Inference for asymptotically independent samples of extremes

Armelle Guillou, Simone A. Padoan, Stefano Rizzelli

Inference for asymptotically independent samples of extremes

Armelle Guillou[(1)], **Simone A. Padoan**[(2)] **and Stefano Rizzelli**[(2)]

*

[(1)] *Institut Recherche Mathématique Avancée, UMR 7501*
Université de Strasbourg et CNRS
7 rue René Descartes
67084 Strasbourg cedex
France
e-mail: armelle.guillou@math.unistra.fr

[(2)] *Department of Decision Sciences,*
Bocconi University of Milan
via Roentgen 1
20136 Milano, Italy
e-mail: simone.padoan@unibocconi.it
stefano.rizzelli@phd.unibocconi.it

Abstract: An important topic of the multivariate extreme-value theory is to develop probabilistic models and statistical methods to describe and measure the strength of dependence among extreme observations. The theory is well established for data whose dependence structure is compatible with that of asymptotically dependent models. On the contrary, in many applications data do not comply with asymptotically dependent models and thus new tools are required. This article contributes to the methodological development of such a context, by considering a componentwise maxima approach. First we propose a statistical test based on the classical Pickands dependence function to verify whether asymptotic dependence or independence holds. Then, we present a new Pickands dependence function to describe the extremal dependence under asymptotic independence. Finally, we propose an estimator of the latter, we establish its main asymptotic properties and we illustrate its performance by a simulation study.

MSC 2010 subject classifications: Primary 62G32, 62G05, 62G20; secondary 60F05, 60G70.
Keywords and phrases: Extremal dependence, extreme-value copula, nonparametric estimation, Pickands dependence function.

1. Introduction

Multivariate extreme-value theory provides the mathematical foundation for performing real data analysis of rare events. To characterize the joint upper tail of a multivariate distribution, two different approaches can be used: either by considering the componentwise maxima, or all the observations above a high threshold. A description of these methodologies can be found for instance in Coles (2001, Ch. 8), Beirlant et al. (2004, Ch. 8-9), de Haan and Ferreira (2006, Ch. 6) and Resnick (2007, Ch. 6), among others. Unfortunately the flexibility of the dependence structures provided by the classical theory of multivariate extreme-values may not be sufficient for statistical modelling (see e.g. Ledford and Tawn, 1996, 1997). To solve this issue, different coefficients of tail dependence or probabilistic models have been introduced. They allow to govern/describe the strength of the extremal dependence. In this paper, we are particularly interested in the notion of asymptotic independence which is common in real data analysis. This concept can be defined as follows.

Let $\boldsymbol{Y}$ be a multivariate random vector of dimension d, with distribution function F and marginals F_j, $1 \leq j \leq d$. We say that F is in the max-domain of attraction of a multivariate extreme-value distribution G, if there exist sequences of constants $\mathbf{a}_n > \mathbf{0}$ and $\mathbf{b}_n \in \mathcal{R}^d$ such that

$$\lim_{n\to\infty} F^n(\mathbf{a}_n \boldsymbol{y} + \mathbf{b}_n) = G(\boldsymbol{y}),$$

for all $\boldsymbol{y} \in \mathcal{R}^d$. Under this condition, a particular case arises when G is equal to the product of its marginal distributions. In this setting, we say that $\boldsymbol{Y}$ satisfies the property of asymptotic independence (or tail independence)

*The first author is supported by a research grant (VKR023480) from VILLUM FONDEN

which is equivalent to say that the elements of $\boldsymbol{Y}$ are asymptotically independent in the upper tail, i.e.

$$\lim_{u\to 1} \Pr(F_j(Y_j) > u | F_i(Y_i) > u) = 0$$

for all $1 \le i \ne j \le d$. On the contrary, if the above limit is positive, then the elements of $\boldsymbol{Y}$ are called asymptotically dependent. The classical theory expects asymptotic dependence and independence as the only two possible scenarios, conceiving extremes as independent in the second case. Many efforts have been made to characterize a residual tail dependence in the data (if there is any) by offering new dependence coefficients or probabilistic and statistical models under asymptotic independence, see Ledford and Tawn (1996), Coles (2001, Ch 8.4), Resnick (2002), Maulik and Resnick (2004), Ramos and Ledford (2009, 2011), Wadsworth and Tawn (2013) and Wadsworth et al. (2017), to name a few. If we restrict our framework to the dimension $d = 2$, several statistical tests for checking asymptotic independence or tail independence have been proposed, among them, Ledford and Tawn (1996), Draisma et al. (2004), Hüsler and Li (2009) and Falk et al. (2010, Ch. 6.5) and the references therein. However, the extension to dimensions higher than 2 are still in its infancy. Recent proposals are based on the kth largest order statistics of the sample. Although these approaches are simple to implement, the performance of the resulting tests depends strongly on the choice of k, see e.g. Kiriliouk et al. (2016).

In this paper, we propose in Section 2 an alternative approach to test asymptotic independence for an arbitrary dimension $d \ge 2$, based on the componentwise maxima. We illustrate the performance of our proposal up to dimension $d = 4$. Then, using again the componentwise maxima approach and in particular the framework proposed by Ramos and Ledford (2011), we introduce, in Section 3, a new dependence function similar to the well-known Pickands dependence function which allows us to measure the residual dependence under asymptotic independence. Finally, we estimate this new dependence function and we establish the main asymptotic properties of the estimator. By means of a simulation study, its good performance is highlighted. A discussion on the methodological assumptions and possible extensions of our work ends the paper. All the proofs are postponed to the appendix.

Throughout the paper, the following notations are used. For any arbitrary dimension d and $f : \mathcal{X} \subset \mathcal{R}^d \to \mathcal{R}$, we set $\|f\|_\infty = \sup_{\boldsymbol{x}\in\mathcal{X}} f(\boldsymbol{x})$. We denote by $\ell^\infty(\mathcal{X})$ the space of all bounded real-valued functions on $\mathcal{X}$. The symbol "$\rightsquigarrow$" stands for convergence in distribution of random vectors, but also for weak convergence of bounded real-valued functions in $\ell^\infty(\mathcal{X})$, the difference will be clear from the context.

2. A test for asymptotic independence

A d-dimensional random vector $\boldsymbol{X} = (X_1, \dots, X_d)$ follows the law of a multivariate extreme-value distribution if the one-dimensional marginal distributions, $G_j(x) = \Pr(X_j \le x)$ for all $x \in \mathcal{R}$, $j = 1, \dots, d$, are Generalized Extreme-Value (GEV) distributions, and the joint distribution takes the form

$$G(\boldsymbol{x}) = C(G_1(x_1), \dots, G_d(x_d)), \qquad \boldsymbol{x} \in \mathcal{R}^d,$$

where C is an extreme-value copula, i.e.,

$$C(\boldsymbol{u}) = \exp\left(-V\left((-\log u_1)^{-1}, \dots, (-\log u_d)^{-1}\right)\right), \qquad \boldsymbol{u} \in (0,1]^d,$$

with $V : (0,\infty]^d \to [0,\infty)$ (see de Haan and Ferreira, 2006, Ch. 1, 6, for details). Consider the map $L : [0,\infty)^d \mapsto [0,\infty)$, defined by $L(\boldsymbol{z}) := V(1/\boldsymbol{z})$ with $\boldsymbol{z} = 1/\boldsymbol{y}$ for $\boldsymbol{y} \in (0,\infty]^d$. The function L is known as the stable tail dependence function. As it is a homogeneous function of order one, i.e. $L(a\boldsymbol{z}) = aL(\boldsymbol{z})$ for all $a > 0$, we have

$$L(\boldsymbol{z}) = (z_1 + \dots + z_d)\, A(\boldsymbol{t}), \qquad \boldsymbol{z} \in [0,\infty)^d,$$

with $t_j = z_j/(z_1 + \dots + z_d)$ for $j = 2, \dots, d$, $t_1 = 1 - t_2 - \dots - t_d$, and A is the restriction of L into the d-dimensional unit simplex,

$$\mathcal{S}_d := \left\{(v_1, \dots, v_d) \in [0,1]^d : v_1 + \dots + v_d = 1\right\}.$$

The function A is well-known as the Pickands dependence function (see Pickands, 1981), and is often used to quantify the dependence among the elements of $\boldsymbol{X}$. Indeed, A satisfies the constraint $1/d \le \max(t_1, \dots, t_d) \le$

$A(\boldsymbol{t}) \leq 1$ for all $\boldsymbol{t} \in \mathcal{S}_d$, with lower and upper bounds corresponding to the complete dependence and independence cases, respectively (see Falk et al., 2010, Ch. 4, for details). Thus, estimating this Pickands dependence function is crucial for analysing multivariate extremes, and it has been an extensively discussed topic in the literature, see Klüppelberg and May (2006), Zhang et al. (2008), Gudendorf and Segers (2011), Capéraà et al. (1997), Berghaus et al. (2013) or Vettori et al. (2017), among others.

2.1. A slightly modified version of the Pickands dependence estimator proposed by Marcon et al. (2017)

This estimator is based on the madogram concept, a notion borrowed from geostatistics in order to capture the spatial structure. Starting from independent and identically distributed (i.i.d.) copies $\boldsymbol{X}_1$, ..., $\boldsymbol{X}_n$, of $\boldsymbol{X}$, our estimator is defined as

$$\widehat{A}_n(\boldsymbol{t}) := \frac{\widehat{\nu}_n(\boldsymbol{t}) + c(\boldsymbol{t})}{1 - \widehat{\nu}_n(\boldsymbol{t}) - c(\boldsymbol{t})} \tag{2.1}$$

where

$$\begin{aligned} \widehat{\nu}_n(\boldsymbol{t}) &:= \frac{1}{n}\sum_{i=1}^{n}\left(\bigvee_{j=1}^{d}\left\{G_{n,j}^{(1)}(X_{i,j})\right\}^{1/t_j} - \frac{1}{d}\sum_{j=1}^{d}\left\{G_{n,j}^{(1)}(X_{i,j})\right\}^{1/t_j}\right) \qquad (2.2)\\ c(\boldsymbol{t}) &:= \frac{1}{d}\sum_{j=1}^{d}\frac{t_j}{1+t_j} \end{aligned}$$

with

$$G_{n,j}^{(a)}(X_{i,j}) := G_{n,j}(X_{i,j})\left(\frac{1+a}{a}\frac{1}{n}\sum_{k=1}^{n}G_{n,j}^{1/a}(X_{k,j})\right)^{-a}, \quad j = 1, \ldots, d, \text{ for } a > 0,$$

and the empirical distribution functions denoted by

$$G_{n,j}(x) := \frac{1}{n}\sum_{i=1}^{n}\mathbb{1}_{\{X_{i,j}\leq x\}}, \quad j = 1, \ldots, d.$$

By convention, here $u^{1/0} = 0$ for $0 < u < 1$. The estimator (2.2) is a slightly modified version of that proposed in Marcon et al. (2017), with $G_{n,j}^{(1)}$ in place of $G_{n,j}$ which ensures that the new Pickands estimator $\widehat{A}_n$ now satisfies $\widehat{A}_n(\boldsymbol{e}_j) = 1$ for all $j = 1, \ldots, d$, where $\boldsymbol{e}_j = (0, \ldots, 0, 1, 0, \ldots, 0)$ is the jth canonical unit vector (see Appendix A.1). This is a necessary condition that a function needs to satisfy in order to be a valid Pickands dependence function (see e.g. Marcon et al., 2017). Although, as established in Appendix A.1, our modified estimator shares the same asymptotic properties as the estimator discussed in Marcon et al. (2017), our modification greatly improves the latter for finite samples.

2.2. Construction of our statistical test

Using our estimator for A, we want now to construct a statistical test to check asymptotic independence in dimensions higher than or equal to two. To this aim, we consider the following system of hypotheses

$$\begin{cases} H_0: & A(\boldsymbol{t}) = 1, \forall\, \boldsymbol{t} \in \mathcal{S}_d \\ H_1: & A(\boldsymbol{t}) < 1, \text{ for some } \boldsymbol{t} \in \mathcal{S}_d. \end{cases}$$

Note that H_0 means that all the components of $\boldsymbol{X}$ are asymptotically independent, whereas under H_1 some elements of $\boldsymbol{X}$ are asymptotically dependent.

Assuming that the extreme-value copula C has continuous partial derivatives over the sets $\{\boldsymbol{u} \in [0,1]^d : 0 < u_i < 1\}$, by Theorem 2.4 in Marcon et al. (2017) and according to Appendix A.1, under H_0 we have that

$$\sqrt{n}\left(\widehat{A}_n(\boldsymbol{t}) - 1\right)_{\boldsymbol{t}\in\mathcal{S}_d} \rightsquigarrow -4\left(\int_0^1 \mathbb{A}(v^{t_1}, \ldots, v^{t_d})\mathrm{d}v\right)_{\boldsymbol{t}\in\mathcal{S}_d}, \quad \text{as } n \to \infty, \tag{2.3}$$

where $\mathbb{A}$ is a centered Gaussian process on $[0,1]^d$ with continuous sample paths and covariance function equal to

$$\mathrm{Cov}(\mathbb{A}(\boldsymbol{u}), \mathbb{A}(\boldsymbol{v})) = \prod_{j=1}^{d} u_j \wedge v_j - \sum_{j=1}^{d} \left(u_j \wedge v_j \prod_{i \neq j} u_i v_i \right) + (d-1) \prod_{j=1}^{d} u_j v_j.$$

As a consequence, by the continuous mapping theorem (see e.g. van der Vaart, 2000, Ch. 2.1) it follows that

$$\widehat{S}_n := \sup_{\boldsymbol{t} \in \mathcal{S}_d} \sqrt{n} \left| \widehat{A}_n(\boldsymbol{t}) - 1 \right| \rightsquigarrow S := \sup_{\boldsymbol{t} \in \mathcal{S}_d} 4 \left| \int_0^1 \mathbb{A}(v^{t_1}, \ldots, v^{t_d}) \, \mathrm{d}v \right|, \quad \text{as } n \to \infty.$$

This convergence result can be used as the cornerstone of our test. Denoting by $Q_S(\alpha)$, $\alpha \in (0,1)$, the $(1-\alpha)$-quantile function for the distribution of the random variable S, H_0 can be rejected with approximately $\alpha\%$ significance level whenever $\widehat{s}_n$, the observed value of $\widehat{S}_n$, exceeds $Q_S(\alpha)$. Unfortunately, there is no closed form for the function $Q_S(\alpha)$, however an approximation can still be computed with a Monte Carlo simulation as follows. Note that for any $u, v \in [0,1]$ and $\boldsymbol{t}, \boldsymbol{w} \in \mathcal{S}_d$, the covariance function of the Gaussian process $\mathbb{A}$ in (2.3), evaluated at the indexes $u^{\boldsymbol{t}}, v^{\boldsymbol{w}} \in [0,1]^d$, is equal to

$$\mathrm{Cov}(\mathbb{A}(v^{t_1}, \ldots, v^{t_d}), \mathbb{A}(u^{w_1}, \ldots, u^{w_d})) = \prod_{j=1}^{d} (v^{t_j} \wedge u^{w_j}) - \sum_{j=1}^{d} (v^{t_j} \wedge u^{w_j}) v^{1-t_j} u^{1-w_j} + (d-1) uv. \tag{2.4}$$

Thus, for any fixed $\alpha \in (0,1)$, an approximation of the quantile $Q_S(\alpha)$ can be obtained by adhering to the following four steps:

1. Divide the unit interval $(0,1)$ and the simplex $\mathcal{S}_d$ in p and m equally spaced points, where p and m are positive integers. Let $v_1, \ldots, v_m$ and $\boldsymbol{t}_1, \ldots, \boldsymbol{t}_p$ be the two sequences of points partitioning $(0,1)$ and $\mathcal{S}_d$, respectively. The sequences $v_1, \ldots, v_m$ and $\boldsymbol{t}_1, \ldots, \boldsymbol{t}_p$ form a finite sequence of positions $v_r^{t_{k,1}}, \ldots, v_r^{t_{k,d}} \in [0,1]^d$, with $r = 1, \ldots, m$ and $k = 1, \ldots, p$, on which the process $\mathbb{A}$ is simulated.
2. Sample n^* realizations

$$x_i(v_1^{t_{1,1}}, \ldots, v_1^{t_{1,d}}), \ldots, x_i(v_m^{t_{p,1}}, \ldots, v_m^{t_{p,d}}), \quad i = 1, \ldots, n^*,$$

   of a zero-mean Gaussian process at $v_r^{t_{k,1}}, \ldots, v_r^{t_{k,d}}$, for $r = 1, \ldots, m$ and $k = 1, \ldots, p$, with a $(mp \times mp)$ variance-covariance matrix defined through the covariance function in (2.4).
3. Simulate samples that approximately follow the distribution of the random variable S, the integral and the sup in S being approximated by a sum and the max for sufficiently large values of m and p. This leads to the realizations

$$\widetilde{s}_i = \max_{1 \le k \le p} 4 \left| \frac{1}{m} \sum_{r=1}^{m} x_i(v_r^{t_k,1}, \ldots, v_r^{t_k,d}) \right|, \quad i = 1, \ldots, n^*.$$

4. An approximation of the quantile $Q_S(\alpha)$, denoted by $\widetilde{Q}_S(\alpha)$, can then be obtained by computing the sample quantile of the realizations $\widetilde{s}_1, \ldots, \widetilde{s}_{n^*}$ for sufficiently large n^*.

