SFB 823

Extreme value copula estimation based on block maxima of a multivariate stationary time series

Axel Bücher, Johan Segers

Discussion Paper

Nr. 45/2013

Extreme value copula estimation based on block maxima of a multivariate stationary time series

Axel Bücher
Ruhr-Universität Bochum
Fakultät für Mathematik
44780 Bochum, Germany
axel.buecher@ruhr-uni-bochum.de

Johan Segers
Université catholique de Louvain
Institut de Statistique, Biostatistique et Sciences Actuarielles
Voie du Roman Pays 20, B-1348 Louvain-la-Neuve, Belgium
johan.segers@uclouvain.be

November 13, 2013

Abstract

The core of the classical block maxima method consists of fitting an extreme value distribution to a sample of maxima over blocks extracted from an underlying series. In asymptotic theory, it is usually postulated that the block maxima are an independent random sample of an extreme value distribution. In practice however, block sizes are finite, so that the extreme value postulate will only hold approximately. A more accurate asymptotic framework is that of a triangular array of block maxima, the block size depending on the size of the underlying sample in such a way that both the block size and the number of blocks within that sample tend to infinity. The copula of the vector of componentwise maxima in a block is assumed to converge to a limit, which, under mild conditions, is then necessarily an extreme value copula. Under this setting and for absolutely regular stationary sequences, the empirical copula of the sample of vectors of block maxima is shown to be a consistent and asymptotically normal estimator for the limiting extreme value copula. Moreover, the empirical copula serves as a basis for rank-based, nonparametric estimation of the Pickands dependence function of the extreme value copula. The results are illustrated by theoretical examples and a Monte Carlo simulation study.

Keywords: extreme value copula, block maxima method, weak convergence, empirical copula process, stationary time series, Pickands dependence function, absolutely regular process

1 Introduction

The block maximum method for extreme value analysis essentially consists of the following procedure: partition a long series of data into blocks; for each block, compute the maximum; fit an extreme value distribution to the sample of block maxima. Often, the blocks correspond to months or years of data, whence the name 'annual maxima series'. The fitted distribution can then be used to compute tail quantiles or 'T-year return levels'. The approach was developed and popularized in the classic monograph of Gumbel (1958). The method is applicable even when the individual 'daily' observations are unavailable or when the time series exhibits seasonality, as long as the block size is a multiple of the period length. The procedure can be extended to multivariate series too: compute or just observe block maxima for each of variables separately, and fit a multivariate extreme value distribution to the sample of vectors of componentwise block maxima.

The method is justified by the extremal types theorem: under broad conditions, the only possible limits of affinely normalized block maxima, as the block length tends to infinity, are the extreme value distributions. The conditions allow for temporal dependence, provided certain mixing conditions hold; see Leadbetter et al. (1983) for the univariate case and Hsing (1989) and Hüsler (1990) for the multivariate case.

Unlike their univariate counterparts, multivariate extreme value distributions do not constitute a parametric family. In statistical applications, a parametric form is often assumed, an early example being Gumbel and Mustafi (1967). In general, the dependence structure or copula should be max-stable. Several representations of max-stable or extreme value copulas exist; see Beirlant et al. (2004, Chapter 8) for an overview. The representation proposed in Pickands (1981) is a popular one and has led to the concept of a Pickands dependence function.

In the large-sample theory for the block maximum method, the data generating process is nearly always specified as independent random sampling from the limiting extreme value distribution. Seminal papers to this view are Prescott and Walden (1980) for the univariate case and Tawn (1988), Tawn (1990) and Deheuvels (1991) for the multivariate case. However, in the light of the above description, this set-up does not correspond to reality for at least two reasons: first, the block maxima are only approximately extreme value distributed, and second, they are only approximately independent.

A first contribution to the mathematical validation of the block maximum method in a more realistic setting is Dombry (2013). The starting point is a single series of independent and identically distributed univariate random variables whose distribution is in the domain of attraction of an extreme value distribution. Consistency is shown for the maximum likelihood estimator for the extreme value index applied to the sample of block maxima extracted from the full sample. The block size tends to infinity so that the extremal types theorem can come into force; at the same time, the block size is of smaller order than the sample size, so that the number of blocks, which determines the size of the sample of block maxima, tends to infinity. In the same set-up, the asymptotic distribution of the probability-weighted moment estimator was addressed by Laurens de Haan at the *8th Conference on Extreme Value Analysis* (Fudan University, Shanghai, July 8–12, 2013).

For multivariate time series, nothing has been done in this direction yet, up to the best of our knowledge. The present paper tries to fill this gap. We focus on the estimation of the limit copula of the vector of componentwise block maxima when the block size tends to infinity. The data generating process is a stationary, multivariate time series. Under weak dependence conditions, the limit copula must then be an extreme value copula (Hsing, 1989). No parametric assumptions are made regarding this extreme value copula. It can be estimated by the empirical copula of the vectors of block maxima. Moreover, the empirical copula can be used as a basis for the nonparametric estimation of the Pickands dependence function of the extreme value copula. For simplicity, we focus on the minimum distance estimator of Bücher et al. (2011) and Berghaus et al. (2013), although alternative procedures could have been considered as well (Gudendorf and Segers, 2011; Peng et al., 2013).

We study the sequence of empirical copula processes constructed from the triangular array of vectors of block maxima as the block size and the number of blocks tend to infinity. We find that if the underlying series is absolutely regular, the limit process is the same Gaussian process as if the block maxima were sampled independently from a distribution whose copula is already equal to the limiting extreme value copula. This result carries over to the estimation of the Pickands dependence function, where we find the same limit process as in Berghaus et al. (2013). This does not mean that the temporal dependence can be neglected, however: because of serial dependence, the limiting extreme value copula is in general different from the extreme value attractor of the copula of the stationary distribution of the series. The results are illustrated by means of Monte Carlo simulations.

The structure of the paper is as follows. The objects of interest are described mathematically in Section 2. The main results on the convergence of the block maxima empirical copula process and the minimum distance estimator for the Pickands dependence function form the subject of Section 3. Section 4 then contains a number of theoretical examples, whereas Section 5 reports on the result of a simulation study. Section 6 concludes. All proofs are collected in the Appendices A and B.

2 Preliminaries, notations, and assumptions

Consider a d-variate stationary time series $\boldsymbol{X}_t = (X_{t,1}, \ldots, X_{t,d})$, $t \in \mathbb{Z}$. For simplicity, assume that the univariate stationary margins are continuous. A sample of size n is divided into k blocks of length m, so that $k = \lfloor n/m \rfloor$, the integer part of n/m, and possibly a remainder block of length $n - km$ at the end. The maximum of the ith block in the jth component is denoted by

$$M_{m,i,j} = \max\{X_{t,j} : t \in (im - m, im] \cap \mathbb{Z}\}.$$

Let $\boldsymbol{M}_{m,i} = (M_{m,i,1}, \ldots, M_{m,i,d})$ be the vector of maxima over the d variables in the ith block. For fixed block length m, the sequence of block maxima $(\boldsymbol{M}_{m,i})_i$ is a stationary process too.

The distributions functions of the block maxima are denoted by

$$F_m(\boldsymbol{x}) = \mathrm{P}[\boldsymbol{M}_{m,1} \le \boldsymbol{x}], \qquad F_{m,j}(x_j) = \mathrm{P}[M_{m,1,j} \le x_j],$$

for $\boldsymbol{x} \in \mathbb{R}^d$ and $j \in \{1, \ldots, d\}$. Observe that F_1 is the distribution function of $\boldsymbol{X}_1$. If the random vectors $\boldsymbol{X}_t$ are serially independent, we have $F_m = F_1^m$. In the general, stationary case, the relation between F_m and F_1 is more complex.

The margins of $\boldsymbol{X}_1$ being continuous, the margins of $\boldsymbol{M}_{m,1}$ are continuous as well. Let C_m be the (unique) copula of F_m, which, in the serially independent case, can be written as $C_m(\boldsymbol{u}) = \{C_1(u_1^{1/m}, \ldots, u_m^{1/m})\}^m$, $\boldsymbol{u} = (u_1, \ldots, u_d) \in [0,1]^d$. In the present context, the domain-of-attraction condition reads as follows.

Condition 2.1. *There exists a copula C_∞ such that*

$$\lim_{m\to\infty} C_m(\boldsymbol{u}) = C_\infty(\boldsymbol{u}) \qquad (\boldsymbol{u} \in [0,1]^d).$$

Typically, the limit C_∞ will be an extreme value copula (Hsing, 1989; Hüsler, 1990). Below we will assume that the time series $(\boldsymbol{X}_t)_t$ is absolutely regular or β-mixing, which, by Theorem 4.2 in Hsing (1989), is already sufficient for the latter statement. However, C_∞ will in general be different from the extreme value attractor of C_1; see for instance Section 4.1. If the copula C_∞ in Condition 2.1 is an extreme value copula, it admits the representation

$$C_\infty(\boldsymbol{u}) \;=\; \exp\left\{ \left(\sum_{j=1}^{d} \log u_j \right) A_\infty \left(\frac{\log u_2}{\sum_{j=1}^{d} \log u_j}, \ldots, \frac{\log u_d}{\sum_{j=1}^{d} \log u_j} \right) \right\} \tag{2.1}$$

for $\boldsymbol{u} \in [0,1]^d$. Here $A_\infty : \Delta_{d-1} \to [0,1]$ is called the Pickands dependence function of C_∞. It is a convex function defined on the unit simplex $\Delta_{d-1} = \{\boldsymbol{t} = (t_1, \ldots, t_{d-1}) \in [0,1]^{d-1} : t_1 + \cdots + t_{d-1} \le 1\}$ and satisfying the bounds $\max\{1 - t_1 - \cdots - t_{d-1}, t_1, \ldots, t_{d-1}\} \le A_\infty(\boldsymbol{t}) \le 1$; see, e.g., Gudendorf and Segers (2010).

Applying the probability integral transform to the block maxima yields

$$U_{m,i,j} = F_{m,j}(M_{m,i,j}), \qquad \boldsymbol{U}_{m,i} = (U_{m,i,1}, \ldots, U_{m,i,d}). \tag{2.2}$$

The random variables $U_{m,i,j}$ are uniformly distributed on $(0,1)$ and the distribution function of the random vector $\boldsymbol{U}_{m,i}$ is the copula C_m. The empirical distribution function of the (unobservable) sample $\boldsymbol{U}_{m,1}, \ldots, \boldsymbol{U}_{m,k}$ is

$$\hat{C}^{\circ}_{n,m}(\boldsymbol{u}) = \frac{1}{k} \sum_{i=1}^{k} I(\boldsymbol{U}_{m,i} \le \boldsymbol{u}), \tag{2.3}$$

where $I(A)$ denotes the indicator variable of the event A.

Since the marginal distributions $F_{m,j}$ are unknown, we replace them in (2.2) by their empirical versions $\hat{F}_{n,m,j}$: for $\boldsymbol{x} = (x_1, \ldots, x_d) \in \mathbb{R}^d$,

$$\hat{F}_{n,m}(\boldsymbol{x}) = \frac{1}{k} \sum_{i=1}^{k} I(\boldsymbol{M}_{m,i} \le \boldsymbol{x}), \quad \hat{F}_{n,m,j}(x_j) = \frac{1}{k} \sum_{i=1}^{k} I(M_{m,i,j} \le x_j). \tag{2.4}$$

The resulting 'pseudo-observations' are

$$\hat{U}_{n,m,i,j} = \hat{F}_{n,m,j}(M_{m,i,j}), \qquad \hat{\boldsymbol{U}}_{n,m,i} = (\hat{U}_{n,m,i,1}, \ldots, \hat{U}_{n,m,i,d}).$$

In analogy to (2.3), the empirical copula is then defined as

$$\hat{C}_{n,m}(\boldsymbol{u}) = \frac{1}{k}\sum_{i=1}^{k} I(\hat{\boldsymbol{U}}_{n,m,i} \le \boldsymbol{u}). \tag{2.5}$$

In practice, it is customary to divide by $k+1$ rather than by k in (2.4); asymptotically, this does not make a difference. An alternative definition of the empirical copula is via

$$\hat{C}^{\mathrm{alt}}_{n,m}(\boldsymbol{u}) = \hat{F}_{n,m}\big(\hat{F}^{\leftarrow}_{n,m,1}(u_1), \ldots, \hat{F}^{\leftarrow}_{n,m,d}(u_d)\big) \tag{2.6}$$

where $H^{\leftarrow}$ denotes the left-continuous generalized inverse function of a distribution function H, defined as

$$H^{\leftarrow}(p) = \begin{cases} \inf\{x \in \mathbb{R} : H(x) \ge p\} & \text{if } p \in (0,1], \\ \sup\{x \in \mathbb{R} : H(x) = 0\} & \text{if } p = 0. \end{cases}$$

In the independent case, it is not difficult to see that the difference between $\hat{C}_{n,m}$ and $\hat{C}^{\mathrm{alt}}_{n,m}$ is bounded in absolute value by d/k almost surely. This difference is asymptotically negligible in view of the $O_p(1/\sqrt{k})$ rate of convergence of $\hat{C}_{n,m}$ that will be established in Theorem 3.5. However, in the case of serial dependence, the situation is more complicated, because with positive probability, there may be ties among the block maxima, even if their distribution is continuous; see for instance the random-repetition process in Subsection 4.2. Nevertheless, we will show in Proposition 3.2 that the difference between $\hat{C}_{n,m}$ and $\hat{C}^{\mathrm{alt}}_{n,m}$ is still $o_p(1/\sqrt{k})$.

