Multivariate extremes over a random number of observations

Enkelejd Hashorva¹ | Simone A. Padoan² | Stefano Rizzelli³

¹Department of Actuarial Science, University of Lausanne
²Department of Decision Sciences, Bocconi University
³Chair of Statistics STAT, École Polytechnique Fédérale de Lausanne

Correspondence
Simone A. Padoan, Department of Decision Sciences, Bocconi University, via Roentgen 1, 20136 Milan, Italy.
Email: simone.padoan@unibocconi.it

Abstract
The classical multivariate extreme-value theory concerns the modeling of extremes in a multivariate random sample, suggesting the use of max-stable distributions. In this work, the classical theory is extended to the case where aggregated data, such as maxima of a random number of observations, are considered. We derive a limit theorem concerning the attractors for the distributions of the aggregated data, which boil down to a new family of max-stable distributions. We also connect the extremal dependence structure of classical max-stable distributions and that of our new family of max-stable distributions. Using an inversion method, we derive a semiparametric composite-estimator for the extremal dependence of the unobservable data, starting from a preliminary estimator of the extremal dependence of the aggregated data. Furthermore, we develop the large-sample theory of the composite-estimator and illustrate its finite-sample performance via a simulation study.

KEYWORDS
extremal dependence, extreme-value copula, inverse problem, multivariate max-stable distribution, nonparametric estimation, Pickands dependence function
1 | INTRODUCTION AND BACKGROUND

The multivariate extreme-value theory aims to quantify the probability of extreme events concerning multiple dependent observations. A commonly employed approach for modeling extremes in high dimensions is the componentwise maxima, where, for each of the involved variables, the partial maximum values are taken into account, for example, yearly maxima (e.g., Falk, Hüsler, & Reiss, 2011, ch. 4). Basic foundations of the componentwise maxima approach are here briefly introduced.

First, however, we specify the notation that we use throughout the article. Given \( \mathcal{X} \subseteq \mathbb{R}^n, n \in \mathbb{N} \), let \( \ell^\infty(\mathcal{X}) \) denote the spaces of bounded real-valued functions on \( \mathcal{X} \). For \( f : \mathcal{X} \mapsto \mathbb{R} \), let \( \|f\|_\infty = \sup_{x \in \mathcal{X}} |f(x)| \). The arrows “\( \rightarrow \)”, “\( \xrightarrow{p} \)”, “\( \xrightarrow{a.s.} \)” denote convergence (outer) almost surely, convergence in (outer) probability, and convergence in distribution of random vectors (see van der Vaart, 2000, ch. 2) or weak convergence of random functions in \( \ell^\infty(\mathcal{X}) \) (see van der Vaart, 2000, ch. 18, 19), the distinction between the two will be clear from the context. For a nondecreasing function \( f \), let \( f^- \) denote the left-continuous inverse of \( f \). The abbreviation \( a \sim b \) stands for \( a \) is asymptotically equivalent to \( b \). Finally, the multiplication, division and maximum operation between vectors is meant componentwise.

Let \( \mathbf{X} = (X_1, \ldots, X_d) \) be a \( d \)-dimensional random vector with distribution \( F_X \) and margins \( F_{X_j} \), \( j = 1, \ldots, d \), and \( X_1, X_2, \ldots \) be independent and identically distributed (iid) copies of \( X \). Assume that \( F_X \) is in the maximum-domain of attraction (simply domain of attraction) of a multivariate extreme-value distribution \( G \), in symbols \( F_X \in D(G) \). This means that there are sequences of constants \( a_n > 0 = (0, \ldots, 0) \) and \( b_n \in \mathbb{R}^d \) such that \( (\max(X_1, \ldots, X_n) - b_n)/a_n \xrightarrow{d} \eta \) as \( n \to \infty \), where the distribution of \( \eta \) is a multivariate extreme-value distribution (e.g., Falk et al., 2011, pp. 147-153) of the form

\[
G(x) = C_G(G_1(x_1), \ldots, G_d(x_d)), \quad x \in \mathbb{R}^d.
\]

Precisely, \( G \)'s are members of the generalized extreme-value (GEV) class of distributions (e.g., Falk et al., 2011, p. 21) and \( C_G \) is an extreme-value copula, that is,

\[
C_G(u) = \exp(-L((-\ln u_1), \ldots, (-\ln u_d))), \quad u \in (0, 1)^d,
\]

where \( L : [0, \infty)^d \mapsto [0, \infty) \) is the so-called stable-tail dependence function (e.g., Falk et al., 2011, pp. 177–179). \( G \) is a max-stable distribution, that is, for \( k = 1, 2, \ldots \), there are norming sequences \( a_k > 0 \) and \( b_k' \in \mathbb{R}^d \) such that \( \frac{G_k(a_k'x + b_k')}{a_k} = G(x) \), for all \( x \in \mathbb{R}^d \). The extreme-value copula expresses the dependence among extremes. Examples of parametric extreme-value copula models are: the logistic or Gumbel, the Hüsler-Reiss, and the extremal-\( t \), to name a few. An extensive list of additional models is available in Joe (2015, ch. 4). Since \( L \) is a homogeneous function of order 1, it can be conveniently represented as

\[
L(z) = (z_1 + \ldots + z_d) \ A(t), \quad z \in [0, \infty)^d,
\]

where \( t_j = z_j/(z_1 + \ldots + z_d) \) for \( j = 1, \ldots, d \). The function \( A \), named Pickands (dependence) function, denotes the restriction of \( L \) on the \( d \)-dimensional unit simplex \( S_d := \{ (v_1, \ldots, v_d) \in [0, 1]^d : v_1 + \ldots + v_d = 1 \} \). It summarizes the extremal dependence among the components of \( \eta \), specifically it holds that \( 1/d \leq \max(t_1, \ldots, t_d) \leq A(t) \leq 1 \), where the
lower and upper bounds represent the cases of complete dependence and independence. A synthesis of the extremal dependence is provided by the extremal coefficient, that is, \( \theta(G) = dA(1/d, \ldots, 1/d) \in [1, d] \). It can be interpreted as the (fractional) number of independent variables in a \( d \)-dimensional random vector with joint distribution \( G \) and common margins. An alternative summary index that measures the dependence among observations falling in the upper tail region is the coefficient of upper tail dependence (e.g., Joe, 2015, ch. 2.13). In the bivariate case, it is equal to

\[
\lambda(F_X) = \lim_{u \uparrow 1} \mathbb{P}(X_1 > F_{X_1}^-(u)|X_2 > F_{X_2}^-(u)) = \lim_{u \uparrow 1} \mathbb{P}(X_2 > F_{X_2}^-(u)|X_1 > F_{X_1}^-(u)) \in [0, 1].
\]

It is said that \( F_X \) exhibits independence or dependence in the upper tail whenever \( \lambda(F_X) = 0 \) or \( \lambda(F_X) > 0 \), respectively, with the case of complete dependence covered when \( \lambda(F_X) = 1 \). The coefficient \( \lambda(F_X) \) is linked to the extremal coefficient by the relationship \( \theta(G) = 2 - \lambda(F_X) \).

Nowadays, applications involving complex phenomena frequently deal with the analysis of aggregated data. This is especially true in big-data problems, where it may be convenient (or unavoidable) to work with aggregated data in order to reduce the computational cost. Examples of aggregated data are the total amounts and maximum amounts obtained on a random number of observations. Such aggregated data are realizations of the random vectors

\[
S_N = \left( \sum_{i=1}^{N} X_{i,1}, \ldots, \sum_{i=1}^{N} X_{i,d} \right), \quad M_N = \left( \max_{1 \leq i \leq N} X_{i,1}, \ldots, \max_{1 \leq i \leq N} X_{i,d} \right),
\]

where \( N \) is a discrete random variable defined on \( \mathbb{N}_+ = \mathbb{N} \setminus \{0\} \). Assume hereafter that \( N \) with distribution \( F_N \) is independent of \( X_i \)'s. For dimension \( d = 1 \), the tail behavior of \( S_N \) has been extensively studied in the literature (e.g., Embrechts, Klüppelberg, & Mikosch, 1997; Robert & Segers, 2008) and only few results are known on the extremes of \( M_N \) (Barakat & El-Shandidy, 1990; Silvestrov & Teugels, 1998).

The first main purpose of this contribution is to extend the classical probabilistic theory on the extreme-values to the case when the latter are computed using replicates of \( M_N \) (the random vector that represents aggregated data). In Section 5 we illustrate an application, in the context of big data on Internet traffic, which would benefit from such new theoretical developments. In the literature there are no results which can point to how different the extremal behavior of \( F_{M_N} \) is with respect to \( F_X \), some preliminary findings are available in Hashorva, Ratovomirija, and Šiauliai-Tamraz (2017). In this work, with Theorem 1, we provide the attractor for the joint distribution of \((M_N, N)\), wherefrom the extremal behavior of \( M_N \) (our main focus) can be deduced, such as the tail behavior of \( F_{M_N} \).

The more interesting results are obtained in the case where \( F_N \) is very heavy-tailed, which appears to be the more relevant one in a context of big-data, as more data are produced and aggregating them is beneficial. Specifically, when \( \mathbb{P}(N > y) = y^{-\alpha} \mathcal{L}(y) \), with \( y > 0 \) and \( \alpha \in (0, 1) \), where \( \mathcal{L} \) is a slowly varying function at infinity, and \( F_X \in D(G) \), we show that \( M_N \) and \( N \) are asymptotically dependent. Furthermore, we find that in this case \( F_{M_N} \in D(G_{\alpha}) \), where \( G_{\alpha} \) is a new max-stable distribution with an extreme-value copula \( C_{G_{\alpha}} \), given in (7), differing from \( C_G \), the extreme-value copula of \( G \). The coefficient \( \alpha \in (0, 1) \) influences the extremal dependence structure of \( G_{\alpha} \). A practical implication of our finding is that the extremal properties of \( F_{M_N} \) can be recovered by knowing the extremal properties of \( F_X \) and the tail behavior of \( N \). Here are two examples. The joint upper-tail probability of \( M_N \) can be approximated as
\[ \mathbb{P} \left( F_{M_N}^{(1)} \left( M_N^{(1)} \right) > 1 - \frac{1}{y_1} \text{ or } \ldots \text{ or } F_{M_N}^{(d)} \left( M_N^{(d)} \right) > 1 - \frac{1}{y_d} \right) \sim L^\alpha(y^{-1/\alpha}), \]

for a large enough \( y > 0 \), where \( M_N^{(j)} = \max_{1 \leq i \leq N} X_{ij}, j = 1, \ldots, d \) and \( L \) is the stable-tail dependence function of \( C_G \) (see Section 2.2). This means that by combining \( L \) and \( \alpha \), we can approximate the probability that at least one component among \( M_N^{(1)} \), \( \ldots \), \( M_N^{(d)} \) exceeds a high percentile of its own distribution. When \( d = 2 \), we also have

\[ \lambda \left( F_{M_N} \right) = 2 - \left[ 2 - \lambda \left( F_X \right) \right]^\alpha. \] (3)

The second purpose of this contribution concerns an inverse statistical problem in the context of extremes of aggregated data, which also motivates the study of the asymptotic joint distribution of \( (M_N, N) \). Precisely, in applications where the variables \( M_N \) and \( N \) are observable in place of \( X \), the interest may however be in inferring the extremal dependence of the distribution \( G \). Theorem 1 provides the mathematical ground to address this issue, as it gives a joint (limiting) statistical model for the sample extremes of \( M_N \) and \( N \). In particular, the sample maxima of the latter random vector and variable can be used to obtain estimators of \( \lambda \left( F_{M_N} \right) \) and \( \alpha \), respectively. Then, by exploiting the relation among extremal properties of \( F_{M_N} \) and \( F_X \) and the tail behavior of \( N \), and solving (3) for \( \lambda(F_X) \), an estimator of the latter can be obtained. More generally, we focus on the Pickands function as it allows to derive the \( \lambda \) and \( \theta \) coefficients and other related quantities. We define a new semi-parametric procedure for inferring \( A \), which combines together preliminary estimators for \( \alpha \) and \( A_\alpha \) (the Pickands function relative to \( C_{G_\alpha} \)). Specifically, we consider a likelihood- and moments-based estimator for \( \alpha \) and we use three nonparametric estimators for \( A_\alpha \) (existing in the literature).

The rest of the article is organized as follows. In Section 2, we present our main theorem providing the attractor for the joint distribution of \( (M_N, N) \). Different representations for the distribution \( G_\alpha \) are derived in Section 3. In Section 4, by an inversion method, we define an estimator for inferring \( A \). We establish its asymptotic properties (Theorem 2) and show its finite-sample performance by a simulation study. Finally, we discuss directions for future research in Section 5, including a real data example which appears as a promising field of application for our theoretical framework. The proofs are reported in the Appendix, whereas some technical details and additional simulation results are included in the supplementary material.

### 2. MAIN RESULTS

First, recall that the members of the GEV distribution are: the \( \alpha \)-Fréchet (heavy-tailed distribution), Gumbel (light-tailed distribution) and Weibull (short-tailed distribution), in symbols, \( \Phi(x) = \exp(-x^{-\alpha}) \), with \( x > 0 \) and \( \alpha > 0 \), \( \Lambda(x) = \exp(-e^{-x}) \) with \( x \in \mathbb{R} \) and \( \Psi_{\alpha}(x) = \exp(-(-x)^{-\alpha}) \) with \( x < 0 \). In the sequel, for a positive random variable, say \( V \), we denote its Laplace transform by \( L_V(s) = \text{E}(e^{-sV}) \), \( s > 0 \). We also recall that a random variable, \( S \), is positive (asymmetric) \( \alpha \)-stable with index parameter \( 0 < \alpha < 1 \) if its Laplace transform is \( L_S(s) = e^{-s^\alpha} \).

Let \( N \) be a random block size and \( M_N \) be a vector of componentwise maxima obtained with a randomly sized block of iid random vectors \( X_1, X_2, \ldots \) with common distribution \( F_X \), defined in (2). Note that
\[ F_{M_N}(x) = \sum_{n=1}^{\infty} F_X^n(x) \mathbb{P}(N = n) = \sum_{n=1}^{\infty} \exp[-n(-\ln F_X(x))] \mathbb{P}(N = n) = \mathbb{L}(\ln F_X(x)). \]

Assuming that \( F_X \in D(G) \) and \( F_N \in D(H) \), where either \( H \equiv \Phi_\alpha \) or \( H \equiv \Lambda \) since \( N \) is positive integer-valued (Robert & Segers, 2008), we establish new limit results concerning the attractor for the joint distribution of the random vector \( (M_N, N) \) and the tail behavior of the random vector \( M_N \).

### 2.1 Domains of attraction

The first limit result provides the attractor for the joint distribution of the random vector \( (M_N, N) \).

**Theorem 1.** Assume that \( F_X \in D(G) \) and \( F_N \in D(H) \) with \( H \equiv \Phi_\alpha \) or \( H \equiv \Lambda \). Then, there exist norming constants \( c_n > 0, \kappa_n > 0, d_n \in \mathbb{R}^d, \theta_n \in \mathbb{R} \) such that

\[
\lim_{n \to \infty} \mathbb{P}^n \left( \frac{M_N - d_n}{c_n} \leq x, \frac{N - \theta_n}{\kappa_n} \leq y \right) = Q(x, y),
\]

where \( Q \) is defined as follows:

1. **if** \( F_N \in D(\Phi_\alpha) \), **then**

\[
-\ln Q(x, y) = \begin{cases} 
      y^{-\alpha} e^{-\alpha(x + a)} + \sigma(x, \alpha)^{1 - \alpha} y (1 - \alpha, y) \sigma(x, \alpha), & \alpha \in (0, 1) \\
      -\ln G(x) + y^{-\alpha}, & \alpha \geq 1
   \end{cases}
\]

   for all \( x \in \mathbb{R}^d \) and \( y > 0 \), where \( \sigma(x, \alpha) = (-\ln G(x))/\Gamma^{1/\alpha}(1 - \alpha) \) and \( \Gamma, \gamma \) denote the Euler gamma and Lower incomplete gamma functions, respectively.

