The limit distribution of the largest interpoint distance for power-tailed spherically decomposable distributions and their affine images

Norbert Henze[1] and Wei Lao

[1] *Institute of Stochastics, Karlsruhe Institute of Technology*
Kaiserstr. 89–93, D-76131 Karlsruhe, Germany

Abstract. Let $X_1, X_2, \ldots$ be i.i.d. copies of a d-variate random vector ('point') X with a spherically decomposable distribution, which means that the Euclidean norm $\|X\|$ of X and the directional part $X/\|X\|$ are independent. We derive the limit distribution of the maximum interpoint distance $D_n := \max_{1 \le i < j \le n} \|X_i - X_j\|$ as $n \to \infty$ under the condition that the distribution function F of $\|X\|$ satisfies $1 - F(s) = s^{-\alpha} L(s)$ as $s \to \infty$ for some $\alpha > 0$ and some slowly varying function L. Whereas D_n, after suitable rescaling, has a Gumbel limit distribution if X follows the (short-tailed) spherically symmetric Kotz distribution [6] and a Weibull limit if X takes values in a sphere [11], the limit distribution of D_n under the condition given above is none of the three types of classical extreme value distributions. The method of proof also covers the case of long-tailed elliptically symmetric distributions, such as the class of symmetric multivariate Pearson type VII laws.

Keywords. Geometric extreme value theory, maximum interpoint distance, Fréchet distribution, multivariate regular variation, Poisson process, elliptical symmetric distribution

AMS 2010 subject classifications: Primary 60G70, secondary 60D05, 60F05, 60G55.

1 Introduction

Let $X_1, X_2, \ldots$ be independent and identically distributed (i.i.d.) copies of a d-dimensional random vector ('point') X. In recent years, the largest Euclidean interpoint distance

$$D_n := \max_{1\le i<j\le n} \|X_i - X_j\|$$

has been an object of ongoing interest, although still little is known on the limit behavior of D_n for a general underlying distribution of X. Matthews and Rukhin [10] derive the limit distribution of D_n, upon suitably rescaling, if X has a standard normal distribution in $\mathbb{R}^d$, and Henze and Klein [6] generalize this result to points from a spherically symmetric Kotz type distribution having a density proportional to $\|x\|^{2(b-1)}\exp(-\kappa\|x\|^2)$, where $2b+d>2$ and $\kappa > 0$ (see Fang, Kotz and Ng [4], Section 3.2). In this case, for some sequence u_n converging to 0 at the rate $\log\log n/\sqrt{\log n}$, we have

$$2\kappa\sqrt{(1/\kappa)\log n}\left(D_n - 2\sqrt{(1/\kappa)\log n} - u_n\right) \xrightarrow{\mathcal{D}} G,$$

where '$\xrightarrow{\mathcal{D}}$' means convergence of random variables and stochastic processes, and G is a random variable having the Gumbel distribution function $\exp(-\exp(-t))$. In particular, we have $D_n - 2\sqrt{(1/\kappa)\log n} \xrightarrow{\mathrm{P}} 0$, where '$\xrightarrow{\mathrm{P}}$' denotes stochastic convergence.

If the distribution of X is supported by the unit d-sphere $\mathcal{B}^d = \{x \in \mathbb{R}^d : \|x\| \le 1\}$, results on the limit behavior have been obtained by Lao [8] for the uniform distribution over $\mathcal{B}^d$ and, independently of Lao [6], by Mayer and Molchanov [11] who used a different method and covered more general underlying distributions supported by $\mathcal{B}^d$. For example, we have (Mayer and Molchanov [11], Lao and Mayer [9]

$$\lim_{n\to\infty} P\left(\tau n^{4/(d-1+4\alpha)}(2-D_n) \le t\right) = 1 - \exp\left(-t^{\frac{d-1}{2}+2\alpha}\right), \quad t>0, \tag{1.1}$$

if $R := \|X\|$ and $U := X/\|X\|$ are independent, the distribution function F of R satisfies $F(1-s) \sim 1 - as^\alpha$ as $s \downarrow 0$ for some positive constants a and α, and the distribution of U has a density with respect to the surface area measure on the boundary of $\mathcal{B}^d$ (the latter distribution, as well as a, α and d, enter into the constant τ figuring in (1.1)). Hence, D_n, after a suitable affine transformation, has a Weibull limit under this setting.

