

A comparison of dependence function estimators in multivariate extremes

Sabrina Vettori¹ · Raphaël Huser¹ · Marc G. Genton¹

Received: 14 March 2016 / Accepted: 8 April 2017 / Published online: 11 May 2017
© Springer Science+Business Media New York 2017

Abstract Various nonparametric and parametric estimators of extremal dependence have been proposed in the literature. Nonparametric methods commonly suffer from the curse of dimensionality and have been mostly implemented in extreme-value studies up to three dimensions, whereas parametric models can tackle higher-dimensional settings. In this paper, we assess, through a vast and systematic simulation study, the performance of classical and recently proposed estimators in multivariate settings. In particular, we first investigate the performance of nonparametric methods and then compare them with classical parametric approaches under symmetric and asymmetric dependence structures within the commonly used logistic family. We also explore two different ways to make nonparametric estimators satisfy the necessary dependence function shape constraints, finding a general improvement in estimator performance either (i) by substituting the estimator with its greatest convex minorant, developing a computational tool to implement this method for dimensions $D \geq 2$ or (ii) by projecting the estimator onto a subspace of dependence functions satisfying such constraints and taking advantage of Bernstein–Bézier polynomials. Implementing the convex minorant method

leads to better estimator performance as the dimensionality increases.

Keywords Asymmetric logistic model · Componentwise maxima · Convexity · Copula · Greatest convex minorant · Nonparametric and parametric estimators · Pickands dependence function

1 Introduction

Classical theory characterising the dependence of multivariate extremes is based on the limiting distribution of componentwise maxima. Let $\{\mathbf{Y}_i = (Y_{i;1}, \dots, Y_{i;D})^\top\}_{i=1}^n$ be a sequence of independent and identically distributed (i.i.d.) copies of the D -dimensional random vector \mathbf{Y} with distribution function (d.f.) F , and let $\mathbf{M}_n = (M_{n;1}, \dots, M_{n;D})^\top = \left(\max_{1 \leq i \leq n} Y_{i;1}, \dots, \max_{1 \leq i \leq n} Y_{i;D} \right)^\top$ denote the vector of multivariate componentwise maxima. If for some sequences of vectors $\mathbf{a}_n = (a_{n;1}, \dots, a_{n;D})^\top \in \mathbb{R}_+^D$ and $\mathbf{b}_n = (b_{n;1}, \dots, b_{n;D})^\top \in \mathbb{R}^D$, the renormalised vector of componentwise maxima $\mathbf{M}_n^* = \mathbf{a}_n^{-1}(\mathbf{M}_n - \mathbf{b}_n)$ converges in distribution to the random vector \mathbf{Z} with limiting joint d.f. G and non-degenerate margins $G_d, d = 1, \dots, D$, i.e.

$$\begin{aligned} \Pr(\mathbf{M}_n^* \leq \mathbf{z}) &= \Pr(M_{n;1}^* \leq z_1, \dots, M_{n;D}^* \leq z_D) \\ &= F^n(\mathbf{a}_n \mathbf{z} + \mathbf{b}_n) \rightarrow G(\mathbf{z}), \quad n \rightarrow \infty, \end{aligned}$$

then G is a D -dimensional extreme-value d.f. and F is said to belong to the max-domain of attraction (MDA) of G . Without loss of generality (Resnick 1987, p. 265), we can assume unit Fréchet marginal distributions, i.e. $G_d(z_d) = \exp(-1/z_d)$ for $z_d > 0$. An interesting characteristic of the class of

Electronic supplementary material The online version of this article (doi:10.1007/s11222-017-9745-7) contains supplementary material, which is available to authorized users.

✉ Sabrina Vettori
sabrina.vettori@kaust.edu.sa

Raphaël Huser
raphael.huser@kaust.edu.sa

Marc G. Genton
marc.genton@kaust.edu.sa

¹ CEMSE Division, King Abdullah University of Science and Technology, Thuwal 23955-6900, Saudi Arabia

extreme-value d.f. is that it coincides with the class of max-stable d.f. (de Haan and Resnick 1977), where we define a d.f. G to be max-stable if, for every positive integer k , there exist vectors $\alpha_k = (\alpha_{k;1}, \dots, \alpha_{k;D})^\top \in \mathbb{R}_+^D$ and $\beta_k = (\beta_{k;1}, \dots, \beta_{k;D})^\top \in \mathbb{R}^D$ such that

$$G^k(\alpha_{k;1}z_1 + \beta_{k;1}, \dots, \alpha_{k;D}z_D + \beta_{k;D}) = G(z_1, \dots, z_D), \quad (z_1, \dots, z_D) \in \mathbb{R}^D.$$

The d.f. G may be represented as

$$G(\mathbf{z}) = \exp\{-V(\mathbf{z})\},$$

$$V(\mathbf{z}) = \int_{S_D} \max_{1 \leq d \leq D} \left(\frac{\omega_d}{z_d} \right) dH(\omega), \quad \mathbf{z} \in \mathbb{R}_+^D, \quad (1)$$

where V is a homogeneous function of order -1 , i.e. $V(c\mathbf{z}) = c^{-1}V(\mathbf{z})$, $c > 0$, called exponent function, and H is a finite spectral measure on the unit simplex $S_D = \{\omega \in [0, 1]^D : \sum_{d=1}^D \omega_d = 1\}$ (de Haan and Resnick 1977; de Haan 1984). The spectral measure H satisfies the mean constraints $\int_{S_D} \omega_d dH(\omega) = 1$, $d = 1, \dots, D$, and it determines the extremal dependence structure. Another representation of G is via the Pickands (1981) dependence function A , i.e.

$$V(\mathbf{z}) = \left(\sum_{d=1}^D z_d^{-1} \right) \times A \left(\frac{z_1^{-1}}{z_1^{-1} + \dots + z_D^{-1}}, \dots, \frac{z_D^{-1}}{z_1^{-1} + \dots + z_D^{-1}} \right),$$

$$\mathbf{z} \in \mathbb{R}_+^D, \quad (2)$$

where $A(\omega)$ for $\omega = (\omega_1, \dots, \omega_D)^\top$, $\omega \in S_D$, is a function on the unit simplex S_D satisfying

- A1 $A(\omega)$ is convex;
- A2 $D^{-1} \leq \max(\omega) \leq A(\omega) \leq 1$ for all $\omega \in S_D$;
- A3 $A(\mathbf{e}_d) = 1$ for the boundary points of S_D , $\mathbf{e}_d = (0, \dots, 0, 1, 0, \dots, 0)^\top$ for $d = 1, \dots, D$;
- A4 $A(\omega)$ is fully D -max decreasing; see Ressel (2013).

The lower and upper bounds of the Pickands function A in condition A2 correspond to the case of perfect dependence and independence between the coordinates, respectively, and the condition A4 ensures that the dependence function A is self-consistent with lower dimensions (see also Schlather and Tawn 2002, 2003; Gudendorf and Segers 2011; Beirlant et al. 2004, p. 257). Although the class of spectral measures or valid Pickands dependence functions cannot be summarised by a finite number of parameters, in practice it may be convenient to restrict ourselves to flexible parametric subclasses. Several models have been proposed in the literature. One of the mostly used multivariate extreme-value distributions is the logistic model (Gumbel 1960a,b) or its asymmetric extension (Tawn 1988, 1990), both described in Appendix

2. For a useful summary of existing models, see Beirlant et al. (2004, p. 304-313).

When inference is performed in a nonparametric fashion, some of the widely used estimators include the ones introduced by Pickands (1981), Capéràa et al. (1997), Naveau et al. (2009), Guillothe (2008), Genest and Segers (2009), Bücher et al. (2011), and Cormier et al. (2014) for the bivariate case and Zhang et al. (2008), Gudendorf and Segers (2011, 2012), Berghaus et al. (2013) and Marcon et al. (2017) for dimensions $D > 2$. Nonparametric estimators of the spectral measure for threshold exceedances, which can be transposed into valid estimators of the dependence function, were introduced, for example, by Einmahl et al. (2001) and Einmahl and Segers (2009). In this paper we conduct a vast simulation experiment comparing several of these estimators with the aim of providing a greater understanding of nonparametric estimators under a unified simulation setting, focusing on componentwise maxima. As most nonparametric estimators do not satisfy the necessary dependence function shape constraints, different modifications have been proposed in the literature. Endpoint corrections were introduced by Deheuvels (1991), Hall and Tajvidi (2000) and Segers (2007), while Pickands (1981) proposed a method based on the greatest convex minorant in order to comply with the dependence function convexity constraint. Recently, Marcon et al. (2017) suggested another modification which uses Bernstein-Bézier polynomial (Sauer 1991; Lorenz 1986) representations owing to their excellent shape-preserving properties (see Camicer and Peña 1993). Other approaches involve spline smoothing under constraints (Hall and Tajvidi 2000; Cormier et al. 2014) and orthogonal projection (Fils-Villetard et al. 2008; Gudendorf and Segers 2012). In addition to satisfying the shape constraints, the estimator obtained by applying the methodology proposed by Gudendorf and Segers (2012) admits an integral representation in terms of the discrete spectral measure and therefore satisfies all constraints A1, A2, A3 and A4. However, in practice, computing the multivariate integrals involved in the projection methodology of Gudendorf and Segers (2012) might be very computationally demanding for dimensions $D > 3$ and the code has not been made available yet. The convex minorant modification proposed by Pickands (1981) is so far the most commonly implemented when the dimension is $D = 2$. One of the main contributions of this paper is to develop a computational tool for allowing the implementation of the convex minorant method in higher dimensions. In particular, we approximate the value of the dependence function estimator at each grid point by linear interpolation of its value at the points forming the convex hull, taking advantage of the barycentric coordinates system to compute the weights to be assigned to each of the points in the convex hull. We then compare the finite-sample properties for dimensions $D \geq 2$ of the adjusted nonparametric estimators

obtained as the convex minorant counterparts or by applying the Bernstein–Bézier polynomial approach. Generally, we expect nonparametric estimators to be more flexible than parametric ones while suffering from the curse of dimensionality; however, this has never been quantified in the context of extremes. Another contribution of this paper is to explore the performance of nonparametric estimators with respect to parametric approaches. For a comparison of the main parametric methods, see, for example, [Huser et al. \(2016\)](#).

The paper is structured as follows. Nonparametric and parametric multivariate inference methods are presented in Sect. 2. The simulation study assessing the performance of various dependence estimators is described in Sect. 3. The main findings are summarised in Sect. 4.

