INSTITUT DE
STATISTIQUE

UNIVERSITÉ CATHOLIQUE DE LOUVAIN

DISCUSSION
PAPER

0809

PROJECTION ESTIMATORS OF PICKANDS DEPENDENCE FUNCTIONS

FILS-VILLETARD, A., GUILLOU, A. and J. SEGERS

The Canadian Journal of Statistics
Vol. ??, No. ?, 200?, Pages ???-???
La revue canadienne de statistique

Projection estimators of Pickands dependence functions

Amélie FILS-VILLETARD, Armelle GUILLOU and Johan SEGERS

Key words and phrases: extreme-value copula; Hilbert space; Pickands dependence function; projection; shape constraints; stable tail dependence function; tangent cone.
MSC 2000: Primary 62G32; secondary 62F30.

Abstract: Shape constraints on a functional parameter can often be formulated in terms of a closed and convex parameter set embedded in a real Hilbert space. This is the case, for instance, if the curve of interest is a Pickands dependence function of an extreme-value copula. The topic of this paper is the estimator that results when an arbitrary initial estimator possibly falling outside the parameter set is projected onto this parameter set. As direct computation of the projection is infeasible, the full parameter set needs to be replaced by a sieve of finite-dimensional subsets. Asymptotic properties of the initial estimator sequence in the Hilbert space topology transfer easily to those of the projected sequence and its finite-dimensional approximations.

Estimateurs par projection de la fonction de dépendance de Pickands

Résumé : Cet article traite de l'estimation d'une fonction sujette à des contraintes de forme données. Souvent l'espace des fonctions satisfaisant ces contraintes, appelé espace des paramètres, peut s'identifier avec un sous-espace convexe fermé d'un espace de Hilbert. Ceci est le cas pour la fonction de dépendance de Pickands d'une copule des valeurs extrêmes, exemple étudié en détail dans l'article. Partant d'un estimateur initial quelconque, éventuellement qui ne satisfait pas les contraintes de forme voulues, les auteurs s'intéressent à l'estimateur par projection sur l'espace des paramètres. Compte tenu du fait que la projection directe de cet estimateur n'est pas possible, l'espace des paramètres devra être remplacé par un sous-ensemble fini-dimensionnel. Il est établi que les propriétés asymptotiques de l'estimateur initial se tranfèrent facilement à celles de l'estimateur projeté ainsi qu'à celles de ses approximations finies-dimensionnelles.

1. INTRODUCTION

A copula is a multivariate distribution function with uniform$(0,1)$ margins (Sklar 1959; Nelsen 1999). The statistical relevance of copulas comes from their role in the margin-free modelling of the dependence structure of a general multivariate distribution.

An interesting class of copulas is that of extreme-value copulas (Ghoudi, Khoudraji & Rivest 1998). The following representation for bivariate extreme-value copulas was discovered in Pickands (1981), building upon de Haan & Resnick (1977); see also Deheuvels (1984). A bivariate copula C is an extreme-value copula if and only if it admits the representation

$$C(u,v) = \exp\left\{\log(uv)A\left(\frac{\log(u)}{\log(uv)}\right)\right\}, \qquad u,v \in (0,1). \tag{1}$$

The above expression defines a genuine copula if and only if A belongs to the class $\mathcal{A}$ defined by

$$\mathcal{A} = \{A : [0,1] \to \mathbb{R} \mid A \text{ is convex and } \max(t, 1-t) \leq A(t) \leq 1 \text{ for all } t \in [0,1]\}. \tag{2}$$

The function A is called a Pickands dependence function.

Estimation of the Pickands dependence function of an extreme-value copula is an important step in the analysis of multivariate extremes (Beirlant, Goegebeur, Segers & Teugels 2004, chapter 9). Non-parametric methods avoid the model risk inherent to parametric methods (Tawn 1988; Smith, Yuen & Tawn 1990), but unfortunately, the estimates they generate typically fail to take values in $\mathcal{A}$. The most obvious way to force an estimator to take values in $\mathcal{A}$ is to cut it off below $\max(t, 1-t)$ and above 1 and then take the biggest convex minorant (Pickands 1981, 1989; Deheuvels 1991; Capéraà & Fougères 2000; Hall & Tajvidi 2000; Jiménez, Villa-Diharce & Flores 2001). An alternative is to approximate an initial estimator by a smoothing spline that is constrained to take values in $\mathcal{A}$ (Hall & Tajvidi 2000; Abdous & Ghoudi 2005). For all these approaches, the constrained estimator is known to be consistent as soon as the initial estimator is so, but the precise effect of such modifications on the performance of the estimator are in general unknown.

The method we propose is the following. We view $\mathcal{A}$ as a closed and convex subset of the real Hilbert space $L^2([0,1], \mathrm{d}x)$ of real-valued, Lebesgue square-integrable functions on $[0,1]$. Given an initial estimator $\hat{A}$ of the unknown $A \in \mathcal{A}$ such that $\hat{A} \notin \mathcal{A}$, we project it onto $\mathcal{A}$, the projection being defined as the minimizer of the L^2 distance between the initial estimator and members of $\mathcal{A}$. Because the latter is closed and convex, the projection operator is well-defined. As projection operators are non-expansive, the estimation error, measured in L^2 distance, of the projection estimator is never larger than the one of the initial estimator. This method is somewhat akin to the methodology described in Mammen, Marron, Turlach & Wand (2001) in the context of nonparametric regression.

The aim of this paper, then, is to develop an abstract framework for the projection estimator, focusing on the relationship between the asymptotic properties of the initial estimator sequence and the corresponding projection estimator sequence. The main tool here is the functional delta method in combination with a result on the Hadamard differentiability of projection operators. The theory is outlined in a general setting and then specialized to Pickands dependence functions.

Unfortunately, computation of the projection involves an infinite-dimensional minimization problem under constraints for which no explicit solution exists. Therefore, we propose a computational tool replacing the latter optimization problem by a sequence of quadratic programs, that is, quadratic optimization problems with linear inequality constraints. The dimension of the quadratic program serves as a tuning parameter. We derive an explicitly computable bound on the distance between the projection estimator and its finite-dimensional approximation as a function of the dimension of the quadratic program. This bound yields guidance on the choice of the tuning parameter both in finite and in large sample settings.

The outline of the paper is as follows. Definitions and elementary properties of the estimators are described in section 2. Finite-dimensional approximations are the subject of section 3, followed by section 4 on asymptotics. The simulation study and data example in section 5 serve to illustrate the finite-sample properties of the projection estimators. All proofs are gathered in the appendix.

2. THE ESTIMATORS

Suppose that it is known that the parameter of interest, f_0, takes values in a non-empty, closed and convex subset $\mathcal{F}$ of a real Hilbert space $\mathbb{H}$ with inner product $\langle \cdot, \cdot \rangle$ and norm $\| \cdot \|$. There is available an estimate $\hat{f}$, but unfortunately $\hat{f}$ does not belong to $\mathcal{F}$.