The serial dependence in the series $(\boldsymbol{X}_t)_t$ is controlled via mixing coefficients. For two σ-fields $\mathcal{F}_1$ and $\mathcal{F}_2$ of a probability space $(\Omega, \mathcal{F}, \mathrm{P})$, let

$$\alpha(\mathcal{F}_1, \mathcal{F}_2) = \sup_{A \in \mathcal{F}_1, B \in \mathcal{F}_2} |\,\mathrm{P}(A \cap B) - \mathrm{P}(A)\,\mathrm{P}(B)|,$$
$$\beta(\mathcal{F}_1, \mathcal{F}_2) = \sup \frac{1}{2} \sum_{i,j \in I \times J} |\,\mathrm{P}(A_i \cap B_j) - \mathrm{P}(A_i)\,\mathrm{P}(B_j)|,$$

where the latter supremum is taken over all finite partitions $(A_i)_{i\in I}$ and $(B_j)_{j \in J}$ of Ω consisting of events that are $\mathcal{F}_1$ and $\mathcal{F}_2$ measurable, respectively. The α- and β-mixing coefficients of a time series $(\boldsymbol{X}_t)_{t\in\mathbb{Z}}$, not necessarily stationary, are defined, for $n \ge 1$, as

$$\alpha(n) = \sup_{t\in\mathbb{Z}} \alpha(\mathcal{F}^{t}_{-\infty}, \mathcal{F}^{\infty}_{t+n}), \qquad \beta(n) = \sup_{t\in\mathbb{Z}} \beta(\mathcal{F}^{t}_{-\infty}, \mathcal{F}^{\infty}_{t+n}), \tag{2.7}$$

where, for $-\infty \le \ell_1 < \ell_2 \le \infty$, $\mathcal{F}^{\ell_2}_{\ell_1}$ denotes the sigma-field generated by those $\boldsymbol{X}_t$ with $t \in [\ell_1, \ell_2] \cap \mathbb{Z}$.

Recall that m is the block size and $k = \lfloor n/m \rfloor$ is the number of blocks. In an asymptotic framework, we consider a block size sequence m_n and the associated block number sequence $k_n = \lfloor n/m_n \rfloor$.

Condition 2.2. *There exists a positive integer sequence ℓ_n such that the following statements hold:*

(i) $m_n \to \infty$ and $m_n = o(n)$;

(ii) $\ell_n \to \infty$ and $\ell_n = o(m_n)$;

(iii) $\sqrt{k_n}\,\ell_n/m_n = o(1)$;
(iv) $k_n\,\alpha(\ell_n) = o(1)$ *and* $(m_n/\ell_n)\,\alpha(\ell_n) = o(1)$;
(v) $\sqrt{k_n}\,\beta(m_n) = o(1)$.

A sufficient condition for *(iv)–(v)* is that $(k_n + m_n/\ell_n)\,\beta(\ell_n) = o(1)$.

3 Main results

The central result of the paper is Theorem 3.5 in Section 3.2, claiming weak convergence of the empirical copula process

$$\mathbb{C}_{n,m} = \sqrt{k}(\hat{C}_{n,m} - C_m). \tag{3.1}$$

To arrive at this result, the case of known margins needs to treated first; this is done in Section 3.1. Weak convergence of $\mathbb{C}_{n,m}$ is applied in Section 3.3 to find a functional central limit theorem for a rank-based, nonparametric estimator of the Pickands dependence function of the limit copula C_∞.

3.1 Block maxima empirical process

Weak convergence of the empirical copula process $\mathbb{C}_{n,m}$ will follow from the functional delta method provided we have a weak convergence result for the process

$$\mathbb{C}^\circ_{n,m} = \sqrt{k}(\hat{C}^\circ_{n,m} - C_m),$$

where $\hat{C}^\circ_{n,m}$ is defined in (2.3). If the random variables $\boldsymbol{U}_{m,i}$ were serially independent, then the weak convergence of $\mathbb{C}^\circ_{n,m}$ would easily follow from Theorem 2.11.9 in van der Vaart and Wellner (1996). The case of serial dependence is reduced to the independence case by a blocking technique and a coupling argument.

Theorem 3.1 (Block maxima empirical process)**.** *Let $(\boldsymbol{X}_t)_{t\in\mathbb{Z}}$ be a stationary multivariate time series with continuous univariate margins. If Conditions 2.1 and 2.2 hold, then*

$$\mathbb{C}^\circ_{n,m} \rightsquigarrow \mathbb{C}^\circ \qquad \text{in } \ell^\infty([0,1]^d),$$

where $\mathbb{C}^\circ$ denotes a tight, centered Gaussian process on $[0,1]^d$ with continuous sample paths and covariance structure

$$\mathrm{E}[\mathbb{C}^\circ(\boldsymbol{u})\mathbb{C}^\circ(\boldsymbol{v})] = C_\infty(\boldsymbol{u}\wedge\boldsymbol{v}) - C_\infty(\boldsymbol{u})\,C_\infty(\boldsymbol{v}).$$

Interestingly, the limiting process $\mathbb{C}^\circ$ is a C_∞-Brownian bridge: the serial dependence between the block maxima has disappeared. The proof of Theorem 3.1 is given in Appendix A.1.

3.2 Block maxima empirical copula process

Recall the two versions of the empirical copula, $\hat{C}_{n,m}$ in (2.5) and $\hat{C}^{\mathrm{alt}}_{n,m}$ in (2.6). By the following proposition, the difference between the two versions is asymptotically negligible. The proofs of Proposition 3.2 and the other results in this section are given in Appendix A.2.

Proposition 3.2. *Under the conditions of Theorem 3.1, we have*

$$\sup_{\boldsymbol{u}\in[0,1]^d} \left|\hat{C}^{\mathrm{alt}}_{n,m}(\boldsymbol{u}) - \hat{C}_{n,m}(\boldsymbol{u})\right| = o_p(1/\sqrt{k}).$$

It follows that in the definition of the empirical copula process in (3.1), we can replace $\hat{C}_{n,m}$ by $\hat{C}^{\mathrm{alt}}_{n,m}$, yielding

$$\mathbb{C}^{\mathrm{alt}}_{n,m} = \sqrt{k}(\hat{C}^{\mathrm{alt}}_{n,m} - C_m)$$

at the cost of an $o_p(1)$ term:

$$\sup_{\boldsymbol{u}\in[0,1]^d} \left|\mathbb{C}^{\mathrm{alt}}_{n,m}(\boldsymbol{u}) - \mathbb{C}_{n,m}(\boldsymbol{u})\right| = o_p(1).$$

Now, let us transfer the weak convergence result on $\mathbb{C}^{\circ}_{n,m}$ to $\mathbb{C}_{n,m}$. Let $\mathcal{D}_\Phi$ denote the set of all cdfs on $[0,1]^d$ whose marginals put no mass at zero. Defining

$$\Phi : \mathcal{D}_\Phi \to \ell^\infty([0,1]^d) : H \mapsto H(H_1^{\leftarrow}, \ldots, H_d^{\leftarrow}) \tag{3.2}$$

as the copula mapping, we can write

$$\mathbb{C}^{\mathrm{alt}}_{n,m} = \sqrt{k}\{\Phi(\hat{C}^{\circ}_{n,m}) - \Phi(C_m)\}.$$

Weak convergence of $\mathbb{C}_{n,m}$ and $\mathbb{C}^{\mathrm{alt}}_{n,m}$ can be shown by the functional delta method (van der Vaart and Wellner, 1996, Section 3.9), provided certain smoothness assumptions on the copulas C_m and C_∞ are made, to be introduced next.

For the limit process to be have continuous trajectories, the following condition (Segers, 2012) is unavoidable and will be assumed throughout.

Condition 3.3. *For any $j = 1, \ldots, d$, the jth first order partial derivative $\dot{C}_{\infty,j} = \partial C_\infty / \partial u_j$ exists and is continuous on $\{\boldsymbol{u} \in [0,1]^d : u_j \in (0,1)\}$.*

In addition, some qualification of the convergence of C_m to C_∞ will be needed. We will impose either (a) or (b) of the following condition. Roughly speaking, (a) says that this convergence is sufficiently fast, every subsequence of $\sqrt{k}(C_m - C_\infty)$ containing a further subsequence that converges uniformly, whereas (b) requires locally uniform convergence of the partial derivatives. For C_m, these partial derivatives are not supposed to exist, however; instead, we will work with the functions

$$\dot{C}_{m,j}(\boldsymbol{v}) = \limsup_{h \searrow 0} h^{-1}\{C_m(\boldsymbol{v} + h\boldsymbol{e}_j) - C_m(\boldsymbol{v})\},$$

with $\boldsymbol{e}_j$ the jth canonical unit vector in $\mathbb{R}^d$, functions which are always defined and which satisfy $0 \le \dot{C}_{m,j} \le 1$.

Condition 3.4.
(a) The sequence $\sqrt{k}(C_m - C_\infty)$ is relatively compact in $\mathcal{C}([0,1]^d)$.
(b) For every $\delta \in (0, 1/2)$,

$$\max_{j=1,\ldots,d} \sup_{\substack{\boldsymbol{u}\in[0,1]^d:\\ u_j\in[\delta,1-\delta]}} \left|\dot{C}_{m,j}(\boldsymbol{u}) - \dot{C}_{\infty,j}(\boldsymbol{u})\right| \to 0 \qquad (n \to \infty).$$

The partial derivatives $\dot{C}_{\infty,j}$ are defined as 0 for $u_j \in \{0,1\}$. For $\boldsymbol{u} \in [0,1]^d$ and $j \in \{1,\ldots,d\}$, write $\boldsymbol{u}^{(j)} = (1,\ldots,1,u_j,1,\ldots,1)$, with u_j appearing at the jth coordinate.

Theorem 3.5 (Block maxima empirical copula process)**.** *Let $(\boldsymbol{X}_t)_{t\in\mathbb{Z}}$ be a stationary multivariate time series with continuous univariate margins. Assume Conditions 2.1, 2.2 and Condition 3.3. If either Condition 3.4(a) or (b) is satisfied, then*

$$\mathbb{C}_{n,m} = \mathbb{C}_{n,m}^{\text{alt}} + o_p(1) \rightsquigarrow \mathbb{C}$$

in $\ell^\infty([0,1]^d)$, where, for $\boldsymbol{u} \in [0,1]^d$,

$$\mathbb{C}(\boldsymbol{u}) = \mathbb{C}^\circ(\boldsymbol{u}) - \sum_{j=1}^{d} \dot{C}_{\infty,j}(\boldsymbol{u})\, \mathbb{C}_j^\circ(\boldsymbol{u}^{(j)}).$$

In Theorem 3.5, the empirical copula process was defined by centering around C_m. Of course, one may also want to center around the limit, C_∞.

Corollary 3.6 (Centering by the limit copula)**.** *Let $(\boldsymbol{X}_t)_{t\in\mathbb{Z}}$ be a stationary multivariate time series with continuous univariate margins. Assume Conditions 2.1, 2.2 and Condition 3.3. If also*

$$\lim_{n\to\infty} \sqrt{k}(C_m - C_\infty) = \Gamma \qquad \text{in } \ell^\infty([0,1]^d), \tag{3.3}$$

then, in $\ell^\infty([0,1]^d)$ and with $\mathbb{C}$ as in Theorem 3.5,

$$\sqrt{k}(\hat{C}_{n,m} - C_\infty) = \sqrt{k}(\hat{C}_{n,m}^{\text{alt}} - C_\infty) + o_p(1) \rightsquigarrow \mathbb{C} + \Gamma.$$

Note that the limit Γ in (3.3) is continuous, being the uniform limit of a sequence of continuous functions.

3.3 Estimating the Pickands dependence function

For strongly mixing sequences, the limit copula C_∞ is an extreme value copula (Hsing, 1989, Theorem 4.2). Inference on the Pickands dependence function A_∞ in (2.1) can then be based on the empirical copula $\hat{C}_{n,m}$ of the block maxima.

Rank-based inference for the Pickands dependence function based on i.i.d. samples whose underlying distribution has an extreme value copula has drawn some attention recently (Genest and Segers, 2009; Bücher et al., 2011; Gudendorf and Segers, 2012; Berghaus et al., 2013; Peng et al., 2013). What the estimators have in common is that they can all be written as weighted integrals with respect to the empirical copula. The asymptotic behavior of all these estimators can then be derived from the weak convergence of the usual empirical copula process. In the following we will exemplarily extend the results on the minimum-distance estimator to the present setting of estimation from block maxima.

For the definition of the estimator, note that, for any probability density p on $(0,1)$ such that the following integral exists, we have

$$A_\infty(\boldsymbol{t}) = \int_0^1 \log\{C_\infty(y^{\boldsymbol{t}})\} \frac{p(y)}{\log(y)}\, dy,$$

where we used the notation $y^{\boldsymbol{t}} = (y^{1-t_1-\dots-t_{d-1}}, y^{t_1}, \dots, y^{t_{d-1}})$. The last display suggests to estimate A_∞ by the sample analogue

$$\widehat{A}_{n,m_n}(\boldsymbol{t}) = \int_0^1 \log\{\tilde{C}_{n,m}(y^{\boldsymbol{t}})\} \frac{p(y)}{\log(y)}\, dy, \tag{3.4}$$

where $\tilde{C}_{n,m} = \max\{k_n^{-\gamma}, \hat{C}_{n,m}(\boldsymbol{u})\}$ with some $\gamma > 1/2$ to be specified later; the latter modification is needed to avoid the logarithm of zero. For the case of i.i.d. samples, the estimator in (3.4) is exactly as defined in Bücher et al. (2011); Berghaus et al. (2013), where it is motivated as a minimum distance estimator.

