

ASYMPTOTIC DOMINATION OF SAMPLE MAXIMA

ENKELEJD HASHORVA AND DIDIER RULLIÈRE

Abstract: For a given random sample from some underlying multivariate distribution F we consider the domination of the component-wise maxima by some independent random vector \mathbf{W} with distribution function G . We show that the probability that certain components of the sample maxima are dominated by the corresponding components of \mathbf{W} can be approximated under the assumptions that both F and G are in the max-domain of attraction of some max-stable distribution functions. We study further some basic probabilistic properties of the dominated components of sample maxima by \mathbf{W} .

Key Words: Max-stable distributions; records; domination of sample maxima; extremal dependence; de Haan representation; infargmax formula;

AMS Classification: Primary 60G15; secondary 60G70

1. INTRODUCTION

Let $\mathbf{Z}_i, i \leq n$ be independent d -dimensional random vectors with common continuous distribution function (df) F and denote by \mathbf{M}_n their component-wise maxima, i.e., $M_{nj} = \max_{1 \leq k \leq n} Z_{kj}, j \leq d$. If \mathbf{W} is another d -dimensional random vector with continuous df G being further independent of \mathbf{M}_n the approximation of the probability that at least one component of \mathbf{W} dominates the corresponding component of \mathbf{M}_n is of interest since it is related to the dependence of the components of \mathbf{M}_n , see e.g., [1]. In the special case that \mathbf{W} has a max-stable df with unit Fréchet marginal df's $\Phi(x) = e^{-1/x}, x > 0$ and \mathbf{M}_n has almost surely positive components, we simply have

$$\mathbb{P}\{\exists i \leq d : W_i > M_{ni}\} = 1 - \mathbb{P}\{\forall i, 1 \leq i \leq d : M_{ni} \geq W_i\} = 1 - \mathbb{E}_{\mathbf{M}_n} \left\{ \exp \left(-\mathbb{E}_{\mathbf{W}} \left\{ \max_{1 \leq i \leq d} \frac{W_i}{M_{ni}} \right\} \right) \right\},$$

where $\mathbf{W} = (W_1, \dots, W_d)$ being independent of \mathbf{M}_n is a spectral random vector of G which exists in view of the well-known de Haan representation, see e.g., [2] and (2.1) below. Note that the assumption that W_i has unit Fréchet df implies that $\mathbb{E}\{W_i\} = 1$.

The above probability is referred to as the marginal domination probability of the sample maxima. If F is also a max-stable df with unit Fréchet marginals, then by definition \mathbf{M}_n/n has for any $n > 0$ df F and consequently

$$(1.1) \quad n[1 - \mathbb{P}\{\forall i, 1 \leq i \leq d : M_{ni} \geq W_i\}] = n \left[1 - \mathbb{E}_{\mathbf{Z}} \left\{ \exp \left(-\frac{1}{n} \mathbb{E}_{\mathbf{W}} \left\{ \max_{1 \leq i \leq d} \frac{W_i}{Z_i} \right\} \right) \right\} \right] \sim \mathbb{E} \left\{ \max_{1 \leq i \leq d} \frac{W_i}{Z_i} \right\},$$

where \sim means asymptotic equivalence as $n \rightarrow \infty$ and $\mathcal{Z} = (\mathcal{Z}_1, \dots, \mathcal{Z}_d)$ has df F being further independent of \mathbf{W} . Under the above assumptions, we have

$$(1.2) \quad p_{n,T}(F, G) = \mathbb{P}\{\forall i, 1 \leq i \leq d : W_i > M_{ni}\} \sim \frac{1}{n} \mathbb{E} \left\{ \min_{1 \leq i \leq d} \frac{W_i}{\mathcal{Z}_i} \right\}, \quad T = \{1, \dots, d\}$$

as $n \rightarrow \infty$, which follows by (1.1) and the inclusion-exclusion formula or directly by [1][Thm 2.5 and Prop 4.2].

Here $p_{n,T}(F, G)$ is referred to as the probability of the complete domination of sample maxima by \mathbf{W} . In the particular case that $F = G$ it is related to the probability of observing a multiple maxima or concurrence probability, see [3-9].

Between these two extreme cases, of interest is also to consider the partial domination of the sample maxima. Let therefore below $T \subset \{1, \dots, d\}$ be non-empty and consider the probability that only the components of \mathbf{W} with indices in T dominate \mathbf{M}_n , i.e.,

$$\mathbb{P}\{\forall i \in T : W_i > M_{ni}, \forall i \in \bar{T} : W_i \leq M_{ni}\} =: p_{n,T}(F, G),$$

where $\bar{T} = \{1, \dots, d\} \setminus T$. Note that $p_{n,T}(F, F)$ relates to the probability of observing a T -record, see [10]. By the continuity of F and G we simply have

$$p_{n,T}(F, G) = \int_{\mathbb{R}^d} \mathbb{P}\{\forall i \in T : W_i > y_i, \forall i \in \bar{T} : W_i \leq y_i\} dF^n(\mathbf{y}),$$

which cannot be evaluated without knowledge of both F and G . In the particular case that F and G are max-stable df's as above, using (1.1) and the inclusion-exclusion formula we obtain

$$(1.3) \quad \lim_{n \rightarrow +\infty} np_{n,T}(F, G) = \mathbb{E} \left\{ \left(\min_{i \in T} \frac{W_i}{\mathcal{Z}_i} - \max_{i \in \bar{T}} \frac{W_i}{\mathcal{Z}_i} \right)_+ \right\}.$$

When $F = G$ the above result is known from [10][Prop 2.2]. Moreover, in the special case that T consists of one element, then the right-hand side of (1.3) is equal to $\mathbb{P}\{C(T) \subset \bar{T}\}$, where $C(T)$ is the tessellation as determined in [11]. If we are not interested on a particular index set T , where the domination of sample maxima by \mathbf{W} occurs but simply on the number of components being dominated, i.e., on the random variable (rv)

$$N_n = \sum_{i=1}^d \mathbf{1}_{\{W_i > M_{ni}\}}$$

a question of interest is if N_n can be approximated as $n \rightarrow \infty$. We have that N_n has the same distribution as

$$\sum_{i=1}^d \mathbf{1}_{\{W_i/n > \mathcal{Z}_i\}},$$

provided that F is max-stable as above and \mathcal{Z} has df F being further independent of \mathbf{W} . Hence if W_i 's are unit Fréchet rv's, then

$$\lim_{n \rightarrow +\infty} n \mathbb{E}\{N_n\} = \sum_{i=1}^d \lim_{n \rightarrow +\infty} n \mathbb{P}\{W_i > n \mathcal{Z}_i\} = \sum_{i=1}^d \lim_{n \rightarrow +\infty} n \left[1 - e^{-\mathbb{E}\left\{\frac{1}{n \mathcal{Z}_i}\right\}} \right] = d.$$