As $\mathcal{F}$ is closed and convex, there exists for every $h \in \mathbb{H}$ a unique $f \in \mathcal{F}$ such that $\|h - f\| = \inf\{\|h - g\| : g \in \mathcal{F}\}$, called the (orthogonal) projection of h on $\mathcal{F}$ and denoted by $f = \Pi(h \mid \mathcal{F})$;

see Stark & Yang (1998, chapter 2) for a review of projections on closed, convex sets. We may therefore define the projection estimator as the projection of the initial estimator $\hat{f}$ on $\mathcal{F}$,

$$\hat{f}^{\mathrm{P}} = \Pi(\hat{f} \mid \mathcal{F}) = \arg\min_{f \in \mathcal{F}} \|f - \hat{f}\|. \tag{3}$$

By definition, the projection estimator always belongs to the parameter set, $\hat{f}^{\mathrm{P}} \in \mathcal{F}$. If the Hilbert space is an L^2 space, then $\hat{f}^{\mathrm{P}} = \arg\min_{f \in \mathcal{F}} \int (f - \hat{f})^2$; therefore, we sometimes also use the phrase 'least-squares estimator'.

In the context of extreme value copulas, the Hilbert space of interest is the space $L^2 = L^2([0,1], \mathrm{d}x)$ of real-valued, Lebesgue square-integrable functions on the unit interval, while the parameter set, $\mathcal{A}$, is the set of Pickands dependence functions described in eq. (2) (with the usual identification of functions that differ on a null set only). Convexity of $\mathcal{A}$ is straightforward to show; closedness of $\mathcal{A}$ follows from the readily verified fact that for Pickands dependence functions, convergence in L^2, pointwise convergence and uniform convergence are all equivalent. Hence, given an initial estimator $\hat{A}_n$, the projection estimator is defined as

$$\hat{A}_n^{\mathrm{P}} = \Pi(\hat{A}_n \mid \mathcal{A}) = \arg\min_{A \in \mathcal{A}} \int (\hat{A}_n - A)^2. \tag{4}$$

Back to the general framework, in geometrical terms, the projection $\Pi(h \mid \mathcal{F})$ of h on $\mathcal{F}$ is the unique point $f \in \mathcal{F}$ such that the hyperplane passing through f and orthogonal to $h - f$ separates h from $\mathcal{F}$. Formally,

for $h \in \mathbb{H}$ and $f \in \mathcal{F}$, the following are equivalent: (5)

(i) $f = \Pi(h \mid \mathcal{F})$;

(ii) $\langle h - f, f - g \rangle \geq 0$ for all $g \in \mathcal{F}$;

(iii) $\|h - f\|^2 + \|f - g\|^2 \leq \|h - g\|^2$ for all $g \in \mathcal{F}$.

From this, it can be deduced that the map $\Pi(\,\cdot \mid \mathcal{F})$ is non-expansive, that is,

$$\|\Pi(g \mid \mathcal{F}) - \Pi(h \mid \mathcal{F})\| \leq \|g - h\|, \qquad g, h \in \mathbb{H}. \tag{6}$$

In particular, projections are continuous. Moreover, if in eq. (6) also $h \in \mathcal{F}$, then $\Pi(h \mid \mathcal{F}) = h$, whence $\|\Pi(g \mid \mathcal{F}) - h\| \leq \|g - h\|$, with strict inequality if $g \notin \mathcal{F}$. Therefore, with respect to the Hilbert space norm, the projection estimator $\hat{f}^{\mathrm{P}}$ is always at least as good as the initial estimator $\hat{f}$:

$$\|\hat{f}^{\mathrm{P}} - f_0\| \leq \|\hat{f} - f_0\|.$$

Hence, from the point of view of statistical decision theory, unless the initial estimator $\hat{f}$ already belongs to $\mathcal{F}$ almost surely, it is inadmissible with respect to all loss functions of the form $L(f, f_0) = \ell(\|f - f_0\|)$ with ℓ increasing.

3. FINITE-DIMENSIONAL APPROXIMATIONS

Unfortunately, the optimization problem in eq. (4) does not admit a direct solution. Therefore, we present here a generally applicable sieve method to approximate the projection estimator up to any desired accuracy.

In the general framework, let $\mathcal{F}_m$ be a finite-dimensional subset of $\mathcal{F}$ of the form

$$\mathcal{F}_m = \left\{ \sum_{i=0}^{m} \lambda_i h_{im} : \boldsymbol{\lambda} = (\lambda_0, \ldots, \lambda_m)' \in \Lambda_m \right\}$$

where the h_{im} are fixed elements in $\mathbb{H}$ and where $\Lambda_m \subset \mathbb{R}^{m+1}$ is closed and convex. The idea is to have a sequence of such finite-dimensional subsets which provide with increasing m increasingly more accurate approximations of $\mathcal{F}$.

The approximate projection estimator based on the initial estimator $\hat{f}$ is then defined as the projection of $\hat{f}$ on $\mathcal{F}_m$,

$$\hat{f}_m^{\mathrm{P}} = \Pi(\hat{f} \mid \mathcal{F}_m). \tag{7}$$

The approximate projection estimator is of the form $\hat{f}_m^{\mathrm{P}} = \sum_i \hat{\lambda}_i h_{im}$, the vector $\hat{\boldsymbol{\lambda}}$ being the solution to the quadratic program

$$\text{minimize } \boldsymbol{\lambda}'\boldsymbol{A}\boldsymbol{\lambda} - 2\boldsymbol{\lambda}'\boldsymbol{b} \text{ subject to } \boldsymbol{\lambda} \in \Lambda_m \tag{8}$$

with the matrix $\boldsymbol{A}$ given by $A_{ij} = \langle h_{im}, h_{jm} \rangle$ and with the vector $\boldsymbol{b}$ given by $b_i = \langle h_{im}, \hat{f} \rangle$. Note that $\boldsymbol{A}$ does not depend on $\hat{f}$, and that if $\mathbb{H}$ is an L^2-space, then the entries of $\boldsymbol{b}$ involve integrals for which numerical quadrature might be required.

In setting of Pickands dependence functions, we propose

$$\mathcal{A}_m = \{A \in \mathcal{A} \mid A \text{ is piecewise linear with knots } k/m,\ k = 0, 1, \ldots, m\}$$

as an approximating subset. (Experiments with higher-order splines did not produce better results.) Each A in $\mathcal{A}_m$ admits the representation $A = \sum_{i=0}^m A(i/m) h_{im}$, where h_{im} is the unique piece-wise linear function on $[0, 1]$ with knots on the given grid such that $h_{im}(i/m) = 1$ and $h_{im}(j/m) = 0$ for integer $j \neq i$. The subset Λ_m of $\mathbb{R}^{m+1}$ to which the vector $\boldsymbol{\lambda}$ in (8) with entries $\lambda_i = A(i/m)$ should belong is defined by the constraints

$$\begin{cases} A(0) = A(1) = 1, \\ A(0) - A(\frac{1}{m}) \leq \frac{1}{m}, \\ A(\frac{i-1}{m}) - 2A(\frac{i}{m}) + A(\frac{i+1}{m}) \geq 0, \qquad i = 1, \ldots, m-1, \\ A(1) - A(\frac{m-1}{m}) \leq \frac{1}{m}. \end{cases}$$

