 3$. By definition of $\bar{\mathbb{C}}_n$ and by (4.12), we can write (note that $\bar{\mathbb{C}}_n(\boldsymbol{u}^{(2)}) \equiv 0$ a.s.)

$$\frac{\bar{\mathbb{C}}_n(\boldsymbol{u})}{g_\omega(\boldsymbol{u})} = \frac{\bar{\mathbb{C}}_n(\boldsymbol{u}) - \bar{\mathbb{C}}_n(\boldsymbol{u}^{(2)})}{(1-u_1)^\omega} = \sum_{p=1}^{4} \bar{T}_{np}(\boldsymbol{u}), \tag{4.15}$$

where

$$\bar{T}_{n1}(\boldsymbol{u}) = \frac{\sqrt{n}\{G_n(\boldsymbol{u}) - G_n(\boldsymbol{u}^{(2)})\}}{(1-u_1)^\omega}, \qquad \bar{T}_{n2}(\boldsymbol{u}) = \frac{\sqrt{n}\{C(\boldsymbol{u}^{(2)}) - C(\boldsymbol{u})\}}{(1-u_1)^\omega},$$

$$\bar{T}_{n3}(\boldsymbol{u}) = -\sum_{j \ne 2} \frac{\dot{C}_j(\boldsymbol{u})\sqrt{n}\{G_{nj}(u_j) - u_j\}}{(1-u_1)^\omega},$$

$$\bar{T}_{n4}(\boldsymbol{u}) = \frac{\sqrt{n}\{\dot{C}_2(\boldsymbol{u}^{(2)}) - \dot{C}_2(\boldsymbol{u})\}\{G_{n2}(u_2) - u_2\}}{(1-u_1)^\omega}.$$

By Lipschitz-continuity of the copula function, we immediately obtain that

$$
\begin{aligned}
|\bar{T}_{n2}(\boldsymbol{u})| &\le \sqrt{n}\sum_{j\neq 2}(1-u_j)(1-u_1)^{-\omega} \\
&\le (d-1)\sqrt{n}(1-u_1)^{1-\omega} = O(n^{1/2-\gamma(1-\omega)}) = o(1).
\end{aligned}
$$

For the estimation of $\bar{T}_{n1}$ we can proceed similarly as in (4.10) and obtain

$$
\begin{aligned}
|\bar{T}_{n1}(\boldsymbol{u})| &\le \frac{\sqrt{n}\sum_{j\neq 2}|1-G_{nj}(u_j)|}{(1-u_1)^{\omega}} \\
&\le \sum_{j\neq 2}\frac{|\sqrt{n}\{G_{nj}(u_j)-u_j\}|}{(1-u_j)^{\omega'}}(1-u_j)^{\omega'-\omega} + (d-1)\sqrt{n}(1-u_1)^{1-\omega} \\
&= O_P(n^{-\gamma(\omega'-\omega)}) + O(n^{1/2-\gamma(1-\omega)}) = o_P(1)
\end{aligned}
$$

uniformly in $\{\boldsymbol{u}\in N(0,n^{-\gamma})\cap(0,1)^d : g_\omega(\boldsymbol{u}) = (1-u_1)^{\omega}\}$, by the choice of γ and by choosing $\omega'\in(\omega,\theta_2)$. Note that the terms with $u_j = 1$ vanish immediately from the definition of G_{nj}. Similarly, using the fact that $|\dot{C}_j(\boldsymbol{u})|\le 1$ for all $j=1,\dots,d$ and for all $\boldsymbol{u}\in[0,1]^d$, we get that

$$
|\bar{T}_{n3}(\boldsymbol{u})| \le \sum_{j\neq 2}\sup_{u_j\in(0,1)}\Big|\frac{\sqrt{n}\{G_{nj}(u_j)-u_j\}}{(1-u_j)^{\omega}}\Big| = O_P(n^{-\gamma(\omega'-\omega)}) = o_P(1)
$$

uniformly in $\{\boldsymbol{u}\in N(0,n^{-\gamma}) : g_\omega(\boldsymbol{u}) = (1-u_1)^{\omega}\}$, by choosing $\omega'\in(\omega,\theta_2)$. Finally, in order to bound the remaining term $\bar{T}_{4n}$, we may use the mean value theorem and Conditions 2.1 and 4.3 to obtain that

$$
\begin{aligned}
|\bar{T}_{n4}(\boldsymbol{u})| &= \frac{\sqrt{n}\sum_{j\neq 2}|\ddot{C}_{2j}(\boldsymbol{\xi})(1-u_j)\{G_{n2}(u_2)-u_2\}|}{(1-u_1)^{\omega}} \\
&\le \sum_{j\neq 2}O_P(1)(1-u_1)^{1-\omega}\{u_2(1-u_2)\}^{\omega'-1} = O_P(n^{-\gamma(\omega'-\omega)}) = o_P(1)
\end{aligned}
$$

uniformly in $\{\boldsymbol{u}\in N(0,n^{-\gamma}) : g_\omega(\boldsymbol{u}) = (1-u_1)^{\omega}\}$, with some intermediate value $\boldsymbol{\xi} = (\xi_1,u_2,\xi_3,\dots\dots,\xi_d)'\in(0,1)^d$ between $\boldsymbol{u}^{(2)}$ and $\boldsymbol{u}$ and where the last estimation follows by choosing ω' in (ω,θ_2).

Treatment of $\bar{R}_{2n\gamma}$. First suppose that $g_\omega(\boldsymbol{u}) = u_1^{\omega}$. Then we write

$$
\begin{aligned}
\Big|\frac{\bar{\mathbb{C}}_n(\boldsymbol{u})}{u_1^{\omega}}\Big| &= \Big|\frac{\bar{\mathbb{C}}_n(\boldsymbol{u}) - \bar{\mathbb{C}}_n(0,u_2,\dots,u_d)}{u_1^{\omega}}\Big| \\
&\le \frac{|\alpha_n(\boldsymbol{u})-\alpha_n(0,u_2,\dots,u_d)|}{u_1^{\omega}} + \sum_{j=1}^{d}\bar{S}_{nj}(\boldsymbol{u}),
\end{aligned}
$$

where $\bar{S}_{nj}(\boldsymbol{u})$ is defined in (4.14). Negligibility of $\bar{S}_{nj}$ in the latter decomposition has been shown subsequent to (4.14). From Condition 4.1, the first term on the right-hand side of the last display can be bounded by $(\{u_1^{\mu-\omega} \vee n^{-\mu}\}u_1^{-\omega})o_P(1) = o_P(\delta_n^{\mu-\omega} \vee n^{-\mu+\gamma\omega})$, which vanishes as $n \to \infty$ if we choose $\mu \in (\omega, \theta_1)$.

Now, suppose $g_\omega(\boldsymbol{u}) = (1-u_1)^\omega$, i.e., $1-u_1 = 1-u_1 \wedge \cdots \wedge \widehat{u_k} \wedge \cdots \wedge u_d$ for some $k \in \{2, \ldots, d\}$. Again, without loss of generality, we may assume that $k = 2$, which implies that $1-u_1 \le 1-u_2$ and $1-u_1 \ge 1-u_j$ for all $j \ge 3$. We decompose

$$\Big|\frac{\bar{\mathbb{C}}_n(\boldsymbol{u})}{g_\omega(\boldsymbol{u})}\Big| = \Big|\frac{\bar{\mathbb{C}}_n(\boldsymbol{u}) - \bar{\mathbb{C}}_n(\boldsymbol{u}^{(2)})}{g_\omega(\boldsymbol{u})}\Big| \le \frac{|\alpha_n(\boldsymbol{u}) - \alpha_n(\boldsymbol{u}^{(2)})|}{(1-u_1)^\omega} + |\bar{T}_{n3}(\boldsymbol{u})| + |\bar{T}_{n4}(\boldsymbol{u})|,$$

where $\bar{T}_{n3}(\boldsymbol{u})$ and $\bar{T}_{n4}(\boldsymbol{u})$ are defined in (4.15). By the same arguments as for their treatment on $N(0, n^{-\gamma})$, we have $|\bar{T}_{n3}(\boldsymbol{u})| = o_P(1)$ and $|\bar{T}_{n4}(\boldsymbol{u})| = o_P(1)$, uniformly in $\boldsymbol{u} \in N(0, \delta_n)$ with $g_\omega(\boldsymbol{u}) = (1-u_1)^\omega$. The remaining term on the right-hand side of the last display can be bounded by an application of Condition 4.1. Choosing $\mu \in (\omega, \theta_1)$, we obtain

$$\frac{|\alpha_n(\boldsymbol{u}) - \alpha_n(\boldsymbol{u}^{(2)})|}{(1-u_1)^\omega} \le \sum_{j \neq 2} \frac{(1-u_j)^\mu \vee n^{-\mu}}{(1-u_1)^\omega} o_P(1) = o_P(\delta_n^{\mu-\omega} \vee n^{-\mu+\omega\gamma}),$$