2 Dependence function estimation

2.1 Generalities

Inference for observed block maxima $\{\mathbf{m}_i = (m_{i;1}, \dots, m_{i;D})^\top\}_{i=1}^M$ involves essentially two aspects: (i) estimation of marginal distributions and standardisation of the margins to a common (here, unit Fréchet) scale and (ii) estimation of the dependence structure. Since in this work the main focus is on the latter, we perform the marginal estimation separately by using the empirical d.f.

$$\hat{G}_d(z_d) = M^{-1} \sum_{i=1}^M \mathbb{I}(m_{i;d} \leq z_d), \quad d = 1, \dots, D, \quad (3)$$

where $\mathbb{I}(E)$ is the indicator function of event E , and then standardise the margins to the unit Fréchet scale by applying the transformation

$$\tilde{m}_{i;d} = -1/\log \left\{ \hat{G}_d(m_{i;d}) \right\}, \quad i = 1, \dots, M, \quad d = 1, \dots, D. \quad (4)$$

In the following sections, we discuss nonparametric and parametric approaches to estimate the extremal dependence via the dependence function (2).

2.2 Nonparametric approach

Table 1 provides a summary of the main characteristics for a quite exhaustive list of widely used nonparametric estimators proposed in the literature for componentwise maxima. We also report variants of these estimators, proposed in order to impose endpoint corrections and extensions to more general frameworks. The finite-sample performances of nonparametric estimators are assessed in the simulation study in Sect. 3. For the sake of clarity and conciseness, we cannot consider all nonparametric estimators presented in Table 1, and therefore, we select the estimators that present dissimilar characteristics and are available in an R package. In particular, we consider the rank-based ([Genest and Segers 2009](#); [Gudendorf and Segers 2012](#)) nonparametric estimators of the [Pickands \(1981\)](#) and [Capéràa et al. \(1997\)](#) class, i.e. \hat{A}_n^{RP} , \hat{A}_n^{RHT} , \hat{A}_n^{RCFG} and \hat{A}_n^{RMCFG} , the estimator \hat{A}_n^{MMD} based on the idea of the madogram ([Matheron 1987](#); [Coo-ley et al. 2006](#)) and the estimator derived by [Cormier et al.](#)

Table 1 List of the most widely used nonparametric estimators for multivariate extremes and their variants

| Reference | $\hat{A}_n(\omega)$ | $D > 2$ | Constraints | R package | Variant of |
|---|---------------------|---------|-------------|-------------|------------|
| Pickands (1981)^a | P | ✓ | ✗ | evd | |
| Deheuvels (1991)^a | D | ✓ | A3 | evd | P |
| Hall and Tajvidi (2000)^a | HT | ✓ | A2,A3 | evd | P |
| Capéràa et al. (1997) | CFG | ✗ | A2 | evd | |
| Jiménez et al. (2001) | JCFG | ✗ | ✓ | ✗ | CFG |
| Segers (2007) | S- | ✗ | + | ✗ | P & CFG |
| Genest and Segers (2009)^b | R- | ✗ | + | copula | P & CFG |
| Gudendorf and Segers (2011)^c | MCFG | ✓ | A2 | copula | CFG |
| Naveau et al. (2009) | MD | ✗ | ✗ | ✗ | |
| Marcon et al. (2017) | MMD | ✓ | A1, A2,A3 | ExtremalDep | MD |
| Cormier et al. (2014) | COBS | ✗ | ✓ | cobs | |
| Bücher et al. (2011) | BDV | ✗ | ✗ | ✗ | |
| Berghaus et al. (2013) | MBDV | ✓ | ✗ | ✗ | BDV |

For each estimator, we report whether it: has been defined or implemented in the multivariate case; satisfies the constraints A1, A2, A3, A4; is contained in an R package; and is a variant or extension of another estimator. The estimators considered in the simulation study in Sect. 3 are highlighted in bold

^a See [Zhang et al. \(2008\)](#) for multivariate version

^b See [Gudendorf and Segers \(2012\)](#) for multivariate version

^c See also [Zhang et al. \(2008\)](#). “+” the constraints satisfied are the same as for the original estimator

(2014) using constrained B-spline smoothing, *i.e.* \hat{A}_n^{COBS} . More details on the selected estimators are available in the original references and in Appendix 1. Most of the aforementioned estimators do not satisfy the dependence function constraints and in the following section we discuss two possible modifications to comply at least with the convexity and boundary constraints A1, A2 and A3.

2.3 Modifications to satisfy the constraints

A first modification was originally proposed by Pickands (1981) and consists in replacing the nonparametric estimator $\hat{A}_n(\omega)$ with its greatest convex minorant in order to comply with the convexity constraint A1; see also Pickands (1989), Deheuvels (1991), Hall and Tajvidi (2000) and Capéràa and Fougères (2000). So far, this method has only been used for bivariate data, perhaps due to the computational difficulties related to its implementation in dimensions $D > 2$. In this paper, we overcome these difficulties and build a methodology to compute the convex minorant of the nonparametric dependence function estimators in higher dimensions by taking advantage of barycentric interpolation, a tool commonly used in many computer graphics applications. The convex hull of a set of $(D - 1)$ -dimensional input points on the simplex S_D is the union of all the “hyper-triangles” determined by at most D vertices. Let $\omega_d \in \mathbb{R}^{D-1}$, $d = 1, \dots, D$, be D vertices of a “hyper-triangle” Q on the simplex S_D . The barycentric coordinates $\lambda_d(\omega_0) \geq 0$, $d = 1, \dots, D$, $\sum_{d=1}^D \lambda_d(\omega_0) = 1$ of an arbitrary point $\omega_0 \in \mathbb{R}^{D-1}$ inside Q represent fictive masses placed at the vertices of Q such that ω_0 is the centre of mass. Therefore, we can evaluate the value of the nonparametric estimator $\hat{A}_n(\omega)$ at any grid point lying inside a $(D - 1)$ -dimensional “hyper-triangle” forming the convex hull by the interpolation of its value at the vertices using the barycentric coordinates as weights, *i.e.*

$$\hat{A}_n^{CONV}(\omega) = \sum_{d=1}^D \lambda_d(\omega) \hat{A}_n(\omega_d). \tag{5}$$

To determine whether a $(D - 1)$ -dimensional input point ω_0 lies inside a specific “hyper-triangle” formed by the vertices

$$\begin{aligned} \omega_1 &= (\omega_{1;1}, \dots, \omega_{1;D-1})^\top, \\ \omega_2 &= (\omega_{2;1}, \dots, \omega_{2;D-1})^\top, \\ &\vdots \\ \omega_D &= (\omega_{D;1}, \dots, \omega_{D;D-1})^\top, \end{aligned}$$

or on its edge, we verify whether the determinants

$$D_0 = \begin{vmatrix} \omega_{1;1} & \cdots & \omega_{1;D-1} & 1 \\ \omega_{2;1} & \cdots & \omega_{2;D-1} & 1 \\ \vdots & \ddots & \vdots & \vdots \\ \omega_{D;1} & \cdots & \omega_{D;D-1} & 1 \end{vmatrix},$$

$$D_1 = \begin{vmatrix} \omega_{0;1} & \cdots & \omega_{0;D-1} & 1 \\ \omega_{2;1} & \cdots & \omega_{2;D-1} & 1 \\ \vdots & \ddots & \vdots & \vdots \\ \omega_{D;1} & \cdots & \omega_{D;D-1} & 1 \end{vmatrix}, \dots, D_D = \begin{vmatrix} \omega_{1;1} & \cdots & \omega_{1;D-1} & 1 \\ \omega_{2;1} & \cdots & \omega_{2;D-1} & 1 \\ \vdots & \ddots & \vdots & \vdots \\ \omega_{0;1} & \cdots & \omega_{0;D-1} & 1 \end{vmatrix}$$

have the same sign. If this is the case, then ω_0 can be expressed as the convex combination of the vertices ω_d , $d = 1, \dots, D$, *i.e.* $\omega_0 = \sum_{d=1}^D \lambda_d(\omega_0) \omega_d$, and the barycentric coordinates can be obtained as $\lambda_d(\omega_0) = D_d / D_0 > 0$. In principle, this method can be applied to any estimator of the dependence function and its performance is tested in Sect. 3.4 for dimensions $D = 3, 4, 5$. The algorithm code (in C++ integrated in R using the package RCPP) is available in the Supplementary Material. To compute the convex hull of a given set of input points, we used the algorithm available in the R package geometry. However, computing the convex hull in dimensions $D > 7$ using the available algorithms is time consuming and requires significant memory space, which makes the application of this method too computationally demanding for the moment. Moreover, the endpoints correction

$$\min \left\{ 1, \max \left(\hat{A}_n^{CONV}(\omega), \omega_1, \dots, \omega_D \right) \right\} \tag{6}$$

can be applied to ensure that the boundary constraints A2, A3 are satisfied.

An alternative method, recently proposed by Marcon et al. (2017), involves the projection of the nonparametric estimator $\hat{A}_n(\omega)$ onto the subspace \mathcal{A} of functions satisfying the constraints A1, A2 and A3 expressible through linear combinations of k th-order Bernstein–Bézier polynomials (Sauer 1991; Lorenz 1986). The Bernstein–Bézier polynomial approximation of a function $A(\omega)$ is

$$B_A(\omega; k) = \sum_{l \in L_k} \beta_l b_l(\omega; k), \quad \beta_l \in \mathbb{R}, \omega \in S_D, \tag{7}$$

where the $b_l(\cdot; k)$ forms the l th element of the basis of Bernstein–Bézier polynomials of degree k , which are defined as continuous functions with values in $[0, 1]$. The projection estimator is based on a first guess \hat{A}_1 and then obtained as the solution of the optimisation problem

$$\hat{A}_n^{BP} = \arg \min_{A \in \mathcal{A}} \|\hat{A}_1 - A\|_2, \tag{8}$$

where the minimum is taken over all possible measurable functions in \mathcal{A} , which is a closed and convex subset of

$L^2(S_D)$, the space of square integrable functions over the unit simplex. Consider a nested sequence of constrained multivariate Bernstein–Bézier polynomial families $\mathcal{A}_k \subseteq \mathcal{A}$; restrictions on the β_l in (7) can be imposed such that each member of \mathcal{A}_k satisfies the necessary dependence function conditions; for more details, see Marcon et al. (2017).