Consequently, the expected number of components of sample maxima being dominated by the components of \mathbf{W} decreases as d/n when n goes to infinity. Moreover, the dependence of both \mathbf{W} and \mathbf{M}_n does not play any role. This is however in general not the case for the expectation of $f(N_n)$, where f is some real-valued function, since the dependence of both \mathbf{M}_n and \mathbf{W} influence the approximation as we shall show in the next section.

From our discussion above the assumptions that F and G are max-stable df's with unit Fréchet marginals lead to tractable asymptotic formulas for various quantities related to the domination of sample maxima \mathbf{M}_n by \mathbf{W} .

In view of [1] we know that both (1.1) and (1.2) are valid in the more general setup that both F and G are in the max-domain of attraction of some max-stable df's (see next section for details). We shall show in this paper that the same assumptions lead to tractable approximations of both $p_{n,T}(F,G)$ and $\mathbb{E}\{f(N_n)\}$ as $n \rightarrow \infty$.

Brief organisation of the paper: Section 2 presents the main results concerning the approximations of the marginal domination probabilities and the expectation of $f(N_n)$. Section 3 is dedicated to properties of \mathbf{W}/\mathbf{Z} which we call the domination spectral vector. All the proofs are relegated to Section 4.

2. MAIN RESULTS

We shall recall first some basic properties of max-stable df's, see [2, 12–14] for details. A d -dimensional df \mathcal{G} is max-stable with unit Fréchet marginals if

$$\mathcal{G}^t(tx_1, \dots, tx_d) = \mathcal{G}(x_1, \dots, x_d)$$

for any $t > 0, x_i \in (0, \infty), 1 \leq i \leq d$. In the light of De Haan representation

$$(2.1) \quad \mathcal{G}(\mathbf{x}) = \exp\left(-\mathbb{E}\left\{\max_{1 \leq j \leq d} \mathcal{W}_j/x_j\right\}\right), \quad \mathbf{x} = (x_1, \dots, x_d) \in (0, \infty)^d,$$

where \mathcal{W}_j 's are non-negative rv's with $\mathbb{E}\{\mathcal{W}_j\} = 1, j \leq d$ and $\mathbf{W} = (\mathcal{W}_1, \dots, \mathcal{W}_d)$ is a spectral vector for \mathcal{G} (which is not unique).

In view of multivariate extreme value theory, see e.g., [14] d -dimensional max-stable df's \mathcal{F} are limiting df's of the component-wise maxima of d -dimensional iid random vectors with some df F . In that case, F is said to be in the max-domain of attraction (MDA) of \mathcal{F} , abbreviated $F \in MDA(\mathcal{F})$. For simplicity we shall assume throughout in the following that F has marginal df's F_i 's such that

$$(2.2) \quad \lim_{n \rightarrow +\infty} F_i^n(nx) = \Phi(x), \quad x \in \mathbb{R}$$

for all $i \leq d$, where we set $\Phi(x) = 0$ if $x \leq 0$. We have thus that $F \in MDA(\mathcal{F})$ if further

$$(2.3) \quad \lim_{n \rightarrow +\infty} \sup_{x_i \in \mathbb{R}, 1 \leq i \leq d} \left| F^n(nx_1, \dots, nx_d) - \mathcal{F}(x_1, \dots, x_d) \right| = 0.$$

In the following \mathcal{F} is a d -dimensional max-stable df of some random vector \mathbf{Z} with unit Fréchet marginals and \mathcal{G} is another max-stable df with unit Fréchet marginals and spectral random vector \mathbf{W} independent of \mathbf{Z} .

Below we extend [15][Prop 1] which considers the case $F = G$.

Proposition 2.1. *If F and G have continuous marginal distributions satisfying (2.2) and $F \in MDA(\mathcal{F}), G \in MDA(\mathcal{G})$, then for any non-empty $T \subset \{1, \dots, d\}$ we have*

$$(2.4) \quad \lim_{n \rightarrow +\infty} np_{n,T}(F,G) = \mathbb{E}\left\{\left(\min_{i \in T} \mathcal{W}_i/\mathcal{Z}_i - \max_{i \in \bar{T}} \mathcal{W}_i/\mathcal{Z}_i\right)_+\right\} =: \lambda_T(\mathcal{F}, \mathcal{G}).$$

Remark 2.2. *Define for a non-empty index set T the rv $K_n = \sum_{j=1}^n \mathbb{1}_{\{\forall i \in T: W_i > M_{ji}, \forall i \in \bar{T}: W_i \leq M_{ji}\}}$. Under the assumptions of Proposition 2.1 we have (see also [16][Corr 3.2]) that*

$$(2.5) \quad \lim_{n \rightarrow +\infty} \frac{\mathbb{E}\{K_n\}}{\ln n} = \lambda_T(\mathcal{F}, \mathcal{G}).$$

Example 2.3 (\mathcal{F} comonotonic and \mathcal{G} a product df). Suppose that \mathcal{F} is comonotonic, i.e., $Z_1 = \dots = Z_d$ almost surely and let \mathcal{G} be a product df with unit Fréchet marginals df's and let N be rv on $\{1, \dots, d\}$ with $\mathbb{P}\{N = i\} = 1/d, i \leq d$. A spectral vector \mathcal{W} for \mathcal{G} can be defined as follows

$$(\mathcal{W}_1, \dots, \mathcal{W}_d) = (d\mathbb{1}_{\{N=1\}}, \dots, d\mathbb{1}_{\{N=d\}}).$$