Theorem 3.7 (Asymptotic normality)**.** *Let $(\boldsymbol{X}_t)_{t\in\mathbb{Z}}$ be a stationary multivariate time series with continuous univariate margins. Suppose that Condition 2.1, 2.2 and 3.3 are met and that $\sqrt{k}(C_m - C_\infty) \to \Gamma$, uniformly. If the weight function $p : (0,1) \to [0,\infty)$ satisfies*

$$\int_0^1 y^{-\lambda} \frac{p(y)}{|\log(y)|}\, dy < \infty \text{ for some } \lambda > 1, \tag{3.5}$$

then, for any $\gamma \in (\frac{1}{2}, \frac{\lambda}{2})$, in the space $\ell^\infty(\Delta_{d-1})$ equipped with the supremum distance,

$$\mathbb{A}_n = \sqrt{k_n}(\widehat{A}_{n,m_n} - A_\infty) \rightsquigarrow \mathbb{A}_\infty,$$

where the limiting process $\mathbb{A}_\infty$ on Δ_{d-1} can be represented as

$$\mathbb{A}_\infty(\boldsymbol{t}) = \int_0^1 \frac{\mathbb{C}(y^{\boldsymbol{t}}) + \Gamma(y^{\boldsymbol{t}})}{C_\infty(y^{\boldsymbol{t}})} \frac{p(y)}{\log(y)}\, dy.$$

In fact, in Appendix A.3 and at no additional cost, we will show a more general result that allows for weight functions inside the integral in (3.4) that may also depend on $\boldsymbol{t}$, see Theorem A.3. In the i.i.d. case in Berghaus et al. (2013), this result proved useful for the development of a test for extreme value dependence.

A useful class of weight functions is given by $p_\kappa(y) = (\kappa+1)^2 \times y^\kappa \times |\log(y)|$ for some $\kappa > 0$, see Example 2.5 in Bücher et al. (2011). Condition (3.5) is obviously satisfied for any $\kappa > 0$.

As it is the case for most of the available estimators for Pickands dependence functions, $\widehat{A}_{n,m_n}$ is itself not a Pickands dependence function. A unifying approach to enforce the necessary and sufficient shape constraints has been proposed in Fils-Villetard et al. (2008) and Gudendorf and Segers (2012). A simple additive boundary correction will be employed in the simulation Section 5, see formula (5.2).

4 Examples

For multivariate Gaussian time series whose cross-correlation function satisfies a certain summability condition, Amram (1985) and Hsing (1989) show that the limit C_∞ is the independence copula. For most popular time series models, however, it is already hard to obtain convenient expressions for the copula C_1 of the stationary distribution, let alone for the one of the block maximum distribution, C_m, and for the limit C_∞. In this section, we work out two simple examples.

4.1 Moving maxima

Consider the discrete-time, d-variate moving maxima process $(\boldsymbol{U}_t)_{t\in\mathbb{Z}}$ of order $p \in \mathbb{N} = \{1, 2, \ldots\}$ given by

$$U_{tj} = \max_{i=0,\ldots,p} W_{t-i,j}^{1/a_{ij}} \qquad (t \in \mathbb{Z};\ j = 1, \ldots, d). \tag{4.1}$$

Here $(\boldsymbol{W}_s)_{s\in\mathbb{Z}}$ is an iid sequence in $(0,1)^d$, the d-variate distribution of $\boldsymbol{W}_s$ being the copula D. Further, the coefficients a_{ij} $(i = 0, \ldots, p;\ j = 1, \ldots, d)$ are nonnegative and satisfy the constraints

$$\sum_{i=0}^{p} a_{ij} = 1 \qquad (j = 1, \ldots, d). \tag{4.2}$$

If $a = 0$ and $w \in (0,1)$, then $w^{1/a} = 0$ by convention. As the notation suggests, the random variables U_{tj} are uniformly distributed on $(0,1)$. A model with arbitrary continuous margins can be considered by defining $X_{tj} = \eta_j(U_{tj})$, where $\eta_1, \ldots, \eta_d$ are strictly increasing functions from $(0,1)$ into $\mathbb{R}$.

Since $\sigma(\boldsymbol{U}_t : t \le 0)$ and $\sigma(\boldsymbol{U}_t : t \ge p+1)$ are independent, Condition 2.2 is satisfied as soon as $n^{1/3} \ll m_n \ll n$; any $\ell_n \to \infty$ such that $n\,\ell_n^2 = o(m_n^3)$ works. Here, for positive sequences a_n, b_n converging to infinity, the notation $a_n \ll b_n$ means that $a_n = o(b_n)$.

Let C_m be the copula of the vector of component-wise maxima $\boldsymbol{M}_m = (M_{m,1}, \ldots, M_{m,d})$ given by $M_{m,j} = \max(U_{1j}, \ldots, U_{mj})$ for $j \in \{1, \ldots, d\}$.

For $m \in \mathbb{N}$, consider the copula, D_m, of the vector of componentwise maxima of m independent random vectors with common distribution D:

$$D_m(\boldsymbol{u}) = \bigl(D(u_1^{1/m}, \ldots, u_d^{1/m})\bigr)^m.$$

We say that D is in the copula domain of attraction of the extreme value copula D_∞ if

$$\lim_{m\to\infty} D_m(\boldsymbol{u}) = D_\infty(\boldsymbol{u}) \qquad (\boldsymbol{u} \in (0,1]^d). \tag{4.3}$$

The limit, D_∞, of C_m is in general different from the copula extreme value attractor of C_1; see (B.3).

Proposition 4.1. *Consider the moving maximum process in* (4.1)–(4.2)*. If* (4.3) *holds, then*

$$\lim_{m\to\infty} C_m(\boldsymbol{u}) = D_\infty(\boldsymbol{u}). \tag{4.4}$$

The proof of Proposition 4.1 is given in Section B.1. By a refinement of the proof of Proposition 4.1, it is actually also possible to derive rates of convergence in (4.4) given a rate of convergence in (4.3). For the sake of brevity, we omit the details.

4.2 Random repetition

Consider independent and identically distributed d-dimensional random vectors $\boldsymbol{X}_0, \boldsymbol{\xi}_1, \boldsymbol{\xi}_2, \ldots$ and, independently of these, iid indicator random variables $I_1, I_2, \ldots$; write $\mathrm{P}(I_t = 1) = \theta \in (0,1]$. For $t = 1, 2, \ldots$, define

$$\boldsymbol{X}_t = \begin{cases} \boldsymbol{\xi}_t & \text{if } I_t = 1, \\ \boldsymbol{X}_{t-1} & \text{if } I_t = 0. \end{cases}$$

Then $\boldsymbol{X}_0, \boldsymbol{X}_1, \ldots$ is a stationary sequence. The process is a simplified version of the doubly stochastic model in Smith and Weissman (1994, Section 3). By stationarity, we can assume without loss of generality that the process is defined for all $t \in \mathbb{Z}$.

Because of the random repetition mechanism, the process $(\boldsymbol{X}_t)_t$ is β-mixing and the mixing coefficients $\beta(n)$ are of the order $O((1-\theta)^n)$ as $n \to \infty$; see Lemma B.1.

Let $\boldsymbol{M}_m = (M_{m,1}, \ldots, M_{m,d})$ with $M_{m,j} = \max(X_{1,j}, \ldots, X_{n,j})$. Further, put $F_m(\boldsymbol{x}) = \mathrm{P}[\boldsymbol{M}_m \le x]$ and $F_{m,j}(x_j) = \mathrm{P}[M_{m,j} \le x_j]$. Assume the margins $F_{m,j}$ are continuous and let C_m be the copula of F_m.

Proposition 4.2. *For $\boldsymbol{u} \in (0,1]^d$ and as $m \to \infty$,*

$$
\begin{aligned}
&C_m(\boldsymbol{u}) \\
&= \{1 + o(1)\} \left[1 - \theta + \theta C_1 \left(1 + \frac{\log(u_1) + o(1)}{\theta(m-1)}, \ldots, 1 + \frac{\log(u_d) + o(1)}{\theta(m-1)} \right) \right]^{m-1}.
\end{aligned}
$$

Consequently, if C_1 is in the copula domain of attraction of an extreme value copula C_∞, then also $C_m \to C_\infty$ as $m \to \infty$.

The proof of Proposition 4.2 is given in Appendix B.2.

5 Numerical results

In this section, we investigate the finite-sample performance of the minimum-distance estimator for the Pickands dependence function A_∞ by means of a small simulation study.

The setup. As a time series model, we consider the bivariate moving maximum process $(U_{t,1}, U_{t,2})_{t \in \mathbb{Z}}$ of order 1 as introduced in Section 4.1, i.e.,

$$
U_{t,1} = \max(W_{t,1}^{1/a}, W_{t-1,1}^{1/(1-a)}), \qquad U_{t,2} = \max(W_{t,2}^{1/a}, W_{t-1,2}^{1/(1-b)}), \tag{5.1}
$$

where $(a, b) \in (0,1)^2$ and $(W_{t,1}, W_{t,2})_{t \in \mathbb{Z}}$ is a bivariate iid sequence whose marginal distributions are uniform on $(0,1)$ and whose joint cdf is denoted by D. In this section, we present results for two different choices for D:

1. The outer power transform of a Clayton copula defined by

$$
D(u, v) = [1 + \{(u^{-\theta} - 1)^\beta + (v^{-\theta} - 1)^\beta\}^{1/\beta}]^{-1/\theta},
$$

   where $\theta > 0$ and $\beta \ge 1$. Independently of $\theta > 0$, the max-attractor copula D_∞ is the Gumbel–Hougaard copula (Charpentier and Segers, 2009), whose Pickands dependence function is given by

$$
A_\infty(t) = \{t^\beta + (1-t)^\beta\}^{1/\beta}, \quad \beta \ge 1.
$$

   In the simulations, we fixed $\theta = 1$.

2. The t-copula with $\nu > 0$ degrees of freedom and correlation parameter $\rho \in (-1, 1)$, given by

$$D(u,v) = \int_{-\infty}^{\mathbf{t}_\nu^{-1}(u)} \int_{-\infty}^{\mathbf{t}_\nu^{-1}(v)} \frac{1}{\pi\nu|P|^{1/2}} \frac{\Gamma(\frac{\nu}{2}+1)}{\Gamma(\frac{\nu}{2})} \left(1 + \frac{x'P^{-1}x}{\nu}\right)^{-\nu/2+1} dx_2\, dx_1,$$

where $\mathbf{t}_\nu$ denote the cdf of the univariate t-distribution with ν degrees of freedom and where P denotes the 2×2 correlation matrix with off-diagonal element ρ. The t-copula lies in the max-domain of attraction of the t-extreme value copula characterized by the Pickands dependence function

$$A_\infty(t) = t \times \mathbf{t}_{\nu+1}(z_t) + (1-t) \times \mathbf{t}_{v+1}(z_{1-t}),$$
$$\text{where} \quad z_t = (1+\nu)^{1/2} \left[\{t/(1-t)\}^{1/\nu} - \rho\right] (1-\rho^2)^{-1/2},$$

see, e.g., Demarta and McNeil (2005). Throughout the simulations we fixed $\nu = 4$.

The remaining parameter of the two models (β and ρ, respectively) are chosen in such a way that the coefficient of upper tail dependence of D varies in the set $\{0.25, 0.5, 0.75\}$. For a and b in (5.1) we consider all possible combinations such that $(a, b) \in \{0.25, 0.5, 0.75\}^2$. Regarding the choice of n, k_n and m_n, we either fix $n = 1,000$ and consider parameters $m_n \in \{1, 2, \ldots, 30\}$, or we fix $m = 30$ (a month, say) and consider block numbers $k_n \in \{12, 24, 36, \ldots, 240\}$ (corresponding to one up to 20 years).

The estimators. In addition to the estimator $\widehat{A}_{n,m_n}$ defined in Section 3.3, we will also consider a simple (additive) boundary correction defined as

$$\widehat{A}^{abc}_{n,m_n}(t) = \widehat{A}_{n,m_n}(t) - (1-t)\{\widehat{A}_{n,m_n}(0) - 1) - t\{\widehat{A}_{n,m_n}(1) - 1\}. \qquad (5.2)$$

Due to the fact that the second and the third summand on the right-hand side of this display are deterministic functions of order $o(k^{-1/2})$, the corrected estimator has the same asymptotic distribution as the uncorrected one.

The estimator $\widehat{A}_{n,m_n}$ depends on a tuning parameter γ and a weight function p. We follow the proposals in Bücher et al. (2011) and consider the choices $\gamma = 2/3$ (the estimator is quite robust with respect to this or larger choices) and $p = p_\kappa(y) = (\kappa+1)^2 y^\kappa |\log(y)|$ with $\kappa = 0.5$, see Example 2.5 in Bücher et al. (2011). The latter choice yields a good compromise between good finite sample behavior and analytical tractability.

The target values. Our simulation study aims at investigating the performance of $\widehat{A}_{n,m_n}$ and $\widehat{A}^{abc}_{n,m_n}$ as estimators for A_∞. For that purpose, we choose 21 points $t_j = j/20$ in the unit interval, $j = 0, 1, \ldots, 20$, and estimate the summed squared bias

$$B^{(sum)} := \sum_{j=1}^{19} \{\mathbb{E}[\widehat{A}_{n,k_n}(j/20) - A_\infty(j/20)]\}^2,$$

the summed variance

$$\mathrm{Var}^{(sum)} := \sum_{j=1}^{19} \mathrm{Var}\{\widehat{A}_{n,k_n}(j/20)\}$$

and the summed mean squared error $\mathrm{MSE}^{(sum)} := B^{(sum)} + \mathrm{Var}^{(sum)}$ by averaging out over $N = 1,000$ repetitions (analogously for $\widehat{A}^{abc}_{n,m_n}$).