Furthermore, the $(m+1)$-by-$(m+1)$ matrix $\boldsymbol{A}$ with elements $A_{ij} = \int h_{im} h_{jm}$ is easily computed to be

$$\boldsymbol{A} = \frac{1}{6m} \begin{pmatrix} 2 & 1 & & & & 0 \\ 1 & 4 & 1 & & & \\ & 1 & 4 & 1 & & \\ & & \ddots & \ddots & \ddots & \\ & & & 1 & 4 & 1 \\ 0 & & & & 1 & 2 \end{pmatrix}.$$

Given an initial estimate $\hat{A}$, computing the vector $\boldsymbol{b}$ involves integrals of the form $b_i = \int h_{im} \hat{A}$. A simple procedure consists of interpolating the integrand on each interval of the form $[(i-1)/m, i/m]$ by a quadratic polynomial with interpolation points $(i-1)/m$, $(i-1/2)/m$ and i/m, resulting in the approximation

$$b_i \approx \begin{cases} \frac{1}{3m} \left\{ \frac{1}{2} \hat{A}(0) + \hat{A}\left(\frac{1}{2m}\right) \right\} & \text{if } i = 0, \\ \frac{1}{3m} \left\{ \hat{A}\left(\frac{2i-1}{2m}\right) + \hat{A}\left(\frac{i}{m}\right) + \hat{A}\left(\frac{2i+1}{2m}\right) \right\} & \text{if } i = 1, \ldots, m-1, \\ \frac{1}{3m} \left\{ \hat{A}\left(\frac{2m-1}{2m}\right) + \frac{1}{2} \hat{A}(1) \right\} & \text{if } i = m. \end{cases}$$

In order to show that this approximation to $\boldsymbol{b}$ is sufficiently accurate, we would need detailed sample-path properties of $\hat{A}$. We do not pursue this issue here and simply take the above formula

as a convenient computational tool. In the further theoretical analysis, we assume that the vector $\boldsymbol{b}$ is computed exactly.

Back to the general framework, an important question is how to choose the tuning parameter m in eq. (7). If the true parameter f_0 is not contained in the finite-dimensional approximation $\mathcal{F}_m$, then the desirable properties (admissibility) of the projection estimator $\hat{f}^{\mathrm{p}} = \Pi(\hat{f} \mid \mathcal{F})$ do not carry over to the approximate projection estimator $\hat{f}^{\mathrm{p}}_m = \Pi(\hat{f} \mid \mathcal{F}_m)$. However, if $\mathcal{F}_m$ is in some sense close to $\mathcal{F}$, then $\hat{f}^{\mathrm{p}}_m$ may still be expected to be close to $\hat{f}^{\mathrm{p}}$. The following lemma serves to get a handle on $\|\hat{f}^{\mathrm{p}}_{n,m_n} - \hat{f}^{\mathrm{p}}_n\|$ (proofs of this and other results are given in the Appendix).

Lemma 1. *If $\mathcal{F}$ and $\mathcal{G}$ are non-empty, closed and convex subsets of $\mathbb{H}$ and if $\mathcal{G} \subset \mathcal{F}$, then for $h \in \mathbb{H}$,*

$$\begin{aligned} \|\Pi(h \mid \mathcal{G}) - \Pi(h \mid \mathcal{F})\| &\leq [\delta\{2\|h - \Pi(h \mid \mathcal{F})\| + \delta\}]^{1/2} \\ &\leq [\delta\{2\|h - \Pi(h \mid \mathcal{G})\| + \delta\}]^{1/2}, \end{aligned}$$

where

$$\delta = \|\Pi(h \mid \mathcal{F}) - \Pi(\Pi(h \mid \mathcal{F}) \mid \mathcal{G})\|$$

is bounded by

$$\sup_{f \in \mathcal{F}} \|f - \Pi(f \mid \mathcal{G})\| = \sup_{f \in \mathcal{F}} \inf_{g \in \mathcal{G}} \|f - g\| = d(\mathcal{F}, \mathcal{G}),$$

the Hausdorff distance between $\mathcal{F}$ and $\mathcal{G}$.

By Lemma 1, the distance between the projection and approximate estimators is bounded by

$$\|\hat{f}^p_m - \hat{f}^p\| \leq [\delta_m\{2\|\hat{f} - \hat{f}^p_m\| + \delta_m\}]^{1/2}, \tag{9}$$

with $\delta_m = \|\hat{f}^p - \Pi(\hat{f}^p \mid \mathcal{F}_m)\| \leq d(\mathcal{F}, \mathcal{F}_m)$. In the case of Pickands dependence functions, we have the following bound on the Hausdorff distance between $\mathcal{A}_m$ and $\mathcal{A}$. We gratefully acknowledge the contribution of the Associate Editor leading to a significant improvement upon the bound in an earlier version of the paper.

Lemma 2. *For $A \in \mathcal{A}$, let $A_m \in \mathcal{A}_m$ be the Pickands dependence function obtained by linearly interpolating A at $\{0, 1/m, 2/m, \ldots, 1\}$. Then*

$$\int (A_m - A)^2 \leq \frac{4}{3} m^{-3}. \tag{10}$$

Combining eqs. (9) and (10), we deduce that for an estimator $\hat{A}$ and its projections $\hat{A}^{\mathrm{p}} = \Pi(\hat{A} \mid \mathcal{A})$ and $\hat{A}^{\mathrm{p}}_m = \Pi(\hat{A} \mid \mathcal{A}_m)$, we have

$$\|\hat{A}^{\mathrm{p}}_m - \hat{A}^{\mathrm{p}}\| \leq (4/3)^{1/4} m^{-3/4} \{2\|\hat{A} - \hat{A}^{\mathrm{p}}_m\| + (4/3)^{1/2} m^{-3/2}\}^{1/2}. \tag{11}$$

The right-hand side of eq. (11) constitutes a readily computable bound for $\|\hat{A}^{\mathrm{p}}_m - \hat{A}^{\mathrm{p}}\|$, which tends to 0 as $m \to \infty$. In practice, such a bound can be computed for several values of m; the selected value of m should be such that this bound is sufficiently small.

Notice that the choice of m is not subject to a bias-variance trade-off: the higher m, the more accurate the approximation A^{p}_m. The only restrictions come from computing time and power. In an asymptotic setting, combination of Lemmas 1 and 2 yields lower bounds to the speed at which m should tend to infinity for the asymptotic behaviors of $\hat{A}^{\mathrm{p}}_m$ and $\hat{A}^{\mathrm{p}}$ to coincide; see eq. (16) below.

4. LIMIT DISTRIBUTIONS

Let $\hat{f}_n$ be a consistent estimator sequence for $f_0 \in \mathcal{F}$ and assume that there exist positive numbers ε_n with $\varepsilon_n \to 0$ as $n \to \infty$ and a random element g in $\mathbb{H}$ such that

$$\frac{\hat{f}_n - f_0}{\varepsilon_n} \rightsquigarrow g, \qquad \text{in } \mathbb{H}; \tag{12}$$

here, the arrow '$\rightsquigarrow$' denotes convergence in distribution in the metric space $\mathbb{H}$. Then what can we say about the limit distributions of the estimator sequences $\hat{f}_n^{\mathrm{P}} = \Pi(\hat{f}_n \mid \mathcal{F})$ and $\hat{f}_{n,m_n}^{\mathrm{P}} = \Pi(\hat{f}_n \mid \mathcal{F}_{m_n})$? Here, m_n is a positive integer sequence tending to infinity. Notice that in the previous display the convergence takes place in the topology of $\mathbb{H}$. In case $\mathbb{H}$ is an L^2 space on a compact interval, this convergence is implied by convergence with respect to the finer topology of uniform convergence.