2. **if** \( F_N \in D(\Lambda) \), **then**

\[
-\ln Q(x, y) = -\ln G(x) + e^{-y}, \quad x \in \mathbb{R}^d, y \in \mathbb{R}.
\]

The margins of \( Q \) are

\[
\begin{cases} 
   G_\alpha(x) := \exp(-(-\ln G(x))^\alpha), & \alpha \in (0, 1) \\
   G(x), & \alpha \geq 1
\end{cases}
\]

and \( \Phi_\alpha(y) \), when \( F_N \in D(\Phi_\alpha), \alpha > 0 \), while they are equal to \( G(x), x \in \mathbb{R}^d \), and \( \Lambda(y), y \in \mathbb{R} \), when \( F_N \in D(\Lambda) \). Specifically, the distribution \( G_\alpha, \alpha \in (0, 1) \), is max-stable with margins \( G_{\alpha,1}, \ldots, G_{\alpha,d} \), which are members of the GEV class, and extreme-value copula

\[
C_{G_\alpha}(u) = \exp \left( -L^\alpha \left( (-\ln u_1)^{1/\alpha}, \ldots, (-\ln u_d)^{1/\alpha} \right) \right), \quad u \in (0, 1]^d, \quad \alpha \in (0, 1),
\]

where \( L \) is the stable-tail dependence function of the max-stable distribution \( G \).

A probabilistic interpretation of the problem addressed in Theorem 1 is as follows. Let \( N_1, N_2, \ldots \) be iid copies of \( N \) and set \( N_n^+ = N_1 + \ldots + N_n, N_n^\gamma = \max(N_1, \ldots, N_n) \), then we have
Loosely speaking, we are concerned with the asymptotic distribution of the random vector \((MN_\alpha^+, N_\nu^+)^\top\) appropriately normalized (a.n.). When \(F_N \in D(\Lambda)\) (light-tailed) or \(F_N \in D(\Phi_\alpha)\) (heavy-tailed), with \(\alpha > 1\), then \(\mu = E(N) < \infty\). Since \(n^{-1}N_\alpha^+\) converges to \(\mu\), then the asymptotic distributions of \(MN_\alpha^+\) and \(M_{[n\mu]}\) a.n. are approximately the same. When \(F_N \in D(\Phi_\alpha)\) (heavy-tailed), with \(\alpha \in (0, 1)\), then \(\kappa_n^{-1}N_\nu^+\) converges in distribution to a positive stable random variable \(S\), where \(\kappa_n := \kappa_n\Gamma^{1/\alpha}(1 - \alpha)\). Consequently, the asymptotic distributions of \(MN_\alpha^+\) and \(M_{[\kappa_n S]}\) a.n. coincide and are equal to \(G_\alpha\) in the first line of (6), see Corollary 1, which is a location-scale mixture of the limiting max-stable distribution \(G\), obtained with a deterministic block size. A similar result can be established in the case of \(\alpha = 1\), which is not explicitly discussed here for the sake of brevity.

**Corollary 1.** Let \(F_X \in D(G)\), \(F_N \in D(\Phi_\alpha)\), \(\alpha \in (0, 1)\) and \(c_n, d_n\) and \(\kappa_n\) as in Theorem 1. Then

\[
\lim_{n \to \infty} \mathbb{P}(M_{[\kappa_n S]} \leq c_n x + d_n) = G_\alpha(x).
\]

As for asymptotic dependence, we point out the following. When \(F_N \in D(\Phi_\alpha)\), \(\alpha \geq 2\), or \(F_N \in D(\Lambda)\), \(N_\alpha^+\) and \(N_\nu^+\) a.n. converge to nondegenerate independent random variables (e.g., Anderson & Turkman, 1995, pp. 1-2). Intuitively, this explains the asymptotic independence between \(MN_\alpha^+\) and \(N_\nu^+\) (and, in turn, the extremal independence between \(M_N\) and \(N\)). In the special case of \(F_N \in D(\Phi_\alpha)\), \(\alpha \in (0, 2)\), \(N_\alpha^+\) and \(N_\nu^+\) a.n. are asymptotically dependent (e.g., Anderson & Turkman, 1995, pp. 1-2). Thus, when \(\alpha \in [1, 2]\) we explain the asymptotic independence between \(MN_\alpha^+\) and \(N_\nu^+\) a.n. in a different way. From the derivations in Lemmas 2-3 and Section A.1.3 one sees that the dependence structure of the limiting distribution \(Q(x, y)\) is determined by the limit of the conditional exceedance probability \(1 - \mathbb{P}(M_N \leq c_n x + d_n | N > \kappa_n y + \rho_n)\), as \(n \to \infty\). Note that, as \(n\) grows, the norming sequences \((c_n, d_n)\) and \((\kappa_n, \rho_n)\) affect this threshold exceedance probability in two opposing ways: the first ones increase the threshold \(c_n x + d_n\); the second ones force \(N\), and thus \(M_N\), to be stochastically larger and larger. Clearly, a nondegenerate limit—that is, in (0, 1)— is obtained only if the two effects offset each other. On one hand

\[
F_{M_{[\kappa_n S]}}(c_n x + d_n) = \mathbb{E}_N(-\ln F_X(c_n x + d_n))
\]

and so \(c_n\) and \(d_n\) are affected by the behavior of \(1 - \mathbb{E}_N(1/y)\) as \(y \to \infty\). On the other hand \(F_{N_\alpha^+}(\kappa_n y + \rho_n) = F_{N_\nu^+}(\kappa_n y + \rho_n)\) and so \(\kappa_n\) and \(\rho_n\) depend on the tail properties of \(F_N\). When \(\alpha \in (0, 1)\), \([1 - F_N(y)]/[1 - \mathbb{E}_N(1/y)]\) converges to a nondegenerate limit as \(y \to \infty\), while it converges to zero when \(1 \leq \alpha < 2\). Accordingly, in the first case, the limit of the conditional probability of exceedances is positive, whereas it is zero in the second case, since \(c_n, d_n\) are “too heavy” relative to \(\kappa_n, \rho_n\). As a result, the marginal distributions of \(MN_\alpha^+\) and \(N_\nu^+\) a.n. converge to nondegenerate limits but their dependence is wiped out as \(n \to \infty\).

The dependence structure of the (d+1)-dimensional distribution \(Q\) defined in (4) and (5) can be synthesized by means of its extremal coefficient.

**Corollary 2.** When the expression of \(Q\) is given in (4), then the extremal coefficient is
\[ \theta(Q) \equiv \theta(G, \alpha) := \begin{cases} 1 - E^{\gamma_1/(1-\alpha)}_{\Gamma_1}(\theta(G)) + (\theta(G))^{\alpha} G^{1-\gamma_1/(1-\alpha)} \theta(G), & \alpha \in (0, 1) \\
\theta(G) + 1, & \alpha \geq 1. \end{cases} \]

where \( E_{\gamma}(\cdot) \) and \( G_{\alpha,b}(\cdot) \) denote the exponential and gamma cumulative distribution functions, with shape and scale parameters \( a \) and \( b \). In particular, for \( \alpha \in (0, 1) \) we have

\[ 1 \leq \theta(G, \alpha) \leq 1 + \theta(G), \quad \lim_{\alpha \to 1} \theta(G, \alpha) = 1 + \theta(G). \]

When the expression of \( Q \) is given in (5), \( \theta(Q) = \theta(G) + 1 \).

From the first result in the left-hand side of (8) we deduce that the extremal coefficient of \( Q \) is larger than or equal to 1 (as expected) and bounded from above by \( 1 + \theta(G) \), representing the case where \( M_N \) and \( N \) have no tail dependence. Moreover, the second result in the right-hand side of (8) highlights a continuous transition of the extremal dependence level from asymptotic dependence between \( M_N \) and \( N \) (i.e., \( \alpha \in (0, 1) \)) to asymptotic independence (i.e., \( \alpha \geq 1 \)), when \( F_N \) belongs to the \( \alpha \)-Fréchet domain.

The extremal coefficient for the distribution \( G_{\alpha} \) with \( \alpha \in (0, 1), \theta(G_{\alpha}) \), is given in (13). In the bivariate case, by (13) and the relationship between the extremal coefficient and the coefficient of upper tail dependence, that is, \( \theta(G) = 2 - \lambda(F_X) \), we obtain

\[ \theta(G_{\alpha}) = [2 - \lambda(F_{M_N})]^\alpha. \]

Since it is also true that \( \theta(G_{\alpha}) = 2 - \lambda(F_{M_N}) \), we obtain the result in (3).

### 2.2 Tail behaviors

The second limit result establishes the tail behaviors of the random vector \( M_N \), which concerns the probability that at least one component of the random vector \( M_N \) exceeds an increasingly large value.

In the sequel, for a given max-stable distribution \( G \) we denote by \( \tilde{G} \) a distribution with the same copula as \( G \) and common unit-Fréchet margins.

**Proposition 1.** Assume that \( F_X \in D(G) \) and \( F_N \in D(H) \) with \( H \equiv \Phi_\alpha \) or \( H \equiv \Lambda \). For \( y \in (0, \infty)^d \) and \( n \in \mathbb{N}_+ \),

1. if \( F_N \in D(\Phi_\alpha) \) with \( 0 < \alpha \leq 1 \)

\[ 1 - \mathbb{P}(M_N \leq U_X(ny)) \sim \begin{cases} \Gamma(1 - \alpha) \mathbb{P} \left( N > \frac{n}{-\ln \tilde{G}(y)} \right), & \alpha \in (0, 1) \\
\{1 - \mathbb{L}_N(1/n)\}\{-\ln \tilde{G}(y)\}, & \alpha = 1 \end{cases}, \quad n \to \infty \]

\[ 1 - \mathbb{P}(M_N \leq U_{M_N}(ny)) \sim \{-\ln \tilde{G}(y^{1/\alpha})\}^\alpha n^{-1}, \quad n \to \infty \]

2. if \( F_N \in D(\Phi_\alpha) \) with \( \alpha > 1 \) or \( F_N \in D(\Lambda) \), then

\[ 1 - \mathbb{P}(M_N \leq U_X(ny)) \sim n^{-1} \mathbb{E}(N)\{-\ln \tilde{G}(y)\}, \quad n \to \infty \]

\[ 1 - \mathbb{P}(M_N \leq U_{M_N}(ny)) \sim n^{-1} \{-\ln \tilde{G}(y)\}, \quad n \to \infty \]
where $U_X(ny), U_{M_n}(ny) \to \infty$ as $n \to \infty$, with $U_X$ and $U_{M_n}$ defined in Appendix A.4.

Set $p_j = (ny_j)^{-1}, j = 1, \ldots, d$, and recall that $L$ denotes the stable-tail dependence function of $G$. As $n \to \infty$, by Proposition 1, the probability that at least one component $M_N^{(i)}$ of $M_N$ exceeds the $1 - p_j$ quantile $U_{M_N^{(i)}}(ny_j)$ of its own distribution is approximately $L((ny)^{-1}) = -\ln \tilde{G}(ny)$, when $E(N) < \infty$, while it is approximately $L^a((ny)^{-1/\alpha})$, when $E(N) = \infty$.

3 | REPRESENTATIONS OF THE MODEL $G_{\alpha}$

In this section we show that there are different constructions that yield a max-stable distribution with the same copula $C_{G_{\alpha}}$ in (7) of the distribution $G_{\alpha}$. Furthermore, we derive the Pickands function corresponding to $C_{G_{\alpha}}$.

Let $S$ be a positive $\alpha$-stable random variable with index parameter $0 < \alpha < 1$. Let $Z$ be a random vector with max-stable distribution $\hat{G}$. Assume $S$ and $Z$ to be independent. Define $R = (SZ_1, \ldots, SZ_d)$, then for every $y > 0$,

$$\mathbb{P}(R \leq y) = E(\hat{G}^S(y)) = L_S (-\ln \tilde{G}(y)) = \exp (-(-\ln \tilde{G}(y))^\alpha) =: \hat{G}_\alpha(y).$$

(9)

The distribution $\hat{G}_\alpha$ is a special case of $G_{\alpha}$, that is, max-stable with extreme-value copula $C_{G_{\alpha}}$ and common $\alpha$-Fréchet margins. In the particular case where the components of $Z$ are independent, the copula of $\hat{G}_\alpha$ is

$$C_{\hat{G}_\alpha}(u) = \exp \left( -((-\ln u_1)^{1/\alpha} + \ldots + (-\ln u_d)^{1/\alpha})^\alpha \right), \quad u \in (0, 1]^d, \quad \alpha \in (0, 1),$$

(10)

which is the well-known symmetric logistic copula (e.g., Joe, 2015, p. 172). Therefore, the elements of $R$ are dependent for any $\alpha \in (0, 1)$, and the dependence level increases (decreases) as $\alpha$ approaches 0 (as $\alpha$ approaches 1). Random scaling constructions similar to this one have been already discussed by Fougères, Nolan, and Rootzén (2009) and Fougères, Mercadier, and Nolan (2013).

The de Haan representation of max-stable processes (de Haan, 1984) provides a Poisson point process construction of a random vector with any max-stable distribution $\hat{G}$. A question that arises here is: What is the spectral representation of a random vector $R$ defined by the random scaling construction? The next result establishes that the findings presented in Robert (2013) indeed provide the spectral representation of $R$.

**Proposition 2.** Let $Z_1, Z_2, \ldots$ be iid copies of $Z$, with distribution $\hat{G}$, independent of $P_1, P_2, \ldots$ that are points of a Poisson process on $(0, \infty)$ with intensity measure $ar^{-(\alpha+1)}dr$, $\alpha \in (0, 1)$. Define

$$R = \frac{1}{\Gamma^{1/\alpha}(1 - \alpha)} \left( \max_{i \geq 1} P_iZ_{i1} \ldots \max_{i \geq 1} P_iZ_{id} \right).$$

Then, the distribution of $R$ is $\hat{G}_\alpha$.

Here we provide an alternative, more general proof than that given in Robert (2013). Specifically, ours does not rely on an unnecessary smoothness assumption. Next, we provide a characterization of $G_{\alpha}$, as the attractor distribution for a general random scaling and centering construction.
Proposition 3. Let \( X_1, \ldots, X_n \) be iid copies of the random vector \( X \), with distribution \( F_X \). Assume \( F_X \in D(G) \). Let \( S \) be a positive \( \alpha \)-stable random variable, \( \alpha \in (0, 1) \). Assume \( S \) is independent of \( X \). Define

\[
  w_n := \frac{a_{\lfloor nS \rfloor}}{a_n}, \quad v_n := b_n - \frac{b_{\lfloor nS \rfloor}}{a_{\lfloor nS \rfloor}} \frac{a_n}{a_{\lfloor nS \rfloor}},
\]

where \( a_n \) and \( b_n \) are the usual norming constants of \( F_X \). Then,

\[
  a_n^{-1} (\max(w_n(X_1 - v_n), \ldots, w_n(X_n - v_n)) - b_n) \overset{d}{\rightarrow} G_\alpha, \quad n \to \infty.
\]

A simple implication of Proposition 3 is the following. Transforming \( X \) into \( Y \), a random vector with common unit-Frêchet marginal distributions, and setting \( R = SY \), implies that \( F_R \in D(\tilde{G}_\alpha) \). To see this, note that \( F_Y \in D(\tilde{G}) \) with norming sequences \( a_n = 1/n \), \( b_n = 0 \), and that, as \( n \to \infty \), \( SY \) and \( w_n Y = (\lfloor nS \rfloor/n)Y \) have approximately the same distribution. As a result, the attractors of \( F_R \) and \( F_{M_\alpha} \), when \( F_N \in D(\Phi_\alpha) \), \( \alpha \in (0, 1) \), share the same extreme-value copula, \( G_\alpha \). We finally derive the explicit form of the Pickands function corresponding to the latter.

Proposition 4. The Pickands function corresponding to the extreme-value copula \( C_{G_\alpha} \) in (7) is

\[
  A_\alpha(t) = ||t||_1^{1/\alpha} A \left( \frac{t/||t||_1^{1/\alpha}}{||t||_1^{1/\alpha}} \right), \quad t \in S_d, \alpha \in (0, 1),
\]

where \( A \) is the Pickands dependence function corresponding to \( C_G \) and

\[
  ||t||_1^{1/\alpha} = \left( \sum_{i=1}^{d} t_i^{1/\alpha} \right)^{\alpha}, \quad t \in S_d, \alpha \in (0, 1),
\]

is the Pickands function corresponding to the Logistic copula.

As a direct consequence of Proposition 4, the following facts ensue. The smaller the parameter \( \alpha \), the more \( A_\alpha \) represents a stronger dependence level than \( A \). Since we have that \( ||(1/d, \ldots, 1/d)||_1 = d^{\alpha - 1} \), then, by the definition of the extremal coefficient in Section 1, we obtain

\[
  \theta(G_\alpha) = (\theta(G))^\alpha.
\]

By solving for \( A \) in Equation (11), we obtain the inverse relation between \( A_\alpha \) and \( A \), that is,

\[
  A^*(t) := A \left( \frac{t/||t||_1^{1/\alpha}}{||t||_1^{1/\alpha}} \right) = (A_\alpha(t)/||t||_1^{1/\alpha})^{1/\alpha}, \quad t \in S_d
\]

and \( A(t) = A^*(t^\alpha/||t^\alpha||_1) \), providing the expression for the Pickands function in (1).