2.4 Parametric approach

Parametric estimation of the dependence function can be performed by maximising the pairwise likelihood under the working assumption of mutual independence between pairs of variables (see, for example, Varin and Vidoni 2005; Varin 2008; Padoan et al. 2010 or Davison and Gholamrezaee 2012). In particular, the log-pairwise likelihood function considered in this paper is of the form

$$\ell_{\text{pair}}(\boldsymbol{\alpha}) = \sum_{i=1}^M \sum_{1 \leq d_1 \leq d_2 \leq D} \log g \left\{ (\tilde{m}_{i;d_1}, \tilde{m}_{i;d_2})^\top; \boldsymbol{\alpha} \right\},$$

where $\boldsymbol{\alpha}$ is the vector of unknown parameters, $(\tilde{m}_{i;d_1}, \tilde{m}_{i;d_2})^\top$ are standardised sample maxima transformed to the Fréchet scale, see (4), and g denotes the bivariate density function:

$$\begin{aligned} g \left\{ (\tilde{m}_{i;d_1}, \tilde{m}_{i;d_2})^\top; \boldsymbol{\alpha} \right\} &= \left[V_1 \left\{ (\tilde{m}_{i;d_1}, \tilde{m}_{i;d_2})^\top; \boldsymbol{\alpha} \right\} V_2 \left\{ (\tilde{m}_{i;d_1}, \tilde{m}_{i;d_2})^\top; \boldsymbol{\alpha} \right\} \right. \\ &\quad \left. - V_{12} \left\{ (\tilde{m}_{i;d_1}, \tilde{m}_{i;d_2})^\top; \boldsymbol{\alpha} \right\} \right] \\ &\quad \times \exp \left[-V \left\{ (\tilde{m}_{i;d_1}, \tilde{m}_{i;d_2})^\top; \boldsymbol{\alpha} \right\} \right], \end{aligned}$$

with V_1 , V_2 and V_{12} denoting the partial and mixed derivatives of the bivariate exponent function V stemming from (1). The marginal and dependence parameters can be estimated either separately or jointly, the latter approach allowing a better estimation of the global uncertainty. An estimate of the Pickands dependence function may then be obtained through (2). Under classical regularity conditions, maximum composite likelihood estimators are strongly consistent and asymptotically Gaussian but present a loss in efficiency with respect to maximum likelihood estimators. In the bivariate case, the log-pairwise likelihood function boils down to the full log-likelihood function, *i.e.* $\ell_{\text{pair}}(\boldsymbol{\alpha}) = \sum_{i=1}^M \log g \left\{ (\tilde{m}_{i;d_1}, \tilde{m}_{i;d_2})^\top; \boldsymbol{\alpha} \right\}$.

3 Simulation study

3.1 Simulation design

We investigate the finite-sample performance of nonparametric and parametric methods on $N = 1000$ independent

multivariate datasets of various dimensions ($D = 2, 3, 4, 5$). For finite-sample sizes, extremes not yet being in their limiting regime (1), it makes sense to assess the subasymptotic bias of the different estimators that is due to being in the *MDA*. Therefore, in this simulation experiment we complement previous comparison studies by simulating the data from distributions in the *MDA* of the multivariate extreme-value distribution G characterised by logistic and asymmetric logistic dependence structures, respectively. A similar setting was considered in Huser et al. (2016), where the goal was to compare different parametric methods. In particular, data are generated from the so-called outer power Clayton copula (Nelsen 2006), *i.e.*

$$C_1(\mathbf{u}) = \varphi \left\{ \sum_{d=1}^D \varphi^{-1}(u_d) \right\}, \tag{9}$$

where $\mathbf{u} = (u_1, \dots, u_D)^\top$, with generator $\varphi(t) = (t^\alpha + 1)^{-1}$. Indeed, it can be shown (Fougères 2004, p. 376) that if $\mathbf{U} = (U_1, \dots, U_D)^\top \sim C_1(\mathbf{u})$, then the distribution of the random vector $\mathbf{X} = \{-1/\log(U_1), \dots, -1/\log(U_D)\}^\top$ is in the *MDA* of the logistic distribution characterised by the dependence function in (24). Moreover, we introduce asymmetry in the dependence structure by generating the data from the copula

$$C_2(\mathbf{u}) = C_1 \left(u_1^{\phi_1}, \dots, u_D^{\phi_D} \right) \prod_{d=1}^D u_d^{1-\phi_d} \tag{10}$$

instead of $C_1(\mathbf{u})$. The construction of such copula is based on Khoudraji’s device proposed by Khoudraji (1995) in a bivariate setting and used by Liebscher (2008) afterwards in higher dimensions; see also Genest et al. (2011) and Li and Genton (2013). Note that if $\mathbf{U}^* = (U_1^*, \dots, U_D^*)^\top \sim C_2(\mathbf{u}^*)$, then the distribution of the random vector $\mathbf{X}^* = \{-1/\log(U_1^*), \dots, -1/\log(U_D^*)\}^\top$ is in the *MDA* of a simplified version of the asymmetric logistic distribution characterised by the dependence function in (25), *i.e.*

$$A(\boldsymbol{\omega}) = \left\{ \sum_{d=1}^D (\omega_d \phi_d)^{1/\alpha} \right\}^\alpha + \sum_{d=1}^D \omega_d (1 - \phi_d), \tag{11}$$

with dependence parameters $0 < \alpha \leq 1$ and asymmetry parameters $0 \leq \phi_d \leq 1, d = 1, \dots, D$. The proof is given in Appendix 3. Moreover, we further specify ϕ_d to be equal to ϕ^{d-1} so as to have only one asymmetry parameter to estimate and meaning that the extent of asymmetry augments as the “distance” between vector components increases. The simulations from the copula (9) and its asymmetric extension (10) were performed using the R package `copula` (Hofert et al. 2014), considering a wide range of dependence scenarios and sample size $n = 10000$ and extracting $M = 100$ multivariate componentwise maxima from each dataset. The extremal

coefficient, used in the simulation study as a summary of extremal dependence (Smith 1990), is defined as

$$\theta = V(\mathbf{1}) = DA(D^{-1}\mathbf{1}) \in [1, D], \tag{12}$$

where $\mathbf{1}$ is a D -dimensional vector of ones. The value of θ decreases as the dependence strength between the margins increases. In particular, independence and perfect dependence occur if and only if $\theta = D$ or $\theta = 1$, respectively. Under each dependence scenario, we estimate the dependence function using the different methods outlined in Sect. 2 and compute a Monte Carlo approximation of the root mean integrated squared error (RMISE) from N replicates. The RMISE of a given estimator $\hat{A}_n(\omega)$, defined as

$$\text{RMISE} = \left(E \left[\int_{S_D} \left\{ \hat{A}_n(\omega) - A(\omega) \right\}^2 d\omega \right] \right)^{1/2}, \tag{13}$$

can be shown to be equal to

$$\text{RMISE} = \left\{ \underbrace{\int_{S_D} \left[E \left\{ \hat{A}_n(\omega) \right\} - A(\omega) \right]^2 d\omega}_{\text{RISB}^2} + \underbrace{\int_{S_D} E \left(\left[\hat{A}_n(\omega) - E \left\{ \hat{A}_n(\omega) \right\} \right]^2 \right) d\omega}_{\text{RIV}^2} \right\}^{1/2} \tag{14}$$

where RISB and RIV denote the root integrated squared bias and the root integrated variance, respectively (Gentle 2002, p. 145). The integrands in (13) and (14) are evaluated on a fine partition of S_D sampled on a regular grid and a sum of these

values is used to approximate the integrals. Computational time for the implementation of both the convex minorant (5) and the projection (8) methods is highly influenced by the number of grid points chosen to discretise S_D , which grows with the dimension D , but it is generally fast up to $D = 4$. In the following sections, we present the results of the simulation study focusing first on the bivariate case and then on higher dimensions.

3.2 Bivariate case: nonparametric estimators comparison

Considering dependence scenarios ranging from strong dependence to independence, for data with distribution in the MDA of the logistic distribution (24) with $D = 2$, we display in Fig. 1 the RMISE, RISB and RIV computed for the convex minorant (5) counterparts of the nonparametric estimators \hat{A}_n^{RP} (17), \hat{A}_n^{RHT} (18), \hat{A}_n^{RCFG} (20) and \hat{A}_n^{MMD} (22), denoted as $\hat{A}_n^{\text{RP-CONV}}$ and the constrained B-spline smoothing estimator \hat{A}_n^{COBS} (23). The RISB is generally quite small; indeed all the estimators considered are asymptotically unbiased. In particular, the RISB of the $\hat{A}_n^{\text{RHT-CONV}}$, $\hat{A}_n^{\text{RCFG-CONV}}$ and $\hat{A}_n^{\text{MMD-CONV}}$ estimators is close to 0 for strong dependence and increases as the dependence strength decreases, i.e. as the bivariate extremal coefficient θ approaches 2. This is consistent with results obtained by Huser et al. (2016) in the parametric framework. By contrast, the RISB of the $\hat{A}_n^{\text{RP-CONV}}$ estimator is only close to 0 when the dependence is rather weak, showing higher RISB for mild dependence and close to independence. This non-monotonic behaviour can be attributed to a change of sign in the bias as we move

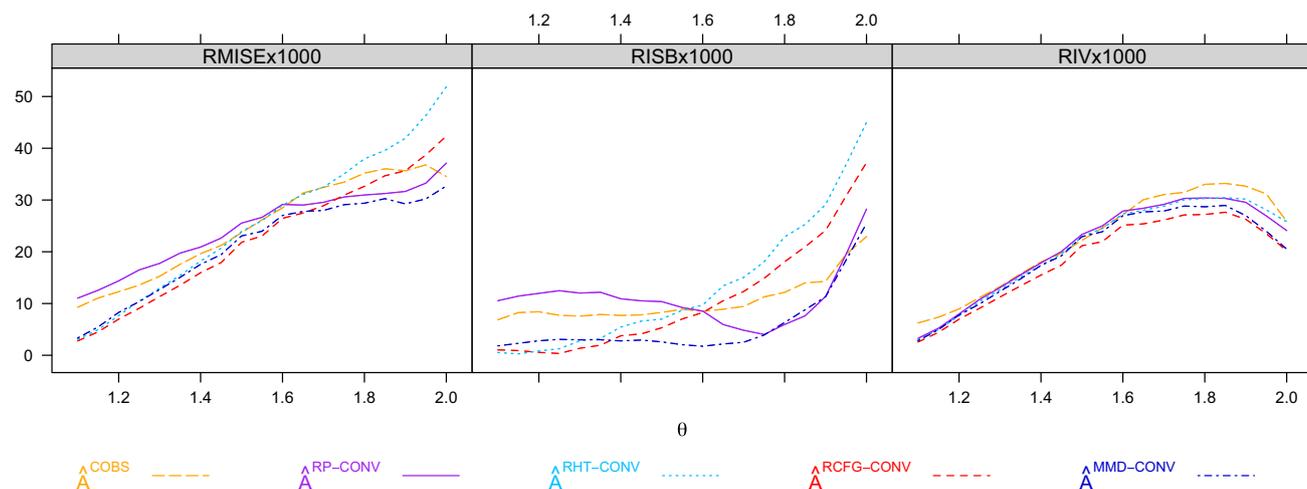


Fig. 1 Root mean integrated squared error (RMISE \times 1000), root integrated squared bias (RISB \times 1000) and root integrated variance (RIV \times 1000) computed for data with distribution in the MDA of the logistic distribution (24) with $D = 2$ are plotted against the extremal