which is $o_P(1)$. This completes the proof. □

Proof of Lemma 4.11. Let us first show (4.5). As in the proof of Lemma 4.9, by a monotonicity argument, it suffices to treat sequences δ_n such that $\delta_n \ge n^{-1/2}$. We split the proof into two cases and begin by considering $\boldsymbol{u} \in N(cn^{-1}, 2\delta_n^{1/2})$. Obviously, $|\boldsymbol{u} - \boldsymbol{u}'| \le \delta_n$ implies $\boldsymbol{u}' \in N(cn^{-1}, 2\delta_n^{1/2} + \delta_n) \subset N(cn^{-1}, 3\delta_n^{1/2})$. Thus, by Lemma 4.9, we obtain

$$\sup_{\boldsymbol{u},\boldsymbol{u}' \in [\frac{c}{n}, 1-\frac{c}{n}]^d, |\boldsymbol{u}-\boldsymbol{u}'| \le \delta_n, \boldsymbol{u} \in N(cn^{-1}, 2\delta_n^{1/2})} \Big|\frac{\mathbb{C}_n(\boldsymbol{u})}{g_\omega(\boldsymbol{u})} - \frac{\mathbb{C}_n(\boldsymbol{u}')}{g_\omega(\boldsymbol{u}')}\Big| = o_P(1).$$

Now, consider the case $\boldsymbol{u} \in N(2\delta_n^{1/2}, 1/2)$. Then, $|\boldsymbol{u} - \boldsymbol{u}'| \le \delta_n$ implies that $\boldsymbol{u}' \in N(2\delta_n^{1/2} - \delta_n, 1/2) \subset N(\delta_n^{1/2}, 1/2)$. Hence, Lemma 4.8 implies that

$$\begin{aligned}
&\sup_{\boldsymbol{u},\boldsymbol{u}' \in [\frac{c}{n}, 1-\frac{c}{n}]^d, |\boldsymbol{u}-\boldsymbol{u}'| \le \delta_n, \boldsymbol{u} \in N(2\delta_n^{1/2}, 1/2)} \Big|\frac{\mathbb{C}_n(\boldsymbol{u})}{g_\omega(\boldsymbol{u})} - \frac{\mathbb{C}_n(\boldsymbol{u}')}{g_\omega(\boldsymbol{u}')}\Big| \\
&\quad \le \sup_{\boldsymbol{u},\boldsymbol{u}' \in [\frac{c}{n}, 1-\frac{c}{n}]^d \cap N(\delta_n^{1/2}, 1/2), |\boldsymbol{u}-\boldsymbol{u}'| \le \delta_n} \Big|\frac{\bar{\mathbb{C}}_n(\boldsymbol{u})}{g_\omega(\boldsymbol{u})} - \frac{\bar{\mathbb{C}}_n(\boldsymbol{u}')}{g_\omega(\boldsymbol{u}')}\Big| + o_P(1).
\end{aligned}$$

Therefore, in order to prove (4.5), it suffices to show that

$$\sup_{\boldsymbol{u},\boldsymbol{u}' \in N(\delta_n^{1/2}, 1/2), |\boldsymbol{u}-\boldsymbol{u}'| \le \delta_n} \Big|\frac{\bar{\mathbb{C}}_n(\boldsymbol{u}) - \bar{\mathbb{C}}_n(\boldsymbol{u}')}{g_\omega(\boldsymbol{u})}\Big| = o_P(1) \tag{4.16}$$

$$\sup_{\boldsymbol{u},\boldsymbol{u}' \in N(\delta_n^{1/2}, 1/2), |\boldsymbol{u}-\boldsymbol{u}'| \le \delta_n} \Big|\bar{\mathbb{C}}_n(\boldsymbol{u}')\Big(\frac{1}{g_\omega(\boldsymbol{u})} - \frac{1}{g_\omega(\boldsymbol{u}')}\Big)\Big| = o_P(1). \tag{4.17}$$

The respective proofs will be given below at the end of this proof.

For the proof of (4.6), note that $\bar{\mathbb{C}}_n(\boldsymbol{u})/\tilde{g}_\omega(\boldsymbol{u}) = 0$ for $g_\omega(\boldsymbol{u}) = 0$. Therefore, we can bound

$$\sup_{|\boldsymbol{u}-\boldsymbol{u}'|\le\delta_n, g_1(\boldsymbol{u})=0, g_1(\boldsymbol{u}')>0} \left| \frac{\bar{\mathbb{C}}_n(\boldsymbol{u})}{\tilde{g}_\omega(\boldsymbol{u})} - \frac{\bar{\mathbb{C}}_n(\boldsymbol{u}')}{\tilde{g}_\omega(\boldsymbol{u}')} \right| \le \sup_{\boldsymbol{u}'\in N(0,\delta_n)} \left| \frac{\bar{\mathbb{C}}_n(\boldsymbol{u}')}{g_\omega(\boldsymbol{u}')} \right| = o_P(1)$$

by Lemma 4.10. The suprema over $\{\boldsymbol{u} : g_1(\boldsymbol{u}) > 0, g_1(\boldsymbol{u}') = 0\}$ or $\{\boldsymbol{u} : g_1(\boldsymbol{u}) = g_1(\boldsymbol{u}') = 0\}$ can be treated analogously, whereas the suprema over $\{\boldsymbol{u} : g_1(\boldsymbol{u}) > 0, g_1(\boldsymbol{u}') > 0\}$ can be handled by (4.16), (4.17) and Lemma 4.10. This proves (4.6).

It remains to be shown that (4.16) and (4.17) are valid.

Proof of (4.16). By Condition 2.1 and 4.1 and the fact that $\dot{C}_j \in [0,1]$ we have, for $\boldsymbol{u}, \boldsymbol{u}' \in N(\delta_n^{1/2}, 1/2)$, $|\boldsymbol{u} - \boldsymbol{u}'| \le \delta_n$ and any $\mu \in (0, \theta_1)$,

$$\left| \frac{\bar{\mathbb{C}}_n(\boldsymbol{u}) - \bar{\mathbb{C}}_n(\boldsymbol{u}')}{g_\omega(\boldsymbol{u})} \right| \le \left| \frac{\alpha_n(\boldsymbol{u}) - \alpha_n(\boldsymbol{u}')}{g_\omega(\boldsymbol{u})} \right| + \sum_{j=1}^d \left| \frac{\dot{C}_j(\boldsymbol{u})\{\alpha_n(\boldsymbol{u}^{(j)}) - \alpha_n(\boldsymbol{u}'^{(j)})\}}{g_\omega(\boldsymbol{u})} \right| + \sum_{j=1}^d \left| \frac{\{\dot{C}_j(\boldsymbol{u}) - \dot{C}_j(\boldsymbol{u}')\}\alpha_n(\boldsymbol{u}'^{(j)})}{g_\omega(\boldsymbol{u})} \right|.$$

The right-hand side can be further bounded by

$$(d+1)\frac{|\boldsymbol{u} - \boldsymbol{u}'|^\mu \vee n^{-\mu}}{g_\omega(\boldsymbol{u})} M_n(\delta_n, \mu) + \sum_{j=1}^d \left| \frac{\{\dot{C}_j(\boldsymbol{u}) - \dot{C}_j(\boldsymbol{u}')\}\alpha_n(\boldsymbol{u}'^{(j)})}{g_\omega(\boldsymbol{u})} \right|.$$

Since $g_\omega(\boldsymbol{u}) \ge \delta_n^{\omega/2}$ for $\boldsymbol{u} \in N(\delta_n^{1/2}, 1/2)$, the first summand on the right of the last display is of order $O_P(\delta_n^{\mu-\omega/2})$, which is $o_P(1)$ if we choose $\mu > \omega/2$. For the second term, we fix j and will consider two cases for each summand separately. First, suppose $1 - u_j' < \delta_n^{1/2}$. In this case, Condition 4.3 yields, for arbitrary $\omega' \in (0, \theta_2)$,

$$\left| \frac{\{\dot{C}_j(\boldsymbol{u}) - \dot{C}_j(\boldsymbol{u}')\}\alpha_n(\boldsymbol{u}'^{(j)})}{g_\omega(\boldsymbol{u})} \right| \le 2\delta_n^{-\omega/2}\{u_j'(1-u_j')\}^{\omega'} \frac{|\alpha_n(\boldsymbol{u}'^{(j)})|}{\{u_j'(1-u_j')\}^{\omega'}} = O_P(\delta_n^{-\omega/2+\omega'/2}).$$

Since we can choose $\omega' \in (\omega, \theta_2)$ the latter is $o_P(1)$.