If the distribution of X is supported by some compact set $K \subset \mathbb{R}^d$ with diameter $\delta(K) := \sup_{x,y\in K}\|x-y\|$, we have $D_n \to \delta(K)$ almost surely (under weak regularity conditions), and it is natural to ask whether $\delta(K) - D_n$, when suitably scaled, has a non-degenerate limit distribution. Result (1.1) is one example within this more general setting. Appel and Russo

[1] state a limit theorem on D_n in the special case $d = 2$ when the distribution of X is uniform and K has a finite number of major axes with no common vertices. Moreover, the boundary of K near the endpoints of these major axes is locally defined by regularly varying functions having indices larger than 1/2. If K is the unit square, we have

$$\sqrt{n}(\sqrt{2} - D_n) \xrightarrow{\mathcal{D}} \min(W_1 + W_2,\ W_3 + W_4),$$

where W_1, W_2, W_3, W_4 are i.i.d. random variables having the Weibull distribution function $1 - \exp(-t^2),\ t > 0$. It should be stressed that, apart from the sphere $\mathcal{B}^d$, nothing is known on the limit behavior of D_n if the compact set K has a smooth boundary. For the case of a proper ellipse in $\mathbb{R}^2$, there is a long standing conjecture that $n^{2/3}(\delta(K) - D_n)$ has a non-degenerate limit distribution (see Appel and Russo [1]), Theorem 3).

In this paper we consider the case that the distribution of X is spherically decomposable, which means that $P(\|X\| > 0) = 1$ and that the norm $\|X\|$ and the directional part $X/\|X\|$ of X are independent. Moreover, we assume that the right tail of the distribution function of $\|X\|$ decays (essentially) like a power law. Notice that this assumption is more general than spherical symmetry, which holds if the distribution of $X/\|X\|$ is uniform over the unit sphere surface $\mathcal{S}^{d-1} = \{x \in \mathbb{R}^d : \|x\| = 1\}$. In Section 4 the results are generalized to the case that the underlying distribution is an affine image of a power-tailed spherically decomposable distribution.

We close this introduction by mentioning that D_n is a functional of the convex hull P_n of $X_1, \ldots, X_n$. For a recent overview over results on other functionals of P_n, see Reitzner [12].

2 Main results

Let $(\Omega, \mathcal{A}, P)$ denote the probability space on which $X, X_1, X_2, \ldots$ are defined. Writing $R = \|X\|$, $U = X/\|X\|$, $R_j = \|X_j\|$ and $U_j = X_j/\|X_j\|$, $j \geq 1$, it follows that $X_j = R_j U_j$ with independent factors R_j and U_j, with U_j having some arbitrary (possibly degenerate) distribution over $\mathcal{S}^{d-1}$. Under this setting, the rate of growth of the maximum interpoint distance D_n depends solely on the right tail behavior of the 'radial' distibution function $F(s) = P(R \leq s)$, $s > 0$ (see Theorem 1). In what follows, we assume that F belongs to the maximum domain of attraction of the Fréchet distribution having distribution function

$\Phi_\alpha(t) = \exp(-t^{-\alpha})$, $t > 0$, for some $\alpha > 0$. In this case, we have

$$(2.1) \qquad 1 - F(s) = s^{-\alpha} \cdot L(s), \qquad s > 0,$$

for some slowly varying function L. In other words, the right tail of F is regularly varying with index $-\alpha$ (see, e.g., Embrechts et al. [3], A3). Together with the independence of the radial part R and the directional part U of X it follows that

$$(2.2) \qquad \lim_{t\to\infty} \frac{P(R > tr, U \in A)}{P(R > t)} = r^{-\alpha} S(A),$$

$r > 0$, $A \in \mathcal{B}(\mathcal{S}^{d-1})$, where S is shorthand for the distribution of U on (the σ-algebra $\mathcal{B}(\mathcal{S}^{d-1})$ of Borel subsets of) $\mathcal{S}^{d-1}$. Condition (2.2) is equivalent to the d-variate regular variation of the distribution function of X (see e.g. Resnick [13]).