4 | INFERRING THE PICKANDS FUNCTION

In this section we introduce a new semiparametric procedure to estimate the Pickands function \( A \) in (1). Several nonparametric estimators are already available for \( A \) when a data sample from the limiting distribution for unaggregated data, \( G \), is observable, see Klüppelberg and May (2006), Zhang, Wells, and Peng (2008), Berghaus, Bücher, and Dette (2013), Cormier, Genest, and
Neslehova (2014), Marcon, Padoan, Naveau, Muliere, and Segers (2017), among others. Unlike the above references, we assume that only replicates of \((M_N, N)\) are observable, from which sample extremes (maxima) are extracted. Then, we construct an estimator for \(A\), exploiting an inversion method via (14). This is a substantial novelty in the extreme-value literature. It is common practice to assume that sample maxima are exactly coming from the limiting model \(Q\) in (4), with \(\alpha \in (0, 1)\), and provide an asymptotic validation of the proposed inferential procedure in such a setting. Via an extensive simulation study in Section 4.2, we show that, in practice, our method provides a good performance with data that are only approximately coming from \(Q\). Extending the asymptotic statistical theory to the latter case goes beyond the scope of the present already quite technical work. Observe that, since \(t \mapsto (t/\|t\|_\alpha)^{1/\alpha}\) is a bijective map, estimating \(A^*\) is equivalent to estimating \(A\), thus, for simplicity, we hereafter focus on the former function.

### 4.1 A semiparametric composite-estimator

Let \((\eta_1, \xi_1), (\eta_2, \xi_2), \ldots\), be iid random vectors with joint distribution in (4) with \(\alpha \in (0, 1)\). Assume that a sample of \(n\) observations from such a sequence is available. An estimate of \(A^*\) is obtained by combining the results of a two-step procedure: we estimate \(\alpha\) and \(A_\alpha\), we plug the estimates in (14). Precisely, \(\xi_1, \ldots, \xi_n\) follows a \(\alpha\)-Fréchet distribution. For estimating \(\alpha\) we consider two well-known estimators: the generalized probability weighted moment (GPWM) (Guillou, Naveau, & Schorgen, 2014) and the maximum likelihood (ML). In the first case the estimator is

\[
\hat{\alpha}_n^{\text{GPWM}} := \left( k - 2 \frac{\hat{\mu}_{1,k}}{\hat{\mu}_{1,k-1}} \right)^{-1},
\]

for \(k \in \mathbb{N}_+\), where

\[
\hat{\mu}_{a,b} = \int_0^1 H_n^\alpha(v)v^a(-\ln v)^b \, dv, \quad a, b \in \mathbb{N},
\]

and

\[
H_n(y) = \frac{1}{n} \sum_{i=1}^n \mathbb{I}(\xi_i \leq y), \quad y > 0.
\]

In the second case the estimator is

\[
\hat{\alpha}_n^{\text{ML}} := \arg \max_{\alpha \in (0, \infty)} \sum_{i=1}^n \ln \Phi_\alpha(\xi_i),
\]

where \(\Phi_\alpha(x) = \partial/\partial x \Phi_\alpha(x), x > 0\).

The sequence \(\eta_1, \ldots, \eta_n\) follows the distribution \(G_\alpha\). For estimating \(A_\alpha\) we consider three well-known estimators: Pickands (P) (Pickands, 1981), Capéraà-Fougère-Genest (CFG) (Capéraà, Fougères, & Genest, 1997), and Madogram (MD) (Marcon et al., 2017). In the first case the estimator is
\[ \hat{A}_{a,n}^p(t) := \left( \frac{1}{n} \sum_{i=1}^{n} \hat{\delta}_i(t) \right)^{-1}, \]

\[ \hat{\delta}_i(t) = \min_{1 \leq j \leq d} \left\{ -\frac{1}{t_j} \ln \left( \frac{n}{n+1} G_{n,j}(\eta_{i,j}) \right) \right\} \quad (18) \]

where for every \( x \in \mathbb{R} \) and \( j \in \{1, \ldots, d\} \)

\[ G_{n,j}(x) = \frac{1}{n} \sum_{i=1}^{n} \mathbb{1}(\eta_{i,j} \leq x). \quad (19) \]

In the second case the estimator is

\[ \hat{A}_{a,n}^{CFG}(t) := \exp \left( -\frac{1}{n} \sum_{i=1}^{n} \ln \hat{\delta}_i(t) - \zeta \right), \quad (20) \]

where \( \zeta \) is the Euler’s constant. Finally, in the third case the estimator is

\[ \hat{A}_{a,n}^{MD}(t) := \frac{\hat{\nu}_n(t) + c(t)}{1 - \hat{\nu}_n(t) - c(t)}, \quad (21) \]

\[ \hat{\nu}_n(t) = \frac{1}{n} \sum_{i=1}^{n} \left( \max_{j=1, \ldots, d} G_{n,j}^{1/t_j}(\eta_{i,j}) - \frac{1}{d} \sum_{j=1}^{d} G_{n,j}^{1/t_j}(\eta_{i,j}) \right), \quad (22) \]

where \( u^{1/0} = 0 \) for \( 0 < u < 1 \) by convention and \( c(t) = \sum_{j=1}^{d} t_j / (1 + t_j) \).

For brevity we denote the estimators of \( \alpha \) and \( A_a \) by \( \hat{\alpha}_a^* \) and \( \hat{A}_{a,n}^* \), respectively, where the symbols “•” and “○” are representative of the labels “GPWM,” “ML” and “P,” “CFG,” “MD”, respectively. Then, plugging the estimators into Equation (14) we obtain the following composite-estimator for \( A_a^* \),

\[ \hat{A}_{a,n}^{\circ \bullet}(t) := \left( \hat{A}_{a,n}(t)/\|t\|_{1/\hat{\alpha}_n^*} \right)^{1/\hat{\alpha}_n^*}, \quad t \in S_d. \quad (23) \]

Next, we establish the asymptotic theory of the composite-estimator in (23) defined by all the combinations of the GPWM and ML estimators for \( \alpha \) with the P, CFG and MD estimators for \( A_a \). Our results rely on the following assumptions.

**Condition 1.** For \( j \in \{1, \ldots, d\} \), let \( U_j = \{ u \in [0,1]^d : 0 < u_j < 1 \} \). Assume that:

(i) for \( j \in \{1, \ldots, d\} \), the first-order partial derivative \( \hat{C}_{G_a,j}(u) := \partial / \partial u_j C_{G_a}(u) \) exists and is continuous in \( U_j \);

(ii) for \( i,j \in \{1, \ldots, d\} \), the second-order partial derivative \( \hat{C}_{G_a,i,j}(u) := \partial / \partial u_i \hat{C}_{G_a,j}(u) \) exists and is continuous in \( U_i \cap U_j \) and

\[ \sup_{u \in U_i \cap U_j} \max(u_i, u_j) |\hat{C}_{G_a,i,j}(u)| < \infty. \]
Theorem 2. For the estimators \( \hat{A}^{\text{MD}*}_{n} \), assume that Condition 1(i) holds true; assume that Condition 1(ii) is also satisfied for the estimators \( \hat{A}^{\text{P}*}_{n} \), \( \hat{A}^{\text{CFG}*}_{n} \). Finally, assume that the choice of \( k \in \mathbb{N}^+ \) in the GPWM-based estimator \( \hat{A}^{*,\text{GPWM}}_{n} \) satisfies \( \alpha > 1/(k - 1) \). Then, as \( n \to \infty \)

\[
\sqrt{n} \left\{ \hat{A}^{*,\text{GPWM}}_{n}(t) - A^*(t) \right\} \stackrel{d}{\to} \left\{ (\phi_{\circ*,\text{CPQ}}(C_Q))(t) \right\} \text{ for } t \in S_d,
\]

in \( \ell^\infty(S_d) \), for an operator \( \phi_{\circ*,\text{CPQ}} \) into \( \ell^\infty(S_d) \) and a zero-mean Gaussian process \( C_Q \), whose covariance function is

\[
\text{Cov}(C_Q(u), C_Q(v)) = C_Q(\min(u, v)) - C_Q(u)C_Q(v), \quad u, v \in [0, 1]^{d+1},
\]

where \( C_Q \) is the copula

\[
C_Q(u, v) = Q(G_{u,1}^{-}(u_1), \ldots, G_{u,d}^{-}(u_d), \Phi_{\circ,\text{CPQ}}^{-}(v))
\]

and the minimum is taken componentwise. Moreover,

\[
\| \hat{A}^{*,\text{GPWM}}_{n} - A^* \|_{\infty} \overset{p}{\to} 0, \quad \| \hat{A}^{*,\text{MD,GPWM}}_{n} - A^* \|_{\infty} \overset{a.s.}{\to} 0, \quad n \to \infty.
\]

Remark 1. For brevity, the explicit definitions of \( \phi_{\circ*,\text{CPQ}} \)'s are postponed to Definition 1(iv), (v), (vii), (viii), and functional limit results are provided all at once. Although each estimator’s combination has its own peculiarities, these can be framed within a fairly general theory, which might be of interest per se. Due to the high degree of technicality, we present such a theory in the appendix, herein focusing on ready-to-use estimators’ examples.

Remark 2. In Gudendorf and Segers (2012) and Marcon et al. (2017) modified versions of the estimators P, CFG, and MD for \( A_\alpha \) are proposed to guarantee that \( \hat{A}^{\circ}_{n}(e_j) = 1 \) for all \( n = 1, 2 \ldots \) and \( j = 1, \ldots, d \) where \( e_j = (0, \ldots, 0, 1, 0, \ldots, 0) \). The results in Theorem 2 are also valid when such adjusted estimators are considered in place of (18), (20) and (21), respectively, due to asymptotic arguments developed in the aforementioned works.

Remark 3. By the identity in (7), Proposition 1 in Gudendorf and Segers (2012) guarantees, if the stable-tail dependence function \( L \) satisfies Assumption 2 therein, that \( C_{G_\alpha} \) satisfies Condition 1.

4.2 Simulation study

We show the finite sample performance of the composite-estimator \( \hat{A}^{*,\text{CPQ}}_{n} \) in (23) through a simulation study. Hereafter we consider the adjusted versions of the P, CFG and MD estimators, mentioned in Remark 2. Since it is not straightforward to simulate from the limit distribution \( Q \), we study the performance of the composite-estimator \( \hat{A}^{*,\text{CPQ}}_{n} \) when it is used with data that are only approximately coming from the limiting distribution \( Q \). Nevertheless, this is a more realistic scenario.

Specifically, we set \( N = \lceil N' \rceil \), where \( N' \) follows a standard Pareto distribution with shape parameter \( \alpha \in (0, 1) \). We simulate \( N \) observations of a two-dimensional random vector \( X \) with a standard bivariate Student-\( t \) distribution with a fixed value of the correlation \( \rho \) and the degrees
of freedom $v$. We recall that a Student-$t$ distribution is in the domain of attraction of a multivariate extreme-value distribution with an extreme-value copula that is the so-called Extremal-$t$ (e.g., Joe, 2015, p. 189). In the bivariate case, the extremal coefficient of the Extremal-$t$ copula is $\theta = 2T_{v+1}^{(v+1)(1-\rho)/(1+\rho))^{1/2}}$, where $T_{v+1}$ is a univariate standard Student-$t$ distribution with $v + 1$ degrees of freedom. Next, with the simulated data we compute the observed value of the componentwise maxima $M_N$ in (2). We repeat these simulation steps $n' = 500$ times generating $n'$ independent observations from the pair $(N, M_N)$ with which we compute an observation from the random variable $\xi = \max(N_1, \ldots, N_n)$ and vector $\eta = \max(M_{N_1}, \ldots, M_{N_N})$, where the later maximum is meant componentwise. We repeat these simulation steps $n$ times, generating a data sample approximately drawn from the distribution $Q$, whose expression is given in the first line of (4) and where the expression of $G$ can be deduced from Joe (2015, p. 189).

Then, we estimate $\alpha$ using the generated sequence of observations $\xi_1, \ldots, \xi_n$ with the GPWM estimator $\hat{\alpha}_n^{\text{GPWM}}$ in Equation (15), with $k = 5$, and the ML estimator $\hat{\alpha}_n^{\text{ML}}$ in (17). Afterwards, we estimate the Pickands dependence function $A_\theta$ using the observations generated sequence of observations $\eta_1, \ldots, \eta_n$ with the P estimator $\hat{A}_n^p$ in (18), CFG estimator $\hat{A}_n^{\text{CFG}}$ in (20), and MD estimator $\hat{A}_n^{\text{MD}}$ in (21). Finally, we estimate $A^*$ using the composite-estimator $\hat{A}_n^{\text{\hat{\alpha}_n^*}}$ in Equation (23).

We repeat the simulation and estimation steps for different values of the model parameters $\alpha$, $\rho$ and $\nu$ and different sample sizes. Precisely, we consider $\alpha = 0.5, 0.633, 0.767, 0.9$ and, for the Student-$t$ distribution, we consider the degrees of freedom $\nu = 1$ and 15 equally spaced values of the correlation $\rho$ in $[-0.99, 0.99]$. With these parameters’ values, the extremal coefficient $\theta$ (related to the Extremal-$t$ copula) takes values in $[1, 2]$, where the lower and upper bounds represent the cases of complete dependence and independence. We also consider the sample sizes $n = 50, 100$. We repeat this experiment (the simulation and estimation considering different values of the parameters and the sample sizes) 1,000 times and we compute a Monte Carlo approximation of the mean integrated squared error (MISE), that is,

$$\text{MISE}(\hat{A}_n^{\alpha^\circ \beta^\circ}, A^*) = \mathbb{E} \left( \int_{S_\delta} \left[ \hat{A}_n^{\alpha^\circ \beta^\circ}(t) - A^*(t) \right]^2 \, dt \right)$$

$$= \int_{S_\delta} \mathbb{E} \left[ \left( \hat{A}_n^{\alpha^\circ \beta^\circ}(t) - A^*(t) \right)^2 \right] \, dt + \int_{S_\delta} \mathbb{E} \left( \hat{A}_n^{\alpha^\circ \beta^\circ}(t) - \mathbb{E} \left( \hat{A}_n^{\alpha^\circ \beta^\circ}(t) \right) \right)^2 \, dt,$$

where the first and second terms in the second line are known as integrated squared bias (ISB) and Integrated Variance (IV) (Gentle, 2009, ch. 6.3).

Figure 1 displays the results obtained with the GPWM-based estimators for the sample size $n = 50$. The MISE, ISB, and IV ($\times 1,000$) of the GPWM-based estimators are reported from the first to the third row. The solid black, dashed green and dotted red lines report the results obtained estimating $A_\theta$ with P, CFG, and MD estimators, respectively. The results for the different values of $\alpha$ are reported along the columns.

For each fixed value of $\alpha$ we see that IV is close to zero at the strongest dependence level ($\theta = 1$), then it increases with the decrease of the dependence level ($\theta$ increases approaching two). On the contrary, ISB takes the largest value at $\theta = 1$ and then it decreases with the decrease of the dependence level, for the cases $\alpha = 0.5, 0.633$. Overall, for the case $\alpha = 0.5$, MISE takes the largest value at $\theta = 1$ and then it decreases with the decrease of the dependence level. For the case $\alpha = 0.633$, MISE does not change much over the whole range of dependence levels, since ISB and IV compensate each other. While, for the cases $\alpha = 0.767, 0.9$, IV grows much more
FIGURE 1 MISE, ISB, and IV for 1,000 samples of size 50 from an approximated distribution $Q$, obtained on the basis of the standard Pareto distribution for $N$ and the bivariate Student-$t$ distribution for $X$, for different values of the parameters $\alpha$ and $\rho, \nu$. The parameter $\alpha$ is estimated with the GPWM estimator in (15). The parameter $\theta$ is the extremal coefficient related to the corresponding extreme-value copula extremal-$t$

than ISB decreases, implying that MISE increases with the decrease of the dependence level. The smallest values of ISB and IV are obtained with the CFG-based and P-based estimator, respectively. Overall, on the basis of the MISE, the best performance is obtained with the CFG-based estimator, although there is little difference with the P-based estimator. As shown in the supplementary material, there is not much difference in the performance of the P-, CFG-, and MD-based estimators already for the sample size $n = 100$.

Although in this experiment we consider synthetic data that only approximately come from the distribution $Q$, the results summarized by ISB, IV, and MISE highlight the robustness of our method to model misspecification (for only approximately max-stable data). In particular, our composite-estimator displays a moderate distortion, despite the fact that the uniform consistency guarantee of Theorem 2 does not directly extend to the present setting. Similar conclusions are obtained with sample size $n = 100$ (available in the supplementary material).