Now, suppose $1 - u_j' \ge \delta_n^{1/2}$. Then, the mean value theorem allows to write

$$\left| \frac{\{\dot{C}_j(\boldsymbol{u}) - \dot{C}_j(\boldsymbol{u}')\}\alpha_n(\boldsymbol{u}'^{(j)})}{g_\omega(\boldsymbol{u})} \right| \le \sum_{\ell=1}^d \left| \frac{\ddot{C}_{j\ell}(\boldsymbol{\xi}_j)\alpha_n(\boldsymbol{u}'^{(j)})(u_\ell - u_\ell')}{g_\omega(\boldsymbol{u})} \right|,$$

where $\boldsymbol{\xi}_j$ denotes an intermediate point between $\boldsymbol{u}$ and $\boldsymbol{u}'$. In particular, the components of $\boldsymbol{\xi}_j = (\xi_{j1}, \ldots, \xi_{jd})$ satisfy $\xi_{j\ell} \geq \sqrt{\delta_n}$ and $1 - \xi_{jj} \geq \sqrt{\delta_n} - \delta_n \geq \sqrt{\delta_n}/2$, for sufficiently large n. Then, by Condition 2.1, the sum on the right-hand side of the last display can be bounded by

$$d \frac{K}{\xi_{jj}(1-\xi_{jj})} |\alpha_n(\boldsymbol{u}'^{(j)})| \delta_n^{1-\omega/2} = O_P(\delta_n^{1/2-\omega/2}) = o_P(1).$$

Proof of (4.17). Note that it is sufficient to bound $|g_\omega(\boldsymbol{u})^{-1} - g_\omega(\boldsymbol{u}')^{-1}|$, because $\sup_{\boldsymbol{u}\in[0,1]^d} |\bar{\mathbb{C}}_n(\boldsymbol{u})| = O_P(1)$. To this end, we first observe that, for $\boldsymbol{u}, \boldsymbol{u}' \in N(\delta_n^{1/2}, 1/2)$ and $|\boldsymbol{u} - \boldsymbol{u}'| \leq \delta_n$, we have

$$|g_\omega(\boldsymbol{u}) - g_\omega(\boldsymbol{u}')| \leq \omega \delta_n^{(\omega-1)/2} |g_1(\boldsymbol{u}) - g_1(\boldsymbol{u}')| = O(\delta_n^{(\omega+1)/2})$$

where we used the mean value theorem and the fact that g_1 is Lipschitz-continuous on $N(\delta_n^{1/2}, 1/2)$. Therefore,

$$\left| \frac{1}{g_\omega(\boldsymbol{u})} - \frac{1}{g_\omega(\boldsymbol{u}')} \right| = \left| \frac{g_\omega(\boldsymbol{u}') - g_\omega(\boldsymbol{u})}{g_\omega(\boldsymbol{u}) g_\omega(\boldsymbol{u}')} \right| = O(\delta_n^{(\omega+1)/2-\omega}) = o(1),$$

which implies (4.17). □

4.4 Proof of Theorem 3.3.

Let $n \geq 2$. Decompose $\sqrt{n}\{R_n - \mathbb{E}[J(\boldsymbol{U})]\} = A_n - r_{n1}$, where

$$A_n = \sqrt{n} \int_{(\frac{1}{2n}, 1-\frac{1}{2n}]^2} J(\boldsymbol{u}) \mathrm{d}(\hat{C}_n - C)(\boldsymbol{u})$$
$$r_{n1} = \sqrt{n} \int_{\{(\frac{1}{2n}, 1-\frac{1}{2n}]^2\}^c} J(\boldsymbol{u}) \mathrm{d}C(\boldsymbol{u}),$$

where A^c denotes the complement of a set A in $(0,1)^2$. From integration by parts for Lebesgue-Stieltjes integrals (see Theorem A.6 in the supplementary material) we have that $A_n = B_n + r_{n2} + r_{n3}$, where

$$B_n = \int_{(\frac{1}{2n}, 1-\frac{1}{2n}]^2} \hat{\mathbb{C}}_n(\boldsymbol{u}) \mathrm{d}J(\boldsymbol{u})$$

where

$$\begin{aligned} r_{n2} &= \Delta(\hat{\mathbb{C}}_n J, \tfrac{1}{2n}, \tfrac{1}{2n}, 1 - \tfrac{1}{2n}, 1 - \tfrac{1}{2n}) \\ &- \int_{(\frac{1}{2n}, 1-\frac{1}{2n}]} \hat{\mathbb{C}}_n(u, 1 - \tfrac{1}{2n}) J(\mathrm{d}u, 1 - \tfrac{1}{2n}) + \int_{(\frac{1}{2n}, 1-\frac{1}{2n}]} \hat{\mathbb{C}}_n(u, \tfrac{1}{2n}) J(\mathrm{d}u, \tfrac{1}{2n}) \\ &- \int_{(\frac{1}{2n}, 1-\frac{1}{2n}]} \hat{\mathbb{C}}_n(1 - \tfrac{1}{2n}, v) J(1 - \tfrac{1}{2n}, \mathrm{d}v) + \int_{(\frac{1}{2n}, 1-\frac{1}{2n}]} \hat{\mathbb{C}}_n(\tfrac{1}{2n}, v) J(\tfrac{1}{2n}, \mathrm{d}v), \end{aligned}$$

with $\Delta(f, a_1, a_2, b_1, b_2) = f(b_1, b_2) - f(a_1, b_2) - f(b_1, a_2) + f(a_1, a_2)$ for $f : (0,1)^2 \to \mathbb{R}$ and $\boldsymbol{a}, \boldsymbol{b} \in (0,1)^2$ and where

$$
\begin{aligned}
r_{n3} = \int_{(\frac{1}{2n}, 1-\frac{1}{2n}]^2} & \nu_n(\{u\} \times (v, 1 - \tfrac{1}{2n}]) + \nu_n((u, 1 - \tfrac{1}{2n}] \times \{v\}) \\
& \qquad\qquad\qquad + \nu_n(\{(u,v)\})\mathrm{d}J(u,v) \\
& + \int_{(\frac{1}{2n}, 1-\frac{1}{2n}]} \nu_n(\{u\} \times (\tfrac{1}{2n}, 1 - \tfrac{1}{2n}])J(\mathrm{d}u, \tfrac{1}{2n}) \\
& \qquad + \int_{(\frac{1}{2n}, 1-\frac{1}{2n}]} \nu_n((\tfrac{1}{2n}, 1 - \tfrac{1}{2n}] \times \{v\})J(\tfrac{1}{2n}, \mathrm{d}v),
\end{aligned}
$$

with ν_n denoting the unique signed measure on $[\frac{1}{2n}, 1-\frac{1}{2n}]$ associated with $\hat{\mathbb{C}}_n$ (see Theorem A.4 in the supplementary material).

For the arguments that follow, we remark that by Proposition 4.4 the conditions of Theorem 2.2 imply those of Theorem 4.5. Thus, all results from the proof of Theorem 4.5 are applicable here.

Regarding weak convergence of B_n, observe that by Theorem 2.2, Lemma 4.10 and the integrability condition in (3.2)

$$
\begin{aligned}
B_n &= \int_{(0,1)^2} \mathbb{1}\{\boldsymbol{u} \in (\tfrac{1}{2n}, 1 - \tfrac{1}{2n}]^2\} \frac{\bar{\mathbb{C}}_n(\boldsymbol{u})}{g_\omega(\boldsymbol{u})} g_\omega(\boldsymbol{u})\,\mathrm{d}J(\boldsymbol{u}) + o_P(1) \\
&= \int_{(0,1)^2} \frac{\bar{\mathbb{C}}_n(\boldsymbol{u})}{g_\omega(\boldsymbol{u})} g_\omega(\boldsymbol{u})\,\mathrm{d}J(\boldsymbol{u}) + o_P(1).
\end{aligned}
$$

Now, the integrability condition in (3.2) implies that the functional $f \mapsto \int_{(0,1)^2} f \tilde{g}_\omega \,\mathrm{d}J$ is continuous when viewed as a map from $(\ell^\infty((0,1)^2), \|\cdot\|_\infty)$ to $\mathbb{R}$, and thus B_n converges weakly to $\int_{(0,1)^2} \mathbb{C}_C(\boldsymbol{u})\mathrm{d}J(\boldsymbol{u})$ by Theorem 2.2 and the continuous mapping theorem. Hence, it remains to be shown that r_{n1}, r_{n2} and r_{n3} are $o_P(1)$.