Results and discussion. The results are reported only partially. Figure 1 is concerned with a fixed sample size $n = 1,000$. We plot $B^{(sum)}$, $\mathrm{Var}^{(sum)}$ and $\mathrm{MSE}^{(sum)}$ against the number of blocks k_n (on a logarithmic scale) for the estimator $\widehat{A}^{abc}_{n,m_n}$ and for both copula models mentioned above with tail dependence coefficients in $\{0.25, 0.5, 0.75\}$ and with fixed $a = 0.25$ and $b = 0.5$. For the sake of brevity, we do not show any results for $\widehat{A}_{n,m_n}$ (they are slightly worse than those for $\widehat{A}^{abc}_{n,m_n}$ in most cases) or for different choices of a and b (they do not reveal any additional qualitative insight compared to the case $a = 0.25$ and $b = 0.5$). From the pictures we see that, as expected, the variance of the estimator is decreasing in k, while the bias is increasing. For $k = n = 1,000$, which corresponds to $m = 1$, i.e., to not forming blocks at all, it can be shown that the estimators are actually consistent for the function

$$A_1^\star(t) = \int_0^1 \log\{C_1(y^{\boldsymbol{t}})\} \frac{p_{0.5}(y)}{\log(y)}\, dy = \frac{9}{4} \int_0^1 \sqrt{y}\, |\log\{C_1(y^{\boldsymbol{t}})\}|\, dy, \tag{5.3}$$

The latter fact may serve as an explanation for the different magnitude of the bias at the right end of the pictures in Figure 1. More precisely, Table 5 states the L_2-distances between $A_1^\star$ and A_∞, which exactly resemble the ordering of the value of the summed squared bias at $k_n = 1,000$ over the respective pictures in Figure 1. Regarding the summed MSE, we observe a rather good and robust performance for values of k_n between 150 and 250, corresponding to block lengths between 4 and 7.

Tail dependence coefficient	0.25	0.5	0.75
Outer power Clayton copula	4.62×10^{-2}	1.62×10^{-2}	1.20×10^{-2}
t_4-copula	2.86×10^{-2}	2.26×10^{-2}	0.80×10^{-2}

Table 1: L_2-distances between A_∞ and A_1^*, the latter being defined in (5.3).

Finally, in Figure 2, we present simulation results on $\mathrm{MSE}^{(sum)}$ in the case of a fixed $m = 30$ and with varying $k_n \in \{12, 24, \ldots, 240\}$, corresponding to monthly blocks over daily data for 1 up to 20 years. The shape of the functions are as expected; in particular we see that $\mathrm{MSE}^{(sum)}$ approximately halves when the number of years doubles. Moreover, the pictures reveal a better performance for increasing strength of dependence. The latter may be explained by the fact that, in the extreme case of perfect dependence, the empirical copula is a deterministic function converging at rate k_n^{-1} rather than $k_n^{-1/2}$.

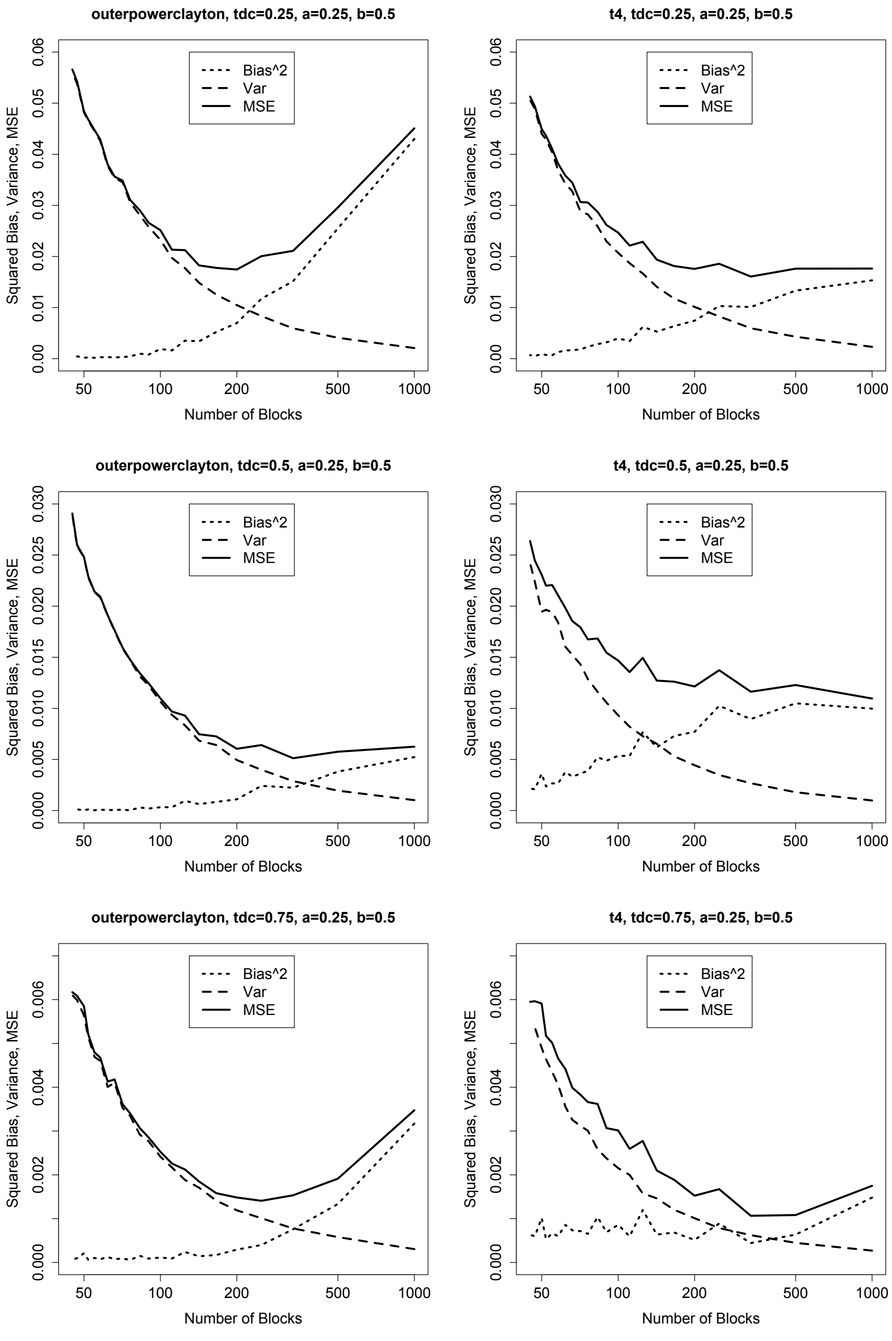

Figure 1: Simulation results on $B^{(sum)}$, $\mathrm{Var}^{(sum)}$ and $\mathrm{MSE}^{(sum)}$ for fixed $n = 1,000$ and varying number of blocks. From top to bottom: tail dependence coefficient $0.25, 0.5$ and 0.75; left: outer power Clayton copula; right: t_4-copula.

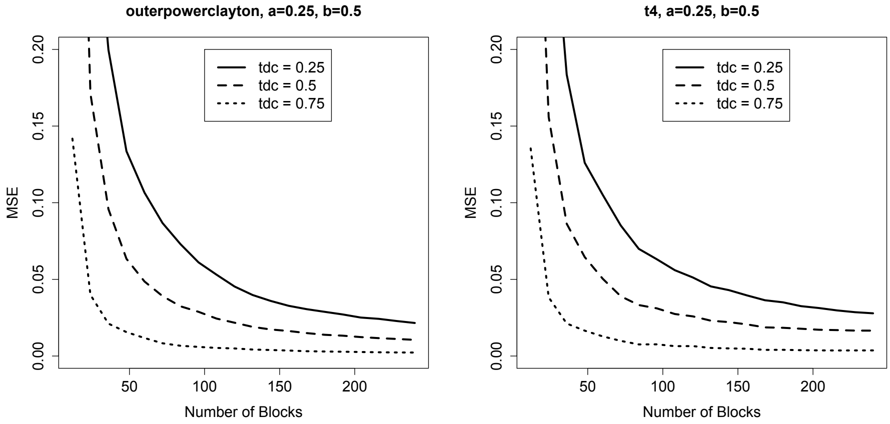

Figure 2: Simulation results for $\mathrm{MSE}^{(sum)}$, with fixed $m = 30$. Left: outer power transform of a Clayton copula, right: t_4-copula, both with parameter chosen in such a way that the coefficient of upper tail dependence is given by 0.25 (solid lines), 0.5 (dashed lines) or 0.75 (dotted lines).

6 Conclusion and discussion

The block maxima method is a time-honoured method in extreme value analysis. In the asymptotic theory, the block maxima are usually modelled as being sampled randomly from an extreme value distribution. In practice, however, the maxima are computed over blocks of finite length. The block length then becomes a tuning parameter, much like the threshold in the peaks-over-threshold method. For large block lengths, the extreme value approximation is accurate, but there are few blocks, leading to large sample variation. Taking smaller blocks augments the number of blocks and thereby reduces the variance of the estimators but at the cost of a potential bias stemming from a bad fit of the extreme value distribution.

The issue is investigated in the context of the nonparametric estimation of the limiting extreme value copula of vectors of componentwise block maxima. The underlying series is supposed to be an absolutely regular, stationary multivariate time series. The sample is partitioned into blocks in such a way that both the block length and the number of blocks tend to infinity. Functional central limit theorems state the asymptotic normality of the empirical copula process and of a rank-based, nonparametric minimum-distance estimator of the Pickands dependence function. The results are illustrated numerically for bivariate moving maximum processes, where the bias-variance trade-off is clearly visible.

The paper leaves ample opportunity for further research into the large-sample theory for the block maxima method for vectors of maxima over blocks of increasing length. We just mention a few possibilities:

– The set-up being nonparametric, a convenient way to calculate standard errors would be via bootstrapping the empirical copula process of block maxima. See for instance Bücher and Dette (2010) for a review of resampling methods for empirical copula processes.

- Often, the extreme value copula is modelled parametrically. In combination with extreme value distributions for the margins, this leads to a parametric model for the block maxima (Tawn, 1988, 1990). The asymptotic theory of estimators for the parameters based on a triangular array of block maxima could then be investigated as well.
- The minimum distance estimator of the Pickands dependence function is itself not a Pickands dependence function. A way of enforcing the proper shape constraints in arbitrary dimensions is via L^2-projection on a parametric sieve (Gudendorf and Segers, 2012).
- Apart from estimation, there are many interesting hypothesis tests that can be investigated: the goodness-of-fit of a parametric model (Genest et al., 2011), the max-stability hypothesis (Kojadinovic et al., 2011), symmetry or other shape constraints (Kojadinovic and Yan, 2012), etc.
- The moving maximum process considered in the paper is a bit artificial. What can one say about the copulas C_m and C_∞ for common time series models? For GARCH type models, even the computation of the copula, C_1, of the stationary distribution is challenging.

Acknowledgments

The research by A. Bücher has been supported in parts by the Collaborative Research Center "Statistical modeling of nonlinear dynamic processes" (SFB 823, Project A7) of the German Research Foundation (DFG) and by the IAP research network Grant P7/06 of the Belgian government (Belgian Science Policy), which is gratefully acknowledged.

J. Segers gratefully acknowledges funding by contract "Projet d'Actions de Recherche Concertées" No. 12/17-045 of the "Communauté française de Belgique" and by IAP research network Grant P7/06 of the Belgian government (Belgian Science Policy).

A Proofs for Section 3

Throughout, simplify notation by writing $m = m_n$, $k = k_n$ and $\ell = \ell_n$.