coefficient θ . The estimators $\hat{A}_n^{\text{RP-CONV}}$, $\hat{A}_n^{\text{RHT-CONV}}$, $\hat{A}_n^{\text{RCFG-CONV}}$ and $\hat{A}_n^{\text{MMD-CONV}}$ based on (17), (18), (20) and (22) modified using the convex minorant method (5) and the estimator \hat{A}_n^{COBS} based on (23)

from complete dependence to independence. In other words, the estimated dependence function tends to lie above the true one for roughly $1 \leq \theta \leq 1.7$ and below it for $1.7 \leq \theta \leq 2$. In the case of independence, however, the true curve is situated on the upper edge of the triangle defined by the vertices $\{(0, 1); (1/2, 1/2); (1, 0)\}$ and all estimated curves are located below. The RMISE of the estimator \hat{A}_n^{COBS} instead is almost constant for strong to mild dependence situations and increases in case of weaker dependence strengths. Generally, the estimators $\hat{A}_n^{\text{RCFG-CONV}}$ and $\hat{A}_n^{\text{RHT-CONV}}$ are the least biased when the dependence is rather strong, whereas the estimator \hat{A}_n^{MMD} is the least biased for mild to weak dependence, together with the estimator \hat{A}_n^{RP} for weak dependence and with the estimator \hat{A}_n^{COBS} only very close to independence. In terms of variability, all estimators show a similar behaviour, although the estimator $\hat{A}_n^{\text{RCFG-CONV}}$ is slightly more efficient. As a result, in terms of RMISE, in cases of strong to mild dependence the estimators $\hat{A}_n^{\text{RCFG-CONV}}$, $\hat{A}_n^{\text{RHT-CONV}}$ and $\hat{A}_n^{\text{MMD-CONV}}$ present very similar performances, though the estimator $\hat{A}_n^{\text{RCFG-CONV}}$ slightly outperforms the others for $1 \leq \theta \leq 1.6$, whereas the estimator $\hat{A}_n^{\text{MMD-CONV}}$ shows the best performance for $1.6 \leq \theta \leq 2$.

Considering the same setting, in Fig. 2 we compare the convex minorant (5) (\hat{A}_n^{CONV}) and Bernstein–Bézier polynomial projection (8) (\hat{A}_n^{BP}) counterparts of the nonparametric estimators \hat{A}_n^{RP} (17), \hat{A}_n^{RHT} (18), \hat{A}_n^{RCFG} (20) and \hat{A}_n^{MMD} (22).

For comparison purposes, we also report the RMISE of the nonparametric estimators to which we imposed only the endpoint corrections in (6).

The issue of choosing an optimal polynomial degree k for the implementation of the Bernstein–Bézier polynomial method was addressed by calculating the value k that minimises the RMISE for each dependence scenario from a range of values from 1 to 24. The optimal k values resulted to be the same for all the estimators considered in this simulation study and are reported in Table 2. As expected, a large k is needed for very strong dependence scenarios, whereas lower values of k yield good approximations when the dependence strength is mild to weak. In practice, the true value of θ is unknown and therefore also the optimal k . However, we found that fixing the Bernstein–Bézier polynomial degree $k = 6$ is generally sufficiently large to obtain a reasonable approximation for a wide range of dependence situations but not too large so as to avoid excessive overfitting. A remedy to the fact that k is unknown has been provided by Marcon et al. (2016) in a Bayesian context, favouring low values in the case of weak dependence and high values in the case of strong dependence in accordance with our findings.

From Fig. 2, we observe that the estimators satisfying the convexity constraint generally show the best performances. Moreover, when the Bernstein–Bézier polynomial method is applied instead of the convex minorant modification, the

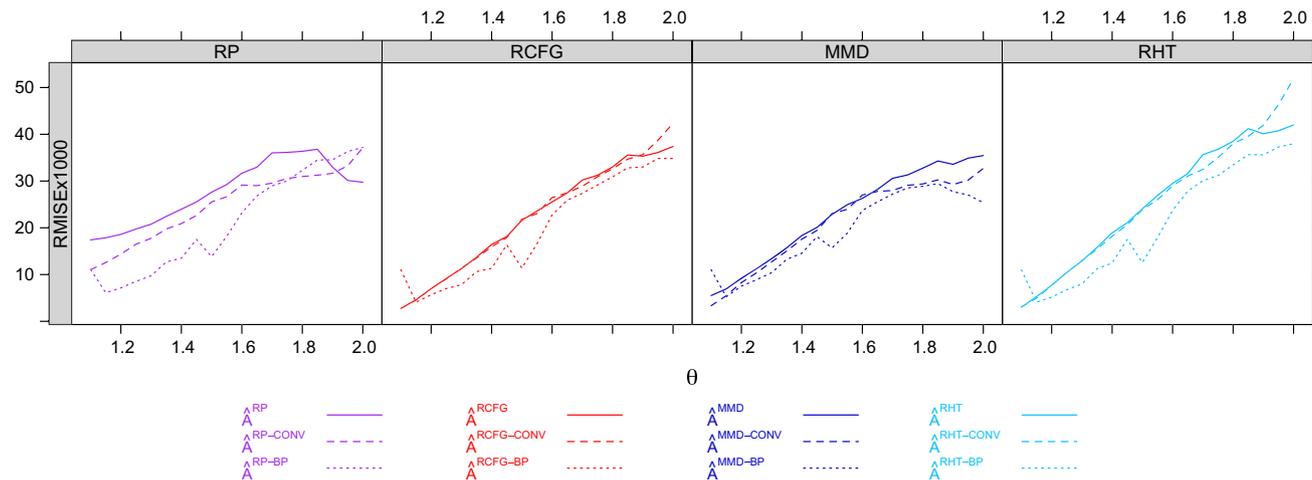


Fig. 2 Root mean integrated squared error (RMISE \times 1000) computed for data with distribution in the MDA of the logistic distribution (24) with $D = 2$ is plotted against the extremal coefficient θ . The estimators \hat{A}_n^{RP} , \hat{A}_n^{RHT} , \hat{A}_n^{RCFG} and \hat{A}_n^{MMD} are based on (17), (18), (20) and (22)

to which we imposed only the endpoint corrections (6). Their convex minorant (\hat{A}_n^{CONV}) and projection (\hat{A}_n^{BP}) counterparts are obtained applying the modifications (5) or (8) with k reported in Table 2, respectively

Table 2 Bernstein–Bézier polynomial degree k (8) that minimises the RMISE considering a range of values from 1 to 24 reported for different values of the extremal coefficient θ

| θ | 1.10 | 1.15 | 1.20 | 1.25 | 1.30 | 1.35 | 1.40 | 1.45 | 1.50 | 1.55 | 1.60 | 1.65 | 1.70 | 1.75 | 1.80 | 1.90 | 1.95 | 2.00 |
|----------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|
| k | 24 | 24 | 18 | 12 | 6 | 6 | 6 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 |

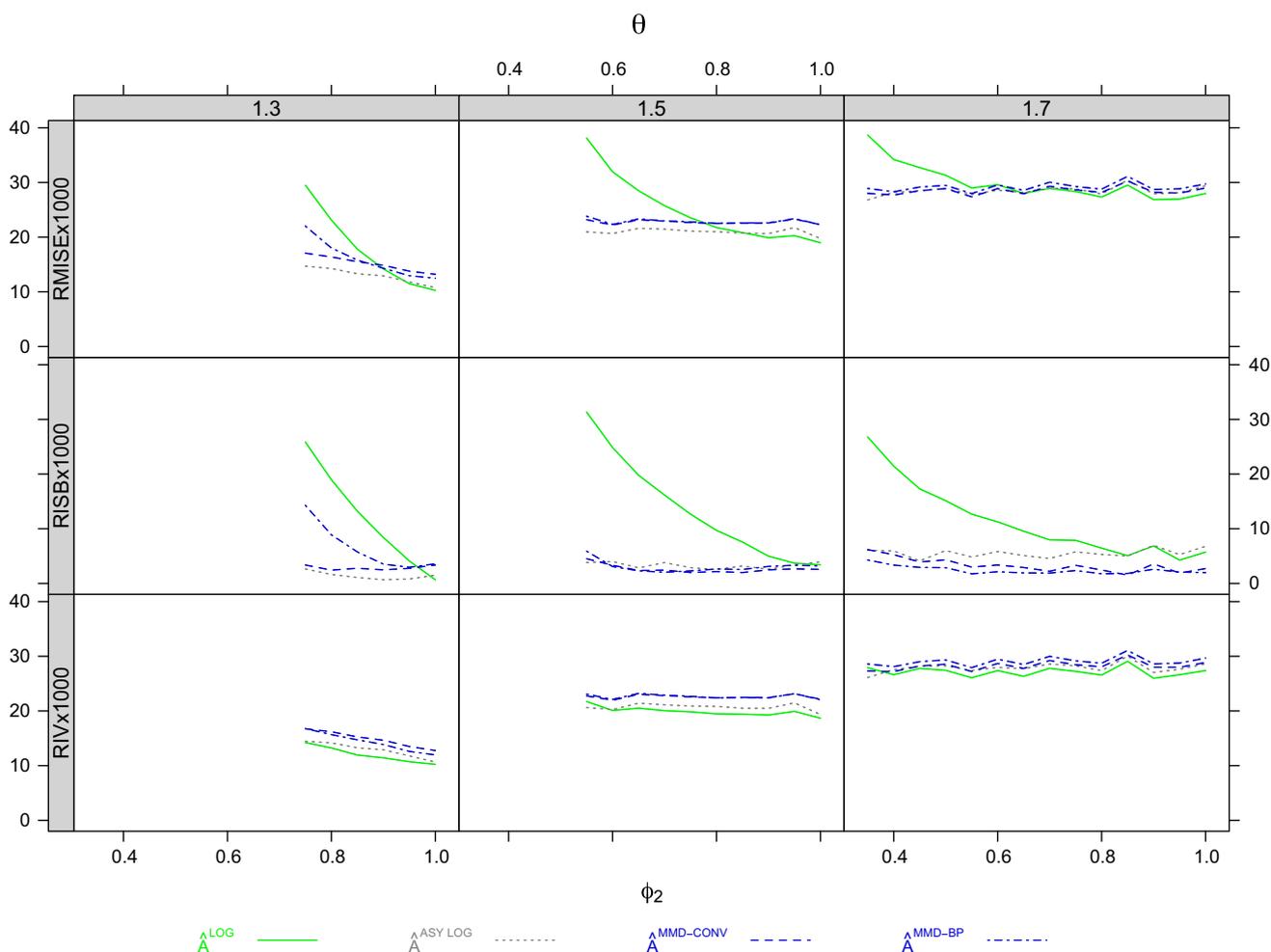


Fig. 3 Root mean integrated squared error (RMISE \times 1000), root integrated squared bias (RISB \times 1000) and root integrated variance (RIV \times 1000) computed for data with distribution in the *MDA* of the asymmetric logistic distribution (25) with $D = 2$ are plotted against the asymmetry parameter ϕ_2 , $\phi_1 = 1$. The columns represent different values of the extremal coefficient θ (12) reflecting the cases of strong ($\theta = 1.3$), mild ($\theta = 1.5$) and weak ($\theta = 1.7$) dependence, respectively.