Regarding r_{n1}, since $|J(u,v)| \le \text{const} \times g_\omega(u,v)^{-1}$, we can bound

$$
|r_{n1}| \le \sqrt{n} \int_{([\frac{1}{2n}, 1-\frac{1}{2n}]^2)^c} g_\omega(u,v)^{-1} \mathrm{d}C(u,v).
$$

The set $\{(\frac{1}{2n}, 1-\frac{1}{2n}]^2\}^c$ consists of vectors where either both components or only one component is close to the boundary of $[0,1]^2$. In order to bound the integral on the right-hand side of the last display, we distinguish these cases and exemplarily consider the integral over $(0, \frac{1}{2n}]^2$ and the one over $(0, \frac{1}{2n}] \times (\frac{1}{2n}, 1 - \frac{1}{2n}]$. Integrals over the remaining subsets can be treated in the same way. First, since $g_\omega(u,v)^{-1} \le u^{-\omega} + v^{-\omega}$ for $u, v \in (0, \frac{1}{2n}]$, we have

$$
\sqrt{n} \int_{(0,\frac{1}{2n}]^2} g_\omega(u,v)^{-1}\,\mathrm{d}C(u,v) \le \sqrt{n} \int_{(0,\frac{1}{2n}]^2} u^{-\omega} + v^{-\omega} \mathrm{d}C(u,v).
$$

Let us only consider the integral over $u^{-\omega}$ on the right-hand side, the one over $v^{-\omega}$ can be treated analogously. We have

$$\begin{aligned}\sqrt{n}\int_{(0,\frac{1}{2n}]^2} u^{-\omega}\mathrm{d}C(u,v) &\le \sqrt{n}\int_{(0,\frac{1}{2n}]\times[0,1]} u^{-\omega}\,\mathrm{d}C(u,v)\\ &= \sqrt{n}\int_{(0,\frac{1}{2n}]} u^{-\omega}\,\mathrm{d}u = O(n^{-1/2+\omega}) = o(1).\end{aligned}$$

Second, on $(0,\frac{1}{2n}]\times(\frac{1}{2n},1-\frac{1}{2n}]$, we have $g_\omega(u,v)^{-1} = u^{-\omega}$, whence, by a similar reasoning,

$$\sqrt{n}\int_{(0,\frac{1}{2n}]\times(\frac{1}{2n},1-\frac{1}{2n}]} g_\omega(u,v)^{-1}\mathrm{d}C(u,v) \le \sqrt{n}\int_{(0,\frac{1}{2n}]} u^{-\omega}\,\mathrm{d}u = O(n^{-1/2+\omega}).$$

Regarding r_{n2}, use Theorem 2.2 and (3.3) and (3.4) to replace $\hat{\mathbb{C}}_n/g_\omega$ by $\bar{\mathbb{C}}_n/g_\omega$ at the cost of a negligible remainder (note that $g_\omega(u,\delta)=\delta^\omega$ for $u\in(\delta,1-\delta]$) . Then, the four integrals in the definition of r_{n2} are $o_P(1)$ by (3.3), (3.4), Lemma 4.10 and Proposition 4.4, while $\Delta(\bar{\mathbb{C}}_n J,\frac{1}{2n},\frac{1}{2n},1-\frac{1}{2n},1-\frac{1}{2n})$ converges to 0 by Lemma 4.10, Proposition 4.4 and the fact that $|J(\boldsymbol{u})|\le \text{const}\times g_\omega(\boldsymbol{u})^{-1}$ for $\boldsymbol{u}\in(0,1)^2$.

Regarding r_{n3}, since $\hat{C}_n$ and C are completely monotone, the (unique) measures in the Jordan decomposition of ν_n are given by $\nu_n^+ = \sqrt{n}\nu_{\hat{C}_n}$ and $\nu_n^- = \sqrt{n}\nu_C$, where $\nu_{\hat{C}_n}$ and ν_C denote the measures corresponding to $\hat{C}_n$ and C, respectively. Thus, continuity of the copula C yields

$$\nu_n(\{u\}\times(v,1-\tfrac{1}{2n}]) = \sqrt{n}\nu_{\hat{C}_n}(\{u\}\times(v,1-\tfrac{1}{2n}]) \le \sqrt{n}\{\hat{C}_n(u,1)-\hat{C}_n(u-,1)\}.$$

Since the last display is bounded by $n^{-1/2}$ times the maximum number of $\hat{U}_{i1}$ that are equal, a reasoning which is similar to the one used to obtain (4.7) yields that, for any $\mu\in(\omega,1/2)$,

$$\nu_n(\{u\}\times(v,1-\tfrac{1}{2n}]) = O_P(n^{-\mu})$$

uniformly in $u,v\in(0,1)^2$. Similar estimations for the remaining terms in r_{n3} imply that $|r_{n3}|$ is of the order

$$O_P(n^{-\mu})\Big\{\int_{(\frac{1}{2n},1-\frac{1}{2n}]^2}|\mathrm{d}J| + \int_{(\frac{1}{2n},1-\frac{1}{2n}]}|J(\mathrm{d}u,\tfrac{1}{2n})| + \int_{(\frac{1}{2n},1-\frac{1}{2n}]}|J(\tfrac{1}{2n},\mathrm{d}v)|\Big\}.$$

By Conditions (3.2)–(3.4), these integrals are of order $O(n^\omega)$ which leads to $|r_{n3}| = O_P(n^{\omega-\mu}) = o_P(1)$. □

5 Auxiliary results

Lemma 5.1. *Suppose Condition 4.1 is met. Then, for $j=1,\ldots,d$ and any $\mu\in[0,\theta_1)$, we have*

$$\sup_{u\in[0,1]}|G_{nj}\{G_{nj}^-(u)\}-u| = o_P(n^{-1/2-\mu}).$$

Proof. From the definition of the (left-continuous) generalized inverse, we have that $\sup_{u\in[0,1]}|H\{H^-(u)\}-u|$ is bounded by the maximum jump heigth of the function H, i.e.,

$$\sup_{u\in[0,1]}|G_{nj}\{G_{nj}^-(u)\}-u|\le\sup_{u\in[0,1]}|G_{nj}(u)-G_{nj}(u-)|$$

Therefore, the assertion follows from (4.7) and Condition 4.1. ☐

Lemma 5.2. *Suppose Condition 4.3 is met. Then, for $j=1,\dots,d$ and any $\gamma\in(0,\{1/[2(1-\theta_2)]\}\wedge\theta_3)$, we have*

$$K_{nj}(\gamma)=\sup\left|\frac{u_j(1-u_j)}{\xi_j(1-\xi_j)}\right|=O_P(1),$$

where the supremum is taken over all $u_j\in[n^{-\gamma},1-n^{-\gamma}]$ and all ξ_j between $G_{nj}^-(u_j)$ and u_j.

Proof of Lemma 5.2. Since

$$K_{nj}(\gamma)\le K_{nj}^{(1)}(\gamma)\times K_{nj}^{(2)}(\gamma):=\sup\left|\frac{u_j}{\xi_j}\right|\times\sup\left|\frac{1-u_j}{1-\xi_j}\right|,$$

it suffices to treat both suprema on the right-hand side separately. In the following, we only consider the first one; the second one can be treated along similar lines. Obviously,

$$K_{nj}^{(1)}(\gamma)\le 1\vee\sup_{u_j\in[n^{-\gamma},1-n^{-\gamma}]}\frac{u_j}{G_{nj}^-(u_j)}$$

Let Ω_n denote the event that $\sup_{u_j\in[n^{-\gamma},1-n^{-\gamma}]}|\{G_{nj}^-(u_j)-u_j\}/u_j|\le 1/2$. Choose $\omega'\in(0\vee(1-\frac{1}{2\gamma}),\theta_2)$ and use Condition 4.3 to conclude that

$$\begin{aligned}\sup_{u_j\in[n^{-\gamma},1-n^{-\gamma}]}\left|\frac{G_{nj}^-(u_j)-u_j}{u_j}\right|&\le\sup_{u_j\in[n^{-\gamma},1-n^{-\gamma}]}\left\{\sqrt{n}\left|\frac{G_{nj}^-(u_j)-u_j}{u_j^{\omega'}}\right|\times\frac{u_j^{\omega'-1}}{\sqrt{n}}\right\}\\&=O_P(n^{-1/2-\gamma(\omega'-1)})=o_P(1).\end{aligned}$$

Thus, $\mathbb{P}(\Omega_n^c)=o(1)$, which implies

$$\begin{aligned}\sup_{u_j\in[n^{-\gamma},1-n^{-\gamma}]}\frac{u_j}{G_{nj}^-(u_j)}&=\sup_{u_j\in[n^{-\gamma},1-n^{-\gamma}]}\left(1+\frac{G_{nj}^-(u_j)-u_j}{u_j}\right)^{-1}\mathbb{1}_{\Omega_n}+o_P(1)\\&\le 2+o_P(1)=O_P(1),\end{aligned}$$

where we used that $1/(1+x)\le 1/(1-|x|)$ for $x\in[-1/2,1/2]$. This yields the assertion. ☐