Since $(\hat{f}_n^{\mathrm{P}} - f_0)/\varepsilon_n = \{\Pi(\hat{f}_n \mid \mathcal{F}) - \Pi(f_0 \mid \mathcal{F})\}/\varepsilon_n$, we can turn (12) into an asymptotic result for $\hat{f}_n^{\mathrm{P}}$ by the functional delta method (van der Vaart 1998, Theorem 20.8). This requires some form of Hadamard differentiability of the projection operator $\Pi(\,\cdot \mid \mathcal{F})$ at f_0. The one-sided derivative turns out to exist and is a projection as well, but now on the tangent cone of $\mathcal{F}$ at f_0, defined as

$$T_{\mathcal{F}}(f_0) = \overline{\mathcal{C}} \quad \text{with} \quad \mathcal{C} = \{\lambda(f - f_0) : \lambda \geq 0,\, f \in \mathcal{F}\};$$

here $\overline{\mathcal{C}}$ denotes the closure of $\mathcal{C}$ in $\mathbb{H}$.

LEMMA 3. *Let $f_0 \in \mathcal{F}$. If $0 < \varepsilon_n \to 0$ and if $g_n \to g$ in $\mathbb{H}$, then*

$$\lim_{n\to\infty} \frac{\Pi(f_0 + \varepsilon_n g_n \mid \mathcal{F}) - f_0}{\varepsilon_n} = \Pi(g \mid T_{\mathcal{F}}(f_0)).$$

THEOREM 1. *If eq. (12) holds, then*

$$\frac{\hat{f}_n^{\mathrm{P}} - f_0}{\varepsilon_n} \rightsquigarrow \Pi\left(g \mid T_{\mathcal{F}}(f_0)\right), \qquad n \to \infty.$$

Moreover, if the integer sequence m_n is such that

$$\|\hat{f}_{n,m_n}^{\mathrm{P}} - \hat{f}_n^{\mathrm{P}}\| = o_p(\varepsilon_n), \qquad n \to \infty,$$

then also

$$\frac{\hat{f}_{n,m_n}^{\mathrm{P}} - f_0}{\varepsilon_n} \rightsquigarrow \Pi\left(g \mid T_{\mathcal{F}}(f_0)\right), \qquad n \to \infty.$$

Since $0 \in T_{\mathcal{F}}(f_0)$, eq. (5)(iii) implies

$$\|g\|^2 \geq \|g - \Pi(g \mid T_{\mathcal{F}}(f_0))\|^2 + \|\Pi(g \mid T_{\mathcal{F}}(f_0))\|^2. \tag{13}$$

In this sense, the limiting random variable for the projection estimator sequence is 'smaller' than the one for the initial estimator sequence.

Let $\hat{A}_n$ be a sequence of estimators for an unknown Pickands dependence function A. Assume that the asymptotics of $\hat{A}_n$ are known and of the form

$$\frac{\hat{A}_n - A}{\varepsilon_n} \rightsquigarrow G \qquad \text{in } L^2, \tag{14}$$

with $0 < \varepsilon_n \to 0$. Typically, the convergence above holds in the stronger topology of uniform convergence. The limit process G is a centered Gaussian process, and ε_n is equal to the reciprocal of the square root of the effective sample size, that is, the number of block maxima (Deheuvels 1991; Capéraà, Fougères & Genest 1997; Segers 2007) or the number of high-threshold exceedances (Capéraà & Fougères 2000; Einmahl, de Haan & Piterbarg 2001).

By Theorem 1, eq. (14) implies that the asymptotic distribution of the projection estimator $\hat{A}_n^{\mathrm{p}} = \Pi(\hat{A}_n \mid \mathcal{A})$ is given by

$$\frac{\hat{A}_n^{\mathrm{p}} - A}{\varepsilon_n} \rightsquigarrow \Pi(G \mid T_{\mathcal{A}}(A)) \qquad \text{in } L^2. \tag{15}$$

Here, $T_{\mathcal{A}}(A)$ is the tangent cone of $\mathcal{A}$ at A, defined as the set of limits (in L^2) of all sequences of the form $\lambda_n(A_n - A)$, with $\lambda_n \in [0, \infty)$ and $A_n \in A$. Furthermore, if m_n is a positive integer sequence such that

$$\varepsilon_n^{2/3} m_n \to \infty, \qquad n \to \infty, \tag{16}$$

then, by Lemmas 1 and 2, denoting $\hat{A}_{n,m}^{\mathrm{p}} = \Pi(\hat{A}_n \mid \mathcal{A}_m)$,

$$\|\hat{A}_n^{\mathrm{p}} - \hat{A}_{n,m_n}^{\mathrm{p}}\| \le O(m_n^{-3/4})\{O_p(\varepsilon_n) + O(m_n^{-3/2})\}^{1/2} = o_p(\varepsilon_n),$$

guaranteeing that eq. (15) remains true if $\hat{A}_n^{\mathrm{p}}$ is replaced by $\hat{A}_{n,m_n}^{\mathrm{p}}$.

According to equation (13), $\int G^2 \ge \int \{\Pi(G \mid T_{\mathcal{A}}(A))\}^2$, suggesting the improved performance of the projection estimator in comparison to the initial estimator. The amount of improvement depends on the tangent cone $T_{\mathcal{A}}(A)$: the smaller this tangent cone, the larger the improvement. In a number of special cases, the tangent cone admits an explicit description.

Lemma 4. *Let $A \in \mathcal{A}$.*

(i) If A is twice differentiable and $\inf_t A''(t) > 0$, then $T_{\mathcal{A}}(A) = L^2$.

(ii) If $A = 1$, then $T_{\mathcal{A}}(A)$ is the set of non-positive, convex functions.

In case (i), the projection estimator has the same asymptotic distribution as the initial estimator on which it is based. In case (ii) of independent margins [$C(u, v) = uv$ in eq. (1)], the integrated squared error of the least-squares estimator will typically be smaller than the one of the corresponding initial estimator.

5. NUMERICAL EXAMPLES

5.1 Simulation study

In order to illustrate the finite-sample properties of the projection estimator, we generated data from bivariate extreme-value distributions with the following Pickands dependence functions:

- the independent copula, $A(t) = 1$ (Figure 1, first row);
- the Gumbel (1960) or logistic model,

$$A(t) = \{t^{1/\alpha} + (1-t)^{1/\alpha}\}^{\alpha}, \qquad t \in [0, 1], \tag{17}$$

  with parameter range $\alpha \in (0, 1]$; here, we took $\alpha = 0.9$ (Figure 1, second row);
- the asymmetric logistic model (Tawn 1988),

$$A(t) = (1-\psi_1)t + (1-\psi_2)(1-t) + [(\psi_1 t)^{1/\alpha} + \{\psi_2(1-t)\}^{1/\alpha}]^{\alpha}, \qquad t \in [0, 1], \tag{18}$$

  with parameter range $\alpha \in (0, 1]$ and $\psi_j \in [0, 1]$ $(j = 1, 2)$; here, we took $(\alpha, \psi_1, \psi_2) = (0.7, 0.1, 0.5)$ (Figure 1, third row).