The estimators \hat{A}_n^{LOG} and $\hat{A}_n^{\text{ASY LOG}}$ are based on the logistic (24) and asymmetric logistic (25) model with $D = 2$, respectively. The estimators $\hat{A}_n^{\text{MMD-CONV}}$ and $\hat{A}_n^{\text{MMD-BP}}$ are based on (22) modified using the convex minorant (5) and the Bernstein–Bézier polynomial projection (8), with $k = 6$, modification, respectively. Notice that there is a lower bound on the value of ϕ_2 for fixed θ

RMISE of the resulting estimator appears noticeably lower. The small jumps present in the curves of the estimators \hat{A}_n^{BP} correspond to a change in the degree of Bernstein–Bézier polynomials in case of mild dependence or suggest that k is insufficiently large in case of very strong dependence (since the Pickands dependence function is nearly non-differentiable at $\omega = 1/2$). Similar results (partly shown in Sect. 3.4 and in the Supplementary Material) are obtained for data with distribution in the *MDA* of the asymmetric logistic distribution.

In conclusion, we established that the estimator $\hat{A}_n^{\text{MMD-CONV}}$ generally performs better than all the other nonparametric estimators, especially in low-dependence scenarios, and its performance slightly improves imposing the convexity constraints by applying the Bernstein–Bézier

polynomial projection method instead of the convex minorant modification. As such, the estimators $\hat{A}_n^{\text{MMD-CONV}}$ and $\hat{A}_n^{\text{MMD-BP}}$ are used as representative nonparametric estimators for the following comparison with parametric methods.

3.3 Bivariate case: parametric versus nonparametric estimators

In order to assess the performance of nonparametric and parametric methods, while avoiding favouring the parametric approach, we compare the nonparametric estimators $\hat{A}_n^{\text{MMD-CONV}}$ and $\hat{A}_n^{\text{MMD-BP}}$ with respect to the parametric estimator based on the logistic model (24) with $D = 2$, \hat{A}_n^{LOG} , for data with distribution in the *MDA* of the asymmetric logistic distribution (25) with $D = 2$. Considering

cases of strong, mild and weak dependence and different levels of asymmetry, in Fig. 3, we display the calculated RMISE, RISB and RIV for the aforementioned estimators and for the estimator based on the asymmetric logistic (25) model with $D = 2$, $\hat{A}_n^{\text{ASY LOG}}$. To compute the estimators $\hat{A}_n^{\text{MMD-BP}}$, we fixed the Bernstein–Bézier polynomial degree $k = 6$ for the reasons discussed in Sect. 3.2. Owing to the constraints A1, A2 and A3, the dependence function cannot be highly asymmetric when the dependence between the margins is strong. Thus, for a fixed value of the extremal coefficient, the asymmetry parameters ϕ_2 represented in the abscissa are bounded from below: if $\theta = 1.3$ or $\theta = 1.5$ then ϕ_2 cannot be lower than ≈ 0.75 or ≈ 0.55 , respectively, with $\phi_1 = 1$.

In accordance with the results detailed in Sect. 3.2, the RMISE generally increases when approaching independence. Although the logistic model becomes increasingly misspecified as ϕ_2 departs from $\phi_1 = \phi_2 = 1$, in terms of RMISE the logistic model estimator slightly outperforms the nonparametric estimators in the cases of mild and weak dependence for mild levels of asymmetry. As expected, the nonparametric estimators are more flexible and show the best performance for highly asymmetric data. Looking at the computed standard deviations, it is clear that the parametric methods are always the most efficient, but nonparametric methods show comparable performances, which is not the

case for higher dimensions (see Sect. 3.4). The logistic model estimator is overall more efficient than its asymmetric extension, especially when the data are close to be symmetric. This can be explained by the reduced number of parameters that need to be estimated when the logistic model is assumed. Consequently, in terms of RMISE, the parametric and nonparametric estimators seem to perform very similarly in the bivariate case, even in the presence of mild asymmetry, whereas, as predicted, the logistic model estimator is insufficiently flexible for describing situations of strong asymmetry.

3.4 Performance of estimators in higher dimensions

We now consider the case of mild dependence strength and asymmetry for data with distribution in the MDA of the asymmetric logistic distribution (11) with dimensions $D = 2, 3, 4, 5$. The RMISE computed for the nonparametric and parametric estimators which can be computed for data dimensions $D > 2$ are reported in Table 3. The Bernstein–Bézier polynomial estimations are obtained fixing the degree to $k = 6$, as $k = 3, 9$ (not shown) do not lead to significant differences in terms of RMISE, suggesting that the dimension does not have significant repercussions on the choice of k . Similarly to Sect. 3.3, the parameter ϕ has a lower bound for fixed θ which depends on the dependence strength and varies

Table 3 Root mean integrated squared error (RMISE \times 1000) computed for data with distribution in the MDA of the asymmetric logistic distribution (11) is reported for different values of the asymmetry parameter ϕ

| Estimators | $D = 2$ | | $D = 3$ | | $D = 4$ | | $D = 5$ |
|---------------------------------|--------------|-------------|--------------|-------------|--------------|-------------|-------------|
| | $\phi = 0.7$ | $\phi = 1$ | $\phi = 0.7$ | $\phi = 1$ | $\phi = 0.7$ | $\phi = 1$ | $\phi = 1$ |
| \hat{A}_n^{RP} | 28.0 | 28.4 | 36.5 | 32.9 | 40.9 | 37.5 | 40.2 |
| \hat{A}_n^{RHT} | 24.7 | 25.1 | 35.9 | 32.2 | 41.1 | 37.7 | 41.7 |
| \hat{A}_n^{RMCFG} | 21.6 | 23.0 | 33.7 | 29.7 | 41.2 | 36.5 | 42.6 |
| \hat{A}_n^{MMD} | 22.3 | 23.0 | 34.5 | 31.4 | 41.6 | 38.0 | 42.8 |
| $\hat{A}_n^{\text{RP-CONV}}$ | 24.9 | 25.2 | 30.7 | 29.7 | 34.0 | 33.5 | 35.5 |
| $\hat{A}_n^{\text{HT-CONV}}$ | 24.6 | 23.3 | 34.3 | 31.1 | 41.4 | 36.3 | 40.2 |
| $\hat{A}_n^{\text{RMCFG-CONV}}$ | 22.7 | 20.9 | 33.0 | 29.4 | 40.7 | 36.1 | 41.6 |
| $\hat{A}_n^{\text{MMD-CONV}}$ | 23.2 | 22.4 | 32.0 | 30.5 | 37.5 | 36.6 | 40.4 |
| $\hat{A}_n^{\text{P-BP}}$ | 23.3 | 22.4 | 32.8 | 30.0 | 40.5 | 36.8 | 38.6 |
| $\hat{A}_n^{\text{RHT-BP}}$ | 24.1 | 23.1 | 32.5 | 29.9 | 40.3 | 37.0 | 40.2 |
| $\hat{A}_n^{\text{RMCFG-BP}}$ | 22.6 | 20.7 | 32.5 | 29.0 | 41.9 | 36.2 | 42.3 |
| $\hat{A}_n^{\text{MMD-BP}}$ | 23.3 | 22.4 | 33.0 | 30.3 | 41.8 | 37.5 | 42.3 |
| \hat{A}_n^{LOG} | 25.8 | 19.9 | 39.7 | 24.6 | 59.6 | 29.0 | 32.8 |
| $\hat{A}_n^{\text{ASY LOG}}$ | 21.4 | 19.3 | 25.5 | 24.7 | 29.6 | 29.6 | 32.5 |

The different columns represent dimensions $D = 2, 3, 4, 5$ and the corresponding extremal coefficient θ is fixed to $\theta = 1.5, 2, 2.5, 3$, respectively, reflecting the case of mild dependence. The estimators \hat{A}_n^{RP} , \hat{A}_n^{RHT} , \hat{A}_n^{RMCFG} and \hat{A}_n^{MMD} are based on (17), (18), (20) and (22) to which we imposed only the endpoint correction (6). Their convex minorant (\hat{A}_n^{CONV}) and projection (\hat{A}_n^{BP}) counterparts are obtained applying the modifications (5) and (8) with $k = 6$, respectively. The parametric estimators \hat{A}_n^{LOG} and $\hat{A}_n^{\text{ASY LOG}}$ are based on the logistic (24) and asymmetric logistic (25) models, respectively. The lower values are represented in bold. Notice that there is a lower bound on the value of ϕ for fixed θ (which depends on dimension). The standard errors of the estimated RMISEs are of the order 10^{-4} .