Indeed $\mathbb{E}\{\mathcal{W}_k\} = d\mathbb{P}\{N = k\} = 1$ for any $k \leq d$ and

$$\mathbb{E}\left\{\max_{1 \leq i \leq d} \mathcal{W}_i / x_i\right\} = \sum_{k=1}^d \mathbb{E}\left\{\max_{1 \leq i \leq d} \mathcal{W}_i / x_i \mathbb{1}_{\{N=k\}}\right\} = \sum_{k=1}^d \mathbb{E}\{\mathcal{W}_k / x_k \mathbb{1}_{\{N=k\}}\} = d \sum_{k=1}^d \mathbb{E}\{\mathbb{1}_{\{N=k\}} / x_k\} = \sum_{k=1}^d 1/x_k$$

for any x_1, \dots, x_d positive. In particular, for a non-empty index set $K \subset \{1, \dots, d\}$ with m elements we have

$$\mathbb{E}\left\{\max_{i \in K} \mathcal{W}_i\right\} = d \sum_{k \in K} \mathbb{E}\{\mathbb{1}_{\{N=k\}}\} = m.$$

Consequently, using further that (see the proof of Proposition 2.1)

$$\lambda_T(\mathcal{F}, \mathcal{G}) = \sum_{j=0}^k (-1)^{j+1} \sum_{J \subset T: |J|=j} \mathbb{E}\left\{\max_{i \in J \cup \bar{T}} \frac{\mathcal{W}_i}{Z_i}\right\}$$

we obtain

$$\lambda_T(\mathcal{F}, \mathcal{G}) = \sum_{j=0}^k (-1)^{j+1} \sum_{J \subset T: |J|=j} \mathbb{E}\left\{\max_{i \in J \cup \bar{T}} \mathcal{W}_i\right\} = \sum_{j=0}^k (-1)^{j+1} \sum_{J \subset T: |J|=j} (j + d - k).$$

If $k = d$, then from above

$$(2.6) \quad \lambda_T(\mathcal{F}, \mathcal{G}) = \sum_{j=0}^d (-1)^{j+1} \sum_{J \subset T: |J|=j} j = d(1-1)^{d-1} = 0.$$

A direct probabilistic proof of (2.6) follows by the properties of \mathcal{W} , namely when $k = d \geq 2$

$$\lambda_T(\mathcal{F}, \mathcal{G}) = \mathbb{E}\left\{\min_{1 \leq i \leq d} \mathcal{W}_i / Z_i\right\} = \mathbb{E}\left\{\min_{1 \leq i \leq d} \mathcal{W}_i\right\} = d\mathbb{E}\left\{\min_{1 \leq i \leq d} \mathbb{1}_{\{N=i\}}\right\} = 0.$$

Now, let us investigate the number N_n of dominations defined as in Introduction by $\sum_{i=1}^d \mathbb{1}_{\{W_i/n > Z_i\}}$.

For a given function $f : \{0, \dots, d\} \rightarrow \mathbb{R}$ we shall be concerned with the behaviour of

$$\mathbb{E}\{f(N_n)\} = \sum_{k=0}^d f(k) \mathbb{P}\{N_n = k\}$$

when n tends to $+\infty$. Throughout in the sequel we set

$$\mathcal{D} = \{1, \dots, d\}.$$

In Proposition 2.4 below, we first express this expectation as a function of minima or maxima of \mathcal{W}_i / Z_i 's.

Proposition 2.4. *If F and G are as in Proposition 2.1, then we have*

$$(2.7) \quad \lim_{n \rightarrow +\infty} n\mathbb{E}\{f(N_n)\} - nf(0) = \sum_{k=1}^d \Delta^k f(0) \sum_{K \subset \mathcal{D}, |K|=k} \mathbb{E}\left\{\min_{i \in K} \mathcal{W}_i / Z_i\right\}$$

or alternatively

$$(2.8) \quad \lim_{n \rightarrow +\infty} n\mathbb{E}\{f(N_n)\} - nf(0) = \sum_{k=1}^d (-1)^{k+1} \Delta^k f(d-k) \sum_{K \subset \mathcal{D}, |K|=k} \mathbb{E}\left\{\max_{i \in K} \mathcal{W}_i / Z_i\right\},$$

where Δ is the difference operator, $\Delta f(x) = f(x+1) - f(x)$.

Proposition 2.5. *If F and G are as in Proposition 2.1, then we have*

$$(2.9) \quad \lim_{n \rightarrow +\infty} n\mathbb{E}\{f(N_n)\} - nf(0) = \sum_{k=1}^d g(k)\mathbb{E}\left\{(\mathcal{W}/\mathcal{Z})_{(k)}\right\} = \sum_{k=0}^d f(k) \left[(\mathcal{W}/\mathcal{Z})_{(d-k+1)} - (\mathcal{W}/\mathcal{Z})_{(d-k)} \right],$$

where $(\mathcal{W}/\mathcal{Z})_{(1)} \leq \dots \leq (\mathcal{W}/\mathcal{Z})_{(d)}$ are the order statistics of $\mathcal{W}_i/\mathcal{Z}_i, i \leq d$ and $g(k) = f(d-k+1) - f(d-k)$, with the convention $(\mathcal{W}/\mathcal{Z})_{(0)} = (\mathcal{W}/\mathcal{Z})_{(d+1)} = 0$.

Remark 2.6 (retrieving simple cases). *For particular cases of f we have:*

- From Proposition 2.4, setting $f(x) = \mathbb{1}_{\{x=d\}}$, one can check that $\Delta^k f(0) = 0$ when $k < d$ and $\Delta^d f(0) = 1$, so that Equation (2.7) implies (1.2). Alternatively, by Proposition 2.5 since $g(1) = f(d) - f(d-1) = 1$ and $g(k) = f(d-k+1) - f(d-k) = 0 - 0 = 0$ if $k > 1$ we have that $\lim_{n \rightarrow +\infty} n\mathbb{E}\{f(N_n)\} - nf(0) = \mathbb{E}\left\{(\mathcal{W}/\mathcal{Z})_{(1)}\right\}$.
- In view of Proposition 2.4, setting $f(x) = \mathbb{1}_{\{x \geq 1\}}$, $\Delta^k f(d-k) = \sum_{i=0}^k \binom{k}{i} (-1)^{k-i} f(d-k+i)$. Thus $\Delta^k f(d-k) = 0$ if $k < d$. If $k = d$, then

$$\Delta^d f(d-d) = \Delta^d f(0) = (1-1)^d - (-1)^d = (-1)^{d+1}$$

and Equation (2.8) implies (1.1). Alternatively, by Proposition 2.5 since if $k < d$, $g(k) = f(d-k+1) - f(d-k) = 1 - 1 = 0$ and $g(d) = f(1) - f(0) = 1$ we obtain $\lim_{n \rightarrow +\infty} n\mathbb{E}\{f(N_n)\} - nf(0) = \mathbb{E}\left\{(\mathcal{W}/\mathcal{Z})_{(d)}\right\}$.