A.1 Proofs for Subsection 3.1

Lemma A.1. *If $\ell = o(m)$ and $(m/\ell)\,\alpha(\ell) \to 0$, then for every $j \in \{1, \ldots, d\}$ and every $u > 0$,*

$$\mathrm{P}\{F_{m,j}(M_{\ell,1,j}) > u\} = O(\ell/m), \qquad n \to \infty.$$

Proof. The result is univariate, so we suppress the index j from the notation. Consider the maxima of consecutive subblocks of size ℓ contained within the first block of length m:

$$M_{\ell,1}, \ldots, M_{\ell,\lfloor m/\ell \rfloor}.$$

Of these blocks, only keep the ones with an odd index. Since the distribution of $M_{m,1}$ is continuous (all variables X_t having a continuous distribution), we find

$$0 < u = \mathrm{P}\{F_m(M_{m,1}) \le u\} \le \mathrm{P}\Big\{ \max_{\substack{1 \le i \le \lfloor m/\ell \rfloor \\ i \text{ is odd}}} F_m(M_{\ell,i}) \le u \Big\}.$$

The odd blocks are separated by a lag ℓ. Therefore, by induction,

$$\Bigg| \mathrm{P}\Big\{ \max_{\substack{1 \le i \le \lfloor m/\ell \rfloor \\ i \text{ is odd}}} F_m(M_{\ell,i}) \le u \Big\} - \prod_{\substack{1 \le i \le \lfloor m/\ell \rfloor \\ i \text{ is odd}}} \mathrm{P}\{F_m(M_{\ell,i}) \le u\} \Bigg| \le (m/\ell)\, \alpha(\ell).$$

The number of indices i in the product is at least equal to $\lfloor m/\ell \rfloor/2$. By stationarity, we obtain

$$[1 - \mathrm{P}\{F_m(M_{\ell,1}) > u\}]^{\lfloor m/\ell \rfloor/2} \ge u + o(1), \qquad n \to \infty.$$

But $m/\ell \to \infty$, and thus

$$\limsup_{n\to\infty} \frac{m}{\ell} \, \mathrm{P}\{F_m(M_{\ell,1}) > u\} < \infty,$$

as required. □

Proof of Theorem 3.1. We first consider convergence of the finite-dimensional distributions of $\mathbb{C}^\circ_{n,m}$. Let $\ell = \ell_n \in \mathbb{N}$ be some sequence of integers as in Condition 2.2. In each block, we clip off smaller blocks of length ℓ:

$$\begin{aligned} M^{[\ell]}_{m,i,j} &= \max\{X_{t,j} : t \in (im - m, im - \ell] \cap \mathbb{Z}\} \\ L^{[\ell]}_{m,i,j} &= \max\{X_{t,j} : t \in (im - \ell, im] \cap \mathbb{Z}\}. \end{aligned}$$

Clearly, $M_{m,i,j} = \max\{M^{[\ell]}_{m,i,j}, L^{[\ell]}_{m,i,j}\}$. Since $\ell = o(m)$, the pieces clipped off, $L^{[\ell]}_{m,i,j}$, can be expected to be small. On the other hand, since $\ell \to \infty$, the clipped blocks $M^{[\ell]}_{m,1,j}, \ldots, M^{[\ell]}_{m,k,j}$ should be approximately independent. Set

$$U^{[\ell]}_{m,i,j} = F_{m,j}(M^{[\ell]}_{m,i,j}), \qquad \boldsymbol{U}^{[\ell]}_{m,i} = (U^{[\ell]}_{m,i,1}, \ldots, U^{[\ell]}_{m,i,d}),$$

and define

$$\hat{C}^{\circ,[\ell]}_{n,m}(\boldsymbol{u}) = \frac{1}{k} \sum_{i=1}^{k} I(\boldsymbol{U}^{[\ell]}_{m,i} \le \boldsymbol{u}), \qquad \mathbb{C}^{\circ,[\ell]}_{n,m}(\boldsymbol{u}) = \sqrt{k}\{\hat{C}^{\circ,[\ell]}_{n,m}(\boldsymbol{u}) - C_m(\boldsymbol{u})\}.$$

First, we are going to show that, for any $\boldsymbol{u} \in [0,1]^d$,

$$|\hat{C}^{\circ,[\ell]}_{n,m}(\boldsymbol{u}) - \hat{C}^{\circ}_{n,m}(\boldsymbol{u})| = o_P(k^{-1/2}). \tag{A.1}$$

Since $M^{[\ell]}_{m,i,j} \le M_{m,i,j}$ and hence $U^{[\ell]}_{m,i,j} \le U_{m,i,j}$, we have

$$\begin{aligned} &\big| I(\boldsymbol{U}_{m,i} \le \boldsymbol{u}) - I(\boldsymbol{U}^{[\ell]}_{m,i} \le \boldsymbol{u}) \big| \\ &\quad = \Big| \prod\nolimits_{j=1}^{d} I(U_{m,i,j} \le u_j) - \prod\nolimits_{j=1}^{d} I(U^{[\ell]}_{m,i,j} \le u_j) \Big| \\ &\quad \le \sum_{j=1}^{d} \big| I(U_{m,i,j} \le u_j) - I(U^{[\ell]}_{m,i,j} \le u_j) \big| = \sum_{j=1}^{d} I(U^{[\ell]}_{m,i,j} \le u_j < U_{m,i,j}) \\ &\quad \le \sum_{j=1}^{d} I\{F_{m,j}(L^{[\ell]}_{m,i,j}) > u_j\}. \end{aligned}$$

Therefore, $\sqrt{k}\,\mathrm{E}\,|\hat{C}^{\circ,[\ell]}_{n,m}(\boldsymbol{u}) - \hat{C}^{\circ}_{n,m}(\boldsymbol{u})|$ can be bounded by

$$\sum_{j=1}^{d} \sqrt{k}\,\mathrm{P}\{F_{m,j}(L^{[\ell]}_{m,i,j}) > u_j\} = O(\sqrt{k}\,\ell/m) = o(1),$$

in view of Lemma A.1 and Condition 2.2(iii)–(iv); note that $L^{[\ell]}_{m,i,j}$ is equal in distribution to $M_{\ell,1,j}$. The assertion in (A.1) is proved.

Now, let $\boldsymbol{u}_1, \dots, \boldsymbol{u}_q$ be a finite collection of vectors in $[0,1]^d$. For the fidi-part of the proof, we have to show that

$$(\mathbb{C}^{\circ}_{n,m}(\boldsymbol{u}_1), \dots, \mathbb{C}^{\circ}_{n,m}(\boldsymbol{u}_q)) \rightsquigarrow (\mathbb{C}^{\circ}(\boldsymbol{u}_1), \dots, \mathbb{C}^{\circ}(\boldsymbol{u}_q)),$$

which, by (A.1), follows if we prove that

$$(\mathbb{C}^{\circ,[\ell]}_{n,m}(\boldsymbol{u}_1), \dots, \mathbb{C}^{\circ,[\ell]}_{n,m}(\boldsymbol{u}_q)) \rightsquigarrow (\mathbb{C}^{\circ}(\boldsymbol{u}_1), \dots, \mathbb{C}^{\circ}(\boldsymbol{u}_q)).$$

By the Cramér–Wold device, the latter is equivalent to

$$Z_n = \sum_{\nu=1}^{q} c_\nu \mathbb{C}^{\circ,[\ell]}_{n,m}(\boldsymbol{u}_\nu) \rightsquigarrow \sum_{\nu=1}^{q} c_\nu \mathbb{C}^{\circ}(\boldsymbol{u}_\nu) = Z,$$

for any $c_1 \dots, c_q \in \mathbb{R}$. Write $Z_n = \sum_{i=1}^{k} Z_{i,n}$, where

$$Z_{i,n} = \frac{1}{\sqrt{k}} \sum_{\nu=1}^{q} c_\nu \{I(\boldsymbol{U}^{[\ell]}_{m,i} \le \boldsymbol{u}_\nu) - C_m(\boldsymbol{u}_\nu)\}.$$

For $t \in \mathbb{R}$, let $\psi_{i,n}(t) = \exp(-\mathrm{i}tZ_{i,n})$ with i the imaginary unit. Note that the characteristic function of Z_n can be written as $t \mapsto \mathrm{E}\left\{\prod_{i=1}^{k} \psi_{i,n}(t)\right\}$. Now, for any $t \in \mathbb{R}$, we can write

$$\begin{aligned}
&\left|\mathrm{E}\left\{\prod\nolimits_{i=1}^{k} \psi_{i,n}(t)\right\} - \prod\nolimits_{i=1}^{k} \mathrm{E}\{\psi_{i,n}(t)\}\right| \\
&\quad \le \left|\mathrm{E}\left\{\prod\nolimits_{i=1}^{k} \psi_{i,n}(t)\right\} - \mathrm{E}\{\psi_{1,n}(t)\}\,\mathrm{E}\left\{\prod\nolimits_{i=2}^{k} \psi_{i,n}(t)\right\}\right| \\
&\quad + |\mathrm{E}\{\psi_{1,n}(t)\}| \left|\mathrm{E}\left\{\prod\nolimits_{i=2}^{k_n} \psi_{i,n}(t)\right\} - \mathrm{E}\{\psi_{2,n}(t)\}\,\mathrm{E}\left\{\prod\nolimits_{i=3}^{k} \psi_{i,n}(t)\right\}\right| \\
&\quad + \dots \\
&\quad + \left|\prod\nolimits_{i=1}^{k-2} \mathrm{E}\{\psi_{i,n}(t)\}\right| \left|\mathrm{E}\left\{\prod\nolimits_{i=k-1}^{k} \psi_{i,n}(t)\right\} - \prod\nolimits_{i=k-1}^{k} \mathrm{E}\{\psi_{i,n}(t)\}\right|.
\end{aligned}$$

Applying $k-1$ times Lemma 3.9 of Dehling and Philipp (2002), we obtain

$$\begin{aligned}
&\left|\mathrm{E}\left\{\prod\nolimits_{i=1}^{k} \psi_{i,n}(t)\right\} - \prod\nolimits_{i=1}^{k} \mathrm{E}\{\psi_{i,n}(t)\}\right| \\
&\qquad \le 2\pi k \max_{1 \le i \le k} \alpha\left(\sigma\left\{\psi_{i,n}(t)\right\}, \sigma\left\{\prod\nolimits_{i'=i+1}^{k} \psi_{i',n}(t)\right\}\right).
\end{aligned}$$

Since the maxima $M^{[\ell]}_{m,i,j}$ over different blocks $i \ne i'$ are based on observations that are at least ℓ observations apart, the right-hand side of the last display is of the order $k\,\alpha(\ell)$. The latter converges to zero by Condition 2.2. Hence, we have shown that the finite-dimensional distributions of $\mathbb{C}_{n,m}^{\circ,[\ell]}$ weakly converge to the same limit as those of the process

$$\tilde{\mathbb{C}}_{n,m}^{\circ,[\ell]}(\boldsymbol{u}) = \sqrt{k}\{\tilde{C}_{n,m}^{\circ,[\ell]}(\boldsymbol{u}) - C_m(\boldsymbol{u})\},$$

with

$$\tilde{C}_{n,m}^{\circ,[\ell]}(\boldsymbol{u}) = \frac{1}{k}\sum_{i=1}^{k} I(\tilde{\boldsymbol{U}}_{m,i}^{[\ell]} \le \boldsymbol{u})$$

and where

$$\tilde{\boldsymbol{U}}_{m,i}^{[\ell]} = (\tilde{U}_{m,i,1}, \dots, \tilde{U}_{m,i,d}), \qquad \tilde{U}_{m,i,j}^{[\ell]} = F_{m,j}(\tilde{M}_{m,i,j}^{[\ell]})$$

and $\tilde{M}_{m,i,j}^{[\ell]} = \max\{\tilde{X}_{t,j} : t \in (im - m, im - \ell] \cap \mathbb{Z}\}$ is based on a sequence $\tilde{\boldsymbol{X}}_1, \dots, \tilde{\boldsymbol{X}}_n$ such that

$$(\tilde{\boldsymbol{X}}_1, \dots, \tilde{\boldsymbol{X}}_m), (\tilde{\boldsymbol{X}}_{m+1}, \dots, \tilde{\boldsymbol{X}}_{2m}), \dots,$$
$$(\tilde{\boldsymbol{X}}_{(k-1)m+1}, \dots, \tilde{\boldsymbol{X}}_{km}), (\tilde{\boldsymbol{X}}_{km+1}, \dots, \tilde{\boldsymbol{X}}_n)$$

are independent and such that each of the $(k+1)$ brackets is equal in law to the same bracket in the original sequence without the tilde. Again applying (A.1) (which is also valid for the corresponding tilde version based on independent blocks), we obtain that the fidis of $\tilde{\mathbb{C}}_{n,m}^{\circ,[\ell]}$ converge to the same limit as those of $\tilde{\mathbb{C}}_{n,m}^{\circ}$, which is defined analogously to $\mathbb{C}_{n,m}^{\circ}$, but with $\boldsymbol{X}_1, \dots, \boldsymbol{X}_n$ replaced by $\tilde{\boldsymbol{X}}_1, \dots, \tilde{\boldsymbol{X}}_n$. Assembling everything, the fidis of $\mathbb{C}_{n,m}^{\circ}$ asymptotically behave as those of $\tilde{\mathbb{C}}_{n,m}^{\circ}$ based on independent blocks. Weak convergence of the latter can easily be deduced from the classical central limit theorem for row-wise independent triangular arrays.

Now, let us prove asymptotic tightness of $\mathbb{C}_{n,m}^{\circ}$. Recall $\beta(n)$ in (2.7). By Berbee's coupling Lemma (Berbee, 1979; Doukhan et al., 1995), we can construct inductively a sequence $(\bar{\boldsymbol{X}}_{im+1}, \dots, \bar{\boldsymbol{X}}_{im+m})_{i\ge 0}$ such that the following three properties hold:

(i) $(\bar{\boldsymbol{X}}_{im+1}, \dots, \bar{\boldsymbol{X}}_{im+m}) \stackrel{d}{=} (\boldsymbol{X}_{im+1}, \dots, \boldsymbol{X}_{im+m})$ for any $i \ge 0$;

(ii) $(\bar{\boldsymbol{X}}_{2im+1}, \dots, \bar{\boldsymbol{X}}_{2im+m})_{i\ge 0}$ and $(\bar{\boldsymbol{X}}_{(2i+1)m+1}, \dots, \bar{\boldsymbol{X}}_{(2i+1)m+m})_{i\ge 0}$ are independent;

(iii) $\mathrm{P}\{(\bar{\boldsymbol{X}}_{im+1}, \dots, \bar{\boldsymbol{X}}_{im+m}) \ne (\boldsymbol{X}_{im+1}, \dots, \boldsymbol{X}_{im+m})\} \le \beta(m)$.