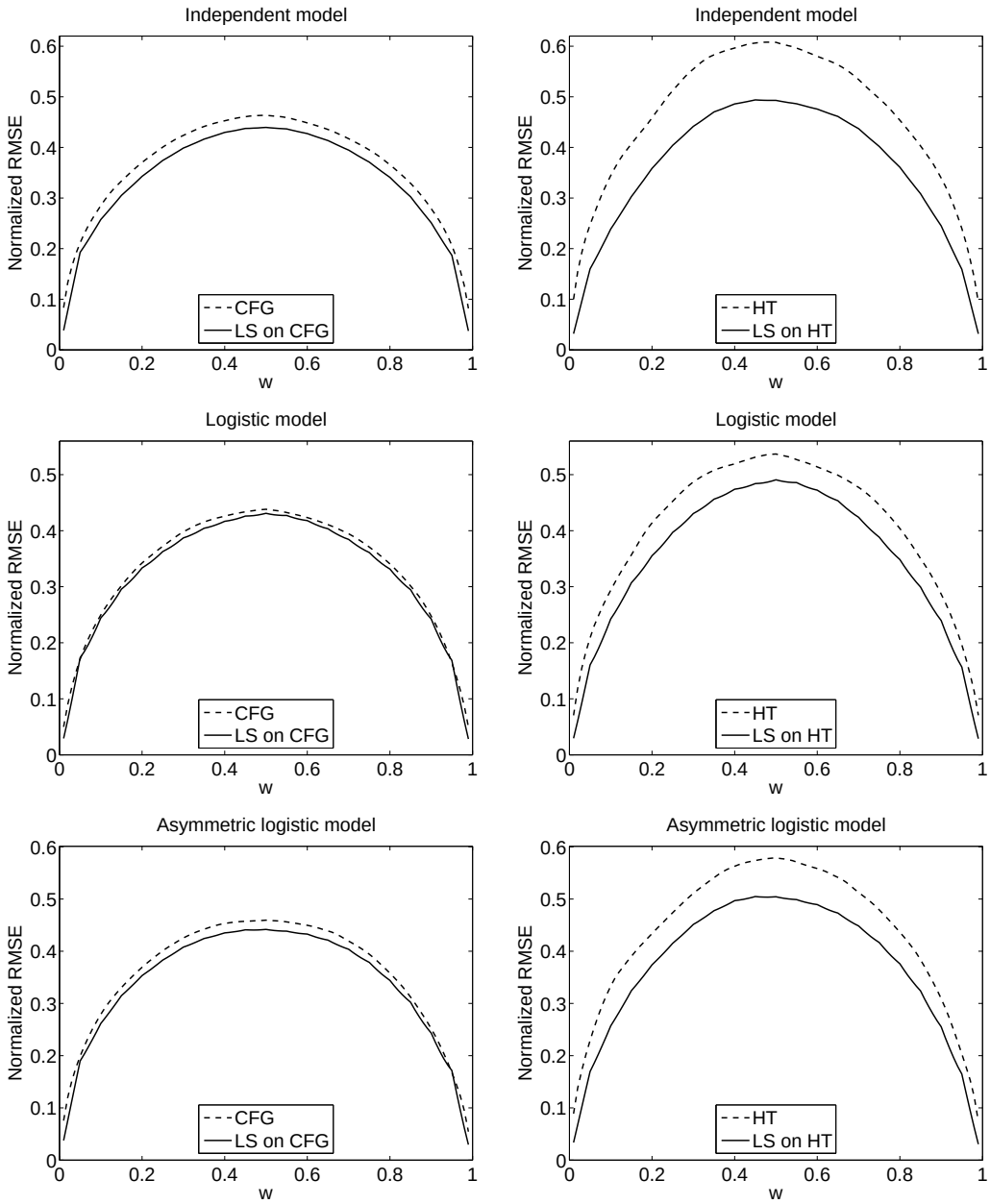


Figure 1: Normalized point-wise RMSE for original Pickands dependence function estimators (dashed line: CFG on the left, HT on the right) and projection versions (full line). First row: independent model; second row: logistic model; third row: asymmetric logistic model.

In each case, 500 samples were generated of size $n = 100$. The performance of the estimators is visualised by their normalized point-wise root mean squared errors (RMSE), defined as

$$\left(\frac{n}{500}\sum_{i=1}^{500}\{A_n^{(i)}(t) - A(t)\}^2\right)^{1/2},$$

this quantity being an estimate of the standard deviation of the limit distribution of $n^{1/2}\{A_n(t) - A(t)\}$, with A_n representing any of the estimators considered.

The initial estimators were the Capéraà–Fougères–Genest (CFG) estimator (Capéraà, Fougères & Genest 1997) with tuning parameter $p(t) = t$ and the Hall–Tajvidi (HT) estimator (Hall & Tajvidi 2000). The marginal distributions were estimated by the empirical distribution functions.

For the actual computations, we implemented the approximation method of section 3. As the rate of convergence of the CFG and HT estimators is $O_p(n^{-1/2})$, their projections on the full class $\mathcal{A}$ and on the subclasses $\mathcal{A}_{m_n}$ will be asymptotically undistinguishable as soon as $n^{-1/3}m_n \to \infty$; see eq. (16). In our simulations, we experimented with several values of m. It turned out that as soon as $m \geq 20$, there was no visible difference anymore between the approximate projection estimators. In Figure 1, only the results for $m = 20$ are shown.

For each model considered and every $t \in [0, 1]$, the least-squares estimator had a smaller root mean squared error than the initial estimator on which it was based. In accordance to Lemma 4(ii), the improvement was largest in case of independence.

5.2 Data example

In Frees & Valdez (1998), copula fitting was illustrated on a data-set of 1 500 insurance company indemnity claims, displayed in Figure 2(a). Each claim consists of an indemnity payment (Loss) and an allocated loss adjustment expense (ALAE), that is, an expense specifically attributable to the settlement of the individual claim such as lawyers' fees and claims investigation expenses. In Genest, Ghoudi & Rivest (1998), the pseudo-maximum likelihood estimate for the parameter of the bivariate extreme-value copula (1) with Gumbel or logistic Pickands dependence function (17) was found to be $\hat{\alpha} = 0.69$ [dotted line in Figure 2(b)].