- By Proposition 2.5, setting $f(x) = x$, we easily retrieve $\lim_{n \rightarrow +\infty} n\mathbb{E}\{N_n\} = \sum_{k=1}^d \mathbb{E}\left\{(\mathcal{W}/\mathcal{Z})_{(k)}\right\} = d$, as seen previously.

Remark 2.7 (Interpretation of $(\mathcal{W}/\mathcal{Z})_{(j)}$). *Let $f(k) = \mathbb{1}_{\{k \geq d-j+1\}}$, for any $j, k \in \mathcal{D}$. Then $g(k) = f(d-k+1) - f(d-k) = \mathbb{1}_{\{k=j\}}$. In this case, $f(0) = 0$ and $\mathbb{E}\{f(N_n)\} = \mathbb{P}\{N \geq d-j+1\}$, thus*

$$\mathbb{E}\left\{(\mathcal{W}/\mathcal{Z})_{(j)}\right\} = \lim_{n \rightarrow +\infty} n\mathbb{P}\{N_n \geq d-j+1\}.$$

3. DOMINATION SPECTRUM

In the previous results, we have considered a particular setting, and we have expressed the domination probability and some expectations relying on number of dominations (see Section 2). We have seen that all these results were expressed as a function of

$$\mathcal{W}/\mathcal{Z} = \left(\frac{\mathcal{W}_i}{\mathcal{Z}_i} \right)_{i \in \mathcal{D}}.$$

By the definition $\mathcal{W}_i/\mathcal{Z}_i$'s are nonnegative, and are such that, by independence, $\mathbb{E}\{\mathcal{W}_i/\mathcal{Z}_i\} = \mathbb{E}\{\mathcal{W}_i\} \mathbb{E}\left\{\frac{1}{\mathcal{Z}_i}\right\} = 1$. Thus in view of the De Haan representation \mathcal{W}/\mathcal{Z} can be viewed as the spectral random vector of some max-stable d -dimensional distribution. Since \mathcal{W}/\mathcal{Z} is related to the domination of M_n by \mathcal{W} , we will refer to it by the term *domination spectrum*. In this section we shall explore some basic properties of the domination spectrum.

Next, assume that \mathcal{W} has a copula $C_{\mathcal{W}}$ and suppose further that \mathcal{Z} has a copula $C_{\mathcal{Z}}$. Note in passing that the latter copula is unique since the marginals of \mathcal{Z} have continuous df.

We shall first study the link between the diagonal sections of both copulas $C_{\mathcal{W}}$ and $C_{\mathcal{Z}}$, defined for all $u \in [0, 1]$ by

$$\delta_{\mathcal{W}}(u) = C_{\mathcal{W}}(u, \dots, u) \quad \text{and} \quad \delta_{\mathcal{Z}}(u) = C_{\mathcal{Z}}(u, \dots, u).$$

We recall that the diagonal section characterizes uniquely many Archimedean copulas (under a condition that is called Frank's condition, see e.g., [17]), some non-parametric estimators of the generator of an Archimedean copulas directly rely on this diagonal section. We consider here the case where the df of \mathcal{Z} has spectral random vector \mathcal{W} . Notice that the upper tail dependence coefficients can be deduced from the regular variation properties of $\delta_{\mathcal{Z}}$ and $\delta_{\mathcal{W}}$, which is straightforward for $\delta_{\mathcal{Z}}$ in the following result.

Proposition 3.1. *Consider a d -dimensional random vector \mathcal{Z} having max-stable df with Fréchet unit marginals and with copula $C_{\mathcal{Z}}$. If the random vector \mathcal{Z} has df $H(\mathbf{y}) = \exp(-\mathbb{E}\{\max_{1 \leq j \leq d} \frac{\mathcal{W}_j}{y_j}\})$, where all \mathcal{W}_j are nonnegative rv's with mean 1, then*

$$\delta_{\mathcal{Z}}(u) = u^{r_{\mathcal{W}}} \quad \text{with} \quad r_{\mathcal{W}} = \mathbb{E} \left\{ \max_{j \in \mathcal{D}} \mathcal{W}_j \right\}.$$

In particular, when $r_{\mathcal{W}} > 1$, this diagonal section $\delta_{\mathcal{Z}}(u)$ is the one of a Gumbel copula with parameter

$$(3.1) \quad \theta = \frac{\ln d}{\ln r_{\mathcal{W}}}.$$

Furthermore, if the components of \mathcal{W} are identically distributed and if $F_{\mathcal{W}_1}$ is invertible, then we have

$$r_{\mathcal{W}} = \int_0^1 F_{\mathcal{W}_1}^{-1}(s) d\delta_{\mathcal{W}}(s).$$

Example 3.2 (From independence to comonotonicity). *Let $\mathcal{W}_j = B d \mathbf{1}_{\{I=j\}} + (1-B)\delta_1$, for all $j \in \mathcal{D}$, where I is a uniformly distributed rv's on \mathcal{D} , B is a Bernoulli rv with $\mathbb{E}\{B\} = \alpha \in (0, 1]$ and δ_1 is a Dirac mass at 1, all these rv's being mutually independent. In this case, $r_{\mathcal{W}} = \mathbb{E} \left\{ \max_{j \in \mathcal{D}} \mathcal{W}_j \right\}$ in Proposition 3.1 becomes $r_{\mathcal{W}} = \alpha d + 1 - \alpha$. As a consequence, $\delta_{\mathcal{Z}}$ is the diagonal of a Gumbel copula which goes from the independence ($\alpha = 1$) to the comonotonicity ($\alpha \rightarrow 0$), with parameter*

$$\theta = \frac{\ln d}{\ln(1 + \alpha(d-1))}.$$

Furthermore, we have when all $t_j > 0$,

$$\mathbb{E} \left\{ \max_{j \in K} \frac{\mathcal{W}_j}{t_j} \right\} = \alpha \sum_{j \in K} \frac{1}{t_j} + (1-\alpha) \frac{1}{\min_{j \in K} t_j}.$$