2.3. Numerical results

We illustrate the performance of our statistical test through a simulation study. Precisely, we estimate some values of the significance level α and the power $1-\beta$ of the test by computing the empirical proportion of simulated samples under the null hypothesis and the alternative hypothesis that rejected the null hypothesis, respectively. For simplicity we focus on the significance levels $\alpha = 0.05$ and 0.01. The study consists of five experiments.

First experiment: In order to perform the first experiment, in a first step we compute the approximated quantile $\widetilde{Q}_S(\alpha)$, for a given α, following our algorithm. The quality of approximation relies on the values of the indexes m, p and n^*. Clearly, the larger their values are, the more accurate the approximation is. We set $n^* = 500\,000$. We consider increasing values of m and p and for each combination we compute $\widetilde{Q}_S$. We stop the search of a

TABLE 1

Estimated significance levels α. From left to right: the dimension of $\boldsymbol{X}$*, the true significance level, the approximate asymptotic* $(1-\alpha)$*-quantile, the empirical proportion of simulated samples under* H_0 *that rejected the null hypothesis and the empirical* $(1-\alpha)$*-quantile. Here* $\psi = 1$.

			$\hat{\alpha}$				$Q_{\widehat{S}_n}(\alpha)$			
			n				n			
d	α	$\widetilde{Q}_S(\alpha)$	25	50	100	200	25	50	100	200
2	0.05	0.960	0.0380	0.0460	0.0512	0.0524	0.9190	0.9393	0.9512	0.9541
	0.01	1.204	0.0060	0.0082	0.0102	0.0102	1.1359	1.1739	1.1926	1.1992
3	0.05	1.300	0.0364	0.0452	0.0508	0.0574	1.2540	1.2755	1.3036	1.3210
	0.01	1.540	0.0056	0.0068	0.0084	0.0092	1.4126	1.4904	1.5295	1.5601
4	0.05	1.480	0.0398	0.0454	0.0548	0.0576	1.5312	1.5508	1.5883	1.5745
	0.01	1.740	0.0064	0.0082	0.0096	0.0126	1.7715	1.7135	1.7849	1.7867

better value for these indexes when the value of $\widetilde{Q}_S(\alpha)$ does not increase anymore, up to the second decimal. The calculation of $\widetilde{Q}_S$ requires a considerable computational effort, therefore we derive its values only for a dimension $d = 2, 3, 4$ of the vector $\boldsymbol{X}$.
In a second step, we compute the rejection rates. We focus on the multivariate logistic extreme-value model introduced by Tawn (1990), with dependence parameter $\psi \in (0, 1]$, $\psi = 1$ corresponding to independent components of $\boldsymbol{X}$, whereas complete dependence can be reached when $\psi \to 0$. We consider 20 equally spaced values of ψ in $(0, 1]$. For each of them, we simulate n independent observations from a logistic extreme-value distribution with unit Fréchet margins. Then we estimate the Pickands dependence function by (2.1) and we compute $\widehat{s}_n$. We repeat this task 5000 times and we compute the proportion of times that $\widehat{s}_n > \widetilde{Q}_S(\alpha)$. This experiment is repeated for different values of the samples size n and different dimension d of $\boldsymbol{X}$. The middle part of Table 1 reports the estimated values of the significance levels α in the case where $\psi = 1$.
We see that accurate estimates of α are already obtained with the sample size $n = 50$, indicating a good performance of our statistical test. Figure 1 displays the estimated powers of the test. In the first and second rows the results obtained with $\alpha = 0.05$ and $\alpha = 0.01$ are reported, respectively. The panels from left to right illustrate the results for the dimensions 2, 3 and 4. Once again, the test shows a good performance already with the sample size $n = 50$. Indeed in the case $d = 2$ we see that the power of the test reaches 1 with mild dependence levels, i.e. $\psi = 0.5$. This figure also outlines that the power of the test improves as the dimension of $\boldsymbol{X}$ increases and that, as expected, for any fixed dimension $d = 2, 3, 4$, it also improves as the sample size increases.

Second experiment: We repeat the second step of the first experiment approximating the test statistic's distribution, under the null hypothesis, via a Monte Carlo approach. Precisely, we simulate n values from d independent univariate Fréchet distributions, then we estimate the Pickands dependence function by (2.1) and we compute $\widehat{s}_n$. We repeat this task 5000 times and we compute the empirical quantile, for a given α, denoted by $Q_{\widehat{S}_n}(\alpha)$. The right-hand side of Table 1 reports the values of them for different values of n and d. We see that the empirical quantiles rapidly approach the asymptotic quantiles, as the sample size increases. Already with the sample size $n = 50$ the two types of quantiles are very close. The third and fourth lines of Figure 1 display the comparison between the estimates of $1-\beta$ obtained with the two types of quantiles, but also the estimates of α since $\psi = 1$ corresponds to independent components and thus the proportion of rejection reported by the figures represents an approximation of α in that case. Since the conclusions are the same for both values $\alpha = 0.01$ and $\alpha = 0.05$, only the latter are reported. It follows that the inferential results obtained with the empirical quantiles are very close to those obtained with the asymptotic quantiles, already with the sample size $n = 50$.

Third experiment: We repeat the second experiment using the Genest-Rémillard (GR) rank-based statistical test (Genest and Rémillard, 2004) and our proposed test with the Capéraà-Fougères-Genest (CFG) estimator (Capéraà et al., 1997; Zhang et al., 2008) of the Pickands dependence function in place of (2.1). Figure 2 shows the estimated powers obtained with the GR test and our test (with both the CFG and the Madogram-based (2.1) estimators). For brevity, we show the results for $\alpha = 0.05$ and the sample sizes $n = 25$ and $n = 50$. We see that our test always outperforms the GR test, with the best results provided with the CFG estimator in the case of a dimension equal to 2, whereas, in higher dimensions, similar results can be reached with either the Madogram-based estimator or the CFG.

Fourth experiment: We repeat the third experiment by sampling from three alternative distributions. In the first case, we draw samples from a three-dimensional random vector with a pair that follows the logistic extreme-value distribution and where the last variable is independent from the other two. In the second case, we consider a four-dimensional random vector with two pairs that follow the logistic extreme-value distribution and where the components of one pair are independent from each component of the second pair. In the last case, we consider a four-dimensional random vector, where one pair with independent components and each of these independent from the components of the other pair that follows the logistic extreme-value distribution. The results are collected in the third (n = 25) and fourth row (n = 50) of Figure 2. In these cases we see that our test loses power and provides inferential results very similar to those provided by the GR test, however the latter outperforms our test in the case of the largest number of independent variables.

Fifth experiment: We consider the multivariate inverted symmetric logistic model (see e.g. Ledford and Tawn, 1997; Wadsworth et al., 2017), with dependence parameter $\psi \in (0, 1]$, $\psi = 1$ corresponding to exact independence of the components of $\boldsymbol{X}$, whereas asymptotic dependence is reached as $\psi \to 0$. This time, we consider 10 equally spaced values of ψ in $(0, 1]$. For each of them, we simulate 366 values (for similarity with annual maxima) from an inverted logistic distribution with exponential margins. Then, we compute the normalized componentwise maxima and we repeat this procedure in order to obtain n normalized maxima from which we estimate the Pickands dependence function using (2.1) and we calculate $\widehat{s}_n$. We repeat this task 5000 times and we compute the proportion of times that $\widehat{s}_n > Q_{S_n}(0.05)$. This procedure has been done for different values of d and n and the results are summarized in Table 2.

TABLE 2

Estimated significance levels α. From left to right: the dimension of $\boldsymbol{X}$, the sample size and the empirical proportion of simulated samples under H_0 that rejected the null hypothesis for different values of ψ.

		ψ									
d	n	1	0.9	0.8	0.7	0.6	0.5	0.4	0.3	0.2	0.1
2	25	0.0522	0.0468	0.0578	0.0580	0.0584	0.0842	0.1294	0.2730	0.6372	0.9920
	50	0.0554	0.0524	0.0542	0.0562	0.0590	0.0980	0.1772	0.4538	0.8932	1.0000
	100	0.0476	0.0470	0.0506	0.0561	0.0594	0.1208	0.2938	0.7128	0.9962	1.0000
	200	0.0486	0.0570	0.0568	0.0560	0.0601	0.1856	0.4934	0.9482	1.0000	1.0000
3	25	0.0470	0.0508	0.0522	0.0582	0.0808	0.1076	0.2190	0.4234	0.8604	0.9990
	50	0.0542	0.0538	0.0528	0.0589	0.8320	0.1604	0.3606	0.7424	0.9950	1.0000
	100	0.0536	0.0468	0.0550	0.0594	0.0922	0.2084	0.5274	0.9486	1.0000	1.0000
	200	0.0540	0.0500	0.0506	0.0652	0.1242	0.3050	0.8096	0.9996	1.0000	1.0000
4	25	0.0488	0.0418	0.0448	0.0504	0.0692	0.1336	0.2676	0.6332	0.9582	1.0000
	50	0.0452	0.0438	0.0536	0.0574	0.1018	0.1854	0.4736	0.8852	0.9996	1.0000
	100	0.0496	0.0484	0.0468	0.0598	0.1130	0.2746	0.7052	0.9920	1.0000	1.0000
	200	0.0486	0.0448	0.0560	0.0770	0.1596	0.4768	0.9440	1.0000	1.0000	1.0000

With $d = 2$, the rejection rates are close to 0.05 whenever ψ is larger than 0.5. Otherwise, the rejection rate is greater than 0.05 and it reaches 1 when ψ approaches 0. In these cases, it can be observed that the normalized maxima show quite a strong dependence, which indeed seems that of an asymptotically dependent model rather than an asymptotically independent one. The strength of the dependence is reduced whenever the normalized maxima are computed on sequences larger than 366, resulting in improvements in the performances of our test. The test performance deteriorates as the dimension of $\boldsymbol{X}$ increases.

In conclusion this study highlights the good performance of our statistical test in detecting the exact independence of sample maxima. However, our test is also useful to detect if multivariate data are asymptotically independent as long as there is a weak dependence within the case of asymptotic independence. On the contrary, in the cases of strong dependence within asymptotic independence our test fails in detecting the data as asymptotically independent. Clearly, these are the most naturally difficult cases to detect and more specific tools are required.

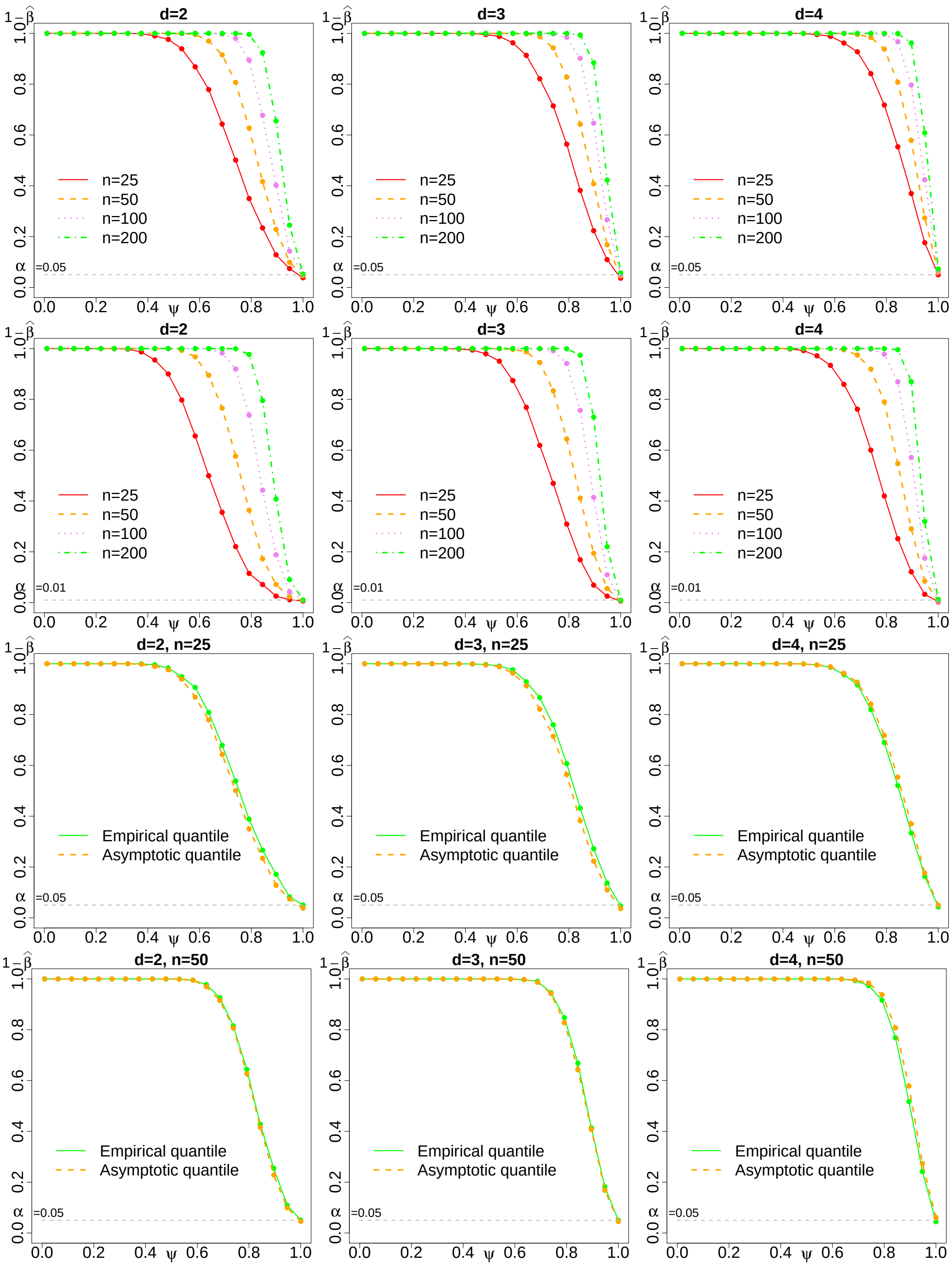

FIG 1. *Estimated power functions. Points report the empirical proportion of simulated samples under H_1 that rejected H_0 as a function of ψ.*