Rather than postulating a parametric model for the copula of the joint distribution of the pair $(X, Y) = (\text{Loss}, \text{ALAE})$, we proceed from the weaker assumption that this copula is in the maximal domain of attraction of some bivariate extreme-value copula. Denoting the marginal distribution functions of X and Y by F and G respectively, this assumption means that the limit

$$\lim_{s\downarrow 0} s^{-1}\Pr[1 - F(X) \leq sx \text{ or } 1 - G(Y) \leq sy] =: l(x, y) \tag{19}$$

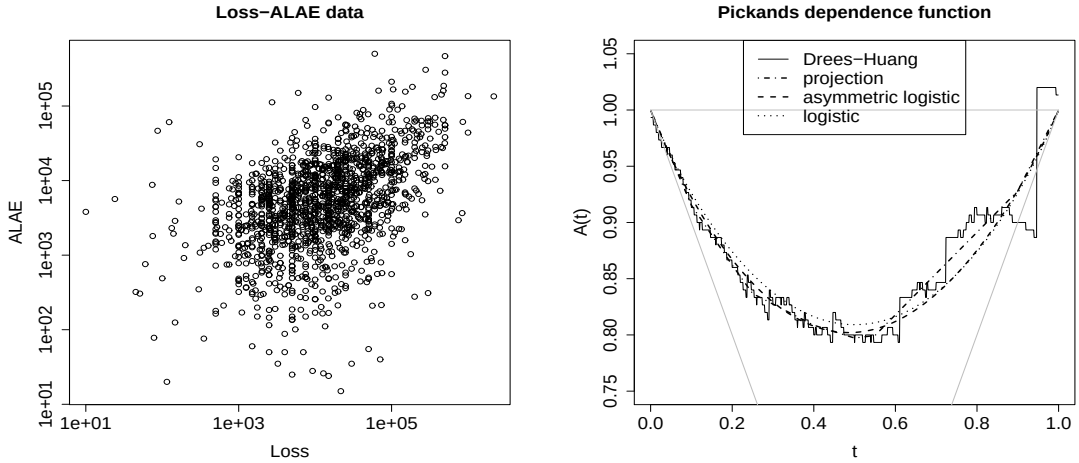
exists for all $(x, y) \in [0, \infty)^2$. The limit l, called stable tail dependence function, is of the form

$$l(x, y) = (x + y)A\left(\frac{x}{x + y}, \frac{y}{x + y}\right) \qquad \text{with } A(t) = l(t, 1 - t) \tag{20}$$

for $(x, y) \in [0, \infty)^2 \setminus \{(0, 0)\}$ and $t \in [0, 1]$, where A is a Pickands dependence function (Huang 1992; Drees & Huang 1998).

Given the data $(X_1, Y_1), \ldots, (X_n, Y_n)$, one can estimate l nonparametrically by replacing the unknown joint and marginal distributions in (19) by their empirical counterparts:

$$\hat{l}_n(x, y) = \frac{1}{k}\sum_{i=1}^{n}\mathbf{1}(R_i > n - kx + 1, S_i > n - ky + 1), \qquad (x, y) \in [0, \infty)^2, \tag{21}$$



(a) (b)

Figure 2: Data example: Loss-ALAE data in Frees & Valdez (1998). (a) Scatterplot of the data (axes on log scale base 10). (b) Estimates for the Pickands dependence function. Full line: nonparametric estimate (22); Dot-dashed line: projection of the nonparametric estimate; Dashed line: parametric estimate based on the asymmetric logistic model (18); Dotted line: parametric estimate based on the Gumbel or logistic model (17).

with R_i the rank of X_i among $X_1, \ldots, X_n$ and S_i the rank of Y_i among $Y_1, \ldots, Y_n$ (Huang 1992; Drees & Huang 1998). Here, $s = k/n$ and $k = k_n$ is a positive integer sequence such that $k \to \infty$ and $k/n \to \infty$ as $n \to \infty$. Combining (20) with (21) yields the nonparametric estimator

$$\hat{A}_n(t) = \hat{l}_n(t, 1-t), \qquad t \in [0,1]. \tag{22}$$

The estimator $\hat{A}_n$ is typically not a Pickands dependence function itself: for the Loss-ALAE data, nor the bounds nor the convexity constraint are fulfilled [full line in Figure 2(b), $k = 150$]. This problem is solved by applying the projection operator (4), yielding the projection estimate $\hat{A}_n^{\mathrm{p}}$ [dot-dashed line in Figure 2(b)]. The shown estimate is in fact the finite-dimensional approximation of section 3 with $m = 20$ computed using the R package quadprog (Weingessel 2007); higher values for m did not change the result visibly. Alternative nonparametric estimators of A (not shown) are proposed in Abdous, Ghoudi & Khoudraji (1999) and Capéraà & Fougères (2000).

If l has continuous first-order partial derivatives and under a certain growth restriction on k determined by the speed of convergence in (19) of the limit on the left, it is known that in the space $D([0,1]^2)$ equipped with the supremum norm, the processes $k^{1/2}(\hat{l}_n - l)$ converge weakly to some mean-zero Gaussian process with continuous sample paths and with covariance function depending only on l (Huang 1992; Drees & Huang 1998; Einmahl, de Haan & Li 2006). As a consequence, the processes $k^{1/2}(\hat{A}_n - A)$ must converge to a mean-zero Gaussian process with continuous sample paths as well. By Theorem 1, $k^{1/2}(\hat{A}_n^{\mathrm{p}} - A)$ converges weakly in $L^2([0,1], \mathrm{d}x)$. Under the conditions of Lemma 4(i), the projection step does not alter the distribution of the limit process.

In Beirlant *et al.* (2004, section 9.3.4), the asymmetric logistic model (18) was fitted to the Loss-ALAE data using the censored likelihood method (Smith 1994; Ledford and Tawn 1996), resulting in the parameter estimates $(\hat{\alpha}, \hat{\psi}_1, \hat{\psi}_2) = (0.66, 1, 0.89)$ [dashed line in Figure 2(b)]. Moving away from the Gumbel or logistic dependence function to the asymmetric logistic model leads to a clear improvement in fit on the left-hand side of the plot.

APPENDIX

Proof of Lemma 1. By eq. (5)(iii), since $\Pi(h \mid \mathcal{G}) \in \mathcal{G} \subset \mathcal{F}$,

$$\|\Pi(h \mid \mathcal{F}) - \Pi(h \mid \mathcal{G})\|^2 \leq \|h - \Pi(h \mid \mathcal{G})\|^2 - \|h - \Pi(h \mid \mathcal{F})\|^2.$$

Since $\Pi(\Pi(h \mid \mathcal{F}) \mid \mathcal{G}) \in \mathcal{G}$ and by an application of the triangle inequality and the definition of δ,

$$\begin{aligned} \|h - \Pi(h \mid \mathcal{G})\| &\leq \|h - \Pi(\Pi(h \mid \mathcal{F}) \mid \mathcal{G})\| \\ &\leq \|h - \Pi(h \mid \mathcal{F})\| + \delta. \end{aligned}$$

Combine the previous two displays to get

$$\begin{aligned} \|\Pi(h \mid \mathcal{F}) - \Pi(h \mid \mathcal{G})\|^2 &\leq \{\|h - \Pi(h \mid \mathcal{F})\| + \delta\}^2 - \|h - \Pi(h \mid \mathcal{F})\|^2 \\ &= 2\delta\|h - \Pi(h \mid \mathcal{F})\| + \delta^2. \end{aligned}$$

Take square roots on both sides of the display to finish the proof. □

Proof of Lemma 2. Put $d_k = m\{A(k/m) - A((k-1)/m)\}$ for $k = 1, \ldots, m$ and put $d_0 = -1$. Furthermore, put $e_k = d_k - d_{k-1}$ for $k = 1, \ldots, m$.