Lemma 5.3. *Under Condition 4.1, Conidtion 4.2 and Condition 4.3 we have for any* $\omega \in (0, \theta_1 \wedge \theta_2)$ *and any* $\gamma > 1/2$

$$\sup_{u_j \in [1-n^{-\gamma}, 1]} |\beta_{nj}(u_j)| = o_P(n^{-\omega/2}).$$

Proof. Since the result is one-dimensional, we drop the index j in the following. Note that all the arguments that follow lead to bounds which are valid uniformly in $u \in [1-n^{-\gamma}, 1]$. Now, fix $u \in [1-n^{-\gamma}, 1]$ and choose $i \in \{0, \ldots, n-1\}$ such that $u \in (\frac{i}{n}, \frac{i+1}{n}]$. Then, $G_n^-(u) = U_{i+1:n}$, where $U_{1:n} \le \cdots \le U_{n:n}$ denote the order statistics of $U_1, \ldots, U_n$. Hence,

$$\begin{aligned} n^{\omega/2}|\beta_n(u)| &\le n^{\omega/2+1/2}\{|U_{i+1:n} - i/n| \vee |U_{i+1:n} - (i+1)/n|\} \\ &\le n^{\omega/2+1/2}|U_{i+1:n} - i/n| + n^{-1/2+\omega/2} \end{aligned}$$

Now, as a consequence of Lemma 5.1, we have $G_n(U_{i+1:n}) = G_n\{G_n^-(u)\} = i/n + \kappa_{i,n}$, where $\max_{i=0}^{n-1} \kappa_{i,n} = o_P(n^{-\mu-1/2})$ with $\mu \in (\omega/2, \theta_1)$. Therefore,

$$n^{\omega/2+1/2}|U_{i+1:n} - i/n| \le n^{\omega/2+1/2}|G_n(U_{i+1:n}) - U_{i+1:n}| + n^{\omega/2+1/2}\kappa_{i,n}$$

The second term on the right-hand side is $o_P(n^{-\mu+\omega/2}) = o_P(1)$. For the first term, we have

$$\begin{aligned} n^{\omega/2+1/2}|G_n(U_{i+1:n}) - U_{i+1:n}| &= \frac{\alpha_n(U_{i+1:n})}{(1-U_{i+1:n})^\omega} n^{\omega/2}(1-U_{i+1:n})^\omega \\ \le \sup_{u \in (0,1)} \frac{|\alpha_n(u)|}{(1-u)^\omega} \times n^{\omega/2}(1-U_{i+1:n})^\omega &= O_P(1) \times \{\sqrt{n}(1-U_{i+1:n})\}^\omega \end{aligned}$$

For the factor on the right, since $u \ge 1 - n^\gamma$, we have, for any $w \in (\frac{i}{n}, \frac{i+1}{n}]$,

$$\begin{aligned} \sqrt{n}(1-U_{i+1:n}) &= \sqrt{n}\{w - G_n^-(w) + 1 - w\} \\ &\le \sup_{v \in [1-n^{-\gamma}, 1]} |\beta_n(v)| + n^{1/2-\gamma} \\ &\le \sup_{v \in [1-n^{-\gamma}, 1]} |\beta_n(v) - \beta_n(1-n^{-1/2})| + |\beta_n(1-n^{-1/2})| + n^{1/2-\gamma}. \end{aligned}$$

The first term in the expression above is $o_P(1)$ by asymptotic equicontinuity of β_n (which follows from weak convergence of β_n to a Gaussian process, this is a consequence of Condition 4.2 and the functional delta method), the second term is $o_P(1)$ by Conidtion 4.3, and the third term vanishes since $\gamma > 1/2$. □

References

Aistleitner, C. and J. Dick (2014). Functions of bounded variation, signed measures, and a general Koksma-Hlawka inequality. *arXiv:1406.0230v2*.

Berghaus, B., A. Bücher, and H. Dette (2013). Minimum distance estimators of the pickands dependence function and related tests of multivariate extreme-value dependence. *Journal de la Société Française de Statistique 154*(1), 116–137.

Beutner, E. and H. Zähle (2010). A modified functional delta method and its application to the estimation of risk functionals. *J. Multivariate Anal. 101*(10), 2452–2463.

Brillinger, D. R. (1975). *Time series.* Holt, Rinehart and Winston, Inc., New York-Montreal, Que.-London. Data analysis and theory, International Series in Decision Processes.

Bücher, A., H. Dette, and S. Volgushev (2011). New estimators of the Pickands dependence function and a test for extreme-value dependence. *Ann. Statist. 39*(4), 1963–2006.

Bücher, A., J. Segers, and S. Volgushev (2014, 08). When uniform weak convergence fails: Empirical processes for dependence functions and residuals via epi- and hypographs. *The Annals of Statistics 42*(4), 1598–1634.

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

Capéraà, P., A.-L. Fougères, and C. Genest (1997). A nonparametric estimation procedure for bivariate extreme value copulas. *Biometrika 84*(3), 567–577.

Chen, X. and Y. Fan (2006). Estimation of copula-based semiparametric time series models. *J. Econometrics 130*(2), 307–335.

Csörgő, M., S. Csörgő, L. Horváth, and D. M. Mason (1986). Weighted empirical and quantile processes. *Ann. Probab. 14*(1), 31–85.

Csörgő, M. and H. Yu (1996). Weak approximations for quantile processes of stationary sequences. *Canad. J. Statist. 24*(4), 403–430.

Deheuvels, P. (1991). On the limiting behavior of the pickands estimator for bivariate extreme-value distributions. *Statistics and Probability Letters 12*, 429–439.

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

Gaenssler, P. and W. Stute (1987). *Seminar on empirical processes*, Volume 9 of *DMV Seminar.* Basel: Birkhäuser Verlag.

Genest, C., K. Ghoudi, and L.-P. Rivest (1995). A semiparametric estimation procedure of dependence parameters in multivariate families of distributions. *Biometrika 82*(3), 543–552.

Genest, C., B. Rémillard, and D. Beaudoin (2009). Goodness-of-fit tests for copulas: a review and a power study. *Insurance Math. Econom. 44*(2), 199–213.

Genest, C. and J. Segers (2009). Rank-based inference for bivariate extreme-value copulas. *Ann. Statist. 37*(5B), 2990–3022.

Genest, C. and J. Segers (2010). On the covariance of the asymptotic empirical copula process. *J. Multivariate Anal. 101*(8), 1837–1845.

Gudendorf, G. and J. Segers (2012). Nonparametric estimation of multivariate extreme-value copulas. *J. Statist. Plann. Inference 142*(12), 3073–3085.

Hallin, M., J.-F. Ingenbleek, and M. L. Puri (1985). Linear serial rank tests for randomness against ARMA alternatives. *Ann. Statist. 13*(3), 1156–1181.

Hallin, M. and M. L. Puri (1988). Optimal rank-based procedures for time series analysis: testing an ARMA model against other ARMA models. *Ann. Statist. 16*(1), 402–432.

Hardy, G. H. (1905). On double *Fourier* series, and especially those which represent the double zeta-function with real and incommensurable parameters. *Quart. J. 37*, 53–79.

Jiménez, J. R., E. Villa-Diharce, and M. Flores (2001). Nonparametric estimation of the dependence function in bivariate extreme value distributions. *J. Multivariate Anal. 76*, 159–191.

Kley, T., S. Volgushev, H. Dette, and M. Hallin (2014). Quantile spectral processes: Asymptotic analysis and inference. *arXiv:1401.8104v1*.

Owen, A. B. (2005). Multidimensional variation for quasi-Monte Carlo. In *Contemporary multivariate analysis and design of experiments*, Volume 2 of *Ser. Biostat.*, pp. 49–74. World Sci. Publ., Hackensack, NJ.

Pickands, III, J. (1981). Multivariate extreme value distributions. In *Proceedings of the 43rd session of the International Statistical Institute, Vol. 2 (Buenos Aires, 1981)*, Volume 49, pp. 859–878, 894–902. With a discussion.

Rio, E. (2000). *Théorie asymptotique des processus aléatoires faiblement dépendants*, Volume 31 of *Mathématiques & Applications (Berlin) [Mathematics & Applications]*. Berlin: Springer-Verlag.

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

Segers, J. (2012). Asymptotics of empirical copula processes under non-restrictive smoothness assumptions. *Bernoulli 18*(3), 764–782.

Shao, Q.-M. and H. Yu (1996). Weak convergence for weighted empirical processes of dependent sequences. *Ann. Probab. 24*(4), 2098–2127.

Shorack, G. R. and J. A. Wellner (1986). *Empirical processes with applications to statistics*. Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics. John Wiley & Sons, Inc., New York.

Stute, W. (1984). The oscillation behavior of empirical processes: the multivariate case. *Ann. Probab. 12*(2), 361–379.

Tsukahara, H. (2005). Semiparametric estimation in copula models. *Canad. J. Statist. 33*(3), 357–375.

Van der Vaart, A. W. and J. A. Wellner (1996). *Weak Convergence and Empirical Processes - Springer Series in Statistics*. New York: Springer.

van der Vaart, A. W. and J. A. Wellner (2007). Empirical processes indexed by estimated functions. In *Asymptotics: particles, processes and inverse problems*, Volume 55 of *IMS Lecture Notes Monogr. Ser.*, pp. 234–252. Beachwood, OH: Inst. Math. Statist.