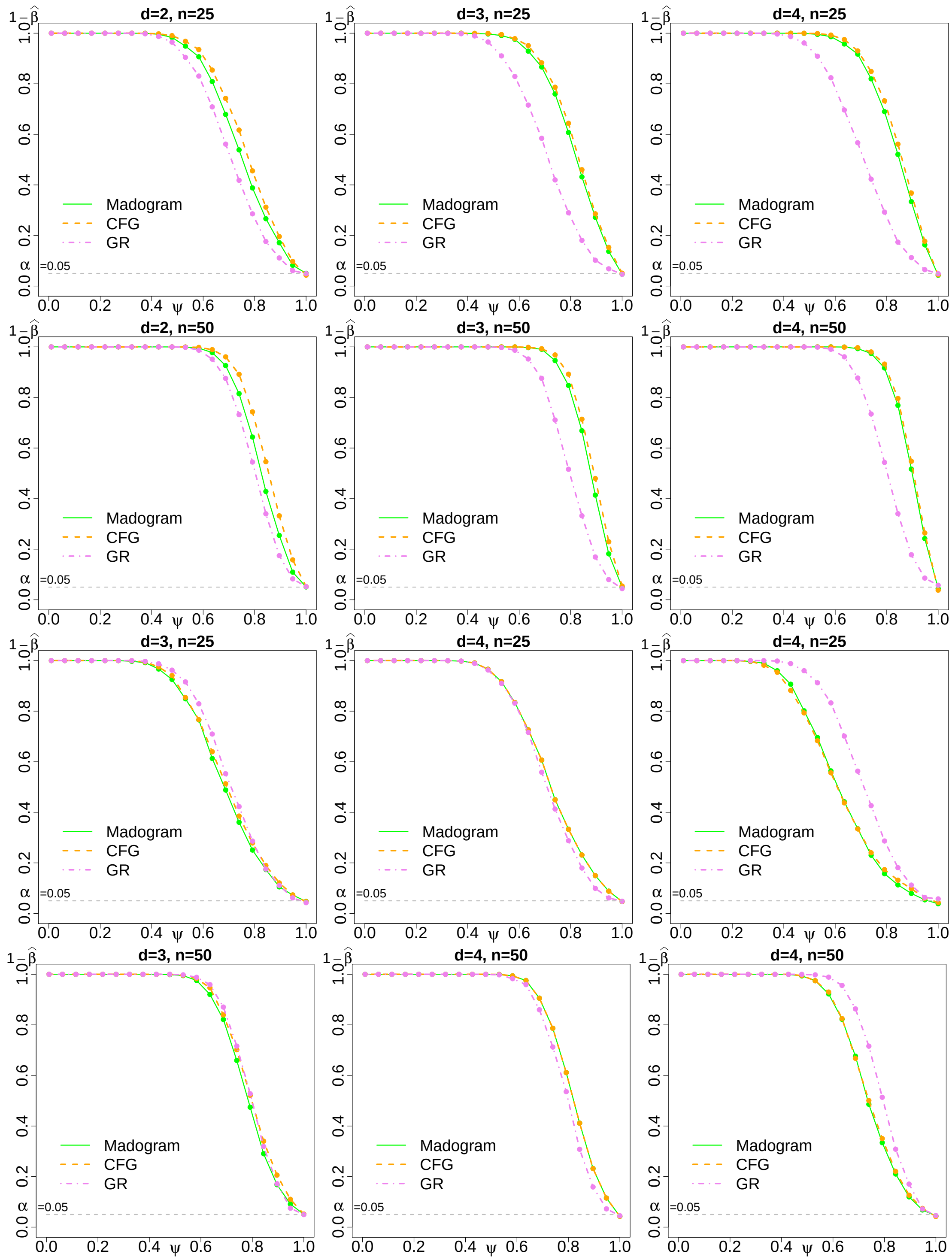

FIG 2. *Estimated power functions. Points are the empirical proportion of simulated samples under H_1 that rejected H_0 as a function of ψ.*

3. Asymptotic independence for componentwise maxima

Being able to test asymptotic independence versus asymptotic dependence is obviously important. However, since asymptotic independence often arises in applications, it is also crucial to develop some general models that accommodate both situations. In this section, we consider the framework of Ramos and Ledford (2009) (see also Ledford and Tawn, 1997). More precisely, if $\boldsymbol{Y}$ is a d-dimensional random vector with common unit Fréchet margins, i.e. $\Pr(Y_j \leq y) = e^{-1/y}$ for every $y > 0$ and $1 \leq j \leq d$, this theory relies on the joint survival function of $\boldsymbol{Y}$ which is assumed to be multivariate regularly varying with index $-1/\eta$, where $\eta \in (0, 1]$, i.e. $\Pr(\boldsymbol{Y} > \boldsymbol{y}) = \tau(\boldsymbol{y})(y_1 \cdot \dots \cdot y_d)^{-1/d\eta}$ with τ a slowly varying function satisfying

$$\lim_{r\to\infty} \frac{\tau(r\,y_1, \dots, r\,y_d)}{\tau(r, \dots, r)} = g(\boldsymbol{y})$$

for all $\boldsymbol{y} \in (0, \infty]^d$. The function g here is homogeneous of order 0, i.e. such that $g(a\,y_1, \dots, a\,y_d) = g(y_1, \dots, y_d)$ for any $a > 0$. This framework implies that the conditional joint survival function admits the following limit representation for every $\boldsymbol{y} \geq \boldsymbol{1}$, the vector of ones,

$$\lim_{u\to\infty} \Pr(\boldsymbol{Y} > u\boldsymbol{y} | \boldsymbol{Y} > u\boldsymbol{1}) = \lim_{u\to\infty} \frac{\Pr(\boldsymbol{Y} > u\boldsymbol{y})}{\Pr(\boldsymbol{Y} > u\boldsymbol{1})} = \eta \int_{S_d} \bigwedge_{j=1}^{d} \left(\frac{w_j}{y_j}\right)^{1/\eta} \mathrm{d}H_\eta(\boldsymbol{w}), \tag{3.1}$$

where H_η is a non-negative measure satisfying the condition

$$\eta \int_{S_d} \bigwedge_{j=1}^{d} w_j^{1/\eta} \mathrm{d}H_\eta(\boldsymbol{w}) = 1. \tag{3.2}$$

This measure H_η is a particular case of the hidden angular measure introduced by Resnick (2002) (see also Maulik and Resnick, 2004) when $\eta < 1$ and it is a rescaled version of the classical angular measure when $\eta = 1$, see Ramos and Ledford (2009) for details. According to Ramos and Ledford (2011) we assume that H_η is a finite measure. We recall that η is the so-called coefficient of tail dependence, which measures the level of dependence within the asymptotic independence framework. Specifically, $\eta = 1$ corresponds to the case of asymptotic dependence, whereas $\eta < 1$ corresponds to the case of asymptotic independence. More precisely, when the coefficient η falls in the following sets: $(1/d, 1)$, $\{1/d\}$ or $(0, 1/d)$, then we say that among the variables there is a positive association, independence or negative association, respectively, within asymptotic independence (see e.g. Ledford and Tawn, 1996).

3.1. A $\eta-$Pickands dependence function

Consider now, n i.i.d. copies $\boldsymbol{Y}_1, \dots, \boldsymbol{Y}_n$ of $\boldsymbol{Y}$ and for a small $\varepsilon > 0$, define $\boldsymbol{M}_{n,\varepsilon} = (M_{n,1,\varepsilon}, \dots, M_{n,d,\varepsilon})$ as the vector of componentwise maxima, precisely

$$M_{n,j,\varepsilon} = \bigvee_{i \in I_n(\varepsilon)} Y_{i,j}, \quad j = 1, \dots, d,$$

with $I_n(\varepsilon) := \{1 \leq i \leq n : \boldsymbol{Y}_i > \boldsymbol{1}\varepsilon\}$. Let b_n be a sequence of normalizing constants defined by the equation $n\Pr(\boldsymbol{Y} > b_n) = 1$. Then, differently from the classical theory (e.g. de Haan and Ferreira, 2006, Ch. 6), here the limiting distribution for the normalized vector of componentwise maxima $\boldsymbol{M}_{n,\varepsilon}$ is obtained as

$$G_\eta(\boldsymbol{y}) := \lim_{\varepsilon\to 0} \lim_{n\to\infty} \Pr(\boldsymbol{M}_{n,b_n\varepsilon} \leq b_n\boldsymbol{y}), \quad \boldsymbol{y} \in (0, \infty]^d, \tag{3.3}$$

see Ramos and Ledford (2011) for details. When a limiting distribution exists with nondegenerate margins, then G_η is called a multivariate η-extreme-value distribution. Specifically, a d-dimensional random vector $\boldsymbol{Z}$ follows the law of a multivariate η-extreme-value distribution, if the one-dimensional marginal distributions are $G_{\eta,j}(y) = \exp(-\sigma_{\eta,j} y^{-1/\eta})$, for all $y > 0$, $j = 1, \dots, d$, and the joint distribution takes the form

$$G_\eta(\boldsymbol{y}) = C_\eta(G_{\eta,1}(y_1), \dots, G_{\eta,d}(y_d)), \qquad \boldsymbol{y} \in (0, \infty]^d, \tag{3.4}$$

where C_η is an η-extreme-value copula, i.e.

$$C_\eta(\boldsymbol{u}) = \exp\left\{-V_\eta\left(\left(\frac{\sigma_{\eta,1}}{-\log u_1}\right)^\eta, \dots, \left(\frac{\sigma_{\eta,d}}{-\log u_d}\right)^\eta\right)\right\} \qquad \boldsymbol{u} \in (0,1]^d$$

with $V_\eta : (0,\infty]^d \to [0,\infty)$ a homogeneous function of order $-1/\eta$ and

$$\sigma_{\eta,j} := V_\eta(\infty, \dots, \infty, 1, \infty, \dots, \infty) = \eta \int_{\mathcal{S}_d} w_j^{1/\eta} \mathrm{d}H_\eta(\boldsymbol{w}). \tag{3.5}$$

Introduce now $L_\eta(\boldsymbol{z}) := V_\eta((\boldsymbol{\sigma}_\eta/\boldsymbol{z})^\eta)$, for all $\boldsymbol{z} = \boldsymbol{\sigma}_\eta/\boldsymbol{y}^{1/\eta}$. This function is called the η-stable tail dependence function and using the homogeneity property, it can be rewritten as

$$L_\eta(\boldsymbol{z}) = (z_1 + \cdots + z_d) A_\eta(\boldsymbol{t}), \qquad \boldsymbol{z} \in [0,\infty)^d,$$

where $t_j = z_j/(z_1 + \cdots + z_d)$ for $j = 2, \dots, d$, $t_1 = 1 - t_2 - \cdots - t_d$. Here, the function A_η is called the η-Pickands dependence function and it satisfies the following properties.

Proposition 3.1. *The η-Pickands dependence function A_η satisfies:*

1. *For all $\eta \in (0,1]$, $A_\eta(\boldsymbol{e}_j) = 1$, $j = 1, ..., d$;*
2. *$A_1(\boldsymbol{t}) = A(\boldsymbol{t})$, for all $\boldsymbol{t} \in \mathcal{S}_d$;*
3. *For every $\eta \in (0,1]$ and $\boldsymbol{t} \in \mathcal{S}_d$,*
$$1/d \le \max(t_1, \dots, t_d) \le A_\eta(\boldsymbol{t}) \le 1.$$
4. *$A_\eta(\boldsymbol{t})$ is convex, i.e. $A_\eta(a\boldsymbol{t}_1 + (1-a)\boldsymbol{t}_2) \le aA_\eta(\boldsymbol{t}_1) + (1-a)A_\eta(\boldsymbol{t}_2)$, for all $a \in [0,1]$ and $\boldsymbol{t}_1, \boldsymbol{t}_2 \in \mathcal{S}_d$.*

Similarly to the classical literature, a $\eta-$madogram function can be defined as the expected distance between the maximum and the mean of the variables $G_{\eta,1}^{1/\eta t_1}(Z_1), \dots, G_{\eta,d}^{1/\eta t_d}(Z_d)$, that is,

$$\nu_\eta(\boldsymbol{t}) = \mathrm{E}\left[\bigvee_{j=1}^d \left\{G_{\eta,j}^{1/\eta t_j}(Z_j)\right\} - \frac{1}{d}\sum_{j=1}^d G_{\eta,j}^{1/\eta t_j}(Z_j)\right].$$

This function can also be linked to the $\eta-$Pickands dependence function as follows.

Proposition 3.2. *Any random vector $\boldsymbol{Z}$ with a η-extreme-value distribution admits a η-Pickands dependence function A_η which satisfies*

$$A_\eta(\boldsymbol{t}) \quad = \quad \frac{1}{\eta}\frac{\nu_\eta(\boldsymbol{t}) + c_\eta(\boldsymbol{t})}{1 - \nu_\eta(\boldsymbol{t}) - c_\eta(\boldsymbol{t})} \tag{3.6}$$

for all $\boldsymbol{t} \in \mathcal{S}_d$, where

$$c_\eta(\boldsymbol{t}) = \frac{1}{d}\sum_{j=1}^d \frac{t_j}{t_j + 1/\eta}. \tag{3.7}$$

This $\eta-$Pickands dependence function can be used to represent the level of dependence among the elements of $\boldsymbol{Z}$, and thus in the next section, we estimate this function and derive the main asymptotic properties of the estimator.

3.2. An estimator of the $\eta-$Pickands dependence function

Let $\boldsymbol{Z}_1, \ldots, \boldsymbol{Z}_n$ be i.i.d. copies of $\boldsymbol{Z}$ with distribution G_η and define

$$H_n(\boldsymbol{y}) = \frac{1}{n}\sum_{i=1}^n \mathbb{1}_{\{\boldsymbol{Z}_i \le \boldsymbol{y}\}}, \quad \boldsymbol{y} \in (0,\infty]^d$$

and its associated empirical process

$$\mathbb{H}_n(\boldsymbol{y}) = \sqrt{n}(H_n(\boldsymbol{y}) - G_\eta(\boldsymbol{y})), \quad \boldsymbol{y} \in (0,\infty]^d.$$

In order to estimate the $\eta-$Pickands dependence function we first assume that we have at our disposal an estimator $\widehat{\eta}_n$ for η satisfying the condition:

Condition 1. *Let $\widehat{\eta}_n$ be an estimator of η satisfying:*

(i) $\widehat{\eta}_n \to \eta$ *a.s. as* $n \to \infty$*;*
(ii) One of the following holds true
  (a) $\sqrt{n}(\widehat{\eta}_n - \eta) = n^{-1/2}\sum_{i=1}^n \rho(\boldsymbol{Z}_i) + o_p(1)$*, where* $\rho : (0,\infty]^d \mapsto \mathcal{R}$ *is a measurable function such that* $E\rho(\boldsymbol{Z}) = 0$ *and* $E\rho^2(\boldsymbol{Z}) < \infty$*;*
  (b) $\sqrt{n}(\widehat{\eta}_n - \eta) = \chi(\mathbb{H}_n) + o_p(1)$*, where* $\chi : \ell^\infty((0,\infty]^d) \mapsto \mathcal{R}$ *is a bounded linear functional.*

In the spirit of (2.1) in Section 2, we propose the following estimator for A_η:

$$\widehat{A}_{\widehat{\eta}_n,n}(\boldsymbol{t}) := \frac{1}{\widehat{\eta}_n}\frac{\widehat{\nu}_{\widehat{\eta}_n,n}(\boldsymbol{t}) + \widehat{c}_{\widehat{\eta}_n,n}(\boldsymbol{t})}{1 - \widehat{\nu}_{\widehat{\eta}_n,n}(\boldsymbol{t}) - \widehat{c}_{\widehat{\eta}_n,n}(\boldsymbol{t})}$$

where

$$\begin{aligned} \widehat{\nu}_{\widehat{\eta}_n,n}(\boldsymbol{t}) &:= \frac{1}{n}\sum_{i=1}^n \left(\bigvee_{j=1}^d \left\{H_{n,j}^{(\widehat{\eta}_n)}(Z_{i,j})\right\}^{1/\widehat{\eta}_n t_j} - \frac{1}{d}\sum_{j=1}^d \left\{H_{n,j}^{(\widehat{\eta}_n)}\left(Z_{i,j}\right)\right\}^{1/\widehat{\eta}_n t_j}\right) \\ \widehat{c}_{\widehat{\eta}_n,n}(\boldsymbol{t}) &:= \frac{1}{n\,d}\sum_{i=1}^n\sum_{j=1}^d \left\{H_{n,j}^{(\widehat{\eta}_n)}\left(Z_{i,j}\right)\right\}^{1/\widehat{\eta}_n t_j} \end{aligned} \tag{3.8}$$

with

$$H_{n,j}^{(a)}(Z_{i,j}) = H_{n,j}(Z_{i,j})\left(\frac{1+a}{a}\frac{1}{n}\sum_{k=1}^n H_{n,j}^{1/a}(Z_{k,j})\right)^{-a}, \quad j = 1,\ldots,d, \quad \text{for } a > 0,$$

and the empirical distribution functions denoted by

$$H_{n,j}(x) := \frac{1}{n}\sum_{i=1}^n \mathbb{1}_{\{Z_{i,j}\le x\}}, \quad j = 1,\ldots,d.$$

Note that (3.8) comes from the fact that c_η defined in Proposition 3.2 can be viewed as

$$c_\eta(\boldsymbol{t}) = E\left(\frac{1}{d}\sum_{j=1}^d \left\{G_{\eta,j}(Z_j)\right\}^{1/\eta t_j}\right)$$

and thus in (3.8) we use the empirical counterpart. Another option would have been to replace η by an estimator in (3.7).
We are now able to state our main result on the convergence of a rescaled version of $\widehat{A}_{\widehat{\eta}_n,n}$.