Let $t > 0$ and suppose that K has cardinal $|K| > 1$. By conditioning over B , we get

$$\mathbb{P} \{ \forall i \in K, \mathcal{W}_i / \mathcal{Z}_i > t \} = (1-\alpha) \mathbb{P} \{ \forall i \in K, \mathcal{Z}_i < 1/t \mid B = 0 \}$$

since $\mathbb{P} \{ \forall i \in K, \mathcal{W}_i / \mathcal{Z}_i > t \mid B = 1 \} = 0$ when $|K| > 1$, because in this case at least one component \mathcal{W}_i , $i \in K$, is zero when $B = 1$. Recall that \mathcal{Z} is independent from \mathcal{W} and B , thus for $t > 0$ and $|K| > 1$

$$\mathbb{P} \left\{ \min_{i \in K} \mathcal{W}_i / \mathcal{Z}_i > t \right\} = (1-\alpha) \exp \left(\mathbb{E} \left\{ -\max_{j \in K} \frac{\mathcal{W}_j}{(1/t)} \right\} \right) = (1-\alpha) \exp(-t(1 + \alpha|K| - \alpha)).$$

When $|K| = 1$, we show similarly that $\mathbb{P} \{ \min_{i \in K} \mathcal{W}_i / \mathcal{Z}_i > t \} = (1-\alpha) \exp(-t) + \alpha \frac{1}{d} \exp(-\frac{t}{d})$. In both cases $|K| = 1$ and $|K| > 1$, the survival function $\mathbb{P} \{ \min_{i \in K} \mathcal{W}_i / \mathcal{Z}_i > t \}$ is a linear combination of exponential functions, and thus can be shown to be a discrete mixture of exponential distributions:

$$\begin{cases} \min_{i \in K} \mathcal{W}_i / \mathcal{Z}_i & \stackrel{d}{=} (1-B) \epsilon_{1+\alpha(|K|-1)} + \mathbf{1}_{\{|K|=1\}} B \mathbf{1}_{\{I=1\}} \epsilon_{1/d} \\ \mathbb{E} \left\{ \min_{i \in K} \mathcal{W}_i / \mathcal{Z}_i \right\} & = \frac{1-\alpha}{1+\alpha(|K|-1)} + \mathbf{1}_{\{|K|=1\}} \alpha, \end{cases}$$

where B is a Bernoulli r.v. of parameter α , $\epsilon_{1+\alpha(|K|-1)}$ and $\epsilon_{1/d}$ are exponentially distributed r.v. with respective parameters $1 + \alpha(|K| - 1)$ and $1/d$, I an uniformly distributed r.v. over \mathcal{D} , all being mutually independent (for simplicity, we denote $\mathbb{1}_{\{|K|=1\}}$ the variable whose value is 1 if $|K| = 1$ or 0 otherwise). Then all results about the limit law of N_n follow immediately, using Equation (2.7) in Proposition 2.4. Notice that one could also determine $r(\mathcal{W}/\mathcal{Z})$ from this, and by application of Proposition 3.1, assess the dependence structure of the random vector whose spectrum is $(\mathcal{W}/\mathcal{Z})$.

4. PROOFS

We first give hereafter some combinatorial results that show how quantities depending on a number of events can be related to quantities involving only intersections or unions of those events. This generalizes inclusion-exclusion formulas that will correspond to very specific functions f and g .

Lemma 4.1 (Inclusion-exclusion relations). *Let $\mathcal{D} = \{1, \dots, d\}$ and let B_i , $i \in \mathcal{D}$ be events. Consider the number of realized events $N = \sum_{i \in \mathcal{D}} \mathbb{1}_{\{B_i\}}$. Then for any function $f : \{0, \dots, d\} \rightarrow \mathbb{R}$*

$$(4.1) \quad \sum_{k=0}^d f(k) \mathbb{P}\{N = k\} = f(0) + \sum_{j=1}^d S_j \Delta^j f(0) = f(0) + \sum_{j=1}^d \bar{S}_j (-1)^{j+1} \Delta^j f(d-j)$$

and similarly for any function $g : \mathcal{D} \rightarrow \mathbb{R}$

$$(4.2) \quad \sum_{k=0}^d g(k) \mathbb{P}\{N \geq k\} = \sum_{j=1}^d S_j \Delta^{j-1} g(1) = \sum_{j=1}^d \bar{S}_j (-1)^{j+1} \Delta^{j-1} g(d-j+1),$$

where $S_j = \sum_{J \subset \mathcal{D}, |J|=j} \mathbb{P}\left\{\bigcap_{i \in J} B_i\right\}$ and $\bar{S}_j = \sum_{J \subset \mathcal{D}, |J|=j} \mathbb{P}\left\{\bigcup_{i \in J} B_i\right\}$.

Proof of Lemma 4.1. The first equality in Equation (4.1) is known in actuarial sciences under the name of Schuette-Nesbitt formula, see [18, section 8.5]. This formula does not require any independence assumption, it is a simple development of $f(N) = (I + \mathbb{1}_{\{B_1\}} \Delta) \cdots (I + \mathbb{1}_{\{B_d\}} \Delta) f(0)$ where I and Δ are the identity and the difference operators respectively. To prove the second equality in Equation (4.1), let us denote $p_J = \mathbb{P}\{\bigcap_{i \in J} B_i\}$ and $\bar{p}_J = \mathbb{P}\{\bigcup_{i \in J} B_i\}$. By inclusion-exclusion principle, we get

$$(4.3) \quad S_k = \sum_{K \subset \mathcal{D}, |K|=k} \sum_{j=1}^k (-1)^{j+1} \sum_{J \subset K, |J|=j} \bar{p}_J = \sum_{j=1}^k (-1)^{j+1} \binom{d-j}{k-j} \bar{S}_j.$$