Supplement to "Weak convergence of the empirical copula process with respect to weighted metrics"

Betina Berghaus, Axel Bücher and Stanislav Volgushev

November 21, 2014

A Bounded variation and Lebesgue-Stieltjes integration for two-variate functions

In this supplement, we briefly recapitulate some results on bounded variation and integration for two-variate functions. We begin by treating the case of functions defined on a compact rectangle in $\mathbb{R}^2$. Of particular interest is the integration by parts formula in Theorem A.6. At the end of this appendix, we consider the case of potentially unbounded functions on open rectangles.

Let A denote some rectangle in $\mathbb{R}^2$ and let f be a real-valued function on A. For $\boldsymbol{x}, \boldsymbol{y} \in A$ such that $\boldsymbol{x} \le \boldsymbol{y}$ we set

$$\Delta(f, x_1, x_2, y_1, y_2) := f(y_1, y_2) - f(x_1, y_2) - f(y_1, x_2) + f(x_1, x_2).$$

For $\boldsymbol{x}, \boldsymbol{y} \in A$ such that $x_1 < y_1$, we set

$$\Delta_1(f, x_1, y_1; x_2) := f(y_1, x_2) - f(x_1, x_2)$$

and finally, for $\boldsymbol{x}, \boldsymbol{y} \in A$ such that $x_2 < y_2$, we set

$$\Delta_2(f, x_2, y_2; x_1) := f(x_1, y_2) - f(x_1, x_2).$$

A function $f : A \to \mathbb{R}$ is called *completely monotone* if $\Delta(f, x_1, x_2, y_1, y_2) \ge 0$ for any $\boldsymbol{x}, \boldsymbol{y} \in A$ such that $\boldsymbol{x} < \boldsymbol{y}$, $\Delta_1(f, x_1, y_1; x_2) \ge 0$ for any $\boldsymbol{x}, \boldsymbol{y} \in A$ such that $x_1 < y_1$ and $\Delta_2(f, x_2, y_2; x_1) \ge 0$ for any $\boldsymbol{x}, \boldsymbol{y} \in A$ such that $x_2 < y_2$.

Definition A.1. (Hardy-Krause variation) Let $\emptyset \ne [\boldsymbol{a}, \boldsymbol{b}] \subset \mathbb{R}^2$ and $f : [\boldsymbol{a}, \boldsymbol{b}] \to \mathbb{R}$. For $\boldsymbol{x} \in [\boldsymbol{a}, \boldsymbol{b}]$ and $\boldsymbol{y} \in [\boldsymbol{a}, \boldsymbol{b}]$, we define

$$\begin{aligned} VHK(f, [\boldsymbol{a}, \boldsymbol{x}], \boldsymbol{y}) = \sup \sum_i \sum_j |\Delta(f, s_i, t_j, s_{i+1}, t_{j+1})| \\ + \sup \sum_i |\Delta_1(f, s_i, s_{i+1}; y_2)| + \sup \sum_j |\Delta_2(f, t_j, t_{j+1}; y_1)|, \end{aligned}$$

as the Hardy-Krause variation of f on $[\boldsymbol{a}, \boldsymbol{x}]$ with anchor-point $\boldsymbol{y}$. Here, the supremum is taken over all decompositions $s_1, \dots, s_n$ and $t_1, \dots, t_m$ with $a_1 = s_1 < \cdots < s_n = x_1$ and $a_2 = t_1 < \cdots < t_m = x_2$, respectively. Let $BVHK([\boldsymbol{a}, \boldsymbol{b}])$ denote the space of all functions such that $VHK(f, [\boldsymbol{a}, \boldsymbol{b}], \boldsymbol{b}) < \infty$.

Example A.2. Consider $f : [0, 1]^2 \to \mathbb{R}$, defined as $f(x, y) = \mathbb{1}(x \geq 1/2, y \geq 1/2)$. Then $VHK(f, [\mathbf{0}, \mathbf{1}], \boldsymbol{y}) = 1$ for any $\boldsymbol{y} \in [0, 1/2)^2$, whereas $VHK(f, [\mathbf{0}, \mathbf{1}], \boldsymbol{y}) = 3$ for $\boldsymbol{y} \in [1/2, 1]^2$.

The following simple properties are collected from Owen (2005) and Aistleitner and Dick (2014): $BVHK([\boldsymbol{a}, \boldsymbol{b}])$ is closed under sums, differences and products. Moreover, for any $\boldsymbol{x}, \boldsymbol{y} \in [\boldsymbol{a}, \boldsymbol{b}]$, we have $VHK(f, [\boldsymbol{a}, \boldsymbol{b}], \boldsymbol{x}) < \infty$ if and only if $VHK(f, [\boldsymbol{a}, \boldsymbol{b}], \boldsymbol{y}) < \infty$, which can be derived directly from the definition. The following theorem is a refinement of Proposition 12 in Owen, 2005.

Theorem A.3. **(Theorem 2 in Aistleitner and Dick, 2014)** *For any function $f \in BVHK([\boldsymbol{a}, \boldsymbol{b}])$ there exist unique functions $f^+, f^- : [\boldsymbol{a}, \boldsymbol{b}] \to \mathbb{R}$ such that (i), (ii) and (iii) hold:*

(i) f^+ *and* f^- *are completely monotone*

(ii) $f^{\pm}(\boldsymbol{a}) = 0$ *and* $f(\boldsymbol{x}) = f(\boldsymbol{a}) + f^+(\boldsymbol{x}) - f^-(\boldsymbol{x})$

(iii) $VHK(f, [\boldsymbol{a}, \boldsymbol{b}], \boldsymbol{a}) = VHK(f^+, [\boldsymbol{a}, \boldsymbol{b}], \boldsymbol{a}) + VHK(f^-, [\boldsymbol{a}, \boldsymbol{b}], \boldsymbol{a})$.

The decomposition in *(ii)* is called the *Jordan decomposition* of f. The explicit form of the functions $f^{\pm}$ is given in the proof of Theorem 2 of Aistleitner and Dick (2014) (see also Hardy, 1905):

$$f^+(\boldsymbol{x}) = \{VHK(f, [\boldsymbol{a}, \boldsymbol{x}], \boldsymbol{a}) + f(\boldsymbol{x}) - f(\boldsymbol{a})\}/2,$$
$$f^-(\boldsymbol{x}) = \{VHK(f, [\boldsymbol{a}, \boldsymbol{x}], \boldsymbol{a}) - f(\boldsymbol{x}) + f(\boldsymbol{a})\}/2.$$

The next theorem shows that, if f is additionally right-continuous, then it defines a unique signed measure on $[\boldsymbol{a}, \boldsymbol{b}]$. Also note that any signed measure ν on $\mathcal{B}([\boldsymbol{a}, \boldsymbol{b}])$ has a unique Jordan decomposition $\nu = \nu^+ - \nu^-$ with two measures ν^+ and ν^-, given by

$$\nu^+(A) = \sup\{\nu(B) : B \subset A, B \in \mathcal{B}([\boldsymbol{a}, \boldsymbol{b}])\},$$
$$\nu^-(A) = -\inf\{\nu(B) : B \subset A, B \in \mathcal{B}([\boldsymbol{a}, \boldsymbol{b}])\}.$$

Theorem A.4. **(Theorem 3 in Aistleitner and Dick, 2014)** *Let $f \in BVHK([\boldsymbol{a}, \boldsymbol{b}])$ be right-continuous. Then there exists a unique signed Borel-measure ν on $\mathcal{B}([\boldsymbol{a}, \boldsymbol{b}])$ such that*

$$f(\boldsymbol{x}) = \nu([\boldsymbol{a}, \boldsymbol{x}]), \quad \boldsymbol{x} \in [\boldsymbol{a}, \boldsymbol{b}]. \tag{A.1}$$

Moreover, if $f(\boldsymbol{x}) = f(\boldsymbol{a}) + f^+(\boldsymbol{x}) - f^-(\boldsymbol{x})$ denotes the Jordan decomposition of f, and if $\nu = \nu^+ - \nu^-$ denotes the Jordan decomposition of ν, then

$$f^+(\boldsymbol{x}) = \nu^+([\boldsymbol{a}, \boldsymbol{x}] \setminus \{\boldsymbol{a}\}), \qquad f^-(\boldsymbol{x}) = \nu^-([\boldsymbol{a}, \boldsymbol{x}] \setminus \{\boldsymbol{a}\}) \tag{A.2}$$

for any $\boldsymbol{x} \in [\boldsymbol{a}, \boldsymbol{b}]$.