Theorem 3.1. *Let $\mathbf{Z}_1, \ldots, \mathbf{Z}_n$ be i.i.d. copies of $\mathbf{Z}$ with distribution G_η, and $\widehat{A}_{\widehat{\eta}_n,n}$ be our proposed estimator for A_η.*

Under Condition 1(i), we have

$$\|\widehat{A}_{\widehat{\eta}_n,n} - A_\eta\|_\infty \to 0 \text{ a.s. as } n \to \infty. \tag{3.9}$$

Under Conditions 1(ii), we have in $\ell^\infty(\mathcal{S}_d)$, as $n \to \infty$,

$$\sqrt{n}(\widehat{A}_{\widehat{\eta}_n,n}(t) - A_\eta(t))_{t\in\mathcal{S}_d} \rightsquigarrow \left(-\frac{(1+\eta A_\eta(\boldsymbol{t}))^2}{\eta}\int_0^1 \mathbb{A}_\eta(v^{\eta t_1}, \ldots, v^{\eta t_d})\,\mathrm{d}v\right)_{t\in\mathcal{S}_d}, \tag{3.10}$$

where $\mathbb{A}_\eta$ is a stochastic process defined as

$$\mathbb{A}_\eta(\boldsymbol{u}) := \mathbb{B}_\eta(\boldsymbol{u}) - \sum_{j=1}^d \dot{C}_{\eta,j}(\boldsymbol{u})\mathbb{B}_\eta(1,\ldots,1,u_j,1,\ldots,1), \quad \boldsymbol{u} \in [0,1]^d,$$

with C_η an η-extreme-value copula such that its partial derivative $\dot{C}_{\eta,j}(\boldsymbol{u}) := \partial C_\eta/\partial u_j(\boldsymbol{u})$ exists and is continuous on $\{\boldsymbol{u} \in [0,1]^d : 0 < u_j < 1\}$, for all $j = 1, \ldots, d$, and $\mathbb{B}_\eta$ a C_η-Brownian bridge, i.e. a zero-mean Gaussian process on $[0,1]^d$ with continuous sample paths and covariance function equal to

$$\mathrm{Cov}\left(\mathbb{B}_\eta(\boldsymbol{u}), \mathbb{B}_\eta(\boldsymbol{v})\right) = C_\eta(\boldsymbol{u}\wedge\boldsymbol{v}) - C_\eta(\boldsymbol{u})C_\eta(\boldsymbol{v}), \qquad \boldsymbol{u}, \boldsymbol{v} \in [0,1]^d.$$

3.3. Examples of estimators satisfying Condition 1

Our η–Pickands dependence function requires an estimator of η which satisfies Condition 1. Below, two examples of such estimators are proposed.

Example 1. Let $Z^* = \max(Z_1, \ldots, Z_d)$, where $\mathbf{Z}$ follows the distribution (3.4). Then, for any $y > 0$, the distribution of Z^* is $G_\eta(y) := G_\eta(y, \ldots, y)$. This distribution can be seen as a two-parameter Fréchet family of distributions. Let $\widehat{\eta}_n$ be the Maximum Likelihood (ML) estimator. By Propositions 3.1 and 3.3 in Bücher and Segers (2017), it follows that the ML estimator satisfies Conditions 1(i) and 1(ii)(a).

Example 2. Let $\widehat{\eta}_n$ be the Generalized Probability Weighted Moment (GPWM) estimator of η introduced by Guillou et al. (2014). The next theorem shows that the GPWM estimator admits a stochastic representation implying that Condition 1(ii)(b) is satisfied. The almost sure consistency of $\widehat{\eta}_n$ is a direct consequence.

Theorem 3.2. *Let $\widehat{\eta}_n$ be the GPWM estimator. For a, b two integers and $Q_\eta(u) := G_\eta^{\leftarrow}(u)$, introduce the parameter*

$$\mu_{a,b} := \int_0^1 Q_\eta(u)u^a(-\log u)^b\mathrm{d}u$$

and on $u \in (0,1)$ the two functions

$$\begin{aligned}\gamma(u) &:= \mu_{1,2}u(-\log u) - \mu_{1,1}u(-\log u)^2\\ \varphi(u) &:= \frac{1}{\eta\, V_\eta^\eta(1,\ldots,1)}u\,(-\log u)^{1+\eta}.\end{aligned}$$

Then,

$$\sqrt{n}(\widehat{\eta}_n - \eta) = -\frac{2}{\mu_{1,1}^2}\int_0^1 \mathbb{H}_n(Q_\eta(u), \ldots, Q_\eta(u))\,\frac{\gamma(u)}{\varphi(u)}\mathrm{d}u + \mathrm{o}(1) \text{ a.s.}. \tag{3.11}$$

Consequently, as $n \to \infty$

$$\begin{aligned}\widehat{\eta}_n &\to \eta \text{ a.s.}\\ \sqrt{n}(\widehat{\eta}_n - \eta) &\rightsquigarrow -\frac{2}{\mu_{1,1}^2}\int_0^1 \mathbb{H}(Q_\eta(u), \ldots, Q_\eta(u))\,\frac{\gamma(u)}{\varphi(u)}\mathrm{d}u\end{aligned}$$

where $\mathbb{H}$ is a tight centered Gaussian process on $(0,\infty]^d$, with covariance function

$$\mathrm{Cov}(\mathbb{H}(\boldsymbol{z}), \mathbb{H}(\boldsymbol{y})) = G_\eta(\boldsymbol{z}\wedge\boldsymbol{y}) - G_\eta(\boldsymbol{z})G_\eta(\boldsymbol{y}), \quad \boldsymbol{z}, \boldsymbol{y} \in (0,\infty]^d.$$

3.4. Simulation

The performance of our estimator $\widehat{A}_{\widehat{\eta}_n,n}$ is illustrated in a simulation study with two different experiments.

First experiment: We consider the bivariate η-asymmetric logistic dependence model introduced by Ramos and Ledford (2011). Such a dependence structure is characterised by the following features. The hidden spectral measure H_η has density

$$h_\eta(w) = \frac{\eta-\psi}{\psi\eta^2 N_\varrho}\left\{(\varrho w)^{-1/\psi} + \left(\frac{1-w}{\varrho}\right)^{-1/\psi}\right\}^{\psi/\eta-2}\{w(1-w)\}^{-(1+1/\psi)}, \quad w\in(0,1), \tag{3.12}$$

where $N_\varrho := \varrho^{-1/\eta} + \varrho^{1/\eta} - (\varrho^{-1/\eta} + \varrho^{1/\eta})^{\psi/\eta}$ and $\psi \in (0,1]$, $\varrho > 0$, $\eta \in (0,1]$ are dependence parameters. This satisfies the condition (3.2), i.e. in case $d = 2$,

$$\eta^{-1} = \int_0^{1/2} \omega^{1/\eta} dH_\eta(w) + \int_{1/2}^1 (1-w)^{1/\eta} dH_\eta(w).$$

The associated limiting distribution in (3.3) takes the form

$$G_\eta(y_1,y_2) = \begin{cases} \exp\left[-N_\varrho^{-1}\left\{(\varrho y_1)^{-1/\psi} + \left(\frac{y_2}{\varrho}\right)^{-1/\psi}\right\}^{\psi/\eta}\right] & \text{for } \psi<\eta \\ 0 & \text{for } \psi\geq\eta, \end{cases} \tag{3.13}$$

where the degenerate case arises when H_η is infinite. In the sequel we focus on the case $\psi < \eta$ and for simplicity we consider $\varrho = 1$. Distribution (3.13) is the attractor for the distribution of normalized componentwise maxima obtained from a random vector whose survival function is

$$\Pr(Y_1>y_1, Y_2>y_2) = \frac{\lambda u^{1/\eta}}{N_\varrho}\left[(\varrho y_1)^{-1/\eta} + \left(\frac{y_2}{\varrho}\right)^{-1/\eta} - \left\{(\varrho y_1)^{-1/\psi} + \left(\frac{y_2}{\varrho}\right)^{-1/\psi}\right\}^{\psi/\eta}\right] \tag{3.14}$$

where $(y_1,y_2) \in [u,\infty)\times[u,\infty)$, with u being a high threshold and λ the joint threshold exceedance probability (see Ramos and Ledford, 2009 for details). The survival function (3.14) satisfies (3.1) and it is an asymptotically independent joint probability model for any $\eta \in (0,1)$, where the strength of the dependence, within asymptotic independence, increases for decreasing values of the parameter ψ. We call (3.13) and (3.14) the η-asymmetric logistic distribution and survival function, respectively.

We simulate n values from the η-asymmetric logistic distribution and we estimate the η-Pickands dependence function with $\widehat{A}_{\widehat{\eta}_n,n}$. We repeat these steps 1000 times and we compute a Monte Carlo approximation of the Mean Integrated Squared Error (MISE), i.e.,

$$\mathrm{MISE}(\widehat{A}_{\widehat{\eta}_n,n}, A_\eta) = \mathrm{E}\left\{\int_{\mathcal{S}_d}\left(\widehat{A}_{\widehat{\eta}_n,n}(\boldsymbol{t}) - A_\eta(\boldsymbol{t})\right)^2 \mathrm{d}\boldsymbol{t}\right\}.$$

This study is done for different values of the sample size n and different values of the dependence parameter ψ. The results are summarized in Table 3. For each value of ψ, between the second and the fifth column the mean of the estimates for η obtained with the GPWM (first row) and ML (second row) estimator are reported for increasing sample size. In parentheses is the standard deviation. Between the sixth and ninth columns the approximated MISE is reported. Accurate estimates are obtained with all the dependence levels. GPWM and ML estimators provide similar results, although those of the former seem slightly better. According to the MISE, the better performances are obtained with stronger dependence strengths. For every dependence level the accuracy of estimates increases with increasing sample size.

Second experiment: We show the performance of the estimator $\widehat{A}_{\widehat{\eta}_n,n}$ under a more realistic scenario. We simulate $n\times 366$ independent observations from a distribution whose survival function is given in (3.14). To do this we use the algorithm described in Theorem 1.1 and Appendix B of Ramos and Ledford (2009). The simulation procedure in Ramos and Ledford (2009) relies on the condition

$$\Pr(\boldsymbol{Y} > \boldsymbol{y}) = \lambda\eta u^{1/\eta}\int_{(0,1)} \min\left(w/y_1, (1-w)/y_2\right)^{1/\eta} \mathrm{d}H_\eta(w), \quad \boldsymbol{y} > u\boldsymbol{1},$$

TABLE 3
Estimates (standard deviation) of η and MISE for the η-Pickands dependence function, based on a bivariate η-asymmetric logistic dependence model with $\eta = 0.7$. The first line corresponds to the GPWM method, whereas the second line is the ML method.

	$\widehat{\eta}_n$				$\mathrm{MISE}(\widehat{A}_{\widehat{\eta}_n,n}, A_\eta)$			
	n				n			
ψ	25	50	100	200	25	50	100	200
0.1	0.661(0.115)	0.678(0.084)	0.690(0.062)	0.695(0.043)	0.0111	0.0037	0.0013	0.0005
	0.800(0.201)	0.763(0.128)	0.741(0.085)	0.728(0.055)	0.0110	0.0036	0.0013	0.0005
0.2	0.667(0.116)	0.679(0.084)	0.688(0.062)	0.692(0.044)	0.0480	0.0195	0.0088	0.0041
	0.807(0.204)	0.761(0.128)	0.740(0.088)	0.724(0.057)	0.0457	0.0187	0.0086	0.0040
0.3	0.665(0.116)	0.680(0.087)	0.692(0.064)	0.696(0.046)	0.1176	0.0542	0.0262	0.0133
	0.811(0.211)	0.768(0.130)	0.745(0.087)	0.730(0.059)	0.1133	0.0527	0.0256	0.0131
0.4	0.673(0.114)	0.687(0.088)	0.694(0.062)	0.697(0.045)	0.2177	0.1021	0.0523	0.0260
	0.810(0.204)	0.770(0.129)	0.744(0.084)	0.729(0.057)	0.2118	0.1000	0.0514	0.0257
0.5	0.670(0.113)	0.684(0.085)	0.692(0.062)	0.695(0.044)	0.3602	0.1795	0.0952	0.0481
	0.805(0.201)	0.766(0.129)	0.742(0.085)	0.728(0.057)	0.3531	0.1765	0.0940	0.0476
0.6	0.670(0.115)	0.685(0.085)	0.691(0.062)	0.696(0.046)	0.4566	0.2252	0.1192	0.0556
	0.822(0.206)	0.778(0.127)	0.751(0.089)	0.734(0.063)	0.4758	0.2275	0.1156	0.0585

TABLE 4
Estimates (standard deviation) of η and MISE for the η-Pickands dependence function, based on componentwise maxima with approximate bivariate η-asymmetric logistic model with $\eta = 0.7$. The first line corresponds to the GPWM method, whereas the second line is the ML method.

	$\widehat{\eta}_n$				$\mathrm{MISE}(\widehat{A}_{\widehat{\eta}_n,n}, A_\eta)$			
	n				n			
ψ	25	50	100	200	25	50	100	200
0.1	0.668(0.115)	0.684(0.089)	0.692(0.061)	0.695(0.044)	0.0108	0.0034	0.0013	0.0005
	0.800(0.204)	0.764(0.128)	0.742(0.086)	0.730(0.060)	0.0106	0.0033	0.0013	0.0005
0.2	0.664(0.116)	0.681(0.086)	0.687(0.061)	0.693(0.045)	0.0456	0.0187	0.0088	0.0040
	0.810(0.213)	0.765(0.133)	0.743(0.091)	0.728(0.064)	0.0442	0.0183	0.0079	0.0039
0.3	0.670(0.120)	0.686(0.089)	0.696(0.063)	0.698(0.045)	0.1088	0.0563	0.0257	0.0119
	0.804(0.194)	0.766(0.119)	0.744(0.078)	0.732(0.055)	0.1080	0.0546	0.0255	0.0117
0.4	0.684(0.119)	0.699(0.091)	0.707(0.063)	0.711(0.045)	0.2265	0.1146	0.0593	0.0279
	0.829(0.209)	0.783(0.125)	0.759(0.082)	0.745(0.054)	0.2207	0.1129	0.0584	0.0276
0.5	0.711(0.119)	0.727(0.088)	0.734(0.063)	0.738(0.044)	0.3820	0.1970	0.1113	0.0657
	0.846(0.197)	0.806(0.122)	0.786(0.080)	0.773(0.049)	0.3819	0.1973	0.1112	0.0661
0.6	0.751(0.125)	0.766(0.094)	0.775(0.070)	0.781(0.050)	0.7105	0.4298	0.3018	0.2333
	0.898(0.201)	0.849(0.121)	0.831(0.082)	0.820(0.053)	0.7293	0.4387	0.3096	0.2381

where u and λ are as in the first experiment. This condition implies that for every $y > u$,

$$
\begin{aligned}
1 - \Pr(Y_1 > y, Y_2 \le u) &= \exp(-1/y) + \lambda\eta \int_{(0,1)} \min\left(\frac{w}{y/u}, 1 - w\right)^{1/\eta} \mathrm{d}H_\eta(w), \\
1 - \Pr(Y_1 \le u, Y_2 > y) &= \exp(-1/y) + \lambda\eta \int_{(0,1)} \min\left(w, \frac{1-w}{y/u}\right)^{1/\eta} \mathrm{d}H_\eta(w).
\end{aligned}
$$

The values of u and λ must be selected in such a way that both functions of y above are monotonically increasing. When the density of H_η is given by (3.12) with $\varrho = 1$ and $\eta = 0.7$, the monotonicity conditions are satisfied for every $\psi < \eta$ by setting $u = 10$ and $\lambda = 1 - \exp(-0.1) - 0.02$.