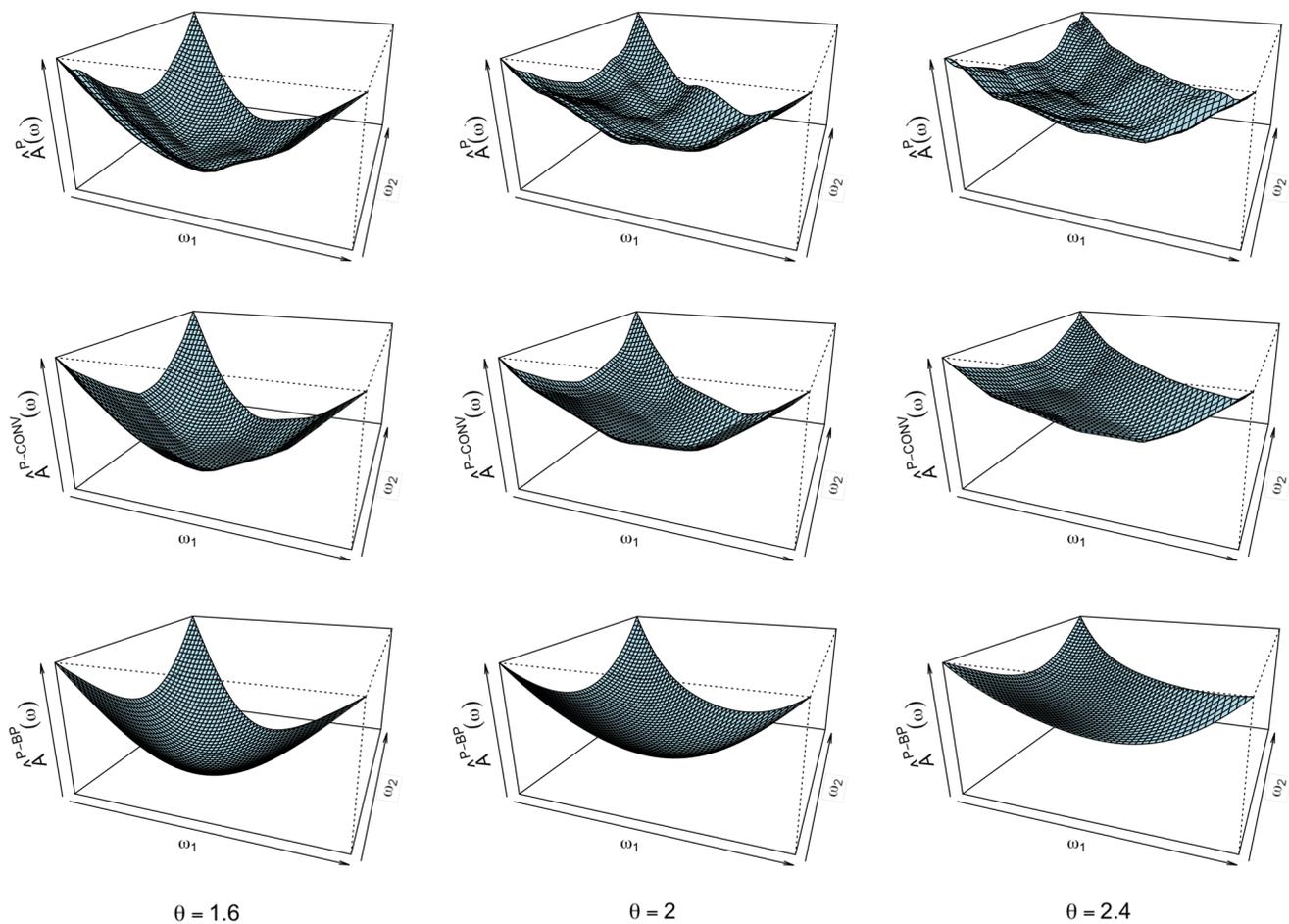


Fig. 4 Dependence functions estimated using the estimator \hat{A}_n^{RP} based on (17) to which we imposed only the endpoint corrections (6) (top row), its convex minorant $\hat{A}_n^{\text{RP-CONV}}$ (middle row) and Bernstein–Bézier polynomial projection $\hat{A}_n^{\text{P-BP}}$ (bottom row) counterparts obtained applying

the modifications (5) and (8) with $k = 6$, respectively, computed for data with distribution in the MDA of the logistic distribution for $D = 3$. The different columns display the cases of strong ($\theta = 1.6$), mild ($\theta = 2$) and weak ($\theta = 2.4$) dependence strength, respectively

across dimensions. By comparing Table 3 with Tables 4 and 5 in the Supplementary Material, the RMISE increases as the dependence strength decreases for all dimensions, which is consistent with the findings for bivariate data in Sects. 3.2 and 3.3. The influence of the dependence strength can be explained by the fact that the dependence function has more degrees of freedom for weak dependence, and therefore, it is more difficult to estimate. Generally, estimator performance decays systematically as the dimension increases. Indeed, for large dimensions D the estimation problem becomes more difficult because the volume of the simplex S_D is larger. The integrals in (13) are also computed on larger domains. Similar results are obtained for the cases of strong and weak dependence strengths; see the Supplementary Material. Overall, in cases of strong dependence we identified the adjusted estimators \hat{A}_n^{RMCFG} as the best across all dimensions (see Table 5 in the Supplementary Material). As we can see from Table 3 and Table 4 in the Supplementary Material, when

the dependence strength is mild or weak, the best performances are provided by the adjusted estimators \hat{A}_n^{RMCFG} and \hat{A}_n^{MMD} , respectively, only for bivariate data, whereas the estimator $\hat{A}_n^{\text{RP-CONV}}$ outperforms all the others in higher dimensions. Imposing the convexity constraint using either the projection or the convex minorant method overall reduces the nonparametric estimators' RMISE. In the bivariate case, smaller RMISE are achieved by applying the Bernstein–Bézier polynomial method. However, when $D > 2$, the two modifications are generally comparable and the estimators $\hat{A}_n^{\text{RP-CONV}}$ and $\hat{A}_n^{\text{MMD-CONV}}$ particularly benefit from the convex minorant modification often outperforming their Bernstein–Bézier polynomial counterparts.

Nonparametric and parametric methods perform very similarly for bivariate data, at least when the assumed parametric model is not strongly misspecified. By contrast, for dimension $D > 2$, the variability of nonparametric estimators increases significantly and parametric methods largely

outperform nonparametric estimators when the parametric model is correctly specified. However, the RISB of the logistic model estimator in dimensions $D > 2$ is very large for slightly asymmetric data causing a significant increase of its RMISE. Nonparametric estimators instead maintain low bias as the dimensions and the extent of asymmetry increase, confirming their greater flexibility with respect to parametric methods.

In Fig. 4, it is possible to visualise how the dependence function estimator \hat{A}_n^{RP} , for dimension $D = 3$, adjusts after the application of the convex minorant or the Bernstein–Bézier polynomial modification. Both methods significantly change the shape of the estimator, making it satisfy the convexity constraint. However, the estimator resulting from the application of the convex minorant method is a piecewise linear function and not smooth on the unit simplex S_D , unlike its Bernstein–Bézier polynomial projection counterpart.

4 Discussion

We investigated the finite-sample performance of some of the mostly used nonparametric estimators for multivariate componentwise maxima considering different strengths and forms of dependence. Generally speaking, estimator performance decays as the dependence strength between the margins decreases and the dimensionality increases. We computed the convex minorant counterparts of the nonparametric estimators for the dependence function in dimensions $D \geq 2$ and highlighted that the nonparametric estimator performance generally improves by imposing the convexity constraint. Applying the Bernstein–Bézier polynomial method leads to better estimator performance for bivariate maxima, whereas the convex minorant modification is a more efficient alternative for higher dimensions and asymmetric datasets. The choice of the Bernstein–Bézier polynomial degree k is mainly influenced by the dependence strength across dimensions. Fixing its value to $k = 6$ should be good enough to avoid over- and under-fitting for most dependence cases, but k should be much larger in the presence of very strong dependence.

Yet, the question of obtaining continuous nonparametric estimators of the spectral measure or the dependence function satisfying the constraint A4 in the multivariate framework ($D > 2$) remains open. The greater flexibility of nonparametric methods with respect to parametric ones was also validated: we have shown that moderate model misspecification may have dramatic consequences on the performance of parametric estimators for dimension $D > 2$ and therefore also on the estimation of joint probabilities for extreme events.

In this simulation study, estimator performance was only assessed for block maxima. Considering other ways of select-

ing the sample of extremes might be of interest. For instance, selecting observations exceeding a specific threshold would offer valuable insight into these popular threshold methods, although the definition of threshold exceedances in the multivariate case is not as clear as in the univariate case. Also, here we focused on a very specific dependence structure, the symmetric and asymmetric logistic families. Future studies would benefit from exploring other types of dependence structures, gaining a greater understanding of nonparametric and parametric methods towards a wider range of applications.

Appendix 1: Nonparametric estimators

If $\mathbf{Z} = (Z_1, \dots, Z_D)^\top$ is a max-stable vector with distribution $G(\mathbf{z})$ (1) and estimated margins \hat{G}_d (3), then the variable

$$\kappa(\boldsymbol{\omega}) = \min_{1 \leq d \leq D} \left[\frac{-\log \{ \hat{G}_d(Z_d) \}}{\omega_d} \right], \quad \boldsymbol{\omega} \in S_D, \quad (15)$$

has survival function

$$\begin{aligned} \Pr \{ \kappa(\boldsymbol{\omega}) > t \} &= \Pr \left[-\log \{ \hat{G}_1(Z_1) \} > \omega_1 t, \dots, \right. \\ &\quad \left. -\log \{ \hat{G}_d(Z_D) \} > \omega_D t \right] \\ &= \exp \{ -t A(\boldsymbol{\omega}) \}, \end{aligned}$$

where $t \geq 0$. Thus, the random variable $\kappa(\boldsymbol{\omega})$ has standard exponential distribution with

$$\begin{aligned} E \{ \kappa(\boldsymbol{\omega}) \} &= A(\boldsymbol{\omega})^{-1}, \quad \text{and} \\ E [\log \{ \kappa(\boldsymbol{\omega}) \}] &= -\log \{ A(\boldsymbol{\omega}) \} - \eta, \end{aligned} \quad (16)$$

where η is the Euler–Mascheroni constant. The most popular nonparametric estimators of $A(\boldsymbol{\omega})$ are based on the equations in (16). For instance, Pickands (1981) proposed to estimate the dependence function by the reciprocal of the sample mean of $\kappa(\boldsymbol{\omega})$, *i.e.*

$$\hat{A}_n^{RP}(\boldsymbol{\omega}) = \left\{ \frac{1}{M} \sum_{i=1}^M \kappa_i(\boldsymbol{\omega}) \right\}^{-1}, \quad \boldsymbol{\omega} \in S_D, \quad (17)$$

where $\kappa_i(\boldsymbol{\omega})$ is the empirical counterpart of (15) using sample maxima \mathbf{m}_i . While $\hat{A}_n^{RP}(\boldsymbol{\omega})$ provides a sensible estimation of $A(\boldsymbol{\omega})$, it does not satisfy any of the conditions A1, A2 and A3. As such, Hall and Tajvidi (2000) proposed a variant which verifies A2 and A3, defined as

$$\hat{A}_n^{\text{RHT}}(\omega) = \left\{ \frac{1}{M} \sum_{i=1}^M \min_{1 \leq d \leq D} \left(\frac{\tilde{m}_{i;d}^{-1} / \bar{m}_{M;d}^{-1}}{\omega_d} \right) \right\}^{-1}, \quad \omega \in S_D, \tag{18}$$

where $\bar{m}_{M;d}^{-1} = M^{-1} \sum_{i=1}^M \tilde{m}_{i;d}^{-1}$, $d = 1, \dots, D$, and \tilde{m} as defined in (4). The second equation in (16) leads to another dependence function estimator

$$\hat{A}^*(\omega) = \exp \left[-\eta - \frac{1}{M} \sum_{i=1}^M \log \{ \kappa_i(\omega) \} \right], \quad \omega \in S_D. \tag{19}$$