Now using Equation (4.3),

$$\sum_{k=1}^d \Delta^k f(0) S_k = \sum_{k=1}^d \Delta^k f(0) \sum_{j=1}^k (-1)^{j+1} \binom{d-j}{k-j} \bar{S}_j = \sum_{j=1}^d \bar{S}_j (-1)^{j+1} \Delta^j (I + \Delta)^{d-j} f(0),$$

and since $(I + \Delta)^{d-j} f(0) = f(d-j)$, the second equality in Equation (4.1) holds. Similarly, the first equality in Equation (4.2) is a known Schuette-Nesbitt formula, see [18, Section 8.5], and one can retrieve the second equality by using Equation (4.3). Alternatively, one can also deduce (4.2) from (4.1) by setting $f(0) = 0$ and $g(k) = \Delta f(k-1)$ for all $k \in \mathcal{D}$. The formulas in Lemma 4.1 generalize a very old formula of Waring which give $\mathbb{P}\{N = k\}$, $k \in \mathcal{D}$. They also generalize the classical inclusion exclusion formula which can be retrieved by setting in (4.1) $f(k) = 1$ if $k \geq 1$, and $f(k) = 0$ otherwise. \square

Proof of Proposition 2.1. By inclusion-exclusion formula for a given index set $T \subset \{1, \dots, d\}$ with $k = |T|$ elements we have

$$\begin{aligned} \mathbb{P}\{\forall i \in \bar{T} : W_i \leq y_i, \exists i \in T : W_i \leq y_i\} &= \sum_{j=1}^k (-1)^{j+1} \sum_{J \subset T: |J|=j} \mathbb{P}\{\forall i \in (J \cup \bar{T}) : W_i \leq y_i\} \\ &= \sum_{j=1}^k (-1)^{j+1} \sum_{J \subset T: |J|=j} G_{J \cup \bar{T}}(\mathbf{y}), \end{aligned}$$

where $G_L(\mathbf{y}) = \mathbb{P}\{\forall i \in L : W_i \leq y_i\}$ is the L -th marginal df of G . In particular, letting $W_i \rightarrow \infty, i \leq d$ yields

$$1 = \sum_{j=1}^k (-1)^{j+1} \sum_{J \subset T: |J|=j} 1.$$

Consequently, for all $n > 1$

$$\begin{aligned} p_{n,T}(F, G) &= \int_{\mathbb{R}^d} \mathbb{P}\{\forall i \in T : W_i \geq y_i, \forall i \in \bar{T} : W_i < y_i\} dF^n(\mathbf{y}) \\ &= \int_{\mathbb{R}^d} \mathbb{P}\{\forall i \in \bar{T} : W_i \leq y_i\} dF^n(\mathbf{y}) - \int_{\mathbb{R}^d} \mathbb{P}\{\forall i \in \bar{T} : W_i \leq y_i, \exists i \in T : W_i \leq y_i\} dF^n(\mathbf{y}) \\ &= 1 - \int_{\mathbb{R}^d} \sum_{j=1}^k (-1)^{j+1} \sum_{J \subset T: |J|=j} G_{J \cup \bar{T}}(\mathbf{y}) dF^n(\mathbf{y}) - \left(1 - \int_{\mathbb{R}^d} G_{\bar{T}}(\mathbf{y}) dF^n(\mathbf{y})\right) \\ &= \sum_{j=0}^k (-1)^{j+1} \sum_{J \subset T: |J|=j} \int_{\mathbb{R}^d} [1 - G_{J \cup \bar{T}}(\mathbf{y})] dF^n(\mathbf{y}) \\ &= \sum_{j=0}^k (-1)^{j+1} \sum_{J \subset T: |J|=j} \int_{\mathbb{R}^{m+i}} [1 - G_{J \cup \bar{T}}(\mathbf{y})] dF_{J \cup \bar{T}}^n(\mathbf{y}). \end{aligned}$$

In view of [1][Prop 4.2] we obtain

$$\lim_{n \rightarrow +\infty} n \int_{\mathbb{R}^{m+|J|}} [1 - G_{J \cup \bar{T}}(\mathbf{y})] dF_{J \cup \bar{T}}^n(\mathbf{y}) = - \int_{\mathbb{R}^{m+|J|}} \ln Q_{J \cup \bar{T}}(\mathbf{y}) dH_{J \cup \bar{T}}(\mathbf{y}).$$

Further by [1][Thm 2.5 and Prop 4.2]

$$- \int_{\mathbb{R}^{m+|J|}} \ln Q_{J \cup \bar{T}}(\mathbf{y}) dH_{J \cup \bar{T}}(\mathbf{y}) = \mathbb{E} \left\{ \max_{i \in J \cup \bar{T}} \frac{W_i}{Z_i} \right\}.$$

Consequently, we have

$$\lim_{n \rightarrow +\infty} n p_{n,T}(F, G) = \sum_{j=0}^k (-1)^{j+1} \sum_{J \subset T: |J|=j} \mathbb{E} \left\{ \max_{i \in J \cup \bar{T}} \frac{W_i}{Z_i} \right\}.$$

In the light of [10][Lem 1] for given constants c_1, \dots, c_d

$$\sum_{j=0}^k (-1)^{j+1} \sum_{J \subset T: |J|=j} \max_{i \in J \cup \bar{T}} c_i = \max \left(\max_{i \in \bar{T}} c_i, \min_{i \in T} c_i \right) - \max_{i \in \bar{T}} c_i = \left(\min_{i \in T} c_i - \max_{i \in \bar{T}} c_i \right)_+$$

implying the claim.