A comparison between the estimation results obtained with the GPWM- and ML-based estimators is reported in Figure 2. Precisely, from the first to the third row, the ratio between the MISE, ISB, and IV computed estimating the function $A^*$ by the GPWM- and ML-based estimators are displayed. For the case $\alpha = 0.5$, on the basis of the ISB, the GPWM- and ML-based estimators perform very similarly, when $1 \leq \theta \leq 1.5$, that is, for strong up to moderate dependence levels. Instead, when $1.5 \leq \theta \leq 2$, that is, for moderate up to weak dependence levels, the ML-based estimators outperform the GPWM-based estimators. However, for very weak dependence levels
Figure 2: Ratio between MISE, ISB, and IV computed estimating the function $A^*$ by the estimator $\hat{A}_{\pi}^{\text{GPWM}}$ and $\hat{A}_{\pi}^{\text{ML}}$ in formula (23). The same setting as Figure 1 is considered.

...(theta close to 2) the GPWM-CFG estimator outperforms the ML-CFG estimator. On the basis of the IV, the ML-based estimators outperform the GPWM-based estimators for all cases. Concerning the configurations with $\alpha > 0.5$, the ML-based estimators considerably outperform GPWM-based estimators in terms of MISE. These conclusions are valid for all three P-, CFG-, and MD-based estimators. Specifically, IV is smaller for the ML-based estimators and their better performances are obtained for weaker extremal dependence structures (when theta approaches 2). The ML-based estimators are much less biased than the GPWM-based estimators and the difference is much more pronounced for increasing values of $\alpha$ and weaker extremal dependence structures (when theta approaches 2), although for the P- and CFG-based estimators such a difference diminishes when theta is close to 2.

The study was performed using the R R Core Team, 2014 packages Copula Kojadinovic & Yan, 2010 and evd Stephenson, 2002.

5 | DISCUSSION

Here we shortly discuss directions for future research, from both an applied and a theoretical viewpoint. First, we illustrate a potential real data problem concerning Internet traffic. In this domain, our theoretical framework can be applied to improve and extend the existing methods...
for the study of Internet traffic data and to perform an extreme-value statistical analysis. Next, we provide some concluding remarks, including probabilistic and methodological extensions of this work.

**Massive Internet traffic data.** The analysis of Internet traffic data is crucial for improving the performance of large networks. Inferring internet congestion at different levels continues to receive increasing attention as new challenges are posed, for example, the booming demand for high-bandwidth contents (e.g., video streaming). Large scale collection and analysis of scientific data on Internet traffic is conducted by renowned research institutions, such as the Center for Applied Internet Data Analysis (CAIDA, http://www.caida.org/home/). CAIDA collects anonymized traffic traces from several monitors connected to commercial Internet backbones and large Internet service providers. A one-minute trace lists a huge amount of IP (Internet Protocol) packets, resulting in several gigabytes of compressed files. A first reduction of the data size is obtained resorting to flow records, that is, measurements pertaining to coherent strings of packets, for example, those stemming from the same traffic source, such as a single user (see van de Meent, 2006, ch. 3). Though, statistical analysis at this level is still computationally too burdensome. Nonetheless, knowledge of the joint extremal behavior of flow size (amount of data transmitted as part of a flow), say $X_1$, and flow duration (time difference between the first and last packet of a flow), say $X_2$, would improve the existing techniques for large network optimization. For instance, it would be of interest to forecast the throughput rate $X_1/X_2$, corresponding to extremely large flow sizes (Markovich, 2007, ch. 1.3.2, 1.3.5).

To infer the extremal dependence between $X_1$ and $X_2$, it is first necessary to reduce the dimension of flow data by appropriately aggregating them. This is obtained by computing maxima of $X_1$ and $X_2$ over the random number of flows, $N$, occurring in a suitable time interval. The survival function can be reasonably expected to display a power-law behavior, that is, $P(N > n) = \mathcal{L}(n) n^{-\alpha}$, with $\alpha$ smaller than one. This is consistent both with earlier theoretical fundings, for example, Bonald, Proutière, Régnié, and Roberts (2001), and with the complexity of modern applications and backbone infrastructures. Differently from the classical probabilistic description of Internet aggregate traffic (e.g., Taqqu, Willinger, & Sherman, 1997; Willinger, Taqqu, Sherman, & Wilson, 1997), in which the number of active sources (flows) is deterministically sent to infinity, we account for randomness in the number of flows. Still, stochastic modeling via Pareto-type tails with $\alpha < 1$ (hence $E(N) = +\infty$) appears a coherent refinement of the standard approach.

In this setting, Theorem 1 provides the mathematical basis for modeling the extremal dependence of aggregated-flow size and duration by means of the componentwise maxima approach. Furthermore, by the statistical inversion method in Section 4.1 we can infer the extremal dependence between the extreme single-flow size and duration. Due to the extremely complex nature of flow data, we defer the actual data analysis to a future specialized work.

**Concluding remarks.** The probabilistic and statistical modeling of aggregated data is an interesting and important topic. The total and maximum amounts derived on a random number of observations described by the random vectors $S_N$ and $M_N$ in (2) are two simple examples of aggregated data. There are not many results available on the extremal behavior of $S_N$ and $M_N$, apart from those on $S_N$ in the univariate case. This contribution makes a first step by establishing the multivariate extreme-value theory for $M_N$. Investigating the joint upper-tail behavior of $M_N$, under the hidden regular variation framework (e.g., Maulik & Resnick, 2004; Mitra & Resnick, 2011; Resnick, 2002) would represent a first extension of this work. Of particular interest would be investigating whether $F_N \in D(\Phi_\alpha)$ with $\alpha \geq 1$ implies a stronger residual dependence
than the case $F_N \in D(\Lambda)$, as no difference emerges in terms of classical extremal dependence (see Theorem 1). In this article we consider for simplicity the same random number $N$ of independent copies for each component of $X = (X_1, \ldots, X_d)$. Of greater generality would be to consider $N = (N_1, \ldots, N_d)$, that is, a different random number of independent copies for each component of $X$. In this context, it can be assumed that $N$ belongs to the domain of attraction of a multivariate extreme value distribution. This provides a second extension of our work. A third extension could be the derivation of nonparametric estimators and their asymptotic results using threshold exceedances (for at least one component). Finally, the extension of our results (probabilistic and inferential) to the case of $S_N$ represents a relevant open problem.

ACKNOWLEDGEMENTS

S.A.P. is supported by the Bocconi Institute for Data Science and Analytics (BIDSA). E.H. is supported by SNSF Grant 200021-175752/1.

ORCID

Simone A. Padoan https://orcid.org/0000-0002-0417-7570

REFERENCES

Anderson, C. W., & Turkman, K. F. (1995). Sums and maxima of stationary sequences with heavy tailed distributions. *Sankhya A*, 57, 1–10.

Barakat, H. M., & El-Shandidy, M. A. (1990). On the limit distribution of the extremes of a random number of independent random variables. *Journal of Statistical Planning and Inference*, 26, 353–361.

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

Bonald, T., Proutière, A., Régnié, G., & Roberts, J. W. (2001). Insensitivity results in statistical bandwidth sharing. *Teletraffic Science and Engineering*, 4, 125–136.

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

Cormier, E., Genest, C., & Neslehova, J. G. (2014). Using B-splines for nonparametric inference on bivariate extreme-value copulas. *Extremes*, 17, 633–659.

de Haan, L. (1984). A spectral representation for max-stable processes. *The Annals of Probability*, 12, 1194–1204.

Embretchts, P., Klüppelberg, C., & Mikosch, T. (1997). *Modelling extremal events, for insurance and finance*. Berlin, Switzerland: Springer-Verlag.

Falk, M., Hüsler, J., & Reiss, R.-D. (2011). *Laws of small numbers: Extremes and rare events*. Basel, Switzerland: Birkhäuser.

Fougères, A.-L., Mercadier, M., & Nolan, P. N. (2013). Dense classes of multivariate extreme value distributions. *Journal of Multivariate Analysis*, 116, 109–129.

Fougères, A.-L., Nolan, P. N., & Rootzén, H. (2009). Models for dependent extremes using stable mixtures. *Scandinavian Journal of Statistics*, 36, 42–59.

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

Gentle, J. E. (2009). *Computational statistics*. New York, NY: Springer.

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

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

Hashorva, E., Ratovomirija, G., & Tamraz, M. (2017). On some new dependence models derived from multivariate collective models in insurance applications. *Scandinavian Actuarial Journal*, 8, 730–750.
Joe, H. (2015). *Dependence modeling with copulas*. Boca Raton, FL: CRC Press.

Klüppelberg, C., & May, A. (2006). Bivariate extreme value distributions based on polynomial dependence functions. *Mathematical Methods in the Applied Sciences*, 29, 1467–1480.

Kojadinovic, I., & Yan, J. (2010). Modeling multivariate distributions with continuous margins using the copula R package. *Journal of Statistical Software*, 34(9), 1–20.

Marcon, G., Padoan, S. A., Naveau, P., Muliere, P., & Segers, J. (2017). Multivariate nonparametric estimation of the Pickands dependence function using Bernstein polynomials. *Journal of Statistical Planning and Inference*, 183, 1–17.

Markovich, N. (2007). *Nonparametric analysis of univariate heavy-tailed data: Research and practice*. New York, NY: John Wiley & Sons Ltd.

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

Mitra, A., & Resnick, S. (2011). Hidden regular variation and detection of hidden risks *Stoch. Models*, 27, 591–614.

Pickands, J., III. (1981). Multivariate extreme value distributions, with a discussion. *Bulletin of the International Statistical*. 49(859–878), 894–902.

R Core Team (2014). R: A language and environment for statistical computing.

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

Resnick, S. I. (2007). *Extreme values, regular variation, and point processes*. New York, NY: Springer.

Robert, C. (2013). Some new classes of stationary max-stable random fields. *Statistics & probability letters*, 83, 1496–1503.

Robert, C. Y., & Segers, J. (2008). Tails of random sums of a heavy-tailed number of light-tailed terms. *Insurance: Mathematics and Economics*, 43, 85–92.

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

Silvestrov, D. S., & Teugels, J. L. (1998). Limit theorems for extremes with random sample sizes. *Advances in Applied Probability*, 30, 777–806.

Stephenson, A. G. (2002). EVD: Extreme value distributions *R* news, Vol 2.

Taqu, M. S., Willinger, W., & Sherman, R. (1997). Proof of a fundamental result in self-similar traffic modeling. *ACM SIGCOMM Computer Communication Review*, 27, 5–23.

Uchaikin, V. V., & Zolotarev, V. M. (1999). *Chance and stability*. Berlin, Boston: De Gruyter.

van de Meent R. (2006). *Network link dimensioning – A measurement and modeling based approach* (Ph.D. thesis).

University of Twente. Retrieved from http://purl.org/utwente/56434

van der Vaart, A. W. (2000). *Asymptotic statistics*. Cambridge, MA: Cambridge University Press.

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

Willinger, W., Taqu, M. S., Sherman, R., & Wilson, D. V. (1997). Self-similarity through high-variability: Statistical analysis of ethernet lan traffic at the source level. *IEEE/ACM Transactions on Networking*, 5, 71–86.

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

**SUPPORTING INFORMATION**

Additional supporting information may be found online in the Supporting Information section at the end of this article.

**How to cite this article:** Hashorva E, Padoan SA, Rizzelli S. Multivariate extremes over a random number of observations. *Scand J Statist*. 2021;48:845–880. https://doi.org/10.1111/sjos.12463
APPENDIX A. PROOFS

A.1 Proof of Theorem 1
We start with some notation. Let \( U_{X_i}(t) := F_{X_i}^{-1}(1 - 1/t), \ t > 1, \ D_j(x) = G_j^{-1}(e^{-1/x}), \ x > 0, \ j = 1, \ldots, d, \) and

\[
F_s(\cdot) = F_X(U_{X_1}(\cdot), \ldots, U_{X_d}(\cdot)), \quad \tilde{G}(y) = G(D_1(y_1), \ldots, D_d(y_d)), \quad y \in (0, \infty)^d.
\]

In extreme-value theory, it is common practice to derive the attractor \( G \) of a distribution \( F_X \) by analyzing separately the behavior of its margins and its dependence structure. The latter is typically investigated by focusing on \( F_s \), which obtains from \( F_X \) by transforming the margins into unit-Pareto. In this way, we have that \( F_s \in D(\tilde{G}) \), where \( \tilde{G} \) has common unit-Fréchet margins and the same extreme-value copula of \( G \). Although different types of common marginal distributions can be considered (e.g., Falk et al., 2011, ch. 4), with Pareto margins we can exploit the theory of regularly varying tails. Precisely, since \( F_X \in D(G) \), we have that by Propositions 5.10, 5.15, and 5.17 in Resnick (2007),

\[
\lim_{n \to \infty} \frac{1 - F_s(ny)}{1 - F_s(n1)} = \frac{-\ln \tilde{G}(y)}{\theta(G)}, \quad y \in (0, \infty)^d,
\]

where \(-\ln \tilde{G}(y)\) is homogeneous of order \(-1\) and \( \theta(G) = \theta(\tilde{G}) = -\ln \tilde{G}(1) \). We also define \( T_s(y) := F_s(y1), \ y > 0, \) a univariate nondecreasing function. Finally, we set \( U_s(t) := T_s^{-1}(1 - 1/t) \) and \( U_N(t) := F_N^{-1}(1 - 1/t), \ t > 1. \)

The proof is organized in three parts: the derivation of the norming constants, two preliminary results and the conclusion. For the sake of brevity, some of the technical derivations are deferred to the supplementary material.

A.1.1 Norming constants
Let the norming sequences \( \kappa_n, \ \theta_n \) for \( F_N \) be defined in the standard way (e.g., Resnick, 2007, pp. 48-54). We recall that \( F^n(x) := F(x), \) \( a_n, b_n \) as \( n \to \infty \) if and only if \( n(1 - F_X(a_n, x + b_n)) \to -\ln G(x) \) as \( n \to \infty \), see Falk et al. (2011), ch. 4, for details. Analogously, here we focus on

\[
n(1 - F_M(na_n + b_n)) = n(1 - L_N(-\ln p_n(x))),
\]

where \( p_n(x) = F_X(c_n + d_n) \). Accordingly, the derivation of the norming constants \( c_n \) and \( d_n \) requires an analysis of the behavior of \( 1 - L_N(s) \), as \( s \downarrow 0 \). Observe that \( L_N(1/x), \ x > 0, \) is monotone nondecreasing. Set \( U_N(x) := (L_N(1/x))^{-1}(1 - 1/x), \ x > 0 \). Define \( z_n := U_n(n/b) \), where \( b := (\theta(G))^\alpha \) when \( F_N \in D(\Phi_\alpha), \ \alpha \in (0, 1), \) and \( b := \theta(G) \) otherwise. Then, setting \( m_n := U_n(z_n) \), as \( n \to \infty \) we have \( m_n \to \infty \) and \( z_n \sim 1/T_s(m_n) \), where \( T_s(y) = 1 - T_s(y), \ y > 0. \) Consequently, we also have

\[
1 - L_N(T_s(m_n)) \sim \frac{b}{n}, \quad n \to \infty.
\]

Next, we explicitly provide some asymptotic approximations which help to understand the results derived in the following subsections. When \( F_N \in D(\Phi_\alpha), \ \alpha \in (0, \infty), \) then \( \tilde{F}_N(y) = 1 - F_N(y), \ y > 0, \) satisfies
\[
\overline{F}_N(y) \sim \mathcal{L}(y) y^{-\alpha}, \quad y \to \infty, \tag{A4}
\]

where \(\mathcal{L}\) is slowly varying. In particular, for any \(\alpha \in (0, 1]\) we have that

\[
1 - L_N(s) \sim \mathcal{L}^*(1/s)s^\alpha, \quad s \downarrow 0, \tag{A5}
\]

for some slowly varying function \(\mathcal{L}^*\), which is equal to \(\Gamma(1 - \alpha)\mathcal{L}\) if \(\alpha \in (0, 1)\), or satisfies \(\lim_{x \to \infty} \mathcal{L}(x)/\mathcal{L}^*(x) = 0\), if \(\alpha = 1\). See Section 3.1 of the supplementary material for details. When \(F_N \in D(\Phi_\alpha), \alpha > 1\), or \(F_N \in D(\Lambda)\), then \(E(N) \in (0, \infty)\) and

\[
1 - L_N(s) \sim E(N)s, \quad s \downarrow 0. \tag{A6}
\]

Therefore, as \(n \to \infty\) we have \(z_n \sim E(N)n/b\) and

\[
\overline{T}_+(m_n) \sim 1/z_n \sim \theta(G)/(nE(N)), \quad n \to \infty. \tag{A7}
\]

We finally derive \(c_n\) and \(d_n\). When \(F_N \in D(\Phi_\alpha), \alpha > 0\),

(i) if \(F_X \in D(\Phi_\beta)\), then we set \(d_{n,j} = 0\) and

\[
c_{n,j} = \begin{cases} U_X(m_n), & \alpha \in (0, 1], \\ U_X(n)^{1/\beta}, & \alpha > 1, \end{cases}
\]

(ii) if \(F_X \in D(\Lambda)\), then we set

\[
c_{n,j} = \begin{cases} \omega_j d_{n,j}, & \alpha \in (0, 1], \\ \omega_j \{Y_j^-(1 - 1/\delta n)\}, & \alpha > 1, \end{cases}
\]

\[
d_{n,j} = \begin{cases} Y_j^-(1 - 1/\delta_j m_n), & \alpha \in (0, 1], \\ c_{n,j} \ln E(N) + Y_j^-(1 - 1/\delta_j n), & \alpha > 1, \end{cases}
\]

where \(Y_j\) is the Von Mises function associated to \(\overline{F}_X\), with \(\overline{F}_X(x) = 1 - F_X(x)\) for \(x \in \mathbb{R}\), \(\omega_j\) is its auxiliary function (e.g., Resnick, 2007, pp. 40-43) and \(\delta_j = \lim_{x \to \infty} \overline{F}_X(x)/\{1 - Y_j(x)\}\);

(iii) \(F_X \in D(\Psi_\beta)\), then we set \(d_{n,j} = x_{0,j}\), where \(x_{0,j} = \sup\{x : F_X(x) < 1\}\), and

\[
c_{n,j} = \begin{cases} \{U_X(m_n)\}^{-1}, & \alpha \in (0, 1], \\ \{U_X(n)E(N)^{1/\beta}\}^{-1}, & \alpha > 1, \end{cases}
\]

where \(\tilde{X}_j \equiv 1/(x_{0,j} - X_j)\) and \(\overline{U}_X = F_X^+(1 - 1/t), t > 0\).