With simulated data we compute $\widehat{b}_{366}$, that is the empirical $(1 - 1/366)$-quantile of the minimum between pairs of all observations. For each block of 366 observations we compute the componentwise maxima using $\varepsilon = Q(0.07)/\widehat{b}_{366}$, where $Q(0.07)$ is the ninety-third percentile of a unit Fréchet distribution, i.e. by retaining only the pairs that are both greater than $\varepsilon\widehat{b}_{366}$. We standardize the maxima by dividing them by $\widehat{b}_{366}$. With the n normalized maxima we estimate the η-Pickands dependence function by $\widehat{A}_{\widehat{\eta}_n,n}$. We repeat these steps 1000 times and we compute an approximation of the MISE. Table 4 collects the results.

We see that the estimates of η and A_η are similar to those obtained in Table 3, indicating a good performance of our estimator. We mention that in each block of 366 observations the componentwise maxima are computed, after the truncation, on average on approximately 17 pairs. Although maxima are obtained with a small number of observations, the estimation results suggest that they are enough to obtain accurate estimates. Estimates are less accurate for $\psi = 0.6$ for the following reason. The simulation method for maxima produces observations that are approximately drawn from the non-degenerate distribution G_η in (3.13), provided that $\psi < \eta$, since G_η is a degenerate distribution for $\psi \geq \eta$. Furthermore, for this example it can be empirically verified that G_η provides a very accurate approximation for the distribution of the simulated maxima when $\psi < 0.6$. Instead, whenever ψ is close to η (a case that resembles the degenerate case), e.g. $0.6 \leq \psi < 0.7 = \eta$, the quality of the approximation deteriorates. In this case the mismatch between G_η and the distribution of simulated maxima is no longer negligible, thus affecting the estimation results.

Finally, note that the asymptotic properties of our estimator established in Theorem 3.1 are no longer valid in this experiment, although our estimator still performs well. Indeed, they should be re-established under the assumption that the data belong to the domain of attraction of G_η. In that case the proofs are much more technical and thus are outside the scope of the present paper.

4. Discussion

The framework for modelling the dependence within asymptotic independence, based on componentwise maxima, relies on the assumption that H_η is a finite measure (see Section 3). Before applying our estimation method, it is desirable to check somehow whether such an assumption is not violated by the available data. To this end, we propose a diagnostic tool. We motivate it on the basis of the following discussion. For simplicity we focus on the bivariate case although our proposal is easily extendable to higher dimensions. Let $\boldsymbol{Y}$ be a two-dimensional random vector defined as in Section 3. Define

$$\begin{aligned}\overline{\overline{F}}_j(y) &:= \lim_{s\to\infty} \Pr(Y_j > sy \,|\, \boldsymbol{Y} > s\mathbf{1}), \quad j = 1,2,\ s > 0,\ y > 1,\\ &= \eta\left(y^{-1/\eta}\int_0^{y/(1+y)} w^{1/\eta}\mathrm{d}H_\eta(w) + \int_{y/(1+y)}^1 (1-w)^{1/\eta}\mathrm{d}H_\eta(w)\right)\\ &\leq 2\eta M y^{-1/\eta},\end{aligned}$$

where $M = H_\eta((0,1)) < \infty$, and

$$\overline{\overline{F}}_{min}(y) := \lim_{s\to\infty} \Pr(\min(Y_1, Y_2) > sy \,|\, \boldsymbol{Y} > s\mathbf{1}).$$

Then, it follows that

$$\begin{aligned}1 &\leq 2\eta M \min\left(y^{-1/\eta}/\overline{\overline{F}}_1(y), y^{-1/\eta}/\overline{\overline{F}}_2(y)\right)\\ &= 2\eta M \min\left(\overline{\overline{F}}_{min}(y)/\overline{\overline{F}}_1(y), \overline{\overline{F}}_{min}(y)/\overline{\overline{F}}_2(y)\right).\end{aligned}$$

Consequently marginal survival functions $\overline{\overline{F}}_j$, $j = 1,2$, heavier than $\overline{\overline{F}}_{min}$, i.e. $\overline{\overline{F}}_{min}(y)/\overline{\overline{F}}_j(y) \to 0$ as $y \to \infty$, $j = 1,2$, provide empirical evidence against the hypothesis that H_η is finite. On the contrary, evidence in favor of a finite H_η is provided by the conditions $\overline{\overline{F}}_{min}(y)/\overline{\overline{F}}_j(y) \to c_j$ as $y \to \infty$, where c_j, $j = 1,2$, are positive constants. On this basis, we suggest implementing the plot

$$\widehat{r}_j(y) := \frac{\widehat{\overline{\overline{F}}}_{min}(y)}{\widehat{\overline{\overline{F}}}_j(y)}, \quad 1 \leq y \leq m^*/s,$$

where

$$\begin{aligned}\widehat{\overline{\overline{F}}}_j(y) &:= \frac{1}{n_s}\sum_{i=1}^{n_s} \mathbb{1}(y_{i,j} > sy), \quad j = 1,2,\ y > 1\\ \widehat{\overline{\overline{F}}}_{min}(y) &:= \frac{1}{n_s}\sum_{i=1}^{n_s} \mathbb{1}(m_i > sy),\end{aligned}$$

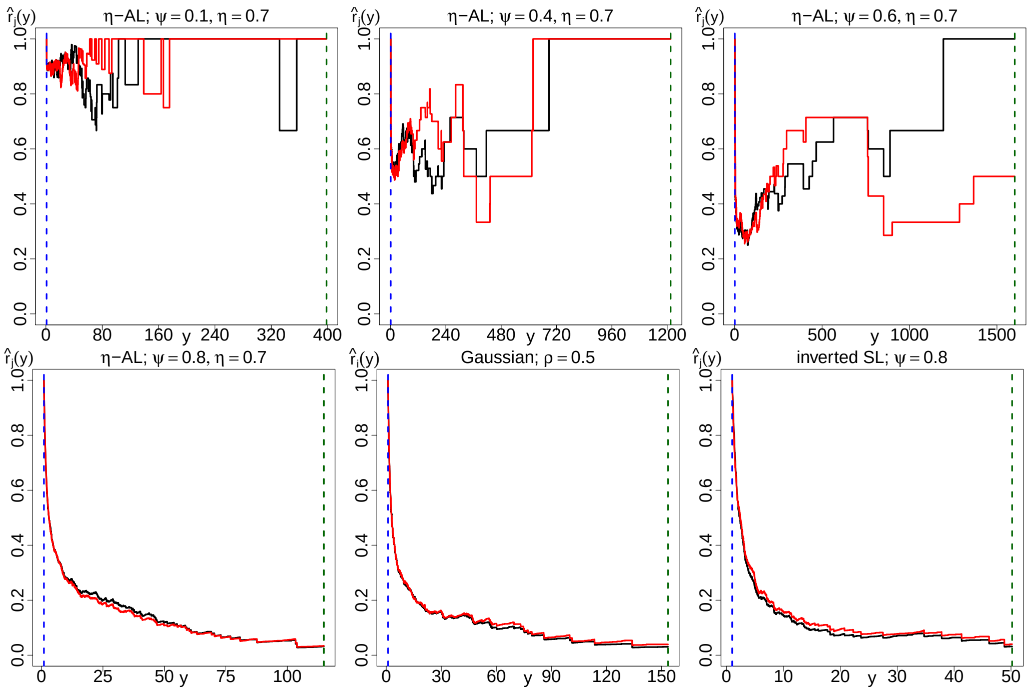

FIG 3. *Diagnostic plots to check the finiteness of H_η. The left-hand vertical dotted line crosses the abscissas at* 1*, while the right-hand one at the value m^*/s. The red line is the case $j = 1$ and the black line the case $j = 2$.*

$y_{i,j}$ are observations larger than a positive threshold s, $m_i = \min(y_{i,1}, y_{i,2})$, n_s is the number of m_i exceeding the threshold s and m^* is the $(n_s$-1)th order statistic of the sample $m_1, \ldots, m_{n_s}$. When H_η is finite, for $j = 1, 2$, $\widehat{r}_j(y)$ approaches a positive constant as $y \to m^*/s$, whereas in the infinite case, it decreases toward zero.

We illustrate the diagnostic tool with some examples. We draw samples of 500×366 values from six different models that satisfy equations (3.1) and (3.2). We consider three models with an η-asymmetric logistic (η-AL) survival function and parameters $\psi = 0.1$, $\psi = 0.4$ and $\psi = 0.6$, respectively, while $u = 10$, $\lambda = 1 - e^{-0.1} - 0.02$, $\varrho = 1$, $\eta = 0.7$ are the same for all the three cases. In these examples H_η is finite and both $\overline{\overline{F}}_{min}$ and $\overline{\overline{F}}_j$, $j = 1, 2$, behave approximately as $y^{-1/\eta}$ for large values of y. Figure 3 displays in the first line the plots of $\widehat{r}_j$, obtained using $s = Q(0.07)$, where $Q(0.07)$ is the ninety-third percentile of unit-Fréchet distribution. As expected, for large values of y, $\widehat{r}_j(y)$ stays away from zero and it approaches 1 when $\psi = 0.1, 0.4$ and a smaller constant when $\psi = 0.6$ (value close to $\psi = 0.7$ with which H_η is infinite). We also consider a η-asymmetric logistic model with $u = 8$, $\lambda = 1 - e^{-1/8} - 0.04$, $\varrho = 1$, $\eta = 0.7$ and $\psi = 0.8$, a bivariate standard Gaussian distribution with $\rho = 0.5$ and a bivariate inverted symmetric logistic (SL) model with $\psi = 0.8$ (see Section 2.3). In the latter two models the marginal distributions of the data are transformed into uni-Fréchet. In these three cases H_η is infinite. Furthermore, for large y, $\overline{\overline{F}}_{min}$ behaves approximately as $y^{-1/\eta}$ with $\eta = 0.7$, $\eta = (1 + \rho)/2 = 0.75$ and $\eta = 2^{-\psi} \simeq 0.57$, respectively. While, $\overline{\overline{F}}_j$, $j = 1, 2$ behaves approximately as $y^{-1/k}$ with $k = \psi = 0.8$, $k = 1 + \rho = 1.5$ and $k = 2^{1-\psi} \simeq 1.15$, respectively. For these three examples the diagnostic plots are displayed in the second line of Figure 3. As expected, $\widehat{r}_j(y)$ goes to zero for large values of y.

The procedure for inferring A_η discussed in Section 3.2, when possible, provides useful means to extrapolate the probability of joint high thresholds exceedances as we describe next. For simplicity we focus on the bivariate case. By (3.1), (3.5) and the definition of the η-Pickands dependence function we have that the approximation

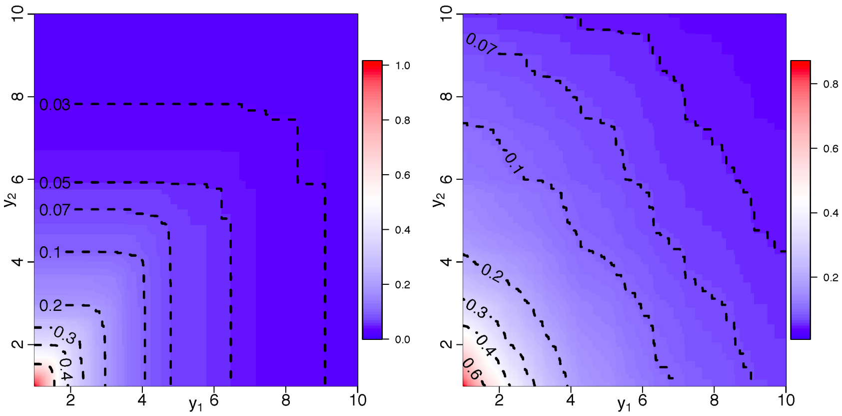

FIG 4. *Estimated probabilities of joint high thresholds exceedances.*

$$
\begin{aligned}
&\Pr(Y_1 > sy_1, Y_2 > sy_2 \,|\boldsymbol{Y} > s\mathbf{1}) \\
&\quad \approx \quad \eta \int_0^1 \min\left(\frac{w}{y_1}, \frac{1-w}{y_2}\right)^{1/\eta} \mathrm{d}H_\eta(w) \\
&\quad = \quad -\left\{\log G_{\eta,1}(y_1) + \log G_{\eta,2}(y_2)\right\}\left\{1 - A_\eta\left(\frac{\log G_{\eta,1}(y_1)}{\log G_{\eta,1}(y_1) + \log G_{\eta,2}(y_2)}, \frac{\log G_{\eta,2}(y_2)}{\log G_{\eta,1}(y_1) + \log G_{\eta,2}(y_2)}\right)\right\}
\end{aligned}
$$

holds for a large threshold s and $y_1, y_2 > 1$. Set $s = \widehat{b}_n$, where $\widehat{b}_n$ is the empirical $(1-1/n)$-quantile of the sequence $\min(y_{i,1}, y_{i,1})$, $i = 1, \ldots, n$, with $y_{i,1}, y_{i,2}$ that are independent realizations of $\boldsymbol{Y}$, see (3.3). Then, the above probability can be approximated by

$$
-\{\log H_{n,1}(y_1) + \log H_{n,2}(y_2)\}\left\{1 - \widehat{A}_{\widehat{\eta}_n,n}\left(\frac{\log H_{n,1}(y_1)}{\log H_{n,1}(y_1) + \log H_{n,2}(y_2)}, \frac{\log H_{n,2}(y_2)}{\log H_{n,1}(y_1) + \log H_{n,2}(y_2)}\right)\right\}, \tag{4.1}
$$

where $H_{n,j}$, $j = 1, 2$, are the empirical distribution functions, see Section 3.2. We illustrate the extrapolation of the probability of high thresholds exceedances with two examples. We simulate 500×366 independent realizations from two distributions with an η-asymmetric logistic survival function and parameters $\psi = 0.1$, $\psi = 0.4$, respectively, while $u = 10$, $\lambda = 1 - e^{-0.1} - 0.02$, $\varrho = 1$, $\eta = 0.7$ are the same for both the cases. Then, we obtain the sample of maxima, using $\widehat{b}_{366}$, $\varepsilon = Q(0.07)/\widehat{b}_{366}$ (see Section 3.1 and the second experiment of Section 3.4 for details) and we estimate the Pickands dependence function with $\widehat{A}_{\widehat{\eta}_{500},500}$, where $\widehat{\eta}_{500}$ is the GPWM estimator of η. For $y_1, y_2 \in [1, 10]$, we extrapolate the probability of joint high thresholds exceedances by applying (4.1). Figure 4 displays the estimated probabilities for the two models. The left and right panels report the results for the cases $\psi = 0.1$ and $\psi = 0.4$, respectively. To go further with this idea, a topic of interest would be to establish the asymptotic properties of the estimator defined in (4.1). This is outside the scope of the present paper but it will lead to further investigations.