However, $\hat{A}^*(\omega)$ does not satisfy any of the constraints A1, A2, A3 either. Capéràa et al. (1997) further modified (16) introducing an estimator meeting the constraint A3 for $D = 2$; its extension to arbitrary dimension, provided by Zhang et al. (2008), see also Gudendorf and Segers (2011), corresponds to the following estimator

$$\begin{aligned} \hat{A}_n^{\text{RMCFG}}(\omega) &= \exp \left[\log \{ A^*(\omega) \} - \sum_{d=1}^D \delta_d(\omega) \log \{ A^*(\mathbf{e}_d) \} \right], \\ \omega &\in S_D, \end{aligned} \tag{20}$$

where the δ_d , $d = 1, \dots, D$, are continuous functions verifying $\delta_d(\mathbf{e}_k) = 1_{(d=k)}$ for all $d, k = 1, \dots, D$. In particular, in the present paper we use $\delta_d(\omega) = \omega_d$. See Gudendorf and Segers (2010) for a more complete explanation of the construction of the above estimators and Segers (2007) for a review of their properties.

Another type of nonparametric estimator of the dependence function can be based on the madogram (Matheron 1987; Cooley et al. 2006). The multivariate madogram is defined as (see Naveau et al. 2009) the expected value of the difference between the maximum and the mean of the rescaled variables $G(\mathbf{Z})^{\omega^{-1}}$, i.e.

$$v(\omega) = E \left[\max_{1 \leq d \leq D} \{ G_d(z_d)^{1/\omega_d} \} - \frac{1}{D} \sum_{d=1}^D G_d(z_d)^{1/\omega_d} \right], \quad \omega \in S_D. \tag{21}$$

The estimator suggested by Marcon et al. (2017) consists of a slight modification of the estimator introduced by Naveau et al. (2009) for $D = 2$ that satisfies the conditions A2 and A3 and applies to high dimensions. It is defined as

$$\hat{A}_n^{\text{MMD}}(\omega) = \frac{\hat{v}_M(\omega) + c(\omega)}{1 - \hat{v}_M(\omega) - c(\omega)}, \quad \omega \in S_D, \tag{22}$$

where $c(\omega) = D^{-1} \sum_{d=1}^D \omega_d / (1 + \omega_d)$ and \hat{v}_M denotes the following estimator of the multivariate madogram (21):

$$\begin{aligned} \hat{v}_M(\omega) &= \frac{1}{M} \sum_{i=1}^M \left[\max_{1 \leq d \leq D} \{ \hat{G}_d(m_{i;d})^{1/\omega_d} \} \right. \\ &\quad \left. - \frac{1}{D} \sum_{d=1}^D \hat{G}_d(m_{i;d})^{1/\omega_d} \right]. \end{aligned}$$

Finally, Cormier et al. (2014) proposed to estimate the Pickands dependence function through a graphical tool. Given the bivariate copula $C(u_1, u_2)$, consider the transformation

$$\begin{aligned} T_i &= \frac{\log \{ \hat{G}(u_{i;1}) \}}{\log \{ \hat{G}(u_{i;1}) \hat{G}(u_{i;2}) \}}, \\ Z_i &= \frac{\log \{ \hat{C}_n \{ \hat{G}(u_{i;1}), \hat{G}(u_{i;2}) \} \}}{\log \{ \hat{G}(u_{i;1}) \hat{G}(u_{i;2}) \}}. \end{aligned}$$

where for each $i = 1, \dots, n$, \hat{C}_n denote the empirical copula. The Pickands dependence function is approximated by fitting to the points $(t_1, z_1), \dots, (t_n, z_n)$ a constrained B-spline smoothing of order $m = 3$ for a suitable sequence of k interior knots, $\beta_1, \dots, \beta_{m+k}$ and B-spline basis $\epsilon_{1;m}, \dots, \epsilon_{m+k;m}$, i.e.

$$\hat{A}_n^{\text{COBS}} = \sum_{j=1}^{m+k} \hat{\beta}_j \epsilon_{m;j}, \tag{23}$$

where $\hat{\beta}_j$ is an estimate of β_j for each $j = 1, \dots, m + k$.

Appendix 2: Parametric models for the dependence function

Several parametric models to describe extremal dependence have been proposed in the literature. The logistic model, introduced by Gumbel (1960a, b), is one of the most widely used thanks to its simplicity. It has dependence function of the form

$$A(\omega) = \left(\sum_{d=1}^D \omega_d^{1/\alpha} \right)^\alpha, \quad \omega \in S_D, \tag{24}$$

with dependence parameter $0 < \alpha \leq 1$. The strength of dependence decreases as α increases. In particular, the cases of independence and perfect dependence correspond to $\alpha = 1$ and $\alpha \downarrow 0$, respectively. When $D = 2$, the model (24) simplifies to $A(\omega) = \{ (1 - \omega)^{1/\alpha} + \omega^{1/\alpha} \}^\alpha$, $\omega \in [0, 1]$.

A limitation of the logistic model resides in its lack of flexibility as its dependence structure relies on only one parameter and it does not allow for asymmetry, i.e. the margins are

exchangeable. An asymmetric generalisation was suggested by Tawn (1988) for dimension $D = 2$ and extended to higher dimensions by Tawn (1990); see also Smith et al. (1990), Coles and Tawn (1991) and Stephenson (2009). The asymmetric logistic model has dependence function

$$A(\omega) = \sum_{C \in S} \left\{ \sum_{d \in C} (\phi_{Cd} \omega_d)^{1/\alpha_C} \right\}^{\alpha_C}, \quad \omega \in S_D, \quad (25)$$

with dependence parameters $0 < \alpha_C \leq 1$ and asymmetry parameters $\phi_{Cd} = 0$ if $d \notin C$, $0 \leq \phi_{Cd} \leq 1$ for $d = 1, \dots, D$, with $\sum_{C \in S} \phi_{Cd} = 1$, where S is the set of all non-empty subsets of $\mathcal{D} = \{1, \dots, D\}$. When $D = 2$, the model (25) simplifies to $A(\omega) = (\phi_2 - \phi_1)\omega + 1 - \phi_2 + [\{\phi_2(1 - \omega)\}^{1/\alpha} + \{\phi_1\omega\}^{1/\alpha}]^\alpha$, $\omega \in [0, 1]$, with dependence parameter $0 < \alpha \leq 1$ and asymmetry parameters $0 \leq \phi_1, \phi_2 \leq 1$. The extent of asymmetry in the bivariate dependence structure decreases as the asymmetry parameters ϕ_1 and ϕ_2 approach unity. The symmetric case arises when $\phi_1 = \phi_2 = 1$. Without much loss of flexibility, ϕ_1 or ϕ_2 could be fixed to 1 (Tawn 1988). Independence corresponds to $\phi_1 = \phi_2 = 1$ and $\alpha = 1$, whereas perfect dependence corresponds to $\phi_1 = 0$ or $\phi_2 = 0$ or $\alpha \downarrow 0$.

Appendix 3: Proof of equation (11)

If $\mathbf{U} = (U_1, \dots, U_D)^\top \sim C_1(u_1, \dots, u_D)$, where C_1 denotes the outer power Clayton copula with generator $\varphi(t) = (t^\alpha + 1)^{-1}$, then the random vector $\mathbf{X} = \{-1/\log(U_1), \dots, -1/\log(U_D)\}^\top$ is distributed according to the same copula C_1 and its marginal distributions are unit Fréchet. According to Fougères (2004, p. 376), taking the sequences $\mathbf{a}_n = (n, \dots, n)^\top$ and $\mathbf{b}_n = (0, \dots, 0)^\top$ defined in Sect. 1, the distribution of \mathbf{X} is in the MDA of the logistic distribution (24). This means that, defining

$$(M_{n;1}, \dots, M_{n;D})^\top = \left[\max_{1 \leq i \leq n} \{-1/\log(U_{i;1})\}, \dots, \max_{1 \leq i \leq n} \{-1/\log(U_{i;D})\} \right]^\top,$$

one has

$$\Pr \left(\frac{M_{n;1}}{n} \leq x_1, \dots, \frac{M_{n;D}}{n} \leq x_D \right) \rightarrow \exp \{-V(x_1, \dots, x_D)\}, \quad n \rightarrow \infty,$$

where $V(x_1, \dots, x_D) = \left(\sum_{d=1}^D x_d^{-1/\alpha} \right)^\alpha$, $\alpha \in (0, 1]$. Now, notice that any vector $\mathbf{U}^* = (U_1^*, \dots, U_D^*)^\top$ distributed as $\mathbf{U}^* \sim C_2(\mathbf{u}^*) = C_1(u_1^{\phi_1}, \dots, u_D^{\phi_D}) \prod_{d=1}^D u_d^{1-\phi_d}$ has uniform margins and may be represented by setting

$U_d^* = \max \left\{ U_d^{1/\phi_d}, \tilde{U}_d^{1/(1-\phi_d)} \right\}$, $d = 1, \dots, D$, with $(U_1, \dots, U_D)^\top \sim C_1(u_1, \dots, u_D)$, independent of $\tilde{U}_d \stackrel{\text{i.i.d.}}{\sim} \text{Unif}(0, 1)$. Therefore, writing $M_{n;d}^* = \max_{1 \leq i \leq n} \{-1/\log(U_{i;d}^*)\}$ and $\tilde{M}_{n;d} = \max_{1 \leq i \leq n} \{-1/\log(\tilde{U}_{i;d})\}$, $d = 1, \dots, D$, one obtains

$$\begin{aligned} & \Pr \left(\frac{M_{n;1}^*}{n} \leq x_1, \dots, \frac{M_{n;D}^*}{n} \leq x_D \right) \\ &= \Pr \left(\frac{M_{n;1}}{n} \leq \frac{x_1}{\phi_1}, \dots, \frac{M_{n;D}}{n} \leq \frac{x_D}{\phi_D} \right) \\ & \quad \times \prod_{d=1}^D \Pr \left(\frac{\tilde{M}_{n;d}}{n} \leq \frac{x_d}{1-\phi_d} \right), \\ & \rightarrow \exp \left[- \left\{ V \left(\frac{x_1}{\phi_1}, \dots, \frac{x_D}{\phi_D} \right) + \sum_{d=1}^D (1-\phi_d)x_d^{-1} \right\} \right], \\ & \quad n \rightarrow \infty, \end{aligned}$$

which is a multivariate extreme-value distribution with exponent function

$$V(x_1, \dots, x_D) = \left\{ \sum_{d=1}^D \left(\frac{x_d}{\phi_d} \right)^{-1/\alpha} \right\}^\alpha + \sum_{d=1}^D (1-\phi_d)x_d^{-1}.$$

Applying (2) we obtain

$$A(\omega) = \left\{ \sum_{d=1}^D (\omega_d \phi_d)^{1/\alpha} \right\}^\alpha + \sum_{d=1}^D \omega_d (1-\phi_d),$$

which is a special case of the asymmetric logistic model dependence function in (25).