When \(F_N \in D(\Lambda), c_n\) and \(d_n\) are set equal to the sequences derived for the case \(F_N \in D(\Phi_\alpha)\), with \(\alpha > 1\). With these norming constants, we obtain the following approximations as \(n \to \infty\)

\[
U_{X_j}^{-}(c_{n,j}x_j + d_{n,j}) \sim m_n D_j^{-}(x_j) = \begin{cases} m_n x_j^\beta, & F_X \in D(\Phi_\beta) \\ m_n x_j^\beta, & F_X \in D(\Lambda) \\ m_n (x_j)^{-\beta}, & F_X \in D(\Psi_\beta) \end{cases} \tag{A8}
\]
and

\[
1 - p_n(x) \sim 1 - F_s(m_n D_1^-(x_1), \ldots, m_n D_d^-(x_d)) \\
\sim \frac{\ln G(x)}{\theta(G)} \\
\to 0.
\]  

(A9)

### A.1.2 Preliminary results

Given the identity in (A2), our first preliminary result provides the attractor of \( F_{M_N} \).

**Lemma 1.** If \( F_N \in D(\Phi_\alpha), \alpha \in (0, 1] \), then we have

\[
\lim_{n \to \infty} n(1 - \ln p_n(x)) = (-\ln G(x))^\alpha, \quad x \in \mathbb{R}^d.
\]  

(A10)

If \( F_N \in D(\Lambda) \) or \( F_N \in D(\Phi_\alpha), \alpha > 1 \), then (A10) holds with \( \alpha = 1 \).

**Proof.** If \( F_N \in D(\Phi_\alpha), \alpha \in (0, 1] \), using sequentially (A9), (A5) together with the properties of slowly varying functions, and (A3), as \( n \to \infty \) we obtain

\[
n(1 - \ln p_n(x)) \sim n(1 - \ln(1 - p_n(x))) \\
\sim n(1 - \ln(T_s(m_n)(-\ln G(x)))/\theta(G))) \\
\sim n(1 - \ln(T_s(m_n))(-\ln G(x)))/b \\
\sim (-\ln G(x))^a.
\]  

(A11)

If \( F_N \in D(\Phi_\alpha), \alpha > 1 \), or \( F_N \in D(\Lambda) \), we have \( E(N) < \infty \), thus from (A9) and (A7) it follows that, in this case, as \( n \to \infty \)

\[
1 - p_n(x) \sim -\ln G(x)/\{nE(N)\}.
\]  

(A12)

Consequently, in view of (A11) and the approximation in (A6), as \( n \to \infty \) we have

\[
n(1 - \ln p_n(x)) \sim nE(N)(1 - p_n(x)) \\
\sim -\ln G(x),
\]  

(A13)

which completes the proof.

Next, we state an auxiliary result (see Section 3.2 of the supplementary material for the proof), that we use to establish our second preliminary result, characterizing the tail dependence between \( M_N \) and \( N \).

**Lemma 2.** If \( F_N \in D(\Phi_\alpha) \) with \( \alpha \in (0, 1) \), then we have (set \( u_n(y) := \kappa_n y + \rho_n, y > 0 \))

\[
\lim_{n \to \infty} u_n(y)T_s(m_n) = \lim_{n \to \infty} \frac{U_N(n/(-\ln H(y)))}{U_{L_N}(b/n)} = y^{\theta(G) \Gamma^{-1/\alpha}(1 - \alpha)}.
\]  

(A14)

\[
\lim_{n \to \infty} \frac{\bar{F}_N(-c/\ln p_n(x))}{\bar{F}_N(u_n(y))} = \lim_{n \to \infty} \frac{\bar{F}_N(c \cdot g(x)U_{L_N}(n/b))}{\bar{F}_N(u_n(y))} = y^a e^{-a^\alpha(x, \alpha)} \quad \forall c > 0.
\]  

(A15)
with \( \sigma(x, \alpha) \) as in (4) and \( g(x) = \theta(G)/(-\ln G(x)) \). If \( F_N \in D(\Phi_a) \) with \( \alpha \geq 1 \) or \( F_N \in D(\Lambda) \), the limits in (A14)-(A15) (with \( y > 0 \) or \( y \in \mathbb{R} \), respectively) are equal to zero.

**Lemma 3.** If \( F_N \in D(\Phi_a) \) with \( \alpha \in (0, 1) \), then we have

\[
\lim_{n \to \infty} \mathbb{P}(M_N \leq c_n x + d_n | N > u_n(y)) = \pi(x, y) - y^\alpha \left[ (-\ln G(x))^\alpha - \sigma^\alpha(x, \alpha) \gamma(1-\alpha, y\sigma(x, \alpha)) \right] < 1,
\]

where \( \pi(x, y) := e^{-y\sigma(x, \alpha)} \). If \( F_N \in D(\Phi_a) \) with \( \alpha \geq 1 \) or \( F_N \in D(\Lambda) \), then the above limit is equal to 1.

**Proof.** Few algebraic steps yield

\[
\mathbb{P}(M_N \leq c_n x + d_n | N > u_n(y)) = p_n^{u_n(y)}(x) - \frac{1}{F_N(u_n(y))} \int_0^{p_n^{u_n(y)}(x)} \frac{\ln v}{\ln p_n(x)} \, dv,
\]

see Section 3.3 of the supplementary material for details. Using (A9) and Lemma 2, as \( n \to \infty \) we obtain

\[
p_n^{u_n(y)}(x) \sim \exp\{-u_n(y)(1-p_n(x))\}
\]

\[
\sim \exp\{-u_n(y)\overline{T}_n(m_n)(-\ln G(x)/\theta(G))\}
\]

\[
\sim \begin{cases} \pi(x, y), & F_N \in D(\Phi_a), \ \alpha \in (0, 1) \\ 1, & F_N \in D(\Phi_a), \ \alpha \geq 1 \ \text{or} \ F_N \in D(\Lambda). \end{cases}
\]

Hence, if \( F_N \in D(\Phi_a) \) with \( \alpha \in (0, 1) \), by uniform convergence (Resnick, 2007, Proposition 0.5) and Lemma 2, as \( n \to \infty \) we also obtain

\[
- \frac{1}{F_N(u_n(y))} \int_0^{p_n^{u_n(y)}(x)} \frac{\ln v}{\ln p_n(x)} \, dv \sim - \frac{\overline{F}_N(-1/\ln p_n(x))}{\overline{F}_N(u_n(y))} \int_0^{\pi(x,y)} (-\ln v)^{-\alpha} \, dv
\]

\[
\sim -y^\alpha \sigma^\alpha(x, \alpha) [\Gamma(1-\alpha) - \gamma(1-\alpha, y\sigma(x, \alpha))]
\]

\[
= -y^\alpha \left[ (-\ln G(x))^\alpha - \sigma^\alpha(x, \alpha) \gamma(1-\alpha, y\sigma(x, \alpha)) \right].
\]

While, if \( F_N \in D(\Phi_a) \) with \( \alpha \geq 1 \) or \( F_N \in D(\Lambda) \), for any arbitrarily small \( \epsilon > 0 \) and large enough \( n \), we have \( p_n^{u_n(y)}(x) \in [1 - \epsilon, 1] \) and \(-\ln v/(-\ln p_n(x)) > u_n(y)\), for all \( v \in (0, p_n^{u_n(y)}(x))\), in particular for \( v \in (0, 1 - \epsilon)\); therefore, an application of Lemma 2 yields that, as \( n \to \infty \),

\[
\overline{F}_N \left( \frac{\ln v}{\ln p_n(x)} \middle| N > u_n(y) \right) = \frac{\overline{F}_N \left( -\frac{\ln v}{\ln p_n(x)} \middle| N > u_n(y) \right)}{\overline{F}_N(u_n(y))} \to 0, \ \forall v \in (0, 1 - \epsilon),
\]

where \( \overline{F}_N(t| N > u_n(y)) := \mathbb{P}(N > t| N > u_n(y)), t > u_n(y) \). Hence, by the dominated convergence theorem, as \( n \to \infty \)

\[
\frac{1}{\overline{F}_N(u_n(y))} \int_0^{p_n^{u_n(y)}(x)} \overline{F}_N \left( \frac{\ln v}{\ln p_n(x)} \middle| N > u_n(y) \right) \, dv = \int_0^{p_n^{u_n(y)}(x)} \overline{F}_N \left( \frac{\ln v}{\ln p_n(x)} \middle| N > u_n(y) \right) \, dv
\]

\[
\leq \int_0^{1-\epsilon} \overline{F}_N \left( \frac{\ln v}{\ln p_n(x)} \middle| N > u_n(y) \right) \, dv + \epsilon
\]

\[
\to \epsilon,
\]
Since \( \epsilon \) is arbitrarily small, the term on the left-hand side must converge to zero. The proof is now complete.

### A.1.3 Conclusion

As \( n \to \infty \), we have

\[
-n \ln P(M_N \leq c_n x + d_n, N \leq \kappa_n y + \varrho_n) \sim n \left[ 1 - P(M_N \leq c_n x + d_n, N \leq \kappa_n y + \varrho_n) \right] \\
= n \left[ 1 - F_{\kappa_n} (c_n x + d_n) + n \bar{F}_N (u_n (y)) \right] \\
P(M_N \leq c_n x + d_n | N > u_n (y)).
\]

In view of (A2), the limit of the first term on the right hand-side of the second line above is established in Lemma 1. The second term is asymptotically equivalent to

\[-\log H(y) P(M_N \leq c_n x + d_n | N > u_n (y)) ,\]

where the limit of the conditional probability is established in Lemma 3. Combining these results we obtain the limiting expressions in (4) and (5) and the proof is now complete.

### A.2 Proof of Corollary 1

Let \( c_n, d_n, \kappa_n, \varrho_n \) be the norming sequences defined in Section A.1 for the case \( F_N \in D(\Phi_\alpha), \alpha \in (0, 1) \). In particular \( \varrho_n = 0, n \bar{F}_N (\kappa_n) \sim 1 \) and \( \kappa_n \sim U_N (n \Gamma (1 - \alpha)) \) as \( n \to \infty \). The first result, that is, \( \kappa_n^{-1} N_n \sim S \) as \( n \to \infty \), now follows from theorem 5.4.2 in Uchaikin and Zolotarev (2011).

We recall that \( p_n (x) = F_X (c_n x + d_n) \). In the proof of Lemma 3 it has been established that

\[ p_n^{\kappa_n} (x) \sim \exp \left( \frac{-\ln G(x)}{\Gamma^{1/\alpha} (1 - \alpha)} \right) , \quad n \to \infty. \]

Consequently, the second result now follows by noting that by the dominated convergence theorem

\[
\lim_{n \to \infty} \frac{\bar{F}_N^{\kappa_n} (x)}{\bar{F}_N^{\kappa_n} (y)} = \lim_{n \to \infty} \int_0^{\infty} \frac{p_n^{\kappa_n} (x)}{p_n^{\kappa_n} (y)} \, dF_S (s) \\
= \lim_{n \to \infty} \int_0^{\infty} \left( e^{-\frac{\ln G(x)}{\Gamma^{1/\alpha} (1 - \alpha)}} \right)^{\alpha^{1/\alpha} (1 - \alpha)} \, dF_S (s) \\
= \left( \ln G(x) \right) G_\alpha (x).
\]

### A.3 Proof of Corollary 2

Let \( Q_j, j = 1, \ldots, d + 1 \), be the one-dimensional marginal distributions of the max-stable distribution \( Q \). We focus on the case \( F_N \in D(\Phi_\alpha) \) with \( \alpha \in (0, 1) \) – the other cases are trivial. From the first line of (6) we have \( Q_j (x_j) = \exp \{ -[-\ln Q_j (x_j)]^\alpha \}, j = 1, \ldots, d \), from which it follows that

\[ Q_j^{-} (u_j) = G_j^{-} (\exp \{ -[-\ln u_j]^{1/\alpha} \}), \quad u_j \in (0, 1). \]

In particular, \( Q_j^{-} (e^{-1}) = G_j^{-} (e^{-1}), j = 1, \ldots, d \). By assumption, \( Q_{d+1} (y) = \Phi_\alpha (y) \) and therefore \( Q_{d+1}^{-} (u_{d+1}) = \Phi_\alpha^{-} (u_{d+1}), \) with \( \Phi_\alpha^{-} (u) = (-\ln u)^{1/\alpha} \) for \( u \in (0, 1) \). Hence \( Q_{d+1}^{-} (e^{-1}) = 1. \)
The extremal-coefficient is defined by
\[ \theta(Q) = -\log Q(Q_1^{-1}(e^{-1}), \ldots, Q_d^{-1}(e^{-1}), Q_{d+1}^{-1}(e^{-1})). \] (A17)

Plugging in the expressions of \( Q_j^{-1}(e^{-1}) \), \( j = 1, \ldots, d + 1 \), into the right-hand side of (A17), the expression of \( \theta(Q) \) is then obtained.

As for the first result in (8), it immediately follows from the inequalities \( \theta(G) \geq 1 \) and
\[ 1 \leq 1 - E_{\Gamma^{1/\alpha}(1-\alpha)}(\theta(G)) + G_{1-\alpha, \Gamma^{1/\alpha}(1-\alpha)}(\theta(G)), \]
for \( \alpha \in (0, 1) \). To obtain the second one, it is sufficient to note that, as \( \alpha \to 1^- \), \( 1 - E_{\Gamma^{1/\alpha}(1-\alpha)}(\theta(G)) \to 1 \) and
\[
(\theta(G))^{\alpha} G_{1-\alpha, \Gamma^{1/\alpha}(1-\alpha)}(\theta(G)) = \frac{(\theta(G))^\alpha}{\Gamma(1-\alpha)} \gamma \left( 1 - \alpha, \frac{\theta(G)}{\Gamma^{1/\alpha}(1-\alpha)} \right) \\
= \frac{(\theta(G))^\alpha}{\Gamma(1-\alpha)} \left[ \frac{\theta(G)}{\Gamma^{1/\alpha}(1-\alpha)} \right]^{1-\alpha} \sum_{k=0}^{\infty} \frac{\theta(G)}{\Gamma^{1/\alpha}(1-\alpha)} k!(1-\alpha+k) \\
= \frac{\theta(G)}{1+o(1)} \left\{ \frac{1}{1+o(1)} + o(1) \right\}
\]

In the above display, we use the following convergence results: as \( \alpha \to 1^- \), \( \Gamma(2-\alpha) \to \Gamma(1) = 1 \), \( \Gamma(1-\alpha) \to \infty \),
\[ \Gamma^{1/\alpha}(1-\alpha) = \exp \left\{ \frac{1}{\alpha} \ln \Gamma(1-\alpha) \right\} \to \infty, \]
\[ \Gamma^{1/\alpha}(1-\alpha) = \exp \left\{ \left( \frac{1}{\alpha} - 1 \right) \ln \Gamma(2-\alpha) - \frac{1}{\alpha}(1-\alpha)\ln(1-\alpha) \right\} \to 1. \]

**A.4 Proof of Proposition 1**

Let \( U_X(t) = (U_{X_1}(t_1), \ldots, U_{X_d}(t_d)) \), with \( t = (t_1, \ldots, t_d) \geq 1 \) and \( U_{X_j}, j = 1, \ldots, d \), as in Section A.1, and set \( F_s(t) = F_X(U_X(t)) \). Consider the case where \( F_N \in D(\Phi_a) \) with \( 0 < \alpha \leq 1 \). By exploiting (A4)-(A5) and arguments similar to those in the proof of Lemma 1, we obtain that for \( y > 0 \) and as \( n \to \infty \)
\[ 1 - \mathbb{P}(M_N \leq U_X(ny)) = 1 - L_N(-\ln F_s(ny)) \\
\sim \{1 - F_s(ny)\}^{\alpha} \mathcal{L}^*(\{1 - F_s(ny)\}^{-1}) \\
\sim \{-\ln \tilde{G}(y)/n\}^{\alpha} \mathcal{L}^*(n\{-\ln \tilde{G}(y)\}^{-1}) \\
\sim \begin{cases} 
\Gamma(1-\alpha)\mathbb{P}(N > -n/\ln \tilde{G}(y)), & \alpha \in (0, 1) \\
(-\ln \tilde{G}(y))(1 - L_N(1/n)), & \alpha = 1
\end{cases} . \]
As a consequence $1 - \mathbb{P}(M_N \leq U_{M_N}(ny)) \sim n^{-1}(-\ln \tilde{G}(y^{1/a}))^a$ as $n \to \infty$.