Note that, by (A.1) and (A.2), for any $\boldsymbol{a} \leq \boldsymbol{x} < \boldsymbol{y} \leq \boldsymbol{b}$,

$$\nu((\boldsymbol{x}, \boldsymbol{y}]) = \Delta(f, x_1, x_2, y_1, y_2), \quad \nu^{\pm}((\boldsymbol{x}, \boldsymbol{y}]) = \Delta(f^{\pm}, x_1, x_2, y_1, y_2).$$

Definition A.5. Let $f \in BVHK([\boldsymbol{a}, \boldsymbol{b}])$ be right-continuous. Let $g : [\boldsymbol{a}, \boldsymbol{b}] \to \mathbb{R}$ denote a measurable function such that either $\int_{[\boldsymbol{a},\boldsymbol{b}]} |g| \mathrm{d}\nu^+ < \infty$ or $\int_{[\boldsymbol{a},\boldsymbol{b}]} |g| \mathrm{d}\nu^- < \infty$. Then

$$\int_{[\boldsymbol{a},\boldsymbol{b}]} g \mathrm{d}f := \int_{[\boldsymbol{a},\boldsymbol{b}]} g \mathrm{d}\nu = \int_{[\boldsymbol{a},\boldsymbol{b}]} g \mathrm{d}\nu^+ - \int_{[\boldsymbol{a},\boldsymbol{b}]} g \mathrm{d}\nu^-$$

denotes the Lebesgue-Stieltjes integral of g with respect to f.

Given $f \in BVHK([\boldsymbol{a}, \boldsymbol{b}])$ and a fixed point $y \in [a_2, b_2]$, we can define a collection of functions $f_{1,y} : [a_1, b_1] \to \mathbb{R}$ through $f_{1,y}(x) = f(x, y)$. We have $f_{1,y} \in BV([a_1, b_1])$, hence, by a one-dimensional analog of the preceding developments, we obtain a unique Jordan decomposition

$$f_{1,y}(x) = f_{1,y}(a_1) + f^+_{1,y}(x) - f^-_{1,y}(x), \quad x \in [a_1, b_1]$$

such that $f^{\pm}_{1,y}$ is non-decreasing with $f^{\pm}_{1,y}(a_1) = 0$ and $V(f_{1,y}, [a_1, b_1]) = V(f^+_{1,y}, [a_1, b_1]) + V(f^-_{1,y}, [a_1, b_1])$, where $V(f, [a_2, b_2])$ denotes the usual total variation of a real-valued function on a compact interval $[a_2, b_2]$. Attached is a unique signed measure $\nu_{1,y}$ such that $f_{1,y}(x) = \nu_{1,y}([a_1, x])$. Moreover, if $\nu_{1,y} = \nu^+_{1,y} - \nu^-_{1,y}$ denotes the Jordan decomposition of $\nu_{1,y}$, then

$$f^+_{1,y}(x) = \nu^+_{1,y}((a_1, x]), \qquad f^-_{1,y}(x) = \nu^-_{1,y}((a_1, x]).$$

Note that the measure $\nu_{1,y}$ is related with ν by

$$\nu_{1,y}([a_1, x]) = \nu([a_1, x] \times [a_2, y]).$$

The same arguments apply for the function $f_{2,x} : [a_2, b_2] \to \mathbb{R}$, defined through $f_{2,x}(y) = f(x, y)$, with $x \in [a_1, b_1]$ fixed. We will write $f(\mathrm{d}x, y)$ and $f(x, \mathrm{d}y)$ for $\nu_{1,y}(\mathrm{d}x)$ and $\nu_{2,x}(\mathrm{d}y)$, respectively.

Theorem A.6. (Integration by parts) *Let μ, ν be finite signed measures on $[\boldsymbol{a}, \boldsymbol{b}]$ and, for $\boldsymbol{x} \in [\boldsymbol{a}, \boldsymbol{b}]$, write $f(\boldsymbol{x}) := \mu([\boldsymbol{a}, \boldsymbol{x}])$, $g(\boldsymbol{x}) := \nu([\boldsymbol{a}, \boldsymbol{x}])$.*

Then, for any $(\boldsymbol{c},\boldsymbol{d}] \subset [\boldsymbol{a},\boldsymbol{b}]$ *with* $\boldsymbol{c} < \boldsymbol{d}$*,*

$$\begin{aligned}\int_{(\boldsymbol{c},\boldsymbol{d}]} f\mathrm{d}g = &\int_{(\boldsymbol{c},\boldsymbol{d}]} g\mathrm{d}f + \Delta(fg, c_1, c_2, d_1, d_2)\\ &- \int_{(c_1,d_1]} g(u,d_2)f(\mathrm{d}u,d_2) + \int_{(c_1,d_1]} g(u,c_2)f(\mathrm{d}u,c_2)\\ &- \int_{(c_2,d_2]} g(d_1,v)f(d_1,\mathrm{d}v) + \int_{(c_2,d_2]} g(c_1,v)f(c_1,\mathrm{d}v)\\ &+ \int_{(\boldsymbol{c},\boldsymbol{d}]} \nu(\{u\}\times(v,d_2]) + \nu((u,d_1]\times\{v\}) + \nu(\{(u,v)\})\mathrm{d}f(u,v)\\ &+ \int_{(c_1,d_1]} \nu(\{u\}\times(c_2,d_2])f(\mathrm{d}u,c_2)\\ &+ \int_{(c_2,d_2]} \nu((c_1,d_1]\times\{v\})f(c_1,\mathrm{d}v)\end{aligned}$$

Before we give a proof of this general result, we recall that to any right-continuous $f,g \in BVHK([\boldsymbol{a},\boldsymbol{b}])$ there correspond unique signed measures μ,ν, respectively. If either f or g is additionally continuous, the last three terms in the representation above vanish and we obtain the following result

Corollary A.7. (Integration by parts) *Let* $f,g \in BVHK([\boldsymbol{a},\boldsymbol{b}])$ *be right-continuous functions with either* f *or* g *continuous. Then, for any* $(\boldsymbol{c},\boldsymbol{d}] \subset [\boldsymbol{a},\boldsymbol{b}]$*,*

$$\begin{aligned}\int_{(\boldsymbol{c},\boldsymbol{d}]} f\mathrm{d}g = &\int_{(\boldsymbol{c},\boldsymbol{d}]} g\mathrm{d}f + \Delta(fg, c_1, c_2, d_1, d_2)\\ &- \int_{(c_1,d_1]} g(u,d_2)f(\mathrm{d}u,d_2) + \int_{(c_1,d_1]} g(u,c_2)f(\mathrm{d}u,c_2)\\ &\qquad - \int_{(c_2,d_2]} g(d_1,v)f(d_1,\mathrm{d}v) + \int_{(c_2,d_2]} g(c_1,v)f(c_1,\mathrm{d}v).\end{aligned}$$

Proof of Theorem A.6. First of all, we use the definition of f,g and obtain

$$\int_{(\boldsymbol{c},\boldsymbol{d}]} f(u,v)\mathrm{d}g(u,v) = \int_{(\boldsymbol{c},\boldsymbol{d}]}\int_{(c_1,u]\times(c_2,v]} \mathrm{d}f(x,y)\mathrm{d}g(u,v) + R_1,$$

where $R_1 = \int_{(\boldsymbol{c},\boldsymbol{d}]} f(u,c_2) + f(c_1,v) - f(c_1,c_2)\mathrm{d}g(u,v)$. Now, Fubini's Theorem yields

$$\begin{aligned}&\int_{(\boldsymbol{c},\boldsymbol{d}]}\int_{(c_1,u]\times(c_2,v]} \mathrm{d}f(x,y)\mathrm{d}g(u,v)\\ =&\int_{(\boldsymbol{c},\boldsymbol{d}]}\int_{[x,d_1]\times[y,d_2]} \mathrm{d}g(u,v)\mathrm{d}f(x,y)\end{aligned}$$

$$= \int_{(\boldsymbol{c},\boldsymbol{d}]} \int_{(x,d_1]\times(y,d_2]} \mathrm{d}g(u,v)\mathrm{d}f(x,y)$$
$$+ \int_{(\boldsymbol{c},\boldsymbol{d}]} \nu(\{x\}\times(y,d_2]) + \nu((x,d_1]\times\{y\}) + \nu(\{(x,y)\})\mathrm{d}f(x,y)$$

and

$$\int_{(\boldsymbol{c},\boldsymbol{d}]} \int_{(x,d_1]\times(y,d_2]} \mathrm{d}g(u,v)\mathrm{d}f(x,y) = \int_{(\boldsymbol{c},\boldsymbol{d}]} g\mathrm{d}f + R_2$$

with $R_2 = \int_{(\boldsymbol{c},\boldsymbol{d}]} g(d_1,d_2) - g(u,d_2) - g(d_1,v)\mathrm{d}f(u,v)$. Summarizing,

$$\int_{(\boldsymbol{c},\boldsymbol{d}]} f\mathrm{d}g = \int_{(\boldsymbol{c},\boldsymbol{d}]} g\mathrm{d}f + R_1 + R_2$$
$$+ \int_{(\boldsymbol{c},\boldsymbol{d}]} \nu(\{u\}\times(v,d_2]) + \nu((u,d_1]\times\{v\}) + \nu(\{(u,v)\})\mathrm{d}f(u,v),$$

whence it remains to consider R_1 and R_2. Observe that

$$\int_{(\boldsymbol{c},\boldsymbol{d}]} f(u,c_2)\mathrm{d}g(u,v) = \int_{(c_1,d_1]} f(u,c_2)g(\mathrm{d}u,d_2) - \int_{(c_1,d_1]} f(u,c_2)g(\mathrm{d}u,c_2)$$

and a similar identity holds for $\int_{(\boldsymbol{c},\boldsymbol{d}]} g(u,d_2)\mathrm{d}f(u,v)$. To see this, note that the identity holds for $f(u,c_2) = I_{(\alpha,\beta]}(u)$ with $c_1 \le \alpha < \beta \le d_1$ arbitrary; the general claim then follows by algebraic induction. Next, observe the following one-dimensional formula for integration by parts

$$\int_{(c_1,d_1]} f(u,c_2)g(\mathrm{d}u,d_2) = -\int_{(c_1,d_1]} g(u,d_2)f(\mathrm{d}u,c_2)$$
$$+ \int_{(c_1,d_1]} \nu(\{u\}\times[a_2,d_2])f(\mathrm{d}u,c_2)$$
$$+ g(d_1,d_2)f(d_1,c_2) - g(c_1,d_2)f(c_1,c_2).$$