Let $\bar{\mathbb{C}}_{n,m}^{\circ}$ and $\bar{\boldsymbol{U}}_{m,i}$ be defined analogously to $\mathbb{C}_{n,m}^{\circ}$ and $\boldsymbol{U}_{m,i}$, respectively, but with $\boldsymbol{X}_1, \dots, \boldsymbol{X}_n$ replaced by $\bar{\boldsymbol{X}}_1, \dots, \bar{\boldsymbol{X}}_n$. Now, write

$$\mathbb{C}_{n,m}^{\circ}(\boldsymbol{u}) = \bar{\mathbb{C}}_{n,m}^{\circ}(\boldsymbol{u}) + \{\mathbb{C}_{n,m}^{\circ}(\boldsymbol{u}) - \bar{\mathbb{C}}_{n,m}^{\circ}(\boldsymbol{u})\}. \tag{A.2}$$

We will show below that the term in brackets on the right-hand side is $o_P(1)$, uniformly in $\boldsymbol{u} \in [0,1]^d$. Then, in order to show asymptotic tightness of $\mathbb{C}_{n,m}^{\circ}$, it suffices to show that $\bar{\mathbb{C}}_{n,m}^{\circ}$ is asymptotically tight. Write $\bar{\mathbb{C}}_{n,m}^{\circ} = \bar{\mathbb{C}}_{n,m}^{\circ,even} + \bar{\mathbb{C}}_{n,m}^{\circ,odd}$, where $\bar{\mathbb{C}}_{n,m}^{\circ,even}$ and $\bar{\mathbb{C}}_{n,m}^{\circ,odd}$ are defined as sums over the even and odd summands of $\bar{\mathbb{C}}^\circ_{n,m}$, respectively. Since both of these sums are based on independent summands by property (ii), they are asymptotically tight by Theorem 2.11.9 in van der Vaart and Wellner (1996).

It remains to consider the term in brackets on the right-hand side of (A.2). We have

$$
\begin{aligned}
|\bar{\mathbb{C}}^\circ_{n,m}(\boldsymbol{u}) - \mathbb{C}^\circ_{n,m}(\boldsymbol{u})| &\le \frac{1}{\sqrt{k}} \sum_{i=1}^{k} \big| I(\bar{\boldsymbol{U}}_{m,i} \le \boldsymbol{u}) - I(\boldsymbol{U}_{m,i} \le \boldsymbol{u}) \big| \\
&\le \frac{1}{\sqrt{k}} \sum_{i=1}^{k} I\{(\bar{\boldsymbol{X}}_{im+1}, \dots, \bar{\boldsymbol{X}}_{im}) \ne (\boldsymbol{X}_{im+1}, \dots, \boldsymbol{X}_{im})\}.
\end{aligned}
$$

Hence, by Markov's inequality and property (iii), for any $\varepsilon > 0$,

$$
\mathrm{P}\left\{ \sup_{\boldsymbol{u} \in [0,1]^d} |\bar{\mathbb{C}}^\circ_{n,m}(\boldsymbol{u}) - \mathbb{C}^\circ_{n,m}(\boldsymbol{u})| > \varepsilon \right\} \le \frac{\sqrt{k}\beta(m)}{\varepsilon}.
$$

By Condition 2.2(v), we obtain that the second summand on the right-hand side of (A.2) is $o_P(1)$ as $n \to \infty$, uniformly in $\boldsymbol{u} \in [0,1]^d$. □

A.2 Proofs for Subsection 3.2

Proposition 3.2 is in fact a corollary to Theorem 3.1 and Lemma A.2 below. We prefer to state the lemma independently of the block maxima set-up, as it might be useful in other contexts involving empirical copulas for serially dependent random vectors. To formulate the lemma, we need a bit of notation.

Let $(\boldsymbol{Y}_{k,i} = (Y_{k,i,1}, \dots, Y_{k,i,d}) : i = 1, \dots, k)_{k \in \mathbb{N}}$ be a triangular array of row-wise stationary, d-dimensional random vectors with continuous marginal distribution functions $G_{k,1}, \dots, G_{k,d}$. Put $\boldsymbol{U}_{k,i} = (U_{k,i,j})_{j=1}^d$ with $U_{k,i,j} = G_{k,j}(Y_{k,i,j})$. Let

$$
\hat{G}_{k,j}(y) = \frac{1}{k} \sum_{i=1}^{k} I(Y_{k,i,j} \le y), \qquad \hat{G}_k(\boldsymbol{y}) = \frac{1}{k} \sum_{i=1}^{k} I(\boldsymbol{Y}_{k,i} \le \boldsymbol{y}),
$$

be the marginal and joint empirical distribution functions, respectively, of the sample $\boldsymbol{Y}_{k,1}, \dots, \boldsymbol{Y}_{k,k}$. Let $\hat{\boldsymbol{U}}_{k,i} = (\hat{U}_{k,i,j})_{j=1}^d$ with $\hat{U}_{k,i,j} = \hat{G}_{k,j}(Y_{k,i,j})$. Finally, let C_k be the copula of $\boldsymbol{Y}_{k,1}$ and consider the following empirical versions:

$$
\begin{aligned}
\hat{C}^\circ_k(\boldsymbol{u}) &= \frac{1}{k} \sum_{i=1}^{k} I(\boldsymbol{U}_{k,i} \le \boldsymbol{u}), \\
\hat{C}_k(\boldsymbol{u}) &= \frac{1}{k} \sum_{i=1}^{k} I(\hat{\boldsymbol{U}}_{k,i} \le \boldsymbol{u}), \\
\hat{C}^{\mathrm{alt}}_k(\boldsymbol{u}) &= \hat{G}_k\big(\hat{G}^{\leftarrow}_{k,1}(u_1), \dots, \hat{G}^{\leftarrow}_{k,d}(u_d)\big).
\end{aligned}
$$

Lemma A.2. *Consider the set-up in the previous paragraph. If $\sqrt{k}(\hat{C}^\circ_k - C_k)$ converges weakly in $\ell^\infty([0,1]^d)$ to a stochastic process with continuous trajectories, then*

$$
\sup_{\boldsymbol{u} \in [0,1]^d} \big| \hat{C}^{\mathrm{alt}}_k(\boldsymbol{u}) - \hat{C}_k(\boldsymbol{u}) \big| = o_p(1/\sqrt{k}).
$$

Proof. We have

$$
\begin{aligned}
\left|\hat{C}_k^{\mathrm{alt}}(\boldsymbol{u}) - \hat{C}_k(\boldsymbol{u})\right| &\overset{(1)}{\le} \sum_{j=1}^{d} \frac{1}{k} \sum_{i=1}^{k} \left| I\{Y_{k,i,j} \le \hat{G}_{k,j}^{\leftarrow}(u_j)\} - I\{\hat{G}_{k,j}(Y_{k,i,j}) \le u_j\}\right| \\
&\overset{(2)}{=} \sum_{j=1}^{d} \frac{1}{k} \sum_{i=1}^{k} \left| I\{Y_{k,i,j} = \hat{G}_{k,j}^{\leftarrow}(u_j)\} - I\{\hat{G}_{k,j}(Y_{k,i,j}) = u_j\}\right| \\
&\overset{(3)}{\le} \sum_{j=1}^{d} \frac{1}{k} \sum_{i=1}^{k} I\{Y_{k,i,j} = \hat{G}_{k,j}^{\leftarrow}(u_j)\}.
\end{aligned}
$$

Explanations:

(1) Write out the definitions of the two versions of the empirical copula and use the inequality $|\prod_j a_j - \prod_j b_j| \le \sum_j |a_j - b_j|$ for numbers $a_j, b_j \in [0,1]$.

(2) Split both indicators into the indicator of a strict inequality and the indicator of an equality. The indicators for the strict inequality are equal, since $x < H^{\leftarrow}(u)$ if and only if $H(x) < u$ for any distribution function H.

(3) If $\hat{G}_{k,j}(Y_{k,i,j}) = u_j$, then $Y_{k,i,j} = \hat{G}_{k,j}^{\leftarrow}(u_j)$. Hence, the second indicator is not larger than the first one.

Fix $j \in \{1, \ldots, d\}$. Let $\hat{C}_{k,j}^{\circ}$ be the jth margin of $\hat{C}_k^{\circ}$ in (2.3), that is,

$$
\hat{C}_{k,j}^{\circ}(u_j) = \frac{1}{k} \sum_{i=1}^{k} I(U_{k,i,j} \le u_j).
$$

Then we can continue the chain of (in)equalities started in the beginning of the proof by

$$
\begin{aligned}
\left|\hat{C}_k^{\mathrm{alt}}(\boldsymbol{u}) - \hat{C}_k(\boldsymbol{u})\right| &\le \sum_{j=1}^{d} \frac{1}{k} \sum_{i=1}^{k} I\{U_{k,i,j} = G_{k,j}(\hat{G}_{k,j}^{\leftarrow}(u_j))\} \\
&\le \sum_{j=1}^{d} \sup_{x \in \mathbb{R}} \frac{1}{k} \sum_{i=1}^{k} I\{U_{k,i,j} = x\} \\
&\le \sum_{j=1}^{d} \sup_{x \in \mathbb{R}} \{\hat{C}_{k,j}^{\circ}(x) - \hat{C}_{k,j}^{\circ}(x - 1/k)\} \\
&\le \frac{d}{\sqrt{k}} \, \omega_k(1/k) + \frac{d}{k}
\end{aligned}
$$

where ω_k is the modulus of continuity of $\mathbb{C}_k^{\circ} = \sqrt{k}(\hat{C}_k^{\circ} - C_k)$, i.e.

$$
\omega_k(\delta) = \sup_{\substack{(\boldsymbol{x},\boldsymbol{y}) \in ([0,1]^d)^2 \\ \max_j |x_j - y_j| \le \delta}} \left|\mathbb{C}_k^{\circ}(\boldsymbol{x}) - \mathbb{C}_k^{\circ}(\boldsymbol{y})\right|.
$$

As $\mathbb{C}_k^{\circ}$ converges weakly in $\ell^{\infty}([0,1]^d)$ to a process with continuous trajectories, it follows that $\omega_k(1/k) = o_p(1)$. □

Proof of Proposition 3.2. By Theorem 3.1, we can apply Lemma A.2 to $\boldsymbol{Y}_{k,i} = \boldsymbol{M}_{m,i}$. □

Proof of Theorem 3.5. We are going to apply the extended continuous mapping theorem (van der Vaart and Wellner, 1996, Theorem 1.11.1). Recall the copula mapping, Φ, in (3.2), with domain $\mathcal{D}_\Phi$. Let $\mathcal{D}_n$ denote the space of all $\alpha_n \in \ell^\infty([0,1]^d)$ for which $C_m + k^{-1/2}\alpha_n \in \mathcal{D}_\Phi$ and define

$$g_n(\alpha_n) = \sqrt{k}\{\Phi(C_m + k^{-1/2}\alpha_n) - \Phi(C_m)\}.$$

Recall that $\mathcal{D}_0 \subset \ell^\infty([0,1]^d)$ denotes the space of all continuous functions that vanish at $(1,\dots,1)$ and at all $\boldsymbol{u} \in [0,1]^d$ with at least one coordinate being equal to 0. Since

$$\sqrt{k}(\hat{C}^{\text{alt}}_{n,m} - C_m) = g_n\{\sqrt{k}(\hat{C}^{\circ}_{n,m} - C_m)\},$$

the assertion will follow from the extended continuous mapping theorem, provided we can show that $g_n(\alpha_n) \to g(\alpha)$ for any $\alpha_n \to \alpha \in \mathcal{D}_0$, where $g : \mathcal{D}_0 \to \ell^\infty([0,1]^d)$ is defined by

$$(g(\alpha))(\boldsymbol{u}) = \alpha(\boldsymbol{u}) - \sum_{j=1}^d \dot{C}_{\infty,j}(\boldsymbol{u})\,\alpha(\boldsymbol{u}^{(j)}).$$

Note that $g = \Phi'_{C_\infty}$, the Hadamard derivative of Φ at C_∞.

Write $\boldsymbol{I}_n(\boldsymbol{u}) = (I_{n1}(u_1),\dots,I_{nd}(u_d))$ where

$$I_{nj}(u_j) = (\mathrm{id}_{[0,1]} + k^{-1/2}\alpha_{nj})^{\leftarrow}(u_j)$$

with $\alpha_{nj}(u_j) = \alpha_n(1,\dots,1,u_j,1\dots,1)$ and with $\mathrm{id}_{[0,1]}$ the identity function on $[0,1]$. Decompose

$$g_n(\alpha_n) = \sqrt{k}\{C_m(\boldsymbol{I}_n) - C_m\} + \alpha_n(\boldsymbol{I}_n) \tag{A.3}$$

It follows from Vervaat's Lemma, see also formula (4.2) in Bücher and Volgushev (2013), that

$$\sup_{u_j\in[0,1]} \left|\sqrt{k}\{I_{nj}(u_j) - u_j\} + \alpha(1,\dots,1,u_j,1,\dots,1)\right| \to 0. \tag{A.4}$$

In particular, by uniform convergence of α_n to α and by uniform continuity of α, this implies that the second term on right-hand side of (A.3) converges to α, uniformly.

It remains to be shown that the first term on the right-hand side of (A.3) converges to the proper limit, i.e., that

$$\sup_{\boldsymbol{u}\in[0,1]^d}\left|\sqrt{k}\{C_m(\boldsymbol{I}_n(\boldsymbol{u})) - C_m(\boldsymbol{u})\} + \sum_{j=1}^d \dot{C}_{\infty,j}(\boldsymbol{u})\,\alpha(\boldsymbol{u}^{(j)})\right| = 0. \tag{A.5}$$

The proof of (A.5) depends on whether we assume Condition 3.4(a) or (b).