Consider the case where $F_N \in D(\Phi_{+})$ with $a > 1$ or $F_N \in D(\Lambda)$. Again, with steps similar to those in Lemma 1 we have that for $y > 0$ and as $n \to \infty$

$$1 - \mathbb{P}(M_N \leq U_X(ny)) \sim \mathbb{E}(N)(1 - F_s(ny))$$

$$\sim \mathbb{E}(N)(1 - F_s(n1)) - \ln \tilde{G}(y)$$

$$\sim n^{-1} \mathbb{E}(N)(-\ln \tilde{G}(y)).$$

In addition, $U_{M_N}(ny) \sim U_X(\mathbb{E}(N)ny)$ as $n \to \infty$. Therefore, $1 - \mathbb{P}(M_N \leq U_{M_N}(ny)) \sim n^{-1}(-\ln \tilde{G}(y))$ as $n \to \infty$.

### A.5 Proof of Proposition 2

Observe that $\mathbb{E}(Z_1^a) = \Gamma(1 - a)$ for every $0 < a < 1$. Setting $k_a = 1/(\Gamma(1 - a)) > 0$ we have

$$\frac{1}{k_a} = \mathbb{E}(Z_1^a) = \int_0^\infty \mathbb{P}\{Z_1 > t^{1/a}\} \, dt$$

$$= \int_0^\infty \left(1 - e^{-t^{1/a}}\right) \, dt = \int_0^\infty \left(1 - e^{-v^{1/a}}\right) v^{-2} \, dv. \quad (A18)$$

For any positive $(y_1 \ldots y_d)$ we have

$$-\ln \mathbb{P}(R_1 \leq y_1, \ldots, R_d \leq y_d) = k_a \mathbb{E}(\max_{1 \leq j \leq d} Z_j^a/y_j^a)$$

$$= k_a \int_0^\infty \mathbb{P}\{\exists j \in \{1, \ldots, d\} : Z_j^a > v y_j^a\} \, dv$$

$$= \int_0^\infty \mathbb{P}\{\exists j \in \{1, \ldots, d\} : v k_a Z_j^a > y_j^a\} v^{-2} \, dv$$

$$= \int_0^\infty \mathbb{P}\{\exists j \in \{1, \ldots, d\} : Z_j > y_j(k_a v)^{-1/a}\} v^{-2} \, dv$$

$$= \int_0^\infty \left[1 - \mathbb{P}\{\forall j \in \{1, \ldots, d\} : Z_j \leq y_j(k_a v)^{-1/a}\}\right] v^{-2} \, dv.$$ 

Now, we have

$$\mathbb{P}\{\forall j \in \{1, \ldots, d\} : Z_j \leq y_j(k_a v)^{-1/a}\} = \exp \left(-L(1/y_1 \ldots 1/y_d) k_a^{1/a} v^{1/a}\right),$$

where $L(z)$, with $z = 1/y$, is the stable-tail dependence function pertaining to $\tilde{G}$. Consequently, by (A18) we obtain the final result
\[-\ln \mathbb{P}(R_1 \leq y_1, \ldots, R_d \leq y_d) = \int_0^\infty \left[ 1 - e^{-L(1/y_1, \ldots, 1/y_d)k_u^{1/\alpha}/u} \right] v^{-2}dv\]
\[= (L(1/y_1 \ldots 1/y_d))^\alpha k_u \int_0^\infty \left[ 1 - e^{-v^\alpha/u} \right] v^{-2}dv\]
\[= L^\alpha (1/y_1 \ldots 1/y_d)\]

establishing the proof.

### A.6 Proof of Proposition 3

Let \(M_n := \max\{X_1, \ldots, X_n\}\). By the max-stability of \(G\), there exist maps \(\mathcal{A} : (0, \infty) \mapsto (0, \infty)^d\) and \(\mathcal{B} : (0, \infty) \mapsto \mathbb{R}^d\), such that \(G^s(x) = G(\mathcal{A}(s)x + \mathcal{B}(s)), s > 0\). By formula (5.18) in Resnick (2007), for every \(s > 0\) we have

\[
\frac{a_{[ns]} (M_n - b_n)}{a_n} \left( \frac{M_n - b_n}{a_n} - \frac{b_n}{a_{[ns]}} \right) \leq 1 \frac{a_{[ns]} (M_n - b_n)}{a_n} \left( \frac{M_n - b_n}{a_n} - \frac{b_n}{a_{[ns]}} \right) \sim G(\mathcal{A}(s) \cdot + \mathcal{B}(s)) = G^s, \quad n \to \infty.
\]

where \(\sim\) denotes asymptotic equivalence in distribution. Consequently, the final result follows from the equality \(G_n(x) = \mathbb{E}(G^s(x)), x \in \mathbb{R}^d\), and the dominated convergence theorem

\[
G_n(x) = \int_0^\infty G^s(x) dF_S(s) = \lim_{n \to \infty} \mathbb{E}\left( \frac{a_{[ns]} (M_n - b_n)}{a_n} \left( \frac{M_n - b_n}{a_n} - \frac{b_n}{a_{[ns]}} \right) \right) \leq x
\]
\[= \lim_{n \to \infty} \mathbb{P}\left( a_n^{-1} \left[ \max\{w_n(X_1 - v_n), \ldots, w_n(X_n - v_n)\} - b_n \right] \leq x \right).
\]

### A.7 Proof of Proposition 4

We recall that for all \(x \in \mathbb{R}^d\) we have \(G(x) = C_G(G_1(x_1), \ldots, G_d(x_d))\), where the extreme-value copula is of the form \(C_G(u) = \exp(-L((-\ln u_1), \ldots, (-\ln u_d)), u \in (0, 1]^d\), \(L\) is the stable tail dependence function of \(G\) and the corresponding Pickands dependence function satisfies (1). We also recall that, for \(\alpha \in (0, 1), G_n(x) = \exp(-L_n(x)^\alpha)\) is a max-stable distribution, with copula \(C_{G_n}\) given in (7).

The copula \(C_{G_n}\) must be of extreme-value type, that is, of the form \(C_{G_n}(u) = \exp\{-L_n((-\ln u_1), \ldots, (-\ln u_d))\}, u \in [0, 1]^d\), for a stable tail-dependence function \(L_n\). We then deduce that

\[
L_n((-\ln u_1), \ldots, (-\ln u_d)) = L^\alpha((-\ln u_1)^{1/\alpha}, \ldots, (-\ln u_d)^{1/\alpha}),
\]

for all \(u \in [0, 1]^d\). By the homogeneity of the stable-tail dependence function, we have

\[
L_n(z_1 \ldots z_d) = (z_1 + \ldots + z_d)A_n(t), \quad z \in [0, \infty)^d,
\]

where \(t_j = z_j/(z_1 + \ldots + z_d)\) for \(j = 1, \ldots, d\) and \(A_n\) is the Pickands function of \(G_n, \alpha \in (0, 1)\). Combining (A19) with (1), we obtain

\[
L_n(z_1 \ldots z_d) = \left( \sum_{j=1}^d z_j^{1/\alpha} \right)^{\alpha} \left\{ A \left( \frac{z_1^{1/\alpha}}{\sum_{j=1}^d z_j^{1/\alpha}} \right) \right\}^\alpha,
\]
where $A$ is the Pickands dependence function of $G$. Therefore, choosing $z \in S_d$ we finally obtain

$$A_a(t) = \left(\frac{\sum_{j=1}^{d} t_j^{1/a}}{\sum_{j=1}^{d} t_j^{1/a}}\right)^{a},$$

which is the result in (11).

### A.8 Proof of Theorem 2

#### A.8.1 Notation and general setting

**Empirical processes.** Recall that $(\eta_1, \xi_1), \ldots, (\eta_n, \xi_n)$, $n = 1, 2, \ldots$, are iid random vectors with distribution $Q$ in (4) with fixed $\alpha \in (0, 1)$. For $i = 1, 2, \ldots$, and $j \in \{1, \ldots, d\}$, let

$$V_i := \Phi_d(\xi_i), \quad U_{ij} := G_{\alpha,j}(\eta_{ij}), \quad \hat{U}_{ij} := G_{\alpha,j}(\eta_{ij}),$$

(A20)

where $\Phi_d$ is the $\alpha$-Fréchet distribution, $G_{\alpha,j}$ is the $j$-th margin of the distributions in the first line of (6) and $G_{\alpha,j}$ is as in (19). Set $\mathbf{U}_i = (U_{i1}, \ldots, U_{id})$ and $\hat{\mathbf{U}}_i = (\hat{U}_{i1}, \ldots, \hat{U}_{id})$. In the sequel, when the index $i$ is omitted, we refer to a single observation. For every $\mathbf{u} \in [0, 1]^d$ and $v \in [0, 1]$, define the random copula functions

$$C_{Q,n}(\mathbf{u}, v) := \frac{1}{n} \sum_{i=1}^{n} 1(U_i \leq \mathbf{u}, V_i \leq v), \quad C_{\alpha,n}(\mathbf{u}) = C_{Q,n}(\mathbf{u}, 1), \quad B_n(v) = C_{Q,n}(1, v),$$

(A21)

where $\mathbf{1} = (1, \ldots, 1)$, and the copula processes

$$C_{Q,n}(\mathbf{u}, v) = \sqrt{n}(C_{Q,n}(\mathbf{u}, v) - C_Q(\mathbf{u}, v)), \quad C_{\alpha,n}(\mathbf{u}) = C_{Q,n}(\mathbf{u}, 1), \quad B_n = C_{Q,n}(\mathbf{1}, v).$$

(A22)

Let $C_{Q,n}(\mathbf{u}) := C_{Q}(\mathbf{u}, 1), \mathbf{u} \in [0, 1]^d$. The covariance function of $C_{\alpha}$ is as in (25), with $C_Q$ replaced by $C_{\alpha}$. Furthermore, for every $\mathbf{u} \in [0, 1]^d$, define the empirical copula function and process

$$\hat{C}_{\alpha,n}(\mathbf{u}) = \frac{1}{n} \sum_{j=1}^{n} 1(\hat{U}_i \leq \mathbf{u}), \quad \hat{C}_{\alpha,n} = \sqrt{n}(\hat{C}_{\alpha,n} - C_{\alpha}).$$

(A23)

In the sequel we view the above empirical processes as random signed measures, when appropriate (e.g., van der Vaart & Wellner, 1996, examples 1.7.4 and 1.10.6). We then use the notation $Mf := \int f \, dM$, for any signed measure $M$ on the measurable space $(\mathcal{X}, \mathcal{A})$ and measurable function $f$. Furthermore, for asymptotically measureable sequences $X_n, X'_n$ in $\ell(\mathcal{X})$, with $n \to \infty$, $X_n(\cdot) = X'_n(\cdot) + o_p(1)$ stands for $\sup_{x \in \mathcal{X}} |X_n(x) - X'_n(x)| = o_p(1)$, and we recall that $X_n(\cdot) \Rightarrow X(\cdot)$ is shorthand for $\{X_n(x)\}_{x \in \mathcal{X}} \Rightarrow \{X(x)\}_{x \in \mathcal{X}}$ in $\ell(\mathcal{X})$, as already stated in Section 1.

**Weighting and selection maps.** For all $f \in \ell([0,1]^{d+1})$, let $g_c : \ell([0,1]^{d+1}) \leftrightarrow \ell([0,1]^{d+1})$ be the weighting-map given by

$$(g_c(f))(z) = \begin{cases} w_c^{-1}(z)f(z), & z \in (0, 1]^{d+1} \setminus \{1\} \\ 0, & \text{otherwise}, \end{cases}$$

(A24)
where, for any fixed $\epsilon \in [0, 1/2)$, $w_\epsilon$ denotes the weighting-function

$$w_\epsilon : [0, 1]^{d+1} \mapsto [0, 1] : z \mapsto \min_{1 \leq j \leq d+1} z_j \left( 1 - \min_{1 \leq j \leq d+1} z_j \right)^\epsilon.$$  

To keep the notation light, when $f(z)$ is computed at $z = (u, 1)$, we occasionally still write $g_\epsilon(f)$. The difference in meaning will be clear from the context. For every $f \in \ell^2([0,1]^{d+1})$, let $\pi_{1,\ldots,d} : \ell^2([0,1]^{d+1}) \mapsto \ell^2([0,1]^d)$ and $\pi_{d+1} : \ell^2([0,1]^{d+1}) \mapsto \ell^2([0,1])$ be the selection maps defined by

$$\pi_{1,\ldots,d}(f)(u) := f(u_1, \ldots, u_d, 1), \quad (\pi_{d+1}(f)) := f(1, \ldots, 1, v). \quad (A25)$$

Then, for every $\alpha \in (0, 1)$, $\epsilon \in [0, 1/2)$ and $u \in (0, 1]^d \setminus \{1\}$, let $\omega_{\epsilon,u} : [0, 1]^d \mapsto \mathbb{R}$ be the weighted-function defined by

$$\omega_{\epsilon,u}(v) := \frac{\mathbb{I}(v \leq u) - C_g(u)}{(\pi_{1,\ldots,d}(w_\epsilon))(u)}. \quad (A26)$$

For every $u, v \in [0, 1]^d$, set

$$\omega^\prime_{\epsilon,u}(v) = \begin{cases} \omega_{\epsilon,u}(v), & u \in (0, 1]^d \setminus \{1\} \\ 0, & \text{otherwise.} \end{cases} \quad (A27)$$

Proof’s overview. By Genest and Segers (2009) and Gudendorf and Segers (2012) we have that as $n \to \infty$

$$g_\epsilon(C_{Q,n}) \sim g_\epsilon(C_Q), \quad g_\epsilon(C_{G_n,n}) \sim g_\epsilon(C_{G_n}), \quad (A28)$$

in $\ell^2([0, 1])^{d+1}$ and $\ell^2([0, 1])^d$, respectively. In particular, $g_\epsilon(C_Q)$ is a zero-mean Gaussian process on $[0, 1]^{d+1}$ with covariance function

$$\text{Cov}\{g_\epsilon(C_Q)(u), g_\epsilon(C_Q)(v)\} = \frac{C_Q(u \wedge v) - C_Q(u)C_Q(v)}{w_\epsilon(u)w_\epsilon(v)} \cdot \omega(1), \quad u, v \in [0, 1]^{d+1},$$

and $g_\epsilon(C_{G_n})$ is a zero-mean Gaussian process on the lower dimensional hypercube $[0, 1]^d$ with covariance function defined analogously. The convergence results in (A28) and, more generally, the asymptotic properties of the copula processes $C_{Q,n}$ and $C_{G_n,n}$, are crucial for the derivation of the results presented in Section A.8. As argued in Appendix A.8.3, the P, CFG and MD estimators of $A_\alpha$ belong to a class of estimators $\hat{A}_{\alpha,n}$ satisfying the following conditions.