Alternatively, we have using again inclusion-exclusion formula

$$\begin{aligned} p_{n,T}(F, G) &= \int_{\mathbb{R}^d} \mathbb{P}\{w_i \geq M_i, i \in T, \quad w_i < M_i, i \in \bar{T}\} dG(\mathbf{w}) \\ &= \int_{\mathbb{R}^d} \mathbb{P}\{M_i \leq w_i, i \in T\} dG(\mathbf{w}) - \int_{\mathbb{R}^d} \mathbb{P}\{M_i \leq w_i, i \in T, \exists i \in \bar{T} : M_i \leq w_i\} dG(\mathbf{w}) \end{aligned}$$

$$\begin{aligned}
&= \int_{\mathbb{R}^d} F_T^n(\mathbf{w}) dG_T(\mathbf{w}) - \int_{\mathbb{R}^d} \sum_{j=1}^m (-1)^{j+1} \sum_{J \subset \bar{T}: |J|=j} F_{J \cup T}^n(\mathbf{w}) dG(\mathbf{w}) \\
&= \sum_{j=0}^{d-k} (-1)^j \sum_{J \subset \bar{T}: |J|=j} \int_{\mathbb{R}^{k+j}} F_{J \cup T}^n(\mathbf{w}) dG_{J \cup T}(\mathbf{w}).
\end{aligned}$$

Applying [1][Thm 2.5 and Prop 4.2] we obtain

$$\lim_{n \rightarrow +\infty} n \int_{\mathbb{R}^{k+i}} F_{J \cup T}^n(\mathbf{y}) dG_{J \cup T}(\mathbf{y}) = \mathbb{E} \left\{ \min_{i \in J \cup T} \frac{\mathcal{W}_i}{\mathcal{Z}_i} \right\}$$

and thus

$$(4.4) \quad \mu_T(H, Q) = \sum_{j=0}^{d-k} (-1)^j \sum_{J \subset \bar{T}: |J|=j} \mathbb{E} \left\{ \min_{i \in J \cup T} \frac{\mathcal{W}_i}{\mathcal{Z}_i} \right\}.$$

By [10][Lem 1] we obtain further

$$\mu_T(H, Q) = \mathbb{E} \left\{ \min_{i \in T} \frac{\mathcal{W}_i}{\mathcal{Z}_i} - \min \left(\min_{i \in T} \frac{\mathcal{W}_i}{\mathcal{Z}_i}, \max_{i \in \bar{T}} \frac{\mathcal{W}_i}{\mathcal{Z}_i} \right) \right\},$$

hence the proof is complete. \square

Proof of Proposition 2.4. In view of the first equality in Equation (4.1)

$$\mathbb{E} \{f(N_n)\} = f(0) + \sum_{k=1}^d \Delta^k f(0) \sum_{K \subset \mathcal{D}, |K|=k} \mathbb{P} \{ \forall i \in K, W_i \geq M_{ni} \}.$$

Alternatively, using the second equality in Equation (4.1)

$$\mathbb{E} \{f(N_n)\} = f(0) + \sum_{k=1}^d (-1)^{k+1} \Delta^k f(d-k) \sum_{K \subset \mathcal{D}, |K|=k} \mathbb{P} \{ \exists i \in K, W_i \geq M_{ni} \}.$$

Thus using (1.1) establishes the claim. \square

Proof of Proposition 2.5. Let us consider $\mathbb{P} \left\{ (\mathcal{W}/\mathcal{Z})_{(k)} \leq x \right\} = \mathbb{P} \{ \text{at least } k \text{ events } [\mathcal{W}_i/\mathcal{Z}_i \leq x] \text{ are realized, } i \in \mathcal{D} \}.$

Using the first equality in Equation (4.2), for any function $g : \{1, \dots, d\} \rightarrow \mathbb{R}$ we obtain

$$\sum_{k=1}^d g(k) \mathbb{P} \left\{ (\mathcal{W}/\mathcal{Z})_{(k)} \leq x \right\} = \sum_{k=1}^d \Delta^{k-1} g(1) \sum_{K \subset \mathcal{D}, |K|=k} \mathbb{P} \{ \forall i \in K, \mathcal{W}_i/\mathcal{Z}_i \leq x \}$$

and hence letting $x \rightarrow \infty$ we have

$$\sum_{k=1}^d g(k) = \sum_{k=1}^d \Delta^{k-1} g(1) \sum_{K \subset \mathcal{D}, |K|=k} 1.$$

Consequently, for any real x

$$\sum_{k=1}^d g(k) \mathbb{P} \left\{ (\mathcal{W}/\mathcal{Z})_{(k)} > x \right\} = \sum_{k=1}^d \Delta^{k-1} g(1) \sum_{K \subset \mathcal{D}, |K|=k} \mathbb{P} \left\{ \max_{i \in K} \mathcal{W}_i/\mathcal{Z}_i > x \right\}.$$

By the assumptions

$$\mathbb{E} \left\{ \max_{1 \leq i \leq d} \mathcal{W}_i/\mathcal{Z}_i \right\} \leq \sum_{i=1}^d \mathbb{E} \{ \mathcal{W}_i/\mathcal{Z}_i \} = d,$$

hence since $\mathcal{W}_i/\mathcal{Z}_i$'s are non-negative it follows that

$$\sum_{k=1}^d g(k) \mathbb{E} \left\{ (\mathcal{W}/\mathcal{Z})_{(k)} \right\} = \sum_{k=1}^d \Delta^{k-1} g(1) \sum_{K \subset \mathcal{D}, |K|=k} \mathbb{E} \left\{ \max_{i \in K} \mathcal{W}_i/\mathcal{Z}_i \right\}.$$

Finally, in order to retrieve Equation (2.8), we must have for any $k \in \{1, \dots, d\}$

$$\Delta^{k-1}g(1) = (-1)^{k+1}\Delta^k f(d-k).$$

Now, assuming that for all $k \in \{1, \dots, d\}$, $g(k) = f(d-k+1) - f(d-k) = \Delta f(d-k)$, then denoting by $T = \Delta + I$ the translation operator

$$\Delta^{k-1}g(1) = \sum_{i=0}^{k-1} \binom{k-1}{i} (-1)^{k-1-i} T^{-i} \Delta f(d-1).$$

This implies

$$\Delta^{k-1}g(1) = (-I + T^{-1})^{k-1} \Delta f(d-1) = (-1)^{k-1} (T^{-1}(T-I))^{k-1} \Delta f(d-1).$$

Thus, for all $k \in \{1, \dots, d\}$ we have

$$\Delta^{k-1}g(1) = (-1)^{k+1} \Delta^k f(d-k)$$

and hence the claim follows. \square

Proof of Proposition 3.1. For the first equality, since \mathbf{Z} has unit Fréchet marginals for any $u > 0$ we have

$$C_{\mathbf{Z}}(u, \dots, u) = H\left(\frac{1}{-\ln u}, \dots, \frac{1}{-\ln u}\right) = \exp\left(\mathbb{E}\left\{\max_{1 \leq j \leq d} \ln(u) \mathcal{W}_j\right\}\right) = u^{\mathbb{E}\left\{\max_{j \in \mathcal{D}} \mathcal{W}_j\right\}}$$

and thus $\delta_{\mathbf{Z}}(u) = u^{r_{\mathbf{Y}}}$. Since the diagonal section of a d -dimensional Archimedean copula with parameter θ is $u^{d^{1/\theta}}$ we obtain the formula for θ . This is consistent with the fact that the Gumbel copula is an Extreme Value Copula (the only Archimedean one, see [19]).