First, we prove (A.5) under Condition 3.4(a). Put

$$\Delta_n = \sqrt{k}(C_m - C_\infty).$$

We have

$$\begin{aligned}&\sqrt{k}\{C_m(\boldsymbol{I}_n(\boldsymbol{u})) - C_m(\boldsymbol{u})\}\\ &\qquad = \sqrt{k}\{C_\infty(\boldsymbol{I}_n(\boldsymbol{u})) - C_\infty(\boldsymbol{u})\} + \{\Delta_n(\boldsymbol{I}_n(\boldsymbol{u})) - \Delta_n(\boldsymbol{u})\}.\end{aligned}$$

It is then sufficient to show that

$$\sup_{\boldsymbol{u}\in[0,1]^d} \left| \sqrt{k}\{C_\infty(\boldsymbol{I}_n(\boldsymbol{u})) - C_\infty(\boldsymbol{u})\} + \sum_{j=1}^d \dot{C}_{\infty,j}(\boldsymbol{u})\,\alpha(\boldsymbol{u}^{(j)}) \right| \to 0, \tag{A.6}$$

$$\sup_{\boldsymbol{u}\in[0,1]^d} \big|\Delta_n(\boldsymbol{I}_n(\boldsymbol{u})) - \Delta_n(\boldsymbol{u})\big| \to 0. \tag{A.7}$$

Convergence in (A.6) essentially follows from Condition 3.3, marginal convergence in (A.4), the fact that $0 \le \dot{C}_j \le 1$, and $\alpha \in \mathcal{D}_0$; the proof is in fact the same as the proof for (A.5) under Condition 3.4(b) for the special case $m = \infty$. The left-hand side of (A.7) is bounded by $\omega(\delta_n)$, where, for $\delta > 0$,

$$\omega(\delta) = \sup_{n\in\mathbb{N}} \sup \left\{ \big|\Delta_n(\boldsymbol{u}) - \Delta_n(\boldsymbol{v})\big| : (\boldsymbol{u},\boldsymbol{v}) \in ([0,1]^d)^2,\ \max_{j=1,\ldots,d} |u_j - v_j| \le \delta \right\},$$

and with

$$\delta_n = \max_{j=1,\ldots,d} \sup_{u_j\in[0,1]} |I_{nj}(u_j) - u_j|.$$

By (A.4), $\delta_n \to 0$ as $n \to \infty$. Since the set $\{\Delta_n : n \in \mathbb{N}\}$ is relatively compact by Condition 3.4(a), the functions Δ_n are uniformly equicontinuous by the Arzelà–Ascoli theorem, which means that $\omega(\delta_n) \to 0$ as $n \to \infty$.

Next, we prove (A.5) under Condition 3.4(b). The margins of C_m being uniform on $(0,1)$, the function C_m is Lipschitz; more precisely,

$$|C_m(\boldsymbol{u}) - C_m(\boldsymbol{v})| \le \sum_{j=1}^d |u_j - v_j|, \qquad (\boldsymbol{u},\boldsymbol{v}) \in ([0,1]^d)^2.$$

As a consequence, the function

$$f : [0,1] \to [0,1] : s \mapsto C_m\big(\boldsymbol{v}_n(\boldsymbol{u},s)\big),$$

where $\boldsymbol{v}_n(\boldsymbol{u},s) = (1-s)\boldsymbol{u} + s\boldsymbol{I}_n(\boldsymbol{u})$, is absolutely continuous, a version of its Radon–Nikodym derivative being

$$f'(s) = \sum_{j=1}^d \dot{C}_{m,j}\big(\boldsymbol{v}_n(\boldsymbol{u},s)\big)\,\{I_{nj}(u_j) - u_j\}.$$

It follows that

$$\begin{aligned}
\sqrt{k}\{C_m(\boldsymbol{I}_n(\boldsymbol{u})) - C_m(\boldsymbol{u})\} &= \sqrt{k}\{f(1) - f(0)\} \\
&= \sqrt{k} \int_0^1 f'(s)\,\mathrm{d}s \\
&= \sum_{j=1}^d \sqrt{k}\{I_{nj}(u_j) - u_j\} \int_0^1 \dot{C}_{m,j}\big(\boldsymbol{v}_n(\boldsymbol{u},s)\big)\,\mathrm{d}s.
\end{aligned}$$

Fix $j = 1,\ldots,d$. We need to show that

$$\sup_{\boldsymbol{u}\in[0,1]^d} \left| \sqrt{k}\{I_{nj}(u_j) - u_j\} \int_0^1 \dot{C}_{m,j}\big(\boldsymbol{v}_n(\boldsymbol{u},s)\big)\,\mathrm{d}s + \alpha(\boldsymbol{u}^{(j)})\,\dot{C}_{\infty,j}(\boldsymbol{u}) \right| \to 0$$

as $n \to \infty$. Fix $\delta \in (0, 1/2)$. In the above supremum, we consider separately the cases $u_j \in [\delta, 1-\delta]$ and $u_j \in [0, \delta] \cup [1-\delta, 1]$.

Given $\varepsilon > 0$, the supremum over those points $\boldsymbol{u} \in [0,1]^d$ such that $u_j \in [0, \delta] \cup [1-\delta, 1]$ can be made smaller than ε for sufficiently large n by using the fact that $0 \le \dot{C}_{m,j} \le 1$ and $0 \le \dot{C}_{\infty,j} \le 1$ together with uniform convergence in (A.4) and the assumption that $\alpha \in \mathcal{D}_0$.

Regarding the supremum over $\boldsymbol{u} \in [0,1]^d$ such that $u_j \in [\delta, 1-\delta]$, note that

$$\begin{aligned}\int_0^1 \dot{C}_{m,j}\big(\boldsymbol{v}_n(\boldsymbol{u}, s)\big)\, \mathrm{d}s = \dot{C}_{\infty,j}(\boldsymbol{u}) & \\ + \int_0^1 \big\{\dot{C}_{\infty,j}\big(\boldsymbol{v}_n(\boldsymbol{u}, s)\big) - \dot{C}_{\infty,j}(\boldsymbol{u})\big\}\, \mathrm{d}s & \\ + \int_0^1 \big\{\dot{C}_{m,j}\big(\boldsymbol{v}_n(\boldsymbol{u}, s)\big) - \dot{C}_{\infty,j}\big(\boldsymbol{v}_n(\boldsymbol{u}, s)\big)\big\}\, \mathrm{d}s. &\end{aligned}$$

All integrands on the right-hand side converging to zero uniformly over $s \in [0,1]$ and $\boldsymbol{u} \in [0,1]^d$ such that $u_j \in [\delta, 1-\delta]$, the proof is complete. □

Proof of Corollary 3.6. As convergence in (3.3) implies relative compactness, Condition 3.4(a) is fulfilled. By Theorem 3.5 and Slutsky's lemma,

$$\sqrt{k}(\hat{C}_{n,m} - C_\infty) = \mathbb{C}_{n,m} + \sqrt{k}(C_m - C_\infty) \rightsquigarrow \mathbb{C} + \Gamma, \qquad n \to \infty,$$

as required. □

A.3 Proofs for Subsection 3.3

For the proof of Theorem 3.7, we introduce the notation

$$\mathbb{W}_{n,\omega}(\boldsymbol{t}) = \sqrt{k_n} \int_0^1 \log\left\{\frac{\tilde{C}_{n,m}(\boldsymbol{y}^{\boldsymbol{t}})}{C_\infty(\boldsymbol{y}^{\boldsymbol{t}})}\right\} \omega(y, \boldsymbol{t})\, \mathrm{d}y,$$

for a measurable weight function $\omega : [0,1] \times \Delta_{d-1} \to \mathbb{R}$ that may depend on y and $\boldsymbol{t}$.

Theorem A.3. *Suppose that Conditions 2.1, 2.2 and 3.3 are met and that* $\sqrt{k}(C_m - C_\infty) \to \Gamma$, *uniformly. Assume that there exists a bounded, measurable function* $\overline{\omega} : [0,1] \to \mathbb{R}$ *such that* $|\omega(y, \boldsymbol{t})| \le \overline{\omega}(y)$ *for all* $y \in [0,1]$ *and all* $\boldsymbol{t} \in \Delta_{d-1}$ *and such that*

$$\int_0^1 \overline{\omega}(y)\, y^{-\lambda} \mathrm{d}y < \infty \text{ for some } \lambda > 1.$$

Then, for any $\gamma \in \left(\frac{1}{2}, \frac{\lambda}{2}\right)$,

$$\mathbb{W}_{n,\omega} \rightsquigarrow \mathbb{W}_{\infty,\omega} \quad \text{in} \quad \ell^\infty(\Delta_{d-1}),$$

as $n \to \infty$, *where the limiting process is given by*

$$\mathbb{W}_{\infty,\omega}(\boldsymbol{t}) = \int_0^1 \frac{\mathbb{C}(\boldsymbol{y}^{\boldsymbol{t}}) + \Gamma(\boldsymbol{y}^{\boldsymbol{t}})}{C_\infty(\boldsymbol{y}^{\boldsymbol{t}})}\, \omega(y, \boldsymbol{t})\, \mathrm{d}y.$$

Proof of Theorem A.3. For $q \in \mathbb{N}$, let

$$\mathbb{W}_{n,\omega,q}(\boldsymbol{t}) = \sqrt{k} \int_{1/q}^{1} \log \left\{ \frac{\tilde{C}_{n,m}(\boldsymbol{y^t})}{C_\infty(\boldsymbol{y^t})} \right\} \omega(y, \boldsymbol{t}) \, \mathrm{d}y,$$
$$\mathbb{W}_{\infty,\omega,q}(\boldsymbol{t}) = \int_{1/q}^{1} \frac{\mathbb{C}(\boldsymbol{y^t}) + \Gamma(\boldsymbol{y^t})}{C_\infty(\boldsymbol{y^t})} \omega(y, \boldsymbol{t}) \, \mathrm{d}y.$$

By Lemma B.1 in Bücher et al. (2011) it suffices to show the following three claims:

(i) $\mathbb{W}_{n,\omega,q} \rightsquigarrow \mathbb{W}_{\infty,\omega,q}$ in $\ell^\infty(\Delta_{d-1})$ as $n \to \infty$;
(ii) $\mathbb{W}_{\infty,\omega,q} \rightsquigarrow \mathbb{W}_{\infty,\omega}$ in $\ell^\infty(\Delta_{d-1})$ as $q \to \infty$;
(iii) $\forall \varepsilon > 0 : \lim_{q\to\infty} \limsup_{n\to\infty} \mathrm{P}\left\{ \sup_{\boldsymbol{t}\in\Delta_{d-1}} |\mathbb{W}_{n,\omega,q}(\boldsymbol{t}) - \mathbb{W}_{n,\omega}(\boldsymbol{t})| > \varepsilon \right\} = 0.$

The proof of all three assertions follows exactly along the lines of the proof of Theorem 6.1 in Berghaus et al. (2013) and is based on the fact that

$$\sqrt{k}(\hat{C}_{n,m} - C_\infty) \rightsquigarrow \mathbb{C} + \Gamma$$

in $\ell^\infty([0,1]^d)$ by Corollary 3.6. The details are omitted for the sake of brevity. □

Proof of Theorem 3.7. Setting $\omega(y, \boldsymbol{t}) = p(y)/\log(y)$, the result is a simple corollary of Theorem A.3. □

B Proofs for Section 4

B.1 Proofs for Subsection 4.1

The cumulative distribution function, F_m, and the copula, C_m, of $\boldsymbol{M}_m$ are given by

$$F_m(\boldsymbol{u}) = \prod_{s=1-p}^{m} D\big((u_j^{\alpha_{mjs}})_{j=1}^d\big), \qquad C_m(\boldsymbol{u}) = \prod_{s=1-p}^{m} D\big((u_j^{\beta_{mjs}})_{j=1}^d\big),$$

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

$$\alpha_{mjs} = \max\{a_{ij} : i = \max(1-s, 0), \ldots, \min(m-s, p)\},$$
$$\alpha_{mj\bullet} = \sum_{s=1-p}^{m} \alpha_{mjs},$$
$$\beta_{mjs} = \frac{\alpha_{mjs}}{\alpha_{mj\bullet}}.$$

The proof is straightforward by direct computation.

For $m = 1$, we have $\alpha_{1js} = a_{1-s,j}$ and thus

$$\alpha_{1j\bullet} = \sum_{s=1-p}^{1} a_{1-s,j} = \sum_{i=0}^{p} a_{ij} = 1$$

by (4.2). We find that

$$C_1(\boldsymbol{u}) = F_1(\boldsymbol{u}) = \prod_{i=0}^{p} D\big((u_j^{a_{ij}})_{j=1}^d\big) \tag{B.1}$$

In particular, the random variables U_{tj} are uniformly distributed on $(0, 1)$.