**Condition 2.** The estimator $\hat{A}_{\alpha,n}$ of $A_\alpha$ allows the subsequent representation:

(i) For a continuous linear map $\phi : \ell^\infty([0,1]^d) \mapsto \ell^\infty(S_d)$ and some $\epsilon \in [0, 1/2)$,

$$\sqrt{n}\{\ln\hat{A}_{\alpha,n}(\cdot) - \ln A_\alpha(\cdot)\} = (\phi \circ g_\epsilon(C_{G_n,n}))(\cdot) \circ \rho(1),$$

where $C_{G_n,n}$ is as in (A22);
(ii) For some \( m \in \mathbb{N_+} \) and \(-\infty < a < b < \infty\) the map \( \phi \) is of the form
\[
(\phi(f))(y) = \sum_{i=0}^{m} \int_{a}^{b} f(\beta_{i,1}(z), \ldots, \beta_{i,d}(z))K_i(z; t) \, dz.
\] (A29)

where, for \( i = 1, \ldots, m, j \in \{1, \ldots, d\} \) and \( t \in S_d, z \mapsto \beta_{i,j}(z, t) \) is a bijective and continuous function, while \( K_1, \ldots, K_m \) are functions satisfying
\[
\sup_{t \in S_d} \max_{0 \leq i \leq m} |K_i(z; t)| \leq K(z), \quad z \in (a, b), \quad -\infty < a < b < \infty,
\]
for an integrable function \( K \).

An estimator \( \hat{A}_{n,a} \) allowing the above representations together with some suitable estimators \( \hat{a}_n \) of \( a \) (that meet some appropriate conditions) enable to deduce a general theory on the weak convergence for composite-estimators \( \hat{A}_{n} \) of \( A \), based on the copula process \( C_{Q,n} \). Such a theory is established in Appendix A.8.2 and then applied to the specific cases of the GPWM and ML estimators of \( a \), and the \( P \), CFG and MD estimators for \( A_{n,a} \), in Appendix A.8.3. Some auxiliary results are deferred to Appendix A.8.4. The arguments therein make extensive use of composition maps related to \( \phi \). Hence, to improve the readability of the remaining part of Section A.8 we conclude this subsection by providing a comprehensive list of such composition maps.

**Definition 1.** Let \( \phi \) be as in Condition 2 and \( a \in (0, 1) \) denote the true parameter value for the distribution \( Q \) in (4). Then:

(i) For a measurable functional \( \tau : L^\infty([0, 1]) \to \mathbb{R} \), the map \( \phi_\tau : L^\infty([0, 1]^{d+1}) \to L^\infty(S_d) \) is defined via
\[
(\phi_\tau(f))(t) = a^{-1}(\phi \circ \pi_{1,\ldots,d}(f))(t) + K_a(t)[\tau \circ \pi_{d+1}(w_c f)],
\] (A30)

for all \( f \in L^\infty([0, 1]^{d+1}) \) and \( t \in S_d \), with
\[
K_a(t) = a^{-2} \left\{ \frac{||t||^{-1/\alpha}}{1/\alpha} \sum_{1 \leq j \leq d; t_j > 0} t_j^{1/\alpha} \ln t_j - \ln A_a(t) \right\},
\] (A31)

\( \pi_{d+1} \) as in (A25) and \( w_c \) as in (A24). In particular, \( \tau \circ \pi_{d+1}(w_c f) \) is a (real-valued) measurable functional;

(ii) \( \phi' : L^\infty([0, 1]^d) \to L^\infty(S_d) \) is defined, for all \( f \in L^\infty([0, 1]^d) \) and \( t \in S_d \), via
\[
(\phi'(f))(t) = a^{-1}(\phi(f))(t) + K_{m+1}(t)f(1),
\]
with \( K_a \) as in (A31) and
\[
K_{m+1}(t) = K_a(t) - a^{-1} \sum_{i=0}^{m} \int_{a}^{b} K_i(z; t) \, dz.
\] (A32)
Letting \( \omega'_{c,u} \) be as in (A27) and \( \delta(u) = 1, \forall u \in [0, 1]^d \), for a measurable function \( \varphi : [0, 1] \rightarrow \mathbb{R} \)
\[
g'_{c,\varphi} : M \mapsto \left\{ \int_{[0,1]^{d+1}} \left[ \omega'_{c,u}(a) + \delta(u)\varphi(b) \right] \ dM(a, b) \right\} \quad \text{for } u \in [0,1]^d \quad (A33)
\]
maps a signed measure \( M \) on \([0, 1]^{d+1}\) to an element of \( \ell^\infty([0, 1]^{d+1}) \). With a little abuse of notation, we also apply this operator to the \( C_Q \)-Brownian bridge \( \mathbb{C}_Q \) in Theorem 2, yielding \( g'_{c,\varphi}(\mathbb{C}_Q) \), the zero-mean Gaussian process on \([0, 1]^d\) with covariance function in (A39);

(iii) \( \phi_{MD} : \ell^\infty([0, 1]^d) \mapsto \ell^\infty(S_d) \) is a special case of the map \( \phi \), defined, for every \( t \in S_d \), by
\[
(\phi_{MD}(f))(t) = \frac{(1 + A_\alpha(t))^2}{A_\alpha(t)} \left( \sum_{j=1}^{d} \int_{0}^{1} \hat{C}_{\alpha,j}(v^i, \ldots, v^d)f(1, \ldots, 1, v^j, 1, \ldots, 1) \ dv \right)
\]
\[
- \int_{0}^{1} f(v^1, \ldots, v^d) \ dv \quad ;
\]

(A34)

(iv) \( \phi_{MD,ML} = A_\alpha \phi'_{MD} \circ g'_{0, \varphi_{ML}} \), where \( \phi'_{MD} \) is a special case of the map \( \phi' \) in (ii) defined via
\[
(\phi'_{MD}(f))(t) = \alpha^{-1}(\phi_{MD}(f))(t) + \frac{1}{A_\alpha(t)} \left( \sum_{j=1}^{d} \int_{0}^{1} \hat{C}_{\alpha,j}(v^i, \ldots, v^d)dv - 1 \right)
\]

for every \( f \in \ell^\infty([0, 1]^d) \) and \( t \in S_d \), with \( \phi_{MD} \) as in (A34), \( \frac{1}{\alpha} \) given by
\[
K_{d+1}^{MD}(t) = \frac{(1 + A_\alpha(t))^2}{A_\alpha(t)} \left( \sum_{j=1}^{d} \int_{0}^{1} \hat{C}_{\alpha,j}(v^i, \ldots, v^d)dv - 1 \right)
\]

and \( A_\alpha \) is as in (A31); \( g'_{0, \varphi_{ML}} \) is as in (ii), for the particular choices of \( \epsilon = 0 \) and \( \varphi = \varphi_{ML} \) in (A41);

(v) \( \phi_{MD,GPWM} : \ell^\infty([0, 1]^{d+1}) \mapsto \ell^\infty(S_d) \) is defined, for all \( f \in \ell^\infty([0, 1]^{d+1}) \) and \( t \in S_d \), via
\[
(\phi_{MD,GPWM}(f))(t) = A_\alpha(t)(\phi_{MD,GPWM}^j(f))(t) = A_\alpha(t)(\phi_{MD,GPWM}^j(f))(t),
\]
where \( \phi_{MD,GPWM} \) is a special case of the map \( \phi_j \) in (A30), with \( \phi = \phi_{MD} \) in (A34) and \( \tau = \tau_{GPWM} \) in (A43);

(vi) \( \phi_o : \ell^\infty([0, 1]^d) \mapsto \ell^\infty(S_d) \) is a special case of the map \( \phi \) defined by
\[
(\phi_o(f))(t) = \int_{0}^{\infty} f(e^{-v^1}, \ldots, e^{-v^d}) \beta_t(e^{-\max(t)}) \ dv
\]
\[
- \sum_{j=1}^{d} \int_{0}^{\infty} \hat{C}_{\alpha,j}(v^i, \ldots, v^d)f(1, \ldots, 1, v^j, 1, \ldots, 1) \beta_t(e^{-v^j})h_o(t; v) \ dv \quad ,
\]

(A35)

with \( \max(t) = \max(t_1, \ldots, t_d) \), \( \beta_t(v) = v^i(1 - v^e) \) for \( v \in (0, 1) \). Here \( h_o \) is either \( h_\varphi(t; v) = -A_\alpha^{-1}(t) \) or \( h_{CFG}(t; v) = 1/v \) for \( v > 0, t \in S_d \);

(vii) \( \phi_{o,ML} = A_\alpha \phi_o \circ g'_{c,\varphi_{ML}} \), where \( g'_{c,\varphi_{ML}} \) is as in (ii), with \( \varphi = \varphi_{ML} \) in (A41), and \( \phi' \) is a special case of the map \( \phi' \) in (ii) defined via
\[(\phi_\circ (f))(t) = \alpha^{-1}(\phi_\circ (f))(t) + K_{d+1}^\circ(t)f(1), \]

for every \(f \in \ell^\infty([0, 1])\) and \(t \in S_d\), with \(\phi_a\) as in (A35) and

\[K_{d+1}^\circ(t) = K_a(t) - \int_0^\infty \beta_c(e^{-v\max(t)}) \frac{h_c(t; v)}{\alpha} \, dv + \sum_{j=1}^d \int_0^\infty \tilde{C}_{G_a,j}(v^1, \ldots, v^d) \beta_c(e^{-v_j}) \frac{h_c(t; v)}{\alpha} \, dv; \]

(viii) \(\phi_{\circ,\text{GPWM}} : \ell^\infty([0, 1]^{d+1}) \mapsto \ell^\infty(S_d)\) is defined, for all \(f \in \ell^\infty([0, 1]^d)\) and \(t \in S_d\), via

\[(\phi_{\circ,\text{GPWM}}(f))(t) = A_\circ(t)(\phi_{\circ,\text{GPWM}} \circ g_\circ(f))(t) \]

where \(\phi_{\circ,\text{GPWM}}\) is a special case of the map \(\phi_\circ\) in (A30), with \(\phi = \phi_a\) in (A35) and \(\tau = \tau_{\text{GPWM}}\) in (A43).

### A.8.2 General preliminary results

The results in this section rely on the following conditions on the estimators \(\hat{\alpha}_n\) of \(\alpha\).

**Condition 3.** Let \(\hat{\alpha}_n\) be an estimator of \(\alpha\) satisfying one of the following properties:

(i) There is a continuous linear map \(\tau : \ell^\infty([0, 1]) \mapsto \mathbb{R}\) such that

\[\sqrt{n}(\hat{\alpha}_n - \alpha) = \tau(\mathbb{B}_n) + o_p(1),\]

where \(\mathbb{B}_n\) is as in (A22);

(ii) There is a measurable function \(\zeta : (0, +\infty) \mapsto \mathbb{R}\) such that \(\Phi_{\alpha, \zeta} = 0, \Phi_{\alpha, \zeta^2} < \infty\) and

\[\sqrt{n}(\hat{\alpha}_n - \alpha) = n^{-1/2}(\zeta(\xi_1) + \ldots + \zeta(\xi_n)) + o_p(1).\]

When a general composite-estimator \(\widehat{A}_{\circ,\alpha}^\circ\) is obtained combining an estimator \(\widehat{A}_{\circ,\alpha}\) of \(A_{\alpha}\) and \(\hat{\alpha}_n\) of \(\alpha\) satisfying Conditions 2(i) and 3(i), respectively, the functional weak convergence of \(\sqrt{n}(\widehat{A}_{\circ,\alpha}^\circ - A_{\alpha})\) can be established by fairly direct arguments. The GPWM-based composite-estimators fall in this category of estimators, see Appendix A.8.3 and Lemma 8. When \(\hat{\alpha}_n\) complies with 3(ii), to study the joint limit behavior of \(\widehat{A}_{\circ,\alpha,n}\) and ultimately determining that of \(\sqrt{n}(\widehat{A}_{\circ,\alpha,n}^\circ - A^\circ)\) requires more complex asymptotic arguments. This is the case, for example, for composite-estimators based on M-estimators of \(\alpha\) (van der Vaart, 2000, ch. 5), such as the MLE. The following propositions derive the required theory. Hereafter, the notations with the superscript “\(^\circ\)” are specific to the second type of asymptotic results.

**Proposition 5.** Let \(\widehat{A}_{\circ,\alpha}^\circ\) be the composite-estimator obtained by the composition of \(\widehat{A}_{\circ,\alpha,n}\) and \(\hat{\alpha}_n\) satisfying Conditions 2(i) and 3(i), respectively. Let \(\phi_\circ\) be as in Definition 1(i) and \(g_\circ\) be as in (A24). Then, in \(\ell^\infty(S_d)\)

\[\sqrt{n}(\widehat{A}_{\circ,\alpha}^\circ(\cdot) - A_{\alpha}(\cdot)) \Rightarrow A_{\circ}(\cdot)(\phi_\circ \circ g_\circ(C_Q))(\cdot), \quad n \to \infty. \quad (A36)\]
Proof. The claim in (A36) relies on the following result.

**Lemma 4.** With \(|t|_{1/h}^1\) being defined as in (12), the map

\[
g : (0, \infty) \mapsto \ell^\infty(S_d) : h \mapsto \left(\ln |t|_{1/h}^1\right)_{t \in S_d}
\]

is Hadamard differentiable at \(\alpha\) with derivative

\[
\{(g_\alpha(h))_\alpha\} \big|_{t \in S_d} = \left\{ -h \alpha^{-2} |t|_{1/\alpha}^{-1/\alpha} \sum_{1 \leq j \leq d} t_j^{1/\alpha} \ln t_j \right\} \big|_{t \in S_d}, \quad 0 < h < \infty.
\]

For the proof see Section 4 of the supplementary material. For simplicity we focus on \(\ln A^*\) and \(\ln \hat{A}^*_n\). For \(t \in S_d\), we have that

\[
\sqrt{n} \{\ln \hat{A}^*_n(t) - \ln A^*(t)\} = \sqrt{n} \left\{ \hat{\alpha}_n^{-1} \ln \hat{A}_{a,n}(t) - \alpha^{-1} \ln A_a(t) \right\} - \sqrt{n} \left( \ln |t|_{1/\hat{\alpha}_n}^1 - \ln |t|_{1/\alpha}^1 \right)
\]

\[
\equiv T_{1,n}(t) + T_{2,n}(t).
\]

By Condition 3(i), the functional version of Slutsky's lemma (van der Vaart & Wellner, 1996, p. 32) and the delta method (van der Vaart, 2000, Theorem 3.1), it follows that

\[
T_{1,n}(\cdot) = \alpha^{-1} \sqrt{n} \{\ln \hat{\alpha}_{a,n}(\cdot) - \ln A_a(\cdot)\} - \alpha^{-2} \ln A_a(\cdot) \sqrt{n}(\hat{\alpha}_n - \alpha) + o_p(1).
\]

By Lemma 4 and the functional delta method (van der Vaart, 2000, ch. 20) it follows that

\[
T_{2,n}(\cdot) = \{K_a(\cdot) - \alpha^{-2} \ln A_a(\cdot)\} \sqrt{n}(\hat{\alpha}_n - \alpha) + o_p(1),
\]

where \(K_a\) is as in (A31). Then, by Conditions 2(i) and 3(i) and the identity \(\tau \circ \pi_{d+1}(\mathbb{C}_{Q,n}) = \tau \circ \pi_{d+1}(w, g_c(\mathbb{C}_{Q,n}))\), with \(\pi_{d+1}\) as in (A25) and \(w_c\) as in (A24), we obtain

\[
T_{1,n}(\cdot) + T_{2,n}(\cdot) = \alpha^{-1} (\phi(\mathbb{C}_{Q,n}))(\cdot) + K_a(\cdot) \sqrt{n}(\hat{\alpha}_n - \alpha) + o_p(1)
\]

\[
= \alpha^{-1} (\phi(\mathbb{C}_{Q,n}))(\cdot) + \alpha(\cdot) \tau(\mathbb{B}_n) + o_p(1)
\]

\[
= (\phi \circ g_c(\mathbb{C}_{Q,n}))(\cdot) + o_p(1), \quad (A37)
\]

Thus, the result in (A36) follows from (A28), by applying the continuous mapping theorem and the functional delta method in the last line of (A37).  

**Proposition 6.** Let \(\hat{A}^*_n\) be the composite-estimator obtained by the composition of \(\hat{A}_{a,n}\) and \(\hat{\alpha}_n\) complying with Condition 2(i)-(ii) and Condition 3(ii), respectively, with \(\varphi = \zeta \circ \Phi^*_\alpha\) satisfying

\[
-\infty < E(\omega'_{\varepsilon, u}(u)\varphi(v)) < \infty, \quad \forall u \in (0, 1]^d \setminus \{1\}, \quad (A38)
\]

where \(\omega'_{\varepsilon, u}\) is given in (A27). Let \(\psi'\) and \(g'_{\varepsilon, \varphi}\) be as in Definition 1(ii). Then, in \(\ell^\infty(S_d)\) as \(n \rightarrow \infty\)
\[ \sqrt{n} \left\{ \hat{A}^*_n(\cdot) - A^*(\cdot) \right\} \overset{d}{\to} A_n(\cdot)(\phi' g_{e,\varphi}(C_Q))(\cdot), \]

where \( g_{e,\varphi}(C_Q) \) is a zero-mean Gaussian process with covariance function defined in (A39).