Appendix A: Proofs

A.1. Some properties of $\widehat{A}_n$

Note that,

$$\widehat{\nu}_n(\boldsymbol{e}_j) = \frac{1}{n}\sum_{i=1}^{n}\left(\frac{G_{n,j}(X_{i,j})}{2\,n^{-1}\sum_{k=1}^{n} G_{n,j}(X_{k,j})} - \frac{1}{d}\frac{G_{n,j}(X_{i,j})}{2\,n^{-1}\sum_{k=1}^{n} G_{n,j}(X_{k,j})}\right) = \frac{1}{2} - \frac{1}{2d}, \quad j = 1,\dots,d.$$

Therefore, $\widehat{A}_n(\boldsymbol{e}_j) = 1$ for all $j = 1,\dots,d$.
The distribution function of the i.i.d. random variables $X_{1,j},\dots,X_{n,j}$, $j = 1,\dots,d$, being continuous, almost surely there are no ties and thus

$$G^{(1)}_{n,j}(X_{i,j}) = G_{n,j}(X_{i,j})\left(\frac{2}{n}\sum_{k=1}^{n} G_{n,j}(X_{k,j})\right)^{-1} = \frac{n}{n+1}G_{n,j}(X_{i,j}).$$

Then, with simple adjustments of the proof of Theorem 2.4 in Marcon et al. (2017), the weak convergence of $\widehat{A}_n$ and its almost sure consistency follow. □

A.2. Proof of Proposition 3.1

Our definition of L_η combining with (6.3) in Ramos and Ledford (2011) entails

$$A_\eta(\boldsymbol{t}) = \eta\int_{\mathcal{S}_d}\max\left(\frac{t_1 w_1^{1/\eta}}{\sigma_{\eta,1}},\dots,\frac{t_d w_d^{1/\eta}}{\sigma_{\eta,d}}\right)\mathrm{d}H_\eta(\boldsymbol{w}), \qquad t \in \mathcal{S}_d.$$

Then, Property 1 follows by the definition of $\sigma_{\eta,j}$ given in (3.5).
When $\eta = 1$, according to Section 3, we have

$$\lim_{n\to\infty}\frac{\Pr(\boldsymbol{Y} > n\boldsymbol{x})}{\Pr(\boldsymbol{Y} > n\boldsymbol{1})} = \int_{\mathcal{S}_d}\bigwedge_{j=1}^{d}\left(\frac{w_j}{x_j}\right)\mathrm{d}H_1(\boldsymbol{w}).$$

Now, this limit can also be rephrased with the classical theory (see e.g. de Haan and Ferreira, 2006, Ch. 6), where

$$\lim_{n\to\infty}\frac{\Pr(\boldsymbol{Y} > n\boldsymbol{x})}{\Pr(\boldsymbol{Y} > n\boldsymbol{1})} = \frac{d\int_{\mathcal{S}_d}\bigwedge_{j=1}^{d}\left(\frac{w_j}{x_j}\right)\mathrm{d}H(\boldsymbol{w})}{R(1,\dots,1)},$$

with H and R defined in pages 218 and 225 in de Haan and Ferreira (2006). Therefore, Property 2 follows from the relations

$$d^{-1}R(1,\dots,1)\mathrm{d}H_1(\boldsymbol{w}) = \mathrm{d}H(\boldsymbol{w}), \quad \boldsymbol{w} \in \mathcal{S}_d$$

and $\sigma_{1,j} = 1/R(1,\dots,1)$, $j = 1,\dots,d$.
For every $\boldsymbol{t} \in \mathcal{S}_d$ we have

$$\eta\int_{\mathcal{S}_d}\max\left(\frac{t_1 w_1^{1/\eta}}{\sigma_{\eta,1}},\dots,\frac{t_d w_d^{1/\eta}}{\sigma_{\eta,d}}\right)\mathrm{d}H_\eta(\boldsymbol{w}) \le \eta\int_{\mathcal{S}_d}\sum_{j=1}^{d}\left(\frac{t_j w_j^{1/\eta}}{\sigma_{\eta,j}}\right)\mathrm{d}H_\eta(\boldsymbol{w}) = 1,$$

from which the upper bound in Property 3 follows. To derive the lower bound, it is sufficient to remark that for every $\boldsymbol{t} \in \mathcal{S}_d$, we have

$$
\begin{aligned}
\eta \int_{\mathcal{S}_d} \max\left(\frac{t_1 w_1^{1/\eta}}{\sigma_{\eta,1}}, \ldots, \frac{t_d w_d^{1/\eta}}{\sigma_{\eta,d}}\right) \mathrm{d}H_\eta(\boldsymbol{w}) &\geq \bigvee_{1\le i<j\le d} \left(\eta \int_{\mathcal{S}_d} \max\left(\frac{t_i w_i^{1/\eta}}{\sigma_{\eta,i}}, \frac{t_j w_j^{1/\eta}}{\sigma_{\eta,j}}\right) \mathrm{d}H_\eta(\boldsymbol{w})\right) \\
&= \bigvee_{1\le i<j\le d} \left(t_i + t_j - \eta \int_{\mathcal{S}_d} \min\left(\frac{t_i w_i^{1/\eta}}{\sigma_{\eta,i}}, \frac{t_j w_j^{1/\eta}}{\sigma_{\eta,j}}\right) \mathrm{d}H_\eta(\boldsymbol{w})\right) \\
&\geq \bigvee_{1\le i<j\le d} \left(t_i + t_j - \min\left(t_i, t_j\right)\right) = \bigvee_{1\le j\le d} t_j.
\end{aligned}
$$

Finally, the convexity in Property 4 can been shown similarly to the convexity of A. □

A.3. Proof of Proposition 3.2

For all $\eta \in (0, 1]$ and $\boldsymbol{t} \in \mathcal{S}_d$, set

$$
\nu_\eta(\boldsymbol{u}; \boldsymbol{t}) := \bigvee_{j=1}^{d} u_j^{1/\eta t_j} - \frac{1}{d} \sum_{j=1}^{d} u_j^{1/\eta t_j}, \quad \boldsymbol{u} \in [0, 1]^d.
$$

By convention $u^{1/\eta t} = 0$ when $t = 0$ and $u \in [0, 1]$. By Lemma A.1 in Marcon et al. (2017) we have

$$
\begin{aligned}
\nu_\eta(\boldsymbol{t}) &= \int_{[0,1]^d} \nu_\eta(\boldsymbol{u}; \boldsymbol{t}) \mathrm{d}C_\eta(\boldsymbol{u}) && \text{(A.1)} \\
&= \frac{1}{d} \sum_{j=1}^{d} \int_0^1 C_\eta(1, \ldots, 1, v^{\eta t_j}, 1, \ldots, 1) \mathrm{d}v - \int_0^1 C_\eta(v^{\eta t_1}, \ldots, v^{\eta t_d}) \mathrm{d}v \\
&= \frac{1}{d} \sum_{j=1}^{d} \int_0^1 v^{\eta t_j} \mathrm{d}v - \int_0^1 v^{\eta A_\eta(\boldsymbol{t})} \mathrm{d}v \\
&= \frac{1}{d} \sum_{j=1}^{d} \frac{1}{1 + \eta t_j} - \frac{1}{1 + \eta A_\eta(\boldsymbol{t})}.
\end{aligned}
$$

The result (3.6) follows by solving the above equality for A_η. □

A.4. Proof of Theorem 3.1

We start with some notation. Let $\widehat{\mathbb{C}}_n := \sqrt{n}(\widehat{C}_n - C_\eta)$, where $\widehat{C}_n$ is the empirical copula defined as

$$
\widehat{C}_n(\boldsymbol{u}) := \frac{1}{n} \sum_{i=1}^{n} \mathbb{1}_{\{\widehat{\boldsymbol{U}}_i \le \boldsymbol{u}\}}, \quad \boldsymbol{u} \in [0, 1]^d,
$$

with $\widehat{\boldsymbol{U}}_i = (H_{n,1}(Z_{i,1}), \ldots, H_{n,d}(Z_{i,d}))$. Define now, for all $\boldsymbol{t} \in \mathcal{S}_d$,

$$
M(\cdot, \boldsymbol{t}) := 1 - \int_0^1 C_\eta(v^{\cdot\, t_1}, \ldots, v^{\cdot\, t_d}) \mathrm{d}v, \tag{A.2}
$$

$$
\widehat{M}_n(\cdot, \boldsymbol{t}) := 1 - \int_0^1 \widehat{C}_n(v^{\cdot\, t_1}, \ldots, v^{\cdot\, t_d}) \mathrm{d}v. \tag{A.3}
$$

We will prove Theorem 3.1 with $H_{n,j}^{(\widehat{\eta}_n)}$ in $\widehat{\nu}_{\widehat{\eta}_n,n}$ and $\widehat{c}_{\widehat{\eta}_n,n}$ replaced by $H_{n,j}$. Indeed, this slight modification has no impact on the convergences (3.9) and (3.10) since

$$H_{n,j}^{(\widehat{\eta}_n)}(Z_{i,j}) = H_{n,j}(Z_{i,j})\left(1+\frac{1+\widehat{\eta}_n}{\widehat{\eta}_n}O\left(\frac{1}{n}\right)\right)^{-\widehat{\eta}_n} =: H_{n,j}(Z_{i,j})\, e_n^{-1},$$

and the terms in (A.2) and (A.3) can be slightly changed by replacing in the integrals $v^{\cdot\, t_j}$ by $v^{\cdot\, t_j}e_n$, $j=1,\dots,d$, without any impact. In view of this remark, we pursue the proof of Theorem 3.1 with $M(\cdot,\boldsymbol{t})$ and $\widehat{M}_n(\cdot,\boldsymbol{t})$ defined in (A.2) and (A.3) without taking care of the adjustment with e_n.
We start to prove (3.10). To this aim, note that from (A.1) we have

$$M(\eta,\boldsymbol{t}) = \frac{\eta A_\eta(\boldsymbol{t})}{1+\eta A_\eta(\boldsymbol{t})}$$

and thus the following decomposition holds

$$\begin{aligned}\sqrt{n}(\widehat{A}_{\widehat{\eta}_n,n}(\boldsymbol{t})-A_\eta(\boldsymbol{t})) &= \sqrt{n}\left(\frac{1}{\widehat{\eta}_n}\frac{\widehat{M}_n(\widehat{\eta}_n,\boldsymbol{t})}{1-\widehat{M}_n(\widehat{\eta}_n,\boldsymbol{t})}-\frac{1}{\eta}\frac{M(\eta,\boldsymbol{t})}{1-M(\eta,\boldsymbol{t})}\right)\\ &= \frac{\sqrt{n}}{\widehat{\eta}_n}\left(\frac{\widehat{M}_n(\widehat{\eta}_n,\boldsymbol{t})}{1-\widehat{M}_n(\widehat{\eta}_n,\boldsymbol{t})}-\frac{M(\eta,\boldsymbol{t})}{1-M(\eta,\boldsymbol{t})}\right)+\frac{M(\eta,\boldsymbol{t})}{1-M(\eta,\boldsymbol{t})}\sqrt{n}\left(\frac{1}{\widehat{\eta}_n}-\frac{1}{\eta}\right)\\ &=: L_n(\boldsymbol{t})+R_n(\boldsymbol{t}),\end{aligned}$$

for all $\boldsymbol{t}\in\mathcal{S}_d$. We derive a tractable expression for L_n by means of the following three results.

Lemma A.1. *We have the following decomposition*

$$\sqrt{n}(\widehat{M}_n(\widehat{\eta}_n,\boldsymbol{t})-M(\eta,\boldsymbol{t})) = \sqrt{n}(\widehat{M}_n(\eta,\boldsymbol{t})-M(\eta,\boldsymbol{t}))+\sqrt{n}(M(\widehat{\eta}_n,\boldsymbol{t})-M(\eta,\boldsymbol{t}))+\mathrm{o_p}(1).$$

Proof. The proof uses arguments from van der Vaart and Wellner (2007). Since

$$\begin{aligned}\sqrt{n}(\widehat{M}_n(\widehat{\eta}_n,\boldsymbol{t})-M(\eta,\boldsymbol{t})) &= \left\{\sqrt{n}(\widehat{M}_n(\widehat{\eta}_n,\boldsymbol{t})-M(\widehat{\eta}_n,\boldsymbol{t}))-\sqrt{n}(\widehat{M}_n(\eta,\boldsymbol{t})-M(\eta,\boldsymbol{t}))\right\}\\ &\quad+\sqrt{n}(\widehat{M}_n(\eta,\boldsymbol{t})-M(\eta,\boldsymbol{t}))+\sqrt{n}(M(\widehat{\eta}_n,\boldsymbol{t})-M(\eta,\boldsymbol{t})),\end{aligned}$$

it remains to show that

$$\|\sqrt{n}(\widehat{M}_n(\widehat{\eta}_n,\boldsymbol{t})-M(\widehat{\eta}_n,\boldsymbol{t}))-\sqrt{n}(\widehat{M}_n(\eta,\boldsymbol{t})-M(\eta,\boldsymbol{t}))\|_\infty = \mathrm{o_p}(1). \tag{A.4}$$

By Condition 1(ii) we have that $\sqrt{n}(\widehat{\eta}_n-\eta)$ is asymptotically tight. Thus, for every $\varepsilon>0$, there exists a compact set $K\equiv K_\varepsilon\subseteq\mathcal{R}$ such that

$$\liminf_{n\to\infty}\Pr(\sqrt{n}(\widehat{\eta}_n-\eta)\in K)>1-\varepsilon.$$

Furthermore, by the compactness of K, there exist $\delta>0$, $p:=p(\delta)\in\mathbb{N}$ and $\{h_1,\dots,h_p\}\subseteq K$ such that $K\subseteq\cup_{1\le s\le p}(h_s-\delta,h_s+\delta)$. Therefore,

$$\begin{aligned}\{\sqrt{n}(\widehat{\eta}_n-\eta)\in K\} &\subseteq \left\{\sqrt{n}(\widehat{\eta}_n-\eta)\subseteq\bigcup_{s=1}^{p}(h_s-\delta,h_s+\delta)\right\}\\ &= \bigcup_{s=1}^{p}\left\{\widehat{\eta}_n\in\left(\eta+n^{-1/2}(h_s-\delta),\eta+n^{-1/2}(h_s+\delta)\right)\right\}.\end{aligned}$$