Proof of Proposition 4.1. Equation (4.4) is a direct consequence of the next sandwich inequality for C_m: For $m \in \mathbb{N}$ such that $m > p$, we have

$$\big(D_{m-p}(\boldsymbol{u})\big)^{\frac{m+p}{m-p}} \le C_m(\boldsymbol{u}) \le \big(D_{m+p}(\boldsymbol{u})\big)^{\frac{m-p}{m+p}}. \tag{B.2}$$

We prove (B.2). For $j = 1, \ldots, d$, put

$$A_j = \max\{a_{ij} : i = 0, \ldots, p\}.$$

Since all a_{ij} are nonnegative and because of (4.2), we have $0 < A_j \le 1$. Note that

$$\alpha_{mjs} = A_j \qquad (s = 1, \ldots, m - p),$$

whereas for the other s, we still have $\alpha_{mjs} \le A_j$. It follows that

$$(m - p)A_j \le \alpha_{mj\bullet} \le (m + p)A_j.$$

In particular, for all $s = 1 - p, \ldots, m$,

$$\beta_{mjs} = \frac{\alpha_{mjs}}{\alpha_{mj\bullet}} \le \frac{A_j}{(m-p)A_j} = \frac{1}{m-p}.$$

We find

$$\begin{aligned} C_m(\boldsymbol{u}) &\ge \prod_{s=1-p}^{m} D(u_1^{1/(m-p)}, \ldots, u_d^{1/(m-p)}) \\ &= \big(D(u_1^{1/(m-p)}, \ldots, u_d^{1/(m-p)})\big)^{m+p} \\ &= \big(D_{m-p}(\boldsymbol{u})\big)^{\frac{m+p}{m-p}}. \end{aligned}$$

On the other hand,

$$\beta_{mjs} \ge \frac{A_j}{(m+p)A_j} = \frac{1}{m+p} \qquad (s = 1, \ldots, m - p),$$

from which

$$C_m(\boldsymbol{u}) \le \prod_{s=1}^{m-p} D(u_1^{1/(m+p)}, \ldots, u_d^{1/(m+p)}) = \big(D_{m+p}(\boldsymbol{u})\big)^{\frac{m-p}{m+p}},$$

proving also the lower bound. This completes the proof of (B.2). □

The limit D_∞ in Proposition 4.1 is in general different from the extreme value attractor of C_1. Indeed, if (4.3) holds, then by (B.1),

$$\begin{aligned} \big(C_1(u_1^{1/m}, \ldots, u_d^{1/m})\big)^{1/m} &= \prod_{i=0}^{p} \big\{D\big((u_j^{a_{ij}/m})_{j=1}^d\big)\big\}^{1/m} \\ &\to \prod_{i=0}^{p} D_\infty\big((u_j^{a_{ij}})_{j=1}^d\big), \qquad m \to \infty. \end{aligned} \tag{B.3}$$

B.2 Proofs for Subsection 4.2

Define the ϕ-mixing coefficient of $(\boldsymbol{X}_t)_t$ as

$$\phi(n) = \sup_{t\in\mathbb{Z}} \sup \left\{ |\operatorname{P}(A \mid B) - \operatorname{P}(A)| : B \in \mathcal{F}_{-\infty}^{t}, A \in \mathcal{F}_{t+n}^{\infty}, \operatorname{P}(B) > 0 \right\}$$

and note that $\beta(n) \le \phi(n)$ (Bradley, 2005). Because of the random repetition mechanism, the process $(\boldsymbol{X}_t)_t$ is ϕ-mixing and the mixing coefficients $\phi(n)$ decay to 0 geometrically.

Lemma B.1. *Let $t \in \mathbb{Z}$ and $n \in \mathbb{N}$, and let $B \in \mathcal{F}_{-\infty}^{t}$ with $\operatorname{P}(B) > 0$. and $A \in \mathcal{F}_{t+n}^{\infty}$. Then*

$$|\operatorname{P}(A \mid B) - \operatorname{P}(A)| \le 2(1-\theta)^n.$$

Proof. Consider the event Q that n consecutive repetitions occur at times $t+1, \ldots, t+n$, that is,

$$Q = \bigcap\nolimits_{i=1}^{n} \{I_{t+i} = 0\}.$$

Note that Q is independent of A and B and that $\operatorname{P}(Q) = (1-\theta)^n$. We have

$$\begin{aligned}
&|\operatorname{P}(A \mid B) - \operatorname{P}(A)| \\
&\qquad \le \operatorname{P}(A \cap Q \mid B) + \operatorname{P}(A \cap Q) + |\operatorname{P}(A \cap Q^c \mid B) - \operatorname{P}(A \cap Q^c)| \\
&\qquad\qquad \le 2\operatorname{P}(Q) = 2(1-\theta)^n.
\end{aligned}$$

The inequality follows from the independence of Q and B and the independence of $A \cap Q^c$ and B. □

Proof of Proposition 4.2. For integer $m \ge 2$, partition the event $\{\boldsymbol{M}_m \le \boldsymbol{x}\}$ into two pieces, according to whether I_m is equal to 1 or not:

$$\begin{aligned}
F_m(\boldsymbol{x}) &= \operatorname{P}[\boldsymbol{M}_m \le \boldsymbol{x}] \\
&= \operatorname{P}[\boldsymbol{M}_{m-1} \le \boldsymbol{x},\, \boldsymbol{X}_m \le \boldsymbol{x}] \\
&= \operatorname{P}[\boldsymbol{M}_{m-1} \le \boldsymbol{x}]\ \operatorname{P}[\boldsymbol{\xi}_m \le \boldsymbol{x}]\,\theta + \operatorname{P}[\boldsymbol{M}_{m-1} \le \boldsymbol{x}]\,(1-\theta) \\
&= F_{m-1}(\boldsymbol{x})\,\{\theta\, F_1(\boldsymbol{x}) + 1 - \theta\}.
\end{aligned}$$

By induction, we find

$$F_m(\boldsymbol{x}) = F_1(\boldsymbol{x})\,[1 - \theta\{1 - F_1(\boldsymbol{x})\}]^{m-1}.$$

For the marginal distributions, we find accordingly

$$F_{m,j}(x_j) = F_{1,j}(x_j)\,[1 - \theta\{1 - F_{1,j}(x_j)\}]^{m-1}$$

for $j = 1, \ldots, d$.

Let $F_{m,j}^{\leftarrow}$ be the left-continuous inverse of $F_{m,j}$. For $u_j \in (0,1]$ and $m \ge 2$, we have

$$\begin{aligned}
u_j &= F_{m,j}\big(F_{m,j}^{\leftarrow}(u_j)\big) \\
&= F_{1,j}(F_{m,j}^{\leftarrow}(u_j))\,[1 - \theta\{1 - F_{1,j}(F_{m,j}^{\leftarrow}(u_j))\}]^{m-1} \\
&\le [1 - \theta\{1 - F_{1,j}(F_{m,j}^{\leftarrow}(u_j))\}]^{m-1},
\end{aligned}$$

and thus

$$\begin{aligned} F_{1,j}\big(F_{m,j}^{\leftarrow}(u_j)\big) &\ge 1 - \theta^{-1}(1 - u_j^{1/(m-1)}) \\ &\ge 1 + \frac{1}{\theta(m-1)} \log u_j \to 1 \qquad (m \to \infty). \end{aligned}$$

Combining the previous two displays, we find

$$u_j = \{1 + o(1)\}\, [1 - \theta\{1 - F_{1,j}(F_{m,j}^{\leftarrow}(u_j))\}]^{m-1}$$

from which

$$\begin{aligned} F_{1,j}\big(F_{m,j}^{\leftarrow}(u_j)\big) &= 1 - \theta^{-1}\big(1 - [u_j\{1 + o(1)\}]^{1/(m-1)}\big) \\ &= 1 + \frac{1}{\theta(m-1)}\{\log(u_j) + o(1)\} \qquad (m \to \infty). \end{aligned}$$

Writing

$$\boldsymbol{F}_m^{\leftarrow}(\boldsymbol{u}) = \big(F_{m,1}^{\leftarrow}(u_1), \ldots, F_{m,d}^{\leftarrow}(u_d)\big),$$

we have, for $\boldsymbol{u} \in (0,1]^d$,

$$F_1\big(\boldsymbol{F}_m^{\leftarrow}(\boldsymbol{u})\big) \ge 1 - \sum_{j=1}^{d} \big\{1 - F_{1,j}\big(F_{m,j}^{\leftarrow}(u_j)\big)\big\} \to 1 \qquad (m \to \infty).$$

The copula, C_m, of F_m in $\boldsymbol{u} \in (0,1]^d$ is given by

$$\begin{aligned} &C_m(\boldsymbol{u}) = F_m\big(\boldsymbol{F}_m^{\leftarrow}(\boldsymbol{u})\big) \\ &= F_1\big(\boldsymbol{F}_m^{\leftarrow}(\boldsymbol{u})\big)\big[1 - \theta + \theta F_1\big(\boldsymbol{F}_m^{\leftarrow}(\boldsymbol{u})\big)\big]^{m-1} \\ &= \{1 + o(1)\} \left[1 - \theta + \theta C_1\left(1 + \frac{\log(u_1) + o(1)}{\theta(m-1)}, \ldots, 1 + \frac{\log(u_d) + o(1)}{\theta(m-1)}\right)\right]^{m-1}. \end{aligned}$$

If C_1 is in the copula domain of attraction of an extreme value copula C_∞ with stable tail dependence function L, then

$$\lim_{h \searrow 0} h^{-1}\{1 - C_1(1 - hx_1, \ldots, 1 - hx_d)\} = L(x_1, \ldots, x_d)$$

locally uniformly in $(x_1, \ldots, x_d) \in [0, \infty)^d$. It follows that

$$C_m(\boldsymbol{u}) \to \exp\{-L(-\log u_1, \ldots, -\log u_d)\} = C_\infty(\boldsymbol{u}) \qquad (m \to \infty)$$

as required. □

References

Amram, F. (1985). Multivariate extreme value distributions for stationary gaussian sequences. *Journal of Multivariate Analysis 16*, 237–240.

Beirlant, J., Y. Goegebeur, J. Segers, and J. Teugels (2004). *Statistics of Extremes: Theory and Applications*. Wiley Series in Probability and Statistics. Chichester: John Wiley & Sons Ltd.

Berbee, H. C. P. (1979). *Random walks with stationary increments and renewal theory*, Volume 112 of *Mathematical Centre Tracts*. Amsterdam: Mathematisch Centrum.

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

Bradley, R. C. (2005). Basic properties of strong mixing conditions. A survey and some open questions. *Probab. Surv. 2*, 107–144. Update of, and a supplement to, the 1986 original.

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

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

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

Charpentier, A. and J. Segers (2009). Tails of multivariate archimedean copulas. *Journal of Multivariate Analysis 100*(7), 1521–1537.

Deheuvels, P. (1991). On the limiting behavior of the Pickands estimator for bivariate extreme-value distributions. *Statistics & Probability Letters 12*(5), 429–439.

Dehling, H. and W. Philipp (2002). Empirical process techniques for dependent data. In H. Dehling, T. Mikosch, and M. Sorensen (Eds.), *Empirical process techniques for dependent data*, pp. 1–113. Boston: Birkhäuser.

Demarta, S. and A. J. McNeil (2005). The *t*-copula and related copulas. *International Statistical Review 73*, 111–129.

Dombry, C. (2013). Maximum likelihood estimators for the extreme value index based on the block maxima method. arXiv:1301.5611v1.

Doukhan, P., P. Massart, and E. Rio (1995). Invariance principles for absolutely regular empirical processes. *Ann. Inst. H. Poincaré Probab. Statist. 31*(2), 393–427.

Fils-Villetard, A., A. Guillou, and J. Segers (2008). Projection estimators of Pickands dependence functions. *The Canadian Journal of Statistics 36*(3), 369–382.

Genest, C., I. Kojadinovic, J. Nešlehová, and J. Yan (2011). A goodness-of-fit test for bivariate extreme-value copulas. *Bernoulli 17*(1), 253–275.

Genest, C. and J. Segers (2009). Rank-based inference for bivariate extreme-value copulas. *The Annals of Statistics 37*, 2990–3022.

Gudendorf, G. and J. Segers (2010). Extreme-value copulas. In P. Jaworski, F. Durante, W. Härdle, and W. Rychlik (Eds.), *Copula theory and its applications (Warsaw, 2009)*, Lecture Notes in Statistics, pp. 127–146. Springer-Verlag. arXiv:0911.1015v2.

Gudendorf, G. and J. Segers (2011). Nonparametric estimation of an extreme-value copula in arbitrary dimensions. *Journal of Multivariate Analysis 102*(1), 37–47.

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

Gumbel, E. J. (1958). *Statistics of extremes*. New York: Columbia University Press.

Gumbel, E. J. and C. K. Mustafi (1967). Some analytical properties of bivariate extremal distributions. *Journal of the American Statistical Association 62*(318), 569–588.

Hsing, T. (1989). Extreme value theory for multivariate stationary sequences. *Journal of Multivariate Analysis 29*(2), 274 – 291.

Hüsler, J. (1990). Multivariate extreme values in stationary random sequences. *Stochastic Processes and their Applications 35*(1), 99 – 108.

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

Kojadinovic, I. and J. Yan (2012). A non-parametric test of exchangeability for extreme-value and left-tail decreasing bivariate copulas. *Scandinavian Journal of Statistics 39*(3), 480–496.

Leadbetter, M. R., G. Lindgren, and H. Rootzén (1983). *Extremes and related properties of random sequences and processes*. Springer Series in Statistics. New York: Springer-Verlag.

Peng, L., L. Qian, and J. Yang (2013). Weighted estimation of the dependence function for an extreme-value distribution. *Bernoulli 19*(2), 492–520.

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.

Prescott, P. and A. T. Walden (1980). Maximum likelihood estimation of the parameters of the generalized extreme-value distribution. *Biometrika 67*(3), 723–724.

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

Smith, R. L. and I. Weissman (1994). Estimating the extremal index. *Journal of the Royal Statistical Society. Series B (Methodological) 56*(3), 525–528.

Tawn, J. A. (1988). Bivariate extreme value theory: models and estimation. *Biometrika 75*(3), 397–415.

Tawn, J. A. (1990). Modelling multivariate extreme value distributions. *Biometrika 77*(2), 245–253.

van der Vaart, A. and J. Wellner (1996). *Weak convergence and empirical processes*. New York: Springer.