For the last equality, setting $\mathcal{W}_j = F_{\mathcal{W}_1}^{-1}(U_j)$, we get $\max_{j \in \mathcal{D}} \mathcal{W}_j = \max_{j \in \mathcal{D}} F_{\mathcal{W}_1}^{-1}(U_j)$. Assuming further that all \mathcal{W}_i 's have a common df $F_{\mathcal{W}_1}$, then $\max_{j \in \mathcal{D}} F_{\mathcal{W}_1}^{-1}(U_j) = F_{\mathcal{W}_1}^{-1}(\max_{j \in \mathcal{D}} U_j)$. Using further

$$\mathbb{P}\left\{\max_{j \in \mathcal{D}} U_j \leq u\right\} = \mathbb{P}\{U_1 \leq u, \dots, U_d \leq u\} = C_{\mathbf{Y}}(u, \dots, u) = \delta_{\mathbf{Y}}(u)$$

we get $\mathbb{E}\{\max_{j \in \mathcal{D}} \mathcal{W}_j\} = \int_0^1 F_{\mathcal{W}_1}^{-1}(s) d\delta_{\mathbf{Y}}(s)$. \square

Acknowledgments: EH is partially supported by SNSF Grant 200021-175752/1 and PSG 1250 grant. We also thank the anonymous reviewer for very helpful comments and suggestions.

REFERENCES

- [1] E. Hashorva, "Domination of sample maxima and related extremal dependence measures," *Depend. Model.*, vol. 6, no. 1, pp. 88–101, 2018.
- [2] L. de Haan, "A spectral representation for max-stable processes," *Ann. Probab.*, vol. 12, no. 4, pp. 1194–1204, 1984.
- [3] A. V. Gnedin, "On a best-choice problem with dependent criteria," *J. Appl. Probab.*, vol. 31, no. 1, pp. 221–234, 1994.
- [4] A. V. Gnedin, "Conical extremes of a multivariate sample," *Journal of Research of National Institute of Standards and Technology*, vol. 99, pp. 511–511, 1994.
- [5] A. V. Gnedin, "Records from a multivariate normal sample," *Statist. Probab. Lett.*, vol. 39, no. 1, pp. 11–15, 1998.

- [6] E. Hashorva and J. Hüsler, “On asymptotics of multivariate integrals with applications to records,” *Stoch. Models*, vol. 18, no. 1, pp. 41–69, 2002.
- [7] E. Hashorva and J. Hüsler, “Multiple maxima in multivariate samples,” *Statist. Probab. Lett.*, vol. 75, no. 1, pp. 11–17, 2005.
- [8] C. Dombry, M. Ribatet, and S. Stoev, “Probabilities of concurrent extremes,” *J. Amer. Statist. Assoc.*, vol. 113, no. 524, pp. 1565–1582, 2018.
- [9] S. Stoev and Y. Wang, “Exchangeable random partitions from max-infinitely-divisible distributions,” *Statist. Probab. Lett.*, vol. 146, pp. 50–56, 2019.
- [10] C. Dombry and M. Zott, “Multivariate records and hitting scenarios,” *Extremes*, vol. 21, no. 2, pp. 343–361, 2018.
- [11] C. Dombry and Z. Kabluchko, “Random tessellations associated with max-stable random fields,” *Bernoulli*, vol. 24, no. 1, pp. 30–52, 2018.
- [12] S. I. Resnick, *Extreme values, regular variation, and point processes*, vol. 4 of *Applied Probability. A Series of the Applied Probability Trust*. New York: Springer-Verlag, 1987.
- [13] J. Beirlant, Y. Goegebeur, J. Teugels, and J. Segers, *Statistics of extremes*. Wiley Series in Probability and Statistics, John Wiley & Sons, Ltd., Chichester, 2004.
- [14] M. Falk, J. Hüsler, and R. D. Reiss, “Laws of Small Numbers: Extremes and Rare Events,” in *DMV Seminar*, vol. 23, Basel: Birkhäuser, third ed., 2010.
- [15] C. Dombry and M. Zott, “Multivariate records and hitting scenarios,” *Extremes*, vol. 21, no. 2, pp. 343–361, 2018.
- [16] C. Dombry, M. Falk, and M. Zott, “On Functional Records and Champions,” *J. Theoret. Probab.*, vol. 32, no. 3, pp. 1252–1277, 2019.
- [17] A. Erdely, J. M. González-Barrios, and M. M. Hernández-Cedillo, “Frank’s condition for multivariate archimedean copulas,” *Fuzzy Sets and Systems*, vol. 240, pp. 131–136, 2014.
- [18] H. U. Gerber, *Life insurance mathematics*. Springer Science & Business Media, 2013.
- [19] C. Genest and L.-P. Rivest, “A characterization of gumbel’s family of extreme value distributions,” *Statistics & Probability Letters*, vol. 8, no. 3, pp. 207–211, 1989.

ENKELEJD HASHORVA, DEPARTMENT OF ACTUARIAL SCIENCE UNIVERSITY OF LAUSANNE,, UNIL-DORIGNY, 1015 LAUSANNE, SWITZERLAND
E-mail address: Enkelejd.Hashorva@unil.ch

DIDIER RULLIÈRE, ECOLE ISFA, LSAF, UNIVERSITÉ LYON 1, 69366 LYON, FRANCE
E-mail address: didier.rulliere@univ-lyon1.fr