References

Beirlant, J., Goegebeur, Y., Segers, J., Teugels, J.: Statistics of Extremes: Theory and Applications. Wiley, Hoboken (2004)

Berghaus, B., Bücher, A., Dette, H.: Minimum distance estimators of the Pickands dependence function and related tests of multivariate extreme-value dependence. *J. de la Soc. de Fr. de Stat.* **154**(1), 116–137 (2013)

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

Camicer, J., Peña, J.: Shape preserving representations and optimality of the Bernstein basis. *Adv. Comput. Math.* **1**, 173–196 (1993)

Capérãa, B.Y.P., Fougères, A.: Estimation of a bivariate extreme value distribution. *Extremes* **3**, 311–329 (2000)

Capérãa, B.Y.P., Fougères, A., Genest, C.: A nonparametric estimation procedure for bivariate extreme value copulas. *Biometrika* **84**(3), 567–577 (1997)

Coles, S.G., Tawn, J.A.: Modelling extreme multivariate events. *J. R. Stat. Soc. B* **53**(2), 377–392 (1991)

- Cooley, D., Naveau, P., Poncet, P.: Variograms for Spatial Max-stable Random Fields, 373–390. Lecture notes in Statistics 187. Springer, New York (2006)
- Cormier, E., Genest, C., Nešlehová, J.G.: Using B-splines for nonparametric inference on bivariate extreme-value copulas. *Extremes* **17**, 633–659 (2014)
- Davison, A.C., Gholamrezaee, M.M.: Geostatistics of extremes. *Proc. R. Soc. A Math. Phys. Eng. Sci.* **468**, 581–608 (2012)
- de Haan, L.: A spectral representation for max-stable processes. *Ann. Probab.* **12**(4), 1194–1204 (1984)
- de Haan, L., Resnick, S.I.: Limit theory for multivariate sample extremes. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete* **40**, 317–337 (1977)
- Deheuvels, P.: On the limiting behavior of the Pickands estimator for bivariate extreme-value distributions. *Stat. Probab. Lett.* **12**, 429–439 (1991)
- Einmahl, J.H., de Haan, L., Piterburg, V.I.: Nonparametric estimation of the spectral measure of an extreme value distribution. *Ann. Stat.* **29**(5), 1401–1423 (2001)
- Einmahl, J.H., Segers, J.: Maximum empirical likelihood estimation of the spectral measure of an extreme-value distribution. *Ann. Stat.* **37**(5), 2953–2989 (2009)
- Fils-Villetard, A., Guillou, A., Segers, J.: Projection estimators of Pickands dependence functions. *Can. J. Stat.* **36**(3), 369–382 (2008)
- Fougères, A.: Multivariate extremes. In: *Extreme Values in Finance, Telecommunications, and the Environment*, chapter pp. 377–388. Chapman and Hall (2004)
- Genest, C., Nešlehová, J., Quessy, J.-F.: Tests of symmetry for bivariate copulas. *Ann. Inst. Stat. Math.* **64**, 811–834 (2011)
- Genest, C., Segers, J.: Rank-based inference for bivariate extreme-value copulas. *Ann. Stat.* **37**(5B), 2990–3022 (2009)
- Gentle, J.E.: *Elements of Computational Statistics*. Springer, New York (2002)
- Gudendorf, G., Segers, J.: Extreme-value copulas. In: Jaworski, W.H.P., Durante, F., Rychlik, T. (eds.) *Proceedings of the Workshop on Copula Theory and its Applications*, pp. 127–146. Springer (2010)
- Gudendorf, G., Segers, J.: Nonparametric estimation of an extreme-value copula in arbitrary dimensions. *J. Multivar. Anal.* **102**(1), 37–47 (2011)
- Gudendorf, G., Segers, J.: Nonparametric estimation of multivariate extreme-value copulas. *J. Stat. Plan. Inference* **142**(12), 3073–3085 (2012)
- Guillotte, S.: A Bayesian estimator for the dependence function of a bivariate extreme-value distribution. *Can. J. Stat.* **36**(3), 383–396 (2008)
- Gumbel, E.J.: Distributions des valeurs extrêmes en plusieurs dimensions. *Publ. de l'Inst. de Stat. de l'Univ. Paris* **9**(294), 171–173 (1960a)
- Gumbel, E.J.: Bivariate exponential distributions. *J. Am. Stat. Assoc.* **55**, 698–707 (1960b)
- Hall, P., Tajvidi, N.: Distribution and dependence-function estimation for bivariate extreme-value distributions. *Bernoulli* **5**, 835–844 (2000)
- Hofert, M., Kojadinovic, I., Mächler, M., Yan, J.: *Copula: multivariate dependence with copulas*. R package version 0.999-12 (2014)
- Huser, R., Davison, A.C., Genton, M.G.: Likelihood estimators for multivariate extremes. *Extremes* **19**, 79–103 (2016)
- Jiménez, J.R., Villa-Diharce, E., Flores, M.: Nonparametric estimation of the dependence function in bivariate extreme value distributions. *J. Multivar. Anal.* **76**, 159–191 (2001)
- Khoudraji, A.: *Contributions à l'étude des copules et à la modélisation de valeurs extrêmes bivariées*. Ph.D. thesis. Université Laval, Québec, Canada (1995)
- Li, B., Genton, M.G.: Nonparametric identification of copula structures. *J. Am. Stat. Assoc.* **108**(502), 666–675 (2013)
- Liebscher, E.: Construction of asymmetric multivariate copulas. *J. Am. Stat. Assoc.* **99**(10), 2234–2250 (2008)
- Lorenz, G.G.: *Bernstein Polynomials*, 2nd edn. Chelsea Publishing Company, New York (1986)
- Marcon, G., Padoan, S.A., Antoniano-Villalobos, I.: Bayesian inference for the extremal dependence. *Electron. J. Stat.* **10**(2), 3310–3337 (2016)
- Marcon, G., Padoan, S.A., Naveau, P., Muliere, P., Segers, J.: Nonparametric estimation of the Pickands dependence function using Bernstein polynomials. *J. Stat. Plan. Inference* **183**, 1–17 (2017)
- Matheron, G.: Suffit-il, pour une covariance, d'être de type positif? *Sci. de la Terre, Série Informatique Géologique* **26**, 51–66 (1987)
- Naveau, P., Guillou, A., Cooley, D., Diebolt, J.: Modelling pairwise dependence of maxima in space. *Biometrika* **96**(1), 1–17 (2009)
- Nelsen, R.: *An Introduction to Copulas*, 2nd edn. Springer, New York (2006)
- Padoan, S.A., Ribatet, M., Sisson, S.A.: Likelihood-based inference for max-stable processes. *J. Am. Stat. Assoc.* **105**(489), 263–277 (2010)
- Pickands, J.: Multivariate extreme value distributions. *Bulletin of the International Statistical Institute (Proceedings of the 43rd Session)*, pp. 859–878 (1981)
- Pickands, J.: *Multivariate Negative Exponential and Extreme Value Distributions*, vol. 51. Springer, Berlin (1989)
- Resnick, S.I.: *Extreme Values, Regular Variation and Point Processes*. Springer, Berlin (1987)
- Ressel, P.: Homogeneous distributions—and a spectral representation of classical mean values and stable tail dependence functions. *J. Multivar. Anal.* **117**, 246–256 (2013)
- Sauer, T.: Multivariate Bernstein polynomials and convexity. *Comput. Aided Geom. Des.* **8**(6), 465–478 (1991)
- Schlather, M., Tawn, J.A.: Inequalities for the extremal coefficients of multivariate extreme value distributions. *Extremes* **5**, 87–102 (2002)
- Schlather, M., Tawn, J.A.: A dependence measure for multivariate and spatial extreme values: properties and inference. *Biometrika* **90**(1), 139–156 (2003)
- Segers, J.: Non-Parametric Inference for Bivariate Extreme-Value Copulas. In: Ahsanullah, M., Kirmani, S.N.U.A. (eds.) *Topics in Extreme Values*, pp. 181–203. Nova Science Publishers Inc, New York (2007)
- Smith, R.L.: Max-stable processes and spatial extremes. Unpublished manuscript. University of North Carolina, Chapel Hill, U. S. (1990)
- Smith, R.L., Tawn, J.A., Yuen, H.K.: Statistics of multivariate extremes. *Int. Stat. Rev.* **58**(1), 47–58 (1990)
- Stephenson, A.G.: High-dimensional parametric modelling of multivariate extreme events. *Aust. N. Z. J. Stat.* **51**(1), 77–88 (2009)
- Tawn, J.A.: Bivariate extreme value theory: models and estimation. *Biometrika* **75**(3), 397–415 (1988)
- Tawn, J.A.: Modelling multivariate extreme value distributions. *Biometrika* **77**(2), 245–253 (1990)
- Varin, B.C., Vidoni, P.: A note on composite likelihood inference and model selection. *Biometrika* **92**(3), 519–528 (2005)
- Varin, C.: On composite marginal likelihoods. *ASTA Adv. Stat. Anal.* **92**(1), 1–28 (2008)
- Zhang, D., Wells, M.T., Peng, L.: Nonparametric estimation of the dependence function for a multivariate extreme value distribution. *J. Multivar. Anal.* **99**(4), 577–588 (2008)