**Proof.** For any \( u_i \in [0, 1]^d, i \leq k, k \in \mathbb{N}_+ \), the random vectors \( (\omega'_{e,u}(U_1), \ldots, \omega'_{e,u}(U_i), \varphi(V_i)), i = 1, \ldots, n \), are iid with zero-mean and finite pairwise covariances, by arguments in Genest and Segers (2009), Gudendorf and Segers (2012), Condition 3(ii), and (A38). Let

\[
(g'_{e}(C_{Q,n}))(u) = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \omega'_{e,u}(U_i), \quad \overline{\varphi}_n(u) = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \varphi(V_i), \quad u \in [0, 1]^d,
\]

where \( \delta(u) = 1 \) for all \( u \in [0, 1]^d \). Note that \( (g'_{e}(C_{Q,n}))(u) = (g_{e}(C_{Q,n}))(u) \), with \( u \in [0, 1]^d \). Then, \( g'_{e}(C_{Q,n}) \) and \( \overline{\varphi}_n \) are asymptotically tight (van der Vaart & Wellner, 1996, Definition 1.3.7) and, by the central limit theorem,

\[
((g'_{e}(C_{Q,n}))(u_1), \ldots, (g'_{e}(C_{Q,n}))(u_k), \overline{\varphi}_n(u_1), \ldots, \overline{\varphi}_n(u_k)) \overset{d}{\to} N(0, \Sigma)
\]
as \( n \to \infty \). By arguments in van der Vaart & Wellner (1996, p. 42 point 3), these facts are sufficient to claim that the class of functions \( \mathcal{Y}_{e,\varphi} := \{(a, b) \mapsto \omega'_{e,u}(a) + \delta(u)\varphi(b) : u \in [0, 1]^d \} \) is \( C_{Q,Donsker} \) (van der Vaart & Wellner, 1996, pp. 80-82). Indeed, introducing the map

\[
g'_{e,\varphi} : M \mapsto \{ Mf : f \in \mathcal{Y}_{e,\varphi} \}
\]
declared on the space of signed measure \( M \) on \( [0, 1]^{d+1} \), we have that \( (g'_{e,\varphi}(C_{Q,n}))(\omega'_{e,u} + \delta(u)\varphi) = (g'_{e}(C_{Q,n}))(u) + \overline{\varphi}_n(u), \forall u \in [0, 1]^d \). Then, as \( n \to \infty \), \( g'_{e,\varphi}(C_{Q,n}) \sim g'_{e,\varphi}(C_Q) \) in \( \ell^\infty(\mathcal{Y}_{e,\varphi}) \), where \( g'_{e,\varphi}(C_Q) \) is a zero-mean Gaussian process with covariance function

\[
\text{Cov} \left\{ (g'_{e,\varphi}(C_Q))(\omega'_{e,u} + \delta(u)\varphi), (g'_{e,\varphi}(C_Q))(\omega'_{e,v} + \delta(v)\varphi) \right\} = \begin{cases} 
E(\omega'_{e,u}(U) + \varphi(V)) \{ \omega_{e,v}(U) + \varphi(V) \}, & \text{if } u, v \in \mathcal{Y} \\
E(\omega'_{e,u}(U) + \varphi(V))\varphi(V), & \text{if } u \in \mathcal{Y}, v \in \mathcal{Y}^c, \quad (A39)
\end{cases}
\]

where \( \mathcal{Y} = (0, 1]^d \setminus \{1\} \). Since each element of \( \mathcal{Y}_{e,\varphi} \) corresponds to a unique \( u \in [0, 1]^d \), we can equivalently think of the codomain of \( g'_{e,\varphi} \) as \( \ell^\infty([0, 1]^d) \) (as done in (A33)) and consider the processes \( g'_{e,\varphi}(C_{Q,n}), g'_{e,\varphi}(C_Q) \) as indexed on \( [0, 1]^d \). From (A29), Condition 3(ii), and the first line of (A37), it follows that

\[
\sqrt{n} \left\{ \ln \hat{A}^*_n(\cdot) - \ln A^*(\cdot) \right\} = (\phi' g_{e,\varphi}(C_{Q,n}))(\cdot) + o_p(1).
\]
The final result is obtained by applying the continuous mapping theorem and the functional delta method to the above expression.

**A.8.3 Main body of the proof**

In what follows, the asymptotic results concerning the ML- and GPWM-based composite-estimators are established by verifying the assumptions of Proposition 6 and
Proposition 5, respectively. First, we address the MD-based case; then, we simultaneously examine P- and CFG-based estimators, due to their similar traits.

We start analyzing the case when \( \alpha \) is estimated with the MD estimator in (17) and \( A_a \) with the MD estimator in (21). We recall that the estimator in (17) is the unique solution of log-likelihood equation \( n^{-1} \sum_{i=1}^{n} \dot{L}_a(\xi_i) = 0 \), where

\[
\dot{L}_a(x) = \partial / \partial \alpha \ln \Phi_a(x) = 1 / \alpha + \ln x^{\alpha - 1}, \quad x > 0,
\]

and \( \Phi_a(x) = \partial \Phi_a(x) / \partial x, \quad x > 0 \). Noting that \( \xi_i^{-1} \) is a Weibull random variable, arguments similar to those in van der Vaart (2000, Theorems 5.41 and 5.42) yield that \( \hat{\alpha}_n^{\text{ML}} \overset{p}{\to} \alpha \) as \( n \to \infty \) and

\[
\sqrt{n} \left( \hat{\alpha}_n^{\text{ML}} - \alpha \right) = i_n^{-1} n^{-1/2} \{ \dot{L}_a(\xi_1) + \ldots + \dot{L}_a(\xi_n) \} + o_p(1),
\]

(A40)

where \( i_n = \alpha^{-2} \{(1 - \zeta^2) + \pi^2 / 6\} \) is the Fisher information and \( \zeta \) is Euler’s constant.

Assuming that Condition 1(i) holds true, then, by Lemma 5, \( \hat{A}_n^{\text{MD}} \) satisfies Condition 2(ii)-(iii) with \( \phi = \phi_{\text{MD}} \) in (A34), and \( \epsilon = 0 \). Furthermore, by (A40), \( \hat{\alpha}_n^{\text{ML}} \) satisfies Condition 3(ii) with \( \zeta = \zeta_{\text{ML}} = i_n^{-1} \dot{L}_a \). Define

\[
\varphi_{\text{ML}}(v) : = \zeta_{\text{ML}} \circ \Phi_a^{-1}(v) = i_n^{-1} \alpha^{-1} \{ 1 + (1 + \ln v) \ln(-\ln v) \}, \quad v \in (0, 1),
\]

(A41)

then (A38) is satisfied with \( \varphi = \varphi_{\text{ML}} \), by Lemma 6. Therefore, from Proposition 6 it follows that, in \( \ell^\infty(S_d) \),

\[
\sqrt{n} \left\{ A_n^{\text{MD,ML}}(\cdot) - A^*(\cdot) \right\} \overset{\mathcal{D}}{\to} (\phi_{\text{MD,ML}}(C_Q)(\cdot)), \quad n \to \infty,
\]

where \( \phi_{\text{MD,ML}} = A_a \phi_{\text{MD}} \circ g'_{0,\varphi_{\text{ML}}} \), with details given in Definition 1(iv), and where \( g'_{0,\varphi_{\text{ML}}}(C_Q) \) is a zero-mean Gaussian process with covariance function

\[
\text{Cov}\{ g'_{0,\varphi_{\text{ML}}}(C_Q)(\mathbf{u}), g'_{0,\varphi_{\text{ML}}}(C_Q)(\mathbf{v}) \} = C_{G_a}(\min(\mathbf{u}, \mathbf{v})) - C_{G_a}(\mathbf{u}) C_{G_a}(\mathbf{v}) + T_a(\mathbf{u}) + T_a(\mathbf{v}) + 1,
\]

for every \( \mathbf{u}, \mathbf{v} \in [0, 1]^d \), with

\[
T_a(\mathbf{u}) = \frac{1}{i_n \alpha} \left( C_{G_a}(\mathbf{u}) - \int_0^1 \frac{\partial}{\partial v} C_{Q_a}(\mathbf{u}, v)(1 + \ln v) \ln(-\ln v) \, dv \right).
\]

Finally, from the weak convergence result and the functional version of Slutsky’s lemma it follows that \( \| A_n^{\text{MD,ML}} - A^* \|_\infty \overset{p}{\to} 0 \) as \( n \to \infty \).

Next, we study the case where \( \alpha \) is estimated with the GPWM estimator in (15) and \( A_a \) with the MD estimator in (21). Here, we additionally assume that \( \alpha > 1/(k - 1) \). By Lemma 8, the estimator \( \hat{\alpha}_n^{\text{GPWM}} \) satisfies Condition 3(i) with \( \tau = \tau_{\text{GPWM}} \) given in (A43). Then, by Proposition 5 it follows that, in \( \ell^\infty(S_d) \),

\[
\sqrt{n} \left\{ A_n^{\text{MD,GPWM}}(\cdot) - A^*(\cdot) \right\} \overset{\mathcal{D}}{\to} (\phi_{\text{MD,GPWM}}(C_Q)(\cdot)), \quad n \to \infty,
\]

where \( \phi_{\text{MD,GPWM}}(f), f \in \ell^\infty([0, 1]^{d+1}) \), is given in Definition 1(v). Furthermore, given the result in Lemma 8 and since \( \| n^{-1/2} \beta_{\mathbb{R}_n} \|_\infty \overset{a.s.}{\to} 0 \), then we have that \( \hat{\alpha}_n^{\text{GPWM}} \overset{a.s.}{\to} \alpha \) as \( n \to \infty \). Consequently, by
Lemma 7,
\[ \left\| \hat{A}_{\text{MD,GPWM}} - A^* \right\|_\infty \xrightarrow{\text{as}} 0, \quad n \to \infty. \]

Now, we study the case where \( \alpha \) is estimated with the ML estimator in (17) and \( A_u \) with the P and CFG estimator in (18) and (20), respectively. Assuming that Conditions 1(i) and 1(ii) hold true, then by Segers (2012, Propositions 3.1 and 4.2) we have
\[ \hat{C}_{G_a,n}(u) = C_{G_a,n}(u) - \sum_{j=1}^{d} \hat{C}_{G_a,n}(1, \ldots, 1, u_j, 1, \ldots, 1) + R_n(u), \]
where \( \hat{C}_{G_a,n} \) is as in (A23) and almost surely \( \|R_n\|_\infty = O(n^{-1/4}(\ln n)^{1/2}(\ln \ln n)^{1/4}) \), \( n \to \infty \). Then, using similar arguments to those in Gudendorf & Segers (2012, pp. 3082–3083) (and the functional delta method for the P estimator) we have that \( \hat{A}_{a,n} \) and \( \hat{A}_{\text{CFG}} \) satisfy Condition 2(i)-(ii) with \( \phi = \phi_p \) and \( \varphi = \phi_{\text{CFG}} \). Precisely, for any fixed \( \epsilon \in (0, 1/2) \)
\[ \sqrt{n} \{ \ln \hat{A}_{a,n}(\cdot) - \ln A_a(\cdot) \} = (\phi_\varphi g_\varphi(C_{G_a,n})(\cdot) + o_p(1), \]
where \( \phi_\varphi \) is given in (A29) and satisfies the representation in (A29). By Lemma 6, the expectation in (A38) is finite for \( \varphi = \varphi_{\text{ML}} \) and any given \( \epsilon \in (0, 1/2) \). Then, by Proposition 6,
\[ \sqrt{n} \left\{ A_{\text{ML}}^* - A^* \right\} \sim (\phi_{\varphi_{\text{ML}}}(C_Q)(\cdot), \quad n \to \infty \]
in \( \ell^\infty(S_d) \), where \( \phi_{\varphi_{\text{ML}}} = A_u \phi_{\varphi_{\text{GPWM}}}^{\prime} g_{\varphi_{\text{GPWM}}}^{\prime} \), with details given in Definition 1(vii), and where \( g^{\prime}_{\varphi_{\text{GPWM}}}(C_Q) \) is a zero-mean Gaussian process with covariance function as in (A39), for \( \varphi = \varphi_{\text{ML}} \).

Finally, from this result and the functional version of Slutsky’s lemma, it follows that \( \|\hat{A}_{\text{MD,GPWM}}^* - A^*\|_\infty \xrightarrow{p} 0 \) and \( \|\hat{A}_{\text{CFG,GPWM}}^* - A^*\|_\infty \xrightarrow{p} 0 \) as \( n \to \infty \).

Concluding, we study the case when \( \alpha \) is estimated with the GPWM estimator in (15) and \( A_u \) with the P and CFG estimators in (18) and (20). Assuming in addition to the previous case that \( \alpha > 1/(k-1) \), then by Proposition 5 we have that in \( \ell^\infty(S_d) \)
\[ \sqrt{n} \left\{ A_{\text{GPWM}}^* - A^* \right\} \sim (\phi_{\varphi_{\text{GPWM}}}(C_Q)(\cdot), \quad n \to \infty, \]
where \( \phi_{\varphi_{\text{GPWM}}}(f), f \in \ell^\infty([0, 1]^{d+1}) \), is given in Definition 1(viii). Ultimately, from this result and the functional version of Slutsky’s lemma, it follows that
\[ \|\hat{A}_{\text{GPWM}}^* - A^*\|_\infty \xrightarrow{p} 0, \quad \|\hat{A}_{\text{CFG,GPWM}}^* - A^*\|_\infty \xrightarrow{p} 0, \quad n \to \infty. \]

A.8.4 Auxiliary results

Lemma 5. Under Condition 1(i) we have
\[ \sqrt{n} \{ \ln \hat{A}_{a,n}^\text{MD} (\cdot) - \ln A_a(\cdot) \} = (\phi_{\text{MD}}(C_{G_a,n})(\cdot) + o_p(1), \]
where \( \phi_{\text{MD}} \) is given in Definition 1(iii).
Lemma 6. Let \((U, V)\) be defined as in (A20). Let \(\omega_{\epsilon,u}'\) be the function defined in (A27) and \(\varphi_{\text{ML}}\) in (A41). Then, for every \(\epsilon \in [0, 1/2)\) and \(u \in (0, 1]^d \setminus \{1\}\), we have \(E(\omega_{\epsilon,u}'(U)\varphi_{\text{ML}}(V)) \in \mathbb{R}\), that is, the expectation is finite.

Lemma 7. Let \(\hat{\alpha}_n\) be an estimator of \(\alpha\) satisfying \(\hat{\alpha}_n \xrightarrow{as} \alpha\) as \(n \to \infty\) and \(\hat{A}_{\alpha,n}^{\text{MD}}\) be the estimator of \(A_n\) given in (21). Let

\[
\hat{A}^{*}_{\alpha,n}(t) = \left(\hat{A}_{\alpha,n}^{\text{MD}}(t)/\|t\|_{1/\hat{\alpha}_n}\right)^{1/\hat{\alpha}_n}, \quad t \in S_d.
\]

Then,

\[
\left\|\hat{A}^{*}_{\alpha,n} - A^{*}\right\|_{\infty} \xrightarrow{as} 0, \quad n \to \infty.
\]

(A42)

Lemma 8. Assume that \(\alpha > 1/(k - 1)\). Then, almost surely as \(n \to \infty\)

\[
\sqrt{n}(\hat{\alpha}_{n}^{\text{GPWM}} - \alpha) = \tau_{\text{GPWM}}(B_n) + o(1),
\]

where \(\tau_{\text{GPWM}} : \mathcal{C}^\infty([0, 1]) \to \mathbb{R}\) is defined as

\[
\tau_{\text{GPWM}}(f) = -2 \int_0^1 f(v) \frac{v(-\ln v)^k\{\mu_{1,k-1} - \mu_{1,k}/(-\ln v)\}}{\Phi_a(\Phi_a^-(v))(k\mu_{1,k-1} - 2\mu_{1,k})^2} dv,
\]

\[
\mu_{a,b} = \int_0^1 \Phi_a^-(v)v^a(-\ln v)^b dv, \quad a, b \in \mathbb{N}
\]

(A43)

and \(\Phi_a(v) = \partial / \partial v \Phi_a(v), v \in (0, 1)\).

For the proofs see Section 4 of the supplementary material.