Consequently, it follows that, with probability at least $1-\varepsilon$,

$$\begin{aligned}
&\| \sqrt{n}(\widehat{M}_n(\widehat{\eta}_n,\boldsymbol{t}) - M(\widehat{\eta}_n,\boldsymbol{t})) - \sqrt{n}(\widehat{M}_n(\eta,\boldsymbol{t}) - M(\eta,\boldsymbol{t}))\|_\infty \\
&\le \sup_{\boldsymbol{t}\in\mathcal{S}_d}\max_{1\le s\le p}\sup_{|h-h_s|<\delta} | \sqrt{n}(\widehat{M}_n(\eta_{n,h},\boldsymbol{t}) - M(\eta_{n,h},\boldsymbol{t})) - \sqrt{n}(\widehat{M}_n(\eta,\boldsymbol{t}) - M(\eta,\boldsymbol{t}))| \\
&\le \sup_{\boldsymbol{t}\in\mathcal{S}_d}\max_{1\le s\le p} | \sqrt{n}(\widehat{M}_n(\eta_{n,h_s},\boldsymbol{t}) - M(\eta_{n,h_s},\boldsymbol{t})) - \sqrt{n}(\widehat{M}_n(\eta,\boldsymbol{t}) - M(\eta,\boldsymbol{t}))| \\
&\ + \sup_{\boldsymbol{t}\in\mathcal{S}_d}\max_{1\le s\le p}\sup_{|h-h_s|<\delta} | \sqrt{n}(\widehat{M}_n(\eta_{n,h_s},\boldsymbol{t}) - M(\eta_{n,h_s},\boldsymbol{t})) - \sqrt{n}(\widehat{M}_n(\eta_{n,h},\boldsymbol{t}) - M(\eta_{n,h},\boldsymbol{t}))| \\
&=: I_{n,1} + I_{n,2}
\end{aligned}$$

where $\eta_{n,\bullet} := \eta + n^{-1/2}\bullet$. Thus to show (A.4) it is sufficient to prove that both $I_{n,1}$ and $I_{n,2}$ tends to 0 in probability, as $n\to\infty$. Using (A.2) and (A.3) we obtain

$$\begin{aligned}
I_{n,1} &= \sup_{\boldsymbol{t}\in\mathcal{S}_d}\max_{1\le s\le p}\left|\int_0^1 \left(\widehat{\mathbb{C}}_n(v^{\eta_{n,h_s}t_1},\dots,v^{\eta_{n,h_s}t_d}) - \widehat{\mathbb{C}}_n(v^{\eta t_1},\dots,v^{\eta t_d})\right)\mathrm{d}v\right| \\
&\le \sup_{\boldsymbol{t}\in\mathcal{S}_d}\max_{1\le s\le p}\sup_{v\in(0,1)}\left|\widehat{\mathbb{C}}_n(v^{\eta_{n,h_s}t_1},\dots,v^{\eta_{n,h_s}t_d}) - \widehat{\mathbb{C}}_n(v^{\eta t_1},\dots,v^{\eta t_d})\right|
\end{aligned}$$

and

$$\begin{aligned}
I_{n,2} &= \sup_{\boldsymbol{t}\in\mathcal{S}_d}\max_{1\le s\le p}\sup_{|h-h_s|<\delta}\left|\int_0^1 \left(\widehat{\mathbb{C}}_n(v^{\eta_{n,h_s}t_1},\dots,v^{\eta_{n,h_s}t_d}) - \widehat{\mathbb{C}}_n(v^{\eta_{n,h}t_1},\dots,v^{\eta_{n,h}t_d})\right)\mathrm{d}v\right| \\
&\le \sup_{\boldsymbol{t}\in\mathcal{S}_d}\max_{1\le s\le p}\sup_{|h-h_s|<\delta}\sup_{v\in(0,1)}\left|\widehat{\mathbb{C}}_n(v^{\eta_{n,h_s}t_1},\dots,v^{\eta_{n,h_s}t_d}) - \widehat{\mathbb{C}}_n(v^{\eta_{n,h}t_1},\dots,v^{\eta_{n,h}t_d})\right|.
\end{aligned}$$

Now, for every $v\in(0,1)$ and small $\epsilon>0$, the map $\varphi:(0,1)\to\ell^\infty([\eta-\epsilon,\eta+\epsilon]) : v\mapsto\varphi(v)$, defined by $(\varphi(v))(x) = v^x$, induces continuously differentiable functions on $[\eta-\epsilon,\eta+\epsilon]$ for every $v\in(0,1)$. The first derivative of such functions is $(\dot\varphi(v))(x) = v^x\log v$, which is bounded above by $\xi_v = v^{\eta-\epsilon}|\log v|$. Therefore, $(\varphi(v))(x)$ is a Lipschitz function and it satisfies the condition

$$|(\varphi(v))(x) - (\varphi(v))(y)| \le \xi_v|x-y|, \quad \forall\, x,y\in[\eta-\epsilon,\eta+\epsilon].$$

Furthermore, there exists a positive constant ξ such that $\sup_{v\in(0,1)}\xi_v<\xi$, and thus for n sufficiently large ensuring that $\eta_{n,h},\eta_{n,h_s}\in[\eta-\epsilon,\eta+\epsilon]$, we have:

$$\begin{aligned}
|v^{\eta_{n,h_s}t_j} - v^{\eta t_j}| &\le \xi\,|\eta-\eta_{n,h_s}| = \xi\,n^{-1/2}|h_s| \to 0 \\
|v^{\eta_{n,h_s}t_j} - v^{\eta_{n,h}t_j}| &\le \xi\,|\eta_{n,h_s}-\eta_{n,h}| = \xi\,n^{-1/2}|h_s-h| \le \xi\,\delta n^{-1/2} \to 0,
\end{aligned}$$

as $n\to\infty$, for every $\boldsymbol{t}\in\mathcal{S}_d$, indexes $s\in\{1,\dots,p\}$, $j\in\{1,\dots,d\}$ and for every $|h-h_s|<\delta$. These results imply that

$$\sup_{\boldsymbol{t}\in\mathcal{S}_d}\max_{1\le s\le p}\sup_{v\in(0,1)}\max_{1\le j\le d}|v^{\eta_{n,h_s}t_j} - v^{\eta t_j}| \to 0, \quad n\to\infty \tag{A.5}$$

and

$$\sup_{\boldsymbol{t}\in\mathcal{S}_d}\max_{1\le s\le p}\sup_{|h-h_s|<\delta}\max_{1\le j\le d}\sup_{v\in(0,1)}|v^{\eta_{n,h_s}t_j} - v^{\eta_{n,h}t_j}| \to 0, \quad n\to\infty. \tag{A.6}$$

Since the first partial derivative of C_η exists and is continuous on $\{\boldsymbol{u}\in[0,1]^d : 0<u_j<1\}$ for all $j=1,\dots,d$, $\widehat{\mathbb{C}}_n \rightsquigarrow \mathbb{A}_\eta$ in $\ell^\infty([0,1]^d)$ as $n\to\infty$ (see e.g. Fermanian et al., 2004; Segers, 2012). Therefore the sequence $\widehat{\mathbb{C}}_n$ is asymptotically uniformly equicontinuous in probability (see Theorem 1.5.7 in van der Vaart and Wellner, 1996). Combining this result with (A.5) and (A.6) we can conclude that $I_{n,1}$ and $I_{n,2}$ tends to 0 in probability, as $n\to\infty$. Therefore (A.4) is established and thus Lemma A.1 follows. □

Lemma A.2. *We have*

$$\sqrt{n}(M(\widehat{\eta}_n,\boldsymbol{t}) - M(\eta,\boldsymbol{t})) = \frac{A_\eta(\boldsymbol{t})}{(\eta A_\eta(\boldsymbol{t})+1)^2}\ \sqrt{n}(\widehat{\eta}_n-\eta) + \mathrm{o}_\mathrm{p}(1).$$

Proof. Let
$$\varphi : ((0,\infty), |\cdot|) \to (\ell^\infty(\mathcal{S}_d), \|\cdot\|_\infty) : a \mapsto M(a,\cdot)$$
be the map defined by
$$M(a,\cdot) = \frac{a\,A_\eta(\cdot)}{1 + a\,A_\eta(\cdot)}.$$
Its Hadamard derivative at $\eta \in (0,1]$ is
$$h \mapsto (\dot\varphi_\eta(h)) = \frac{h\,A_\eta}{(\eta A_\eta + 1)^2}.$$
Indeed, for every $\epsilon_n \downarrow 0$ and $h_n \to h \in (0,\infty)$, as $n \to \infty$, such that $\eta + \epsilon_n h_n \in (0,\infty)$, we have
$$\begin{aligned}
&\lim_{n\to\infty} \sup_{\boldsymbol{t}\in\mathcal{S}_d} \left| \frac{(\varphi(\eta+\epsilon_n h_n))(\boldsymbol{t}) - (\varphi(\eta))(\boldsymbol{t})}{\epsilon_n} - (\dot\varphi_\eta(h))(\boldsymbol{t}) \right| \\
&= \lim_{n\to\infty} \sup_{\boldsymbol{t}\in\mathcal{S}_d} \left| \frac{1}{\epsilon_n}\left(\frac{(\eta+\epsilon_n h_n)A_\eta(\boldsymbol{t})}{(\eta+\epsilon_n h_n)A_\eta(\boldsymbol{t})+1} - \frac{\eta A_\eta(\boldsymbol{t})}{\eta A_\eta(\boldsymbol{t})+1} \right) - \frac{hA_\eta(\boldsymbol{t})}{(\eta A_\eta(\boldsymbol{t})+1)^2} \right| \\
&= \lim_{n\to\infty} \sup_{\boldsymbol{t}\in\mathcal{S}_d} \left| \frac{A_\eta(\boldsymbol{t})}{\eta A_\eta(\boldsymbol{t})+1} \right| \left| \frac{h_n}{(\eta+\epsilon_n h_n)A_\eta(\boldsymbol{t})+1} - \frac{h}{\eta A_\eta(\boldsymbol{t})+1} \right| \\
&\le \lim_{n\to\infty} d^2\, \frac{|h_n - h| + |h h_n|\epsilon_n}{(d+\eta)(d+\eta+\epsilon_n h_n)} = 0.
\end{aligned}$$
Lemma A.2 now follows from Theorem 20.8 in van der Vaart (2000) and under our Condition 1(ii). □

Lemma A.3. *We have*
$$\sqrt{n}\left(\frac{\widehat{M}_n(\widehat{\eta}_n,\boldsymbol{t})}{1-\widehat{M}_n(\widehat{\eta}_n,\boldsymbol{t})} - \frac{M(\eta,\boldsymbol{t})}{1-M(\eta,\boldsymbol{t})} \right) = (1+\eta A_\eta(\boldsymbol{t}))^2\ \sqrt{n}(\widehat{M}_n(\widehat{\eta}_n,\boldsymbol{t}) - M(\eta,\boldsymbol{t})) + \mathrm{o_p}(1).$$

Proof. The proof of this lemma is based on an application of the functional delta method after proving the Hadamard differentiability of the map $\varphi : \ell^\infty(\mathcal{S}_d) \mapsto \ell^\infty(\mathcal{S}_d) : f \mapsto f/(1-f)$, with f in $\ell^\infty(\mathcal{S}_d)$, and the existence of the weak limit of $\sqrt{n}(M(\widehat{\eta}_n,\cdot) - M(\eta,\cdot))$ in $\ell^\infty(\mathcal{S}_d)$.
First, we start showing that the Hadamard derivative of φ at $M := M(\eta,\cdot)$ is
$$h \mapsto (\dot\varphi_M(h)) = \frac{h}{(1-M)^2},$$
with h in $\ell^\infty(\mathcal{S}_d)$. Indeed, for every sequence $\epsilon_n \downarrow 0$ and $h_n \to h$ as $n \to \infty$, such that $M + \epsilon_n h_n$ in $\ell^\infty(\mathcal{S}_d)$, we have
$$\begin{aligned}
&\lim_{n\to\infty} \sup_{\boldsymbol{t}\in\mathcal{S}_d} \left| \frac{(\varphi(M+\epsilon_n h_n))(\boldsymbol{t}) - (\varphi(M))(\boldsymbol{t})}{\epsilon_n} - (\dot\varphi_M(h))(\boldsymbol{t}) \right| \\
&= \lim_{n\to\infty} \sup_{\boldsymbol{t}\in\mathcal{S}_d} \left| \frac{1}{\epsilon_n}\left(\frac{M(\eta,\boldsymbol{t}) + \epsilon_n h_n(\boldsymbol{t})}{1 - M(\eta,\boldsymbol{t}) - \epsilon_n h_n(\boldsymbol{t})} - \frac{M(\eta,\boldsymbol{t})}{1-M(\eta,\boldsymbol{t})} \right) - \frac{h(\boldsymbol{t})}{(1-M(\eta,\boldsymbol{t}))^2} \right| \\
&= \lim_{n\to\infty} \sup_{\boldsymbol{t}\in\mathcal{S}_d} (1+\eta A_\eta(\boldsymbol{t}))^2 \left| \frac{h_n(\boldsymbol{t}) - h(\boldsymbol{t}) + h(\boldsymbol{t})\epsilon_n h_n(\boldsymbol{t})(1+\eta A_\eta(\boldsymbol{t}))}{1 - \epsilon_n h_n(\boldsymbol{t})(1+\eta A_\eta(\boldsymbol{t}))} \right| \\
&\le \lim_{n\to\infty} (1+\eta)^2\, \frac{\|h_n - h\|_\infty + \epsilon_n \|h_n\, h\|_\infty (1+\eta)}{1 - \epsilon_n \|h_n\|_\infty (1+\eta)} = 0.
\end{aligned}$$

Then, combining Lemmas A.1, A.2 with Proposition 3.1 in Segers (2012), under Condition 1(ii)(b) we have that
$$\sqrt{n}(\widehat{M}_n(\widehat{\eta}_n,\cdot) - M(\eta,\cdot)) = T_{n,1}(\cdot) + T_{n,2}(\cdot) + \mathrm{o_p}(1),$$

where for all $\boldsymbol{t} \in \mathcal{S}_d$

$$T_{n,1}(\boldsymbol{t}) := -\int_0^1 \left(\mathbb{C}_n(v^{t_1\eta}, \ldots, v^{t_d\eta}) - \sum_{j=1}^d \dot{C}_{\eta,j}(v^{t_1\eta}, \ldots, v^{t_d\eta}) \mathbb{C}_n(1, \ldots, 1, v^{t_j\eta}, 1, \ldots, 1) \right) \mathrm{d}v$$

and

$$T_{n,2} := \frac{A_\eta}{(1+\eta A_\eta)^2} \chi(\mathbb{H}_n).$$

For any $\boldsymbol{u} \in [0,1]^d$, $\mathbb{C}_n(\boldsymbol{u}) = \mathbb{H}_n(G_{\eta,1}^{\leftarrow}(u_1), \ldots, G_{\eta,d}^{\leftarrow}(u_d))$, so both terms can be expressed as continuous transformations of the empirical process $\mathbb{H}_n$. Therefore, the weak convergence of $T_{n,1} + T_{n,2}$ follows from the continuous mapping theorem. A similar reasoning can be obtained if Condition 1(ii)(b) is replaced by Condition 1(ii)(a). In that case, we have the following decomposition

$$\sqrt{n}(\widehat{M}_n(\widehat{\eta}_n, \cdot) - M(\eta, \cdot)) =: T_{n,1} + \widetilde{T}_{n,2} + \mathrm{o}_{\mathrm{p}}(1),$$

where

$$T_{n,1}(\boldsymbol{t}) = \frac{1}{\sqrt{n}} \sum_{i=1}^n (W_{i,\boldsymbol{t}} - \mathrm{E}(W_{i,\boldsymbol{t}})), \quad \widetilde{T}_{n,2}(\boldsymbol{t}) := \frac{1}{\sqrt{n}} \sum_{i=1}^n \widetilde{W}_{i,\boldsymbol{t}}, \quad \boldsymbol{t} \in \mathcal{S}_d$$

and

$$\begin{aligned} W_{i,\boldsymbol{t}} &= \bigvee_{j=1}^d G_\eta^{1/\eta t_j}(Z_{i,j}) + \sum_{j=1}^d \int_0^1 \dot{C}_{\eta,j}(v^{t_1\eta}, \ldots, v^{t_d\eta}) \mathbb{1}_{\{v > G_\eta^{1/\eta t_j}(Z_{i,j})\}} \mathrm{d}v, \\ \widetilde{W}_{i,\boldsymbol{t}} &= \frac{A_\eta(\boldsymbol{t})}{(1+\eta A_\eta(\boldsymbol{t}))^2} \rho(\boldsymbol{Z}_i). \end{aligned}$$

Note that the new expression for $T_{n,1}$ is obtained by applying Fubini's theorem. The pair $(T_{n,1}, \widetilde{T}_{n,2})$ is asymptotically tight and so to show that its weak limit exists, it remains to prove that all its finite dimensional distributions converge. This can be done by applying the central limit theorem since, for all $k = 1, 2, \ldots$, the i.i.d. random vectors

$$\left(W_{i,\boldsymbol{t}_1}, \ldots, W_{i,\boldsymbol{t}_k}, \widetilde{W}_{i,\boldsymbol{t}_1}, \ldots, \widetilde{W}_{i,\boldsymbol{t}_k} \right), \quad i = 1, \ldots, n,$$

have finite second order moments under the assumptions of our Theorem 3.1 (see Nelsen, 2006, Theorem 2.2.7). This achieves the proof of Lemma A.3. □