Let $t \in (0, 1]$ and let $k = 1, \ldots, m$ be such that $(k-1)/m < t \leq k/m$. By convexity and the fact that $A(t) \geq 1 - t$,

$$A(t) \begin{cases} \leq & A\left(\dfrac{k-1}{m}\right) + \left(t - \dfrac{k-1}{m}\right) d_k = A_m(t), \\ \geq & A\left(\dfrac{k-1}{m}\right) + \left(t - \dfrac{k-1}{m}\right) d_{k-1}, \end{cases}$$

whence

$$0 \leq A_m(t) - A(t) \leq \left(t - \frac{k-1}{m}\right) e_k.$$

Integrate over t to obtain

$$\int (A_m - A)^2 \leq \frac{1}{3m^3} \sum_{k=1}^{m} e_k^2.$$

Since $-1 = d_0 \leq d_1 \leq \ldots \leq d_m \leq 1$, we have $0 \leq e_k \leq 2$ for all $k = 1, \ldots, m$, as well as $\sum_{k=1}^{m} e_k = d_m - d_0 \leq 1 - (-1) = 2$. Therefore,

$$\sum_{k=1}^{m} e_k^2 \leq \left(\sum_{k=1}^{m} e_k\right)^2 \leq 2^2 = 4.$$

Combine the final two displays to conclude the proof. □

Proof of Lemma 3. By eq. (6), we have

$$\left\| \frac{\Pi(x + \varepsilon_n y_n \mid \mathcal{F}) - x}{\varepsilon_n} - \frac{\Pi(x + \varepsilon_n y \mid \mathcal{F}) - x}{\varepsilon_n} \right\| \leq \|y_n - y\|.$$

Hence, without loss of generality we may take assume $y_n = y$ for all n.

First, assume that $y \in T_{\mathcal{F}}(x)$. If $y = 0$ there is nothing to prove, so suppose $y \neq 0$. There exists a sequence $z_n = \lambda_n(x_n - x)$ with $\lambda_n > 0$ and $x_n \in \mathcal{F}$ such that $z_n \to y$ as $n \to \infty$. Without loss of generality we may assume that $\lambda_n \varepsilon_n \leq 1$. We have

$$\begin{aligned}
&\left\| \frac{\Pi(x + \varepsilon_n y \mid \mathcal{F}) - x}{\varepsilon_n} - y \right\| \\
&\leq \left\| \frac{\Pi(x + \varepsilon_n y \mid \mathcal{F}) - x}{\varepsilon_n} - \frac{\Pi(x + \varepsilon_n z_n \mid \mathcal{F}) - x}{\varepsilon_n} \right\| \\
&\quad + \left\| \frac{\Pi(x + \varepsilon_n z_n \mid \mathcal{F}) - x}{\varepsilon_n} - z_n \right\| + \|z_n - y\|.
\end{aligned}$$

By eq. 6, the first term on the right is at most $\|y - z_n\|$. Moreover, $x + \varepsilon_n z_n = x + \varepsilon_n \lambda_n (x_n - x) \in \mathcal{F}$, so the middle term on the right-hand side of the previous display vanishes. Hence

$$\left\| \frac{\Pi(x + \varepsilon_n y \mid \mathcal{F}) - x}{\varepsilon_n} - y \right\| \leq 2\|z_n - y\| \to 0,$$

as required.

Next, take a general $y \in \mathbb{H}$ and put $v = \Pi(y \mid T_{\mathcal{F}}(x))$ and $v_n = \varepsilon_n^{-1}\{\Pi(x + \varepsilon_n v \mid \mathcal{F}) - x\}$. We have

$$\begin{aligned}
\left\| y - \frac{\Pi(x + \varepsilon_n y \mid \mathcal{F}) - x}{\varepsilon_n} \right\| &= \varepsilon_n^{-1} \|(x + \varepsilon_n y) - \Pi(x + \varepsilon_n y \mid \mathcal{F})\| \\
&\leq \varepsilon_n^{-1} \|(x + \varepsilon_n y) - \Pi(x + \varepsilon_n v \mid \mathcal{F})\| \\
&= \|y - v_n\|.
\end{aligned}$$

By the previous paragraph, $v_n \to v$, whence

$$\limsup_{n \to \infty} \left\| y - \frac{\Pi(x + \varepsilon_n y \mid \mathcal{F}) - x}{\varepsilon_n} \right\| \leq \|y - v\|.$$

On the other hand, since $\varepsilon_n^{-1}\{\Pi(x + \varepsilon_n y \mid \mathcal{F}) - x\} \in T_{\mathcal{F}}(x)$ and $v = \Pi(y \mid T_{\mathcal{F}}(x))$, eq. (5)(iii) implies

$$\left\| y - \frac{\Pi(x + \varepsilon_n y \mid \mathcal{F}) - x}{\varepsilon_n} \right\|^2 \geq \|y - v\|^2 + \left\| v - \frac{\Pi(x + \varepsilon_n y \mid \mathcal{F}) - x}{\varepsilon_n} \right\|^2.$$

Combine the previous two displays to obtain that the second term on the right-hand side of the last display converges to zero, as required. □

Proof of Lemma 4. (i) Fix $f \in L^2$. By the dominated convergence theorem, $\|f - f\mathbf{1}_{[1/n, 1-1/n]}\|_2 \to 0$ as $n \to \infty$. Define $g_n \in L^2$ by $g_n(t) = f(t)t^{-2}(1-t)^{-2}\mathbf{1}_{[1/n,1-1/n]}(t)$. Since the set of polynomials is dense in L^2, there exist polynomials p_n such that $\|g_n - p_n\|_2 \to 0$ as $n \to \infty$. Define $q_n \in L^2$ by $q_n(t) = t^2(1-t)^2 p_n(t)$. Then

$$\begin{aligned}
\|f - q_n\|_2 &\leq \|f - f\mathbf{1}_{[1/n,1-1/n]}\|_2 + \|f\mathbf{1}_{[1/n,1-1/n]} - q_n\|_2 \\
&\leq \|f - f\mathbf{1}_{[1/n,1-1/n]}\|_2 + \|g_n - p_n\|_2 \\
&\to 0, \qquad n \to \infty.
\end{aligned}$$

Now, let $\lambda_n > 0$ be large enough so that

$$\lambda_n^{-1} \sup\{|q_n''(t)| : t \in [0,1]\} \leq \inf\{A''(t) : t \in [0,1]\}.$$

Define $A_n = A + \lambda_n^{-1} q_n$. Clearly, $\|f - \lambda_n(A_n - A)\|_2 \to 0$ as $n \to \infty$. Moreover, $A_n \in \mathcal{A}$, as follows from $A_n(0) = A(0) = 1$, $A_n(1) = A(1) = 1$, $A_n'(0) = A'(0) \in [-1, 0]$, $A_n'(1) = A'(1) \in [0, 1]$, and $A_n'' = A'' + \lambda_n^{-1} q_n'' \geq 0$. Together, $f \in T_{\mathcal{A}}(A)$.