This formula can be proved by an application of the Fubini Theorem after writing $f(u,c_2) = \int_{(c_1,u]} f(\mathrm{d}x,c_2) + f(c_1,c_2)$ and by observing that $\nu_{1,d_2}(\{u\}) = \nu(\{u\}\times[a_2,d_2])$. Thus we have (apply a similar formula for integration by parts to all four integrals below)

$$R_1 = \int_{(c_1,d_1]} f(u,c_2)g(\mathrm{d}u,d_2) - \int_{(c_1,d_1]} f(u,c_2)g(\mathrm{d}u,c_2)$$
$$+ \int_{(c_2,d_2]} f(c_1,v)g(d_1,\mathrm{d}v) - \int_{(c_2,d_2]} f(c_1,v)g(c_1,\mathrm{d}v)$$
$$- f(\boldsymbol{c})\Delta(g,c_1,c_2,d_1,d_2)$$
$$= -\int_{(c_1,d_1]} g(u,d_2)f(\mathrm{d}u,c_2) + g(d_1,d_2)f(d_1,c_2) - g(c_1,d_2)f(c_1,c_2)$$

$$
\begin{aligned}
&+\int_{(c_1,d_1]} \nu(\{u\}\times[a_2,d_2])f(\mathrm{d}u,c_2)\\
&+\int_{(c_1,d_1]} g(u,c_2)f(\mathrm{d}u,c_2)-g(d_1,c_2)f(d_1,c_2)+g(c_1,c_2)f(c_1,c_2)\\
&-\int_{(c_1,d_1]} \nu(\{u\}\times[a_2,c_2])f(\mathrm{d}u,c_2)\\
&-\int_{(c_2,d_2]} g(d_1,v)f(c_1,\mathrm{d}v)+g(d_1,d_2)f(c_1,d_2)-g(d_1,c_2)f(c_1,c_2)\\
&+\int_{(c_2,d_2]} \nu([a_1,d_1]\times\{v\})f(c_1,\mathrm{d}v)\\
&+\int_{(c_2,d_2]} g(c_1,v)f(c_1,\mathrm{d}v)-g(c_1,d_2)f(c_1,d_2)+g(c_1,c_2)f(c_1,c_2)\\
&-\int_{(c_2,d_2]} \nu([a_1,c_1]\times\{v\})f(c_1,\mathrm{d}v)\\
&-f(\boldsymbol{c})\Delta(g,c_1,c_2,d_1,d_2).
\end{aligned}
$$

For R_2 we obtain

$$
\begin{aligned}
R_2 = &-\int_{(c_1,d_1]} g(u,d_2)f(\mathrm{d}u,d_2)+\int_{(c_1,d_1]} g(u,d_2)f(\mathrm{d}u,c_2)\\
&-\int_{(c_2,d_2]} g(d_1,v)f(d_1,\mathrm{d}v)+\int_{(c_2,d_2]} g(d_1,v)f(c_1,\mathrm{d}v)\\
&+g(d_1,d_2)\Delta(f,c_1,c_2,d_1,d_2).
\end{aligned}
$$

The result follows after collecting terms. □

Definition A.8. (Locally bounded variation and Lebesgue-Stieltjes integration) Consider $f:(\boldsymbol{a},\boldsymbol{b})\to\mathbb{R}$ which is potentially unbounded. We say that f is of *locally bounded Hardy-Krause variation*, notationally $f\in BVHK_{loc}((\boldsymbol{a},\boldsymbol{b}))$ if and only if $f|_{[\boldsymbol{c},\boldsymbol{d}]}\in BVHK([\boldsymbol{c},\boldsymbol{d}])$ for any $\boldsymbol{a}<\boldsymbol{c}<\boldsymbol{d}<\boldsymbol{b}$. In the following, f is assumed to be right-continuous. Let $\boldsymbol{a}_n,\boldsymbol{b}_n$ be two sequences converging to $\boldsymbol{a}$ and $\boldsymbol{b}$, respectively, and such that $\boldsymbol{a}<\boldsymbol{a}_{n+1}<\boldsymbol{a}_n<\boldsymbol{b}_n<\boldsymbol{b}_{n+1}<\boldsymbol{b}$. Since $f|_{[\boldsymbol{a}_n,\boldsymbol{b}_n]}\in BVHK([\boldsymbol{a}_n,\boldsymbol{b}_n])$, we can define unique measures ν_n^+ and ν_n^- on $\mathcal{B}([\boldsymbol{a}_n,\boldsymbol{b}_n])$ as in Theorem A.4. Now, for $A\in\mathcal{B}((\boldsymbol{a},\boldsymbol{b}))$ set

$$
\nu^{\pm}(A):=\lim_{n\to\infty}\nu_n^{\pm}(A\cap(\boldsymbol{a}_n,\boldsymbol{b}_n]).
$$

By monotone convergence, ν^+ and ν^- are $[0,\infty]$-valued measures on $\mathcal{B}((\boldsymbol{a},\boldsymbol{b}))$. Moreover, by Proposition A.9 below, the definition of $\nu^{\pm}$ is independent of the choice of the sequences $\boldsymbol{a}_n$ and $\boldsymbol{b}_n$. Finally, for $\boldsymbol{a}<\boldsymbol{c}<\boldsymbol{d}<\boldsymbol{b}$, the proposition implies that

$$
\nu((c,d]):=\nu^+((\boldsymbol{c},\boldsymbol{d}])-\nu^-((\boldsymbol{c},\boldsymbol{d}])=\Delta(f,c_1,c_2,d_1,d_2).
$$

Note that ν is not necessarily a signed measure on $\mathcal{B}((\boldsymbol{a},\boldsymbol{b}))$, since expressions of the form "$\infty - \infty$" are possible in principle. Still, for a measurable function $g : (\boldsymbol{a},\boldsymbol{b}) \to \mathbb{R}$ such that $\int |g| \mathrm{d}\nu^+ < \infty$ or $\int |g| \mathrm{d}\nu^- < \infty$, we may define the Lebesgue-Stieltjes integral

$$\int_{(\boldsymbol{a},\boldsymbol{b})} g \mathrm{d}f := \int_{(\boldsymbol{a},\boldsymbol{b})} g \mathrm{d}\nu := \int_{(\boldsymbol{a},\boldsymbol{b})} g \mathrm{d}\nu^+ - \int_{(\boldsymbol{a},\boldsymbol{b})} g \mathrm{d}\nu^-$$

Proposition A.9. *Let $f \in BVHK([\boldsymbol{a},\boldsymbol{b}])$ be right-continuous and let $\boldsymbol{a} < \boldsymbol{c} < \boldsymbol{d} < \boldsymbol{b}$. Set $g := f|_{[\boldsymbol{c},\boldsymbol{d}]}$. Then, for any $A \in \mathcal{B}((\boldsymbol{c},\boldsymbol{d}])$,*

$$\nu_f^{\pm}(A) = \nu_g^{\pm}(A),$$

where $\nu_f^{\pm}$ and $\nu_g^{\pm}$ denote the unique measures associated to the unique signed measure ν_f of f and ν_g of g, respectively.

Proof. It suffices to show the identity on sets of the form $(\boldsymbol{x},\boldsymbol{y}] \subset (\boldsymbol{c},\boldsymbol{d}]$. By (A.1), we have

$$\nu_f((\boldsymbol{x},\boldsymbol{y}]) = \Delta(f, x_1, x_2, y_1, y_2) = \Delta(g, x_1, x_2, y_1, y_2) = \nu_g((\boldsymbol{x},\boldsymbol{y}]).$$

Uniqueness of the Jordan decomposition implies the assertion. □