We come back now to the proof of Theorem 3.1. Combining the three previous lemmas with the definition of $M(\eta, \boldsymbol{t})$, we have

$$\begin{aligned} L_n + R_n &= \frac{(1+\eta A_\eta(\boldsymbol{t}))^2}{\widehat{\eta}_n} \sqrt{n}(\widehat{M}_n(\eta, \boldsymbol{t}) - M(\eta, \boldsymbol{t})) + \frac{A_\eta(\boldsymbol{t})}{\widehat{\eta}_n} \sqrt{n}(\widehat{\eta}_n - \eta) + \eta A_\eta(\boldsymbol{t}) \sqrt{n} \left(\frac{1}{\widehat{\eta}_n} - \frac{1}{\eta} \right) + \mathrm{o}_{\mathrm{p}}(1) \\ &= \frac{(1+\eta A_\eta(\boldsymbol{t}))^2}{\widehat{\eta}_n} \sqrt{n}(\widehat{M}_n(\eta, \boldsymbol{t}) - M(\eta, \boldsymbol{t})) + \mathrm{o}_{\mathrm{p}}(1) \\ &= -\frac{(1+\eta A_\eta(\boldsymbol{t}))^2}{\eta} \int_0^1 \widehat{\mathbb{C}}_n(v^{\eta t_1}, \ldots, v^{\eta t_d}) \mathrm{d}v + \mathrm{o}_{\mathrm{p}}(1). \end{aligned}$$

As in the proof of Lemma A.1, using again the convergence $\widehat{\mathbb{C}}_n \rightsquigarrow \mathbb{A}_\eta$ in $\ell^\infty([0,1]^d)$ as $n \to \infty$, (3.10) follows from the continuous mapping theorem and Slutsky's lemma.
It remains now to prove (3.9). Note that

$$\begin{aligned} \|\widehat{A}_{\widehat{\eta}_n,n} - A_\eta\|_\infty &= \sup_{\boldsymbol{t} \in \mathcal{S}_d} \left| \frac{1}{\widehat{\eta}_n} \frac{\widehat{M}_n(\widehat{\eta}_n, \boldsymbol{t})}{1 - \widehat{M}_n(\widehat{\eta}_n, \boldsymbol{t})} - \frac{1}{\eta} \frac{M(\eta, \boldsymbol{t})}{1 - M(\eta, \boldsymbol{t})} \right| \\ &= \sup_{\boldsymbol{t} \in \mathcal{S}_d} \left| \frac{1}{\widehat{\eta}_n \eta \{1 - \widehat{M}_n(\widehat{\eta}_n, \boldsymbol{t})\}\{1 - M(\eta, \boldsymbol{t})\}} \right| \times \sup_{\boldsymbol{t} \in \mathcal{S}_d} \left| \eta \{1 - M(\eta, \boldsymbol{t})\} \widehat{M}_n(\widehat{\eta}_n, \boldsymbol{t}) - \widehat{\eta}_n \{1 - \widehat{M}_n(\widehat{\eta}_n, \boldsymbol{t})\} M(\eta, \boldsymbol{t}) \right| \\ &=: T_{n,1} \times T_{n,2}. \end{aligned}$$

Since $\widehat{\eta}_n \to \eta$ a.s., for a small $\varepsilon > 0$ and large n, we have almost surely that

$$T_{n,1} \leq \frac{1+1/\eta}{\widehat{\eta}_n \int_0^1 \widehat{C}_n(v^{1+\varepsilon},\ldots,v^{1+\varepsilon})\mathrm{d}v} \longrightarrow \frac{1+1/\eta}{\eta \int_0^1 C_\eta(v^{1+\varepsilon},\ldots,v^{1+\varepsilon})\mathrm{d}v} < \infty.$$

Now, using the Lipschitz property of order $k > 0$ of C_η, we have

$$\begin{aligned}
T_{n,2} &\leq \|\eta\{1 - M(\eta,\boldsymbol{t})\} - \widehat{\eta}_n\{1 - \widehat{M}(\widehat{\eta}_n,\boldsymbol{t})\}\|_\infty + \|\{1 - \widehat{M}_n(\widehat{\eta}_n,\boldsymbol{t})\}\{1 - M(\eta,\boldsymbol{t})\}\|_\infty |\widehat{\eta}_n - \eta| \\
&\leq |\widehat{\eta}_n - \eta| \|1 - M(\eta,\boldsymbol{t})\|_\infty + \widehat{\eta}_n \|M(\eta,\boldsymbol{t}) - \widehat{M}_n(\widehat{\eta}_n,\boldsymbol{t})\|_\infty + |\widehat{\eta}_n - \eta| \\
&\leq 2|\widehat{\eta}_n - \eta| + \widehat{\eta}_n \|M(\widehat{\eta}_n,\boldsymbol{t}) - \widehat{M}_n(\widehat{\eta}_n,\boldsymbol{t})\|_\infty + \widehat{\eta}_n \|M(\eta,\boldsymbol{t}) - M(\widehat{\eta}_n,\boldsymbol{t})\|_\infty \\
&\leq 2|\widehat{\eta}_n - \eta| + \widehat{\eta}_n \|\widehat{C}_n - C_\eta\|_\infty + \widehat{\eta}_n k \int_0^1 \|v^{\widehat{\eta}_n t_1} - v^{\eta t_1},\ldots,v^{\widehat{\eta}_n t_d} - v^{\eta t_d}\|_\infty \mathrm{d}v.
\end{aligned}$$

Each term on the right-hand side of this inequality tend to 0 a.s. under our assumptions and according to similar arguments to those used in Lemma A.1 for the last term. Thus (3.9) is established and the proof of Theorem 3.1 is now complete. □

A.5. Proof of Theorem 3.2

According to Guillou et al. (2014), η can be rewritten as

$$\eta = 2\left(1 - \frac{\mu_{1,2}}{\mu_{1,1}}\right).$$

A natural estimator can thus be obtained by replacing $Q_\eta(u)$ by the empirical version $G_n^{\leftarrow}(u)$ where $G_n(u) := G_n(u,\ldots,u)$. This entails

$$\widehat{\eta}_n = 2\left(1 - \frac{\widehat{\mu}_{1,2}}{\widehat{\mu}_{1,1}}\right),$$

where

$$\widehat{\mu}_{a,b} := \int_0^1 Q_n(u) u^a (-\log u)^b \mathrm{d}u.$$

Consequently, we can decompose the left-hand side of (3.11) as

$$\sqrt{n}(\widehat{\eta}_n - \eta) = 2\sqrt{n}\left(\frac{\mu_{1,2}}{\mu_{1,1}} - \frac{\widehat{\mu}_{1,2}}{\widehat{\mu}_{1,1}}\right) = 2\frac{\int_0^1 \mathbb{Q}_n(u)\gamma(u)\mathrm{d}u}{n^{-1/2}\mu_{1,1}\int_0^1 \mathbb{Q}_n(u)u(-\log u)\mathrm{d}u + \mu_{1,1}^2} =: 2\frac{N_n}{D_n}$$

with

$$\mathbb{Q}_n(u) := \sqrt{n}(Q_n(u) - Q_\eta(u)).$$

We start to study the numerator N_n. To this aim, we define the empirical and quantile processes:

$$\begin{aligned}
\widetilde{\mathbb{H}}_n(u) &:= \sqrt{n}(\widetilde{G}_n(u) - u), \quad u \in (0,1), \\
\widetilde{\mathbb{Q}}_n(u) &:= \sqrt{n}(\widetilde{Q}_n(u) - u), \quad u \in (0,1),
\end{aligned}$$

where for i.i.d. copies $U_1,\ldots,U_n$ of $U = G_\eta(\max(Z_1,\ldots,Z_d))$, we denote

$$\widetilde{G}_n(u) := \frac{1}{n}\sum_{i=1}^n \mathbb{1}(U_i \leq u), \quad u \in (0,1),$$

and as before $\widetilde{Q}_n := \widetilde{G}_n^{\leftarrow}$. Let $\dot{G}_\eta(y)$ and $\ddot{G}_\eta(y)$ be the first and second derivatives of $G_\eta(y)$ with respect to $y > 0$. The function defined in Theorem 3.2 is then equal to

$$\varphi(u) = \dot{G}_\eta(Q_\eta(u)), \quad u \in (0,1).$$

We can easily check that G_η satisfies the conditions of Theorem 3 in Csörgő and Révész (1978), whence

$$\sup_{u\in(0,1)} |\varphi(u)\,\mathbb{Q}_n(u) - \widetilde{\mathbb{Q}}_n(u)| = \mathrm{o}(1) \text{ a.s.} \tag{A.7}$$

and by Bahadur-Kiefer theorem (see e.g. Einmahl, 1996) we have

$$\sup_{u\in(0,1)} |\widetilde{\mathbb{Q}}_n(u) + \widetilde{\mathbb{H}}_n(u)| = \mathrm{o}(1) \text{ a.s..} \tag{A.8}$$

As by direct computations $\int_0^1 \left|\frac{\gamma(u)}{\varphi(u)}\right| \mathrm{d}u < \infty$, (A.7) and (A.8) entail

$$N_n = -\int_0^1 \widetilde{\mathbb{H}}_n(u)\frac{\gamma(u)}{\varphi(u)}\mathrm{d}u + \mathrm{o}(1) \text{ a.s..}$$

A similar reasoning implies that almost surely

$$D_n = -n^{-1/2}\mu_{1,1}\int_0^1 \widetilde{\mathbb{H}}_n(u)\frac{u(-\log u)}{\varphi(u)}\mathrm{d}u + \mu_{1,1}^2 + o(1) = \mu_{1,1}^2 + o(1).$$

Assembling N_n and D_n, we deduce that

$$\sqrt{n}(\widehat{\eta}_n - \eta) \quad = \quad -\frac{2}{\mu_{1,1}^2}\int_0^1 \mathbb{H}_n(Q_\eta(u),\ldots,Q_\eta(u))\frac{\gamma(u)}{\varphi(u)}\mathrm{d}u + \mathrm{o}(1) \text{ a.s.,}$$

where we used the fact that $\widetilde{\mathbb{H}}_n(u) = \mathbb{H}_n(Q_\eta(u),\ldots,Q_\eta(u))$. Thus (3.11) is established. The other statements of the theorem are direct consequences. □

Acknowledgements

We sincerely thank the editor, associate editor and the referees for their helpful comments and suggestions that led to substantial improvement of the paper. In addition, we are also gratefull to Carlo Baldassi for his help in parallelizing the code of the first simulation study, to Alexandra Ramos for her suggestions on the second simulation study and to Aad van der Vaart for having kindly replied to our technical questions.

References

Beirlant, J., Goegebeur, Y., Segers, J. and Teugels, J. (2004). *Statistics of Extremes: Theory and Applications* John Wiley & Sons Ltd., Chichester. MR2108013

Berghaus, B., Bücher, A. and Dette, H. (2013). Minimum distance estimators of the Pickands dependence function and related tests of multivariate extreme-value dependence *J. SFdS* **154** 116–137. MR3089619

Bücher, A. and Segers, J. (2017). On the maximum likelihood estimator for the generalized extreme-value distribution *Extremes* **20** 839–872. MR3737387

Capéraà, P., Fougères, A.-L. and Genest, C. (1997). A nonparametric estimation procedure for bivariate extreme value copulas *Biometrika* **84** 567–577. MR1603985

Genest, C. and Rémillard, B. (2004). Test of independence and randomness based on the empirical copula process *Test* **13** 335–369. MR2154005

Coles, S. G. (2001). *An Introduction to Statistical Modelling of Extreme Values* Springer, London. MR1932132

Csörgő, M. and Révész, P. (1978). Strong approximations of the quantile process *Ann. Statist.* **6** 882–894. MR0501290

de Haan, L. and Ferreira, A. (2006). *Extreme Value Theory: An Introduction* Springer, New York. MR2234156

Deheuvels, P. (1991). On the limiting behavior of the Pickands estimator for bivariate extreme-value distributions *Statist. Probab. Lett.* **12** 429–439. MR1142097

Draisma, G., Drees, H., Ferreira, A. and de Haan, L. (2004). Bivariate tail estimation: dependence in asymptotic independence *Bernoulli* **10** 251–280. MR2046774

Einmahl, J. H. J. (1996). A short and elementary proof of the main Bahadur-Kiefer theorem *Ann. Probab.* **24** 526–531. MR1387649

Falk, L., Hüsler, J. and Reiss, R. D. (2010). *Laws of Small Numbers: Extremes and Rare Events* (third ed.) Birkhäuser Boston. MR2732365

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

Gudendorf, G. and Segers, J. (2011). Nonparametric estimation of an extreme-value copula in arbitrary dimensions *J. Multivariate Anal.* **102** 37–47. MR2729418

Guillou, A., Naveau, P. and Schorgen, A. (2014). Madogram and asymptotic independence among maxima *REVSTAT* **12** 119–134. MR3228763

Hüsler, J. and Li, D. (2009). Testing asymptotic independence in bivariate extremes *J. Statist. Plann. Inference* **139** 990–998. MR2479843

Kiriliouk, A., Segers, J. and Warchoł, M. (2016). Nonparametric estimation of extremal dependence *In Extreme Value Modeling and Risk Analysis: Methods and Applications* 353–376.

Klüppelberg, C. and May, A. (2006). Bivariate extreme value distributions based on polynomial dependence functions *Math. Meth. Appl. Sci.* **29** 1467–1480. MR2247312

Ledford, A. W. and Tawn, J. A. (1996). Statistics for near independence in multivariate extreme values *Biometrika* **83** 169–187. MR1399163

Ledford, A. W. and Tawn, J. A. (1997). Modelling dependence within joint tail regions *J. Roy. Statist. Soc. Ser. B* **59** 475–499. MR1440592

Marcon, G., Padoan, S. A., Naveau, P., Muliere, P. and Segers, J. (2017). Multivariate nonparametric estimation of the Pickands dependence function using Bernstein polynomials *J. Statist. Plann. Inference* **183** 1-17. MR3589492

Maulik, K. and Resnick, S. (2004). Characterizations and examples of hidden regular variation *Extremes* **7** 31–67. MR2201191

Nelsen, R. B. (2006). *An introduction to copulas* Springer, New York. MR2197664

Pickands III, J. (1981). Multivariate extreme value distributions, with a discussion *Bull. Inst. Internat. Statist.* **49** 859–878, 894–902. MR0820979

Ramos, A. and Ledford, A. (2009). A new class of models for bivariate joint tails *J. R. Stat. Soc. Ser. B Stat. Methodol.* **71** 219–241. MR2655531

Ramos, A. and Ledford, A. (2011). An alternative point process framework for modeling multivariate extreme values *Comm. Statist. Theory Methods* **40** 2205–2224. MR2862707

Resnick, S. (2002). Hidden regular variation, second order regular variation and asymptotic independence *Extremes* **5** 303–336. MR2002121

Resnick, S. I. (2007). *Heavy-tail phenomena: Probabilistic and statistical modeling* Springer, New York. MR2271424

Segers S. (2012). Asymptotics of empirical copula processes under non-restrictive smoothness assumptions *Bernoulli* **18** 764–782. MR2948900

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

van der Vaart, A. W. (2000). *Asymptotic Statistics* Cambridge University Press. MR1652247

van der Vaart, A. W. and Wellner, J. A. (1996). *Weak Convergence and Empirical Processes. With Applications to Statistics* Springer Verlag, New York. MR1385671

van der Vaart, A. W and Wellner, J. A. (2007). Empirical processes indexed by estimated functions *IMS Lecture Notes Monogr. Ser.* **55** 234–252. MR2459942

Vettori, S., Huser, R. and Genton, M. G. (2017). A comparison of dependence function estimators in multivariate extremes *Stat. Comput.* DOI:10.1007/s11222-017-9745-7.

Wadsworth, J. L. and Tawn, J. A. (2013). A new representation for multivariate tail probabilities *Bernoulli* **19** 2689–2714. MR3160568

Wadsworth, J. L., Tawn, J. A., Davison, A. C. and Elton, D. M. (2017). Modelling across extremal dependence classes *J. R. Stat. Soc. Ser. B Stat. Methodol.* **79** 149–175. MR3597968

Zhang, D., Wells, M. T. and Peng, L. (2008). Nonparametric estimation of the dependence function for a multivariate extreme value distribution *J. Multivariate Anal.* **99** 577–588. MR2406072