(ii) Let f be a convex, non-positive function on $[0, 1]$. For positive integer n, define

$$f_n(x) = \begin{cases} nxf(1/n) & \text{if } 0 \leq x \leq 1/n, \\ f(x) & \text{if } 1/n \leq x \leq 1 - 1/n, \\ n(1-x)f(1-1/n) & \text{if } 1 - 1/n \leq x \leq 1. \end{cases}$$

The function f_n is convex, and since $f \leq f_n \leq 0$ and $f_n(t) \to f(t)$ for all $0 < t < 1$, also $\|f_n - f\|_2 \to 0$ as $n \to \infty$ by dominated convergence. Let $\lambda_n > 0$ be such that $\lambda_n \geq n \max\{|f(1/n)|, |f(1-1/n)|\}$ and define $A_n = 1 + \lambda_n^{-1} f_n$. Then $\|f - \lambda_n(A_n - 1)\|_2 \to 0$ as $n \to \infty$. Moreover, $A_n \in \mathcal{A}$ as $A_n(0) = 1$, $A_n(1) = 1$, $A_n'(0) = \lambda_n^{-1} n f(1/n) \in [-1, 0]$, and $A_n'(1) = \lambda_n^{-1} n |f(1 - 1/n)| \in [0, 1]$. Hence $f \in T_{\mathcal{A}}(1)$.

Conversely, let $f \in T_{\mathcal{A}}(1)$. Then $\|f - \lambda_n(A_n - 1)\|_2 \to 0$ for some sequence λ_n of positive numbers and some sequence A_n in $\mathcal{A}$. Each $f_n = \lambda_n(A_n - 1)$ is non-positive and convex. Then, along some subsequence $(n_k)_k$, we have $f_{n_k}(t) \to f(t)$ as $k \to \infty$ for almost every $t \in [0, 1]$ (Doob 1994, p. 68 and p. 90). Hence f is almost everywhere equal to some non-positive, convex function. □

ACKNOWLEDGEMENTS

The research by Johan Segers was supported by the IAP research network grant nr. P6/03 of the Belgian government (Belgian Science Policy). Exceptionally constructive comments of the Editor and an anonymous Associate Editor and Referee are gratefully acknowledged.

REFERENCES

B. Abdous, K. Ghoudi & A. Khoudraji (1999). Non-parametric estimation of the limit dependence function of multivariate extremes. *Extremes*, 2, 245–268.

B. Abdous & K. Ghoudi (2005). Non-parametric estimators of multivariate extreme dependence functions. *Nonparametric Statistics*, 17, 915–935.

J. Beirlant, Y. Goegebeur, J. Segers & J. Teugels (2004). *Statistics of Extremes: Theory and Applications.* Wiley, New York.

P. Capéraà & A.-L. Fougères (2000). Estimation of a bivariate extreme value distribution. *Extremes*, 3, 311–329.

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

P. Deheuvels (1984). Probabilistic aspects of multivariate extremes. In: *Statistical Extremes and Applications* (editor: J. Tiago de Oliveira), 117–130. Reidel.

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

J. L. Doob (1994). *Measure theory.* Springer-Verlag, New York.

H. Drees & X. Huang (1998). Best attainable rates of convergence for estimators of the stable tail dependence function. *Journal of Multivariate Analysis*, 64, 25–47.

J. H. J. Einmahl, L. de Haan & D. Li (2006). Weighted approximations of tail copula processes with application to testing the bivariate extreme value condition. *The Annals of Statistics*, 34, 1987–2014.

J. H. J. Einmahl, L. de Haan & V. I. Piterbarg (2001). Nonparametric estimation of the spectral measure of an extreme value distribution. *The Annals of Statistics*, 29, 1401–1423.

E. W. Frees & E. A. Valdez (1998). Understanding relationships using copulas. *North American Actuarial Journal*, 2, 1–25

C. Genest, K. Ghoudi & L.-P. Rivest (1998). Discussion of "Understanding Relationships Using Copulas," by Edward Frees and Emiliano Valdez, January 1998. *North American Actuarial Journal* 2, 143–149.

K. Ghoudi, A. Khoudraji & L.-P. Rivest (1998). Propriétés statistiques des copules de valeurs extrêmes bidimensionnelles. *The Canadian Journal of Statistics*, 26, 187–197.

E. J. Gumbel (1960). Bivariate exponential distributions. *Journal of the American Statistical Association*, 55, 698-707.

L. de Haan & S. I. Resnick (1977). Limit theory for multivariate sample extremes. *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete*, 40, 317–337.

P. Hall & N. Tajvidi (2000). Distribution and dependence-function estimation for bivariate extreme-value distributions. *Bernoulli*, 6, 835–844.

X. Huang (1992). *Statistics of bivariate extremes.* Ph.D. dissertation, Erasmus University Rotterdam, Tinbergen Institute Series no. 22.

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

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

E. Mammen, J. S. Marron, B. A. Turlach & M. P. Wand (2001). A general projection framework for constrained smoothing. *Statistical Science*, 16, 232–248.

R. B. Nelsen (1999). *An introduction to copulas.* Lecture Notes in Statistics 139, Springer Verlag, New York.

J. Pickands III (1981). Multivariate extreme value distributions. *Bulletin of the International Statistical Institute*, 859–878.

J. Pickands III (1989). Multivariate negative exponential and extreme value distributions. In: *Extreme Value Theory* (editors: J. Hüsler and R. D. Reiss), *Lecture Notes in Statistics*, 51.

J. Segers (2007). Non-parametric inference for bivariate extreme-value copulas. In: *Extreme Value Distributions* (editors: M. Ahsanullah, S. Kirmani) Nova Science Publishers, 181–203.

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

R. L. Smith (1994). Multivariate threshold methods. In: *Extreme Value Theory and Applications* (editors: J. Galambos, J. Lechner & E. Simiu), 225–248. Kluwer.

R. L. Smith, J. A. Tawn & H. K. Yuen (1990). Statistics of multivariate extremes. *International Statistical Review*, 58, 47–58.

H. Stark & Y. Yang (1998). *Vector Space Projections.* Wiley, New York.

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

A. W. van der Vaart (1998). *Asymptotic Statistics.* Cambridge University Press, Cambridge.

A. Weingessel (2007). quadprog: Functions to solve quadratic programming problems (S original by B. A. Turlach). R package version 1.4-11.

Received ???
Accepted ???

Amélie FILS-VILLETARD: fils@ccr.jussieu.fr
Université Pierre et Marie Curie
Laboratoire de Statistique Théorique et Appliquée
175, rue du Chevaleret
75013 Paris, FRANCE

Armelle GUILLOU: guillou@math.u-strasbg.fr
Université de Strasbourg, IRMA
Département de Mathématique
7, rue René Descartes
67084 Strasbourg cedex, FRANCE

Johan SEGERS: johan.segers@uclouvain.be
Université catholique de Louvain
Institut de statistique
20, Voie du Roman Pays
B-1348 Louvain-la-Neuve, BELGIQUE