To state the main result, we adopt the notation of Resnick [14], Chapter 3. Consider the punctured compact space $E = [-\infty, \infty]^d \setminus \{0\}$, in which compact subsets are those closed sets bounded away from the origin $0 \in \mathbb{R}^d$, equipped with the Borel σ-algebra $\mathcal{E}$ of E. Write ϵ_x for the Dirac measure centred on a point $x \in E$, and let $M_P(E)$ be the set of point measures m of the form $m = \sum_{j=1}^{\infty} \epsilon_{x_j}$, where $\{x_j : j \geq 1\}$ is a countable collection of not necessarily distinct points of E. We assume that m is Radon, which means that m takes finite values on compact subsets of E. The set $M_P(E)$ is equipped with the smallest σ-algebra (denoted by $\mathcal{M}_P(E)$) such that the evaluation maps $M_P(E) \ni m \mapsto m(B) \in [0, \infty]$, $B \in \mathcal{E}$, are measurable.

A point process on E is a random element of $M_P(E)$ that is defined on some probability space. A Poisson process with (Radon) intensity measure μ is a point process ξ such that, for $k \in \{0, 1, 2, \ldots\}$

$$P(\xi(B) = k) = \exp(-\mu(B)) \cdot \frac{\mu(B)^k}{k!}, \quad \text{if} \quad \mu(B) < \infty$$

and $P(\xi(B) = k) = 0$, otherwise. Moreover, $\xi(B_1), \ldots, \xi(B_l)$ are independent for any choice of $l \geq 2$ and mutually disjoint sets $B_1, \ldots, B_l \in \mathcal{E}$. As a last bit of notation, *conv* M stands for the convex hull and $\delta(M) = \sup\{\|x - y\| : x, y \in M\}$ for the diameter of a bounded subset M of $\mathbb{R}^d$.

The following result shows that the rate of growth to infinity of the largest interpoint distance depends solely on the radial part of the underlying distribution.

Theorem 2.1 *Let $c_n = \inf\{t \in \mathbb{R} : F(t) \geq 1 - 1/n\}$. Under the standing assumptions and condition (2.1), we have*

$$c_n^{-1} D_n \xrightarrow{\mathcal{D}} \delta(conv\{\mathcal{P}_k : k \geq 1\}) \quad as \quad n \to \infty,$$

where $\{\mathcal{P}_k : k \geq 1\}$ *are the points of a Poisson process on* $(E, \mathcal{E})$ *with intensity measure* ν *defined by*

$$(2.3) \qquad \nu(\{x : \|x\| > r, x/\|x\| \in A\}) \;=\; r^{-\alpha} S(A)$$

$(r > 0,\ A \in \mathcal{B}(\mathcal{S}^{d-1}))$.

Proof: Define point processes N_n, $n \geq 1$, and N on E by

$$(2.4) \qquad N_n = \sum_{j=1}^{n} \epsilon_{c_n^{-1} X_j}, \qquad N = \sum_{k=1}^{\infty} \epsilon_{\mathcal{P}_k},$$

with $\{\mathcal{P}_k : k \geq 1\}$ given above. From Proposition 3.21 of Resnick [14], we have $N_n \xrightarrow{\mathcal{D}} N$ as $n \to \infty$. Let $C(\mathcal{K}^d)$ denote the collection of non-empty convex compact subsets of $\mathbb{R}^d$, equipped with the Hausdorff metric $\rho(K, H) = \inf\{\varepsilon > 0 : K \subset H^\varepsilon,\ H \subset K^\varepsilon\}$, where $M^\varepsilon = \{x : \inf_{y \in M} \|x - y\| \leq \varepsilon\}$, rendering $(C(\mathcal{K}^d), \rho)$ a complete separable metric space. Consider the function $T : M_P(E) \to C(\mathcal{K}^d)$, defined by

$$T(m) = T\left(\sum_{k \geq 1} \epsilon_{x_k}\right) = conv\{x_k : k \geq 1\}.$$

Since $\nu(E \setminus \mathbb{R}^d) = 0$, the mapping T is a.s. continuous (see Davis et al. [2], p. 8-9). From $N_n \xrightarrow{\mathcal{D}} N$ and the continuous mapping theorem (see, e.g., Kallenberg [7], p.76) we therefore have

$$(2.5) \qquad conv\{c_n^{-1} X_1, \ldots, c_n^{-1} X_n\} \;\xrightarrow{\mathcal{D}}\; conv\{\mathcal{P}_k : k \geq 1\}$$

in $C(\mathcal{K}^d)$. Note that $\nu(E \setminus \mathbb{R}^d) = 0$ ensures that $conv\{\mathcal{P}_k : k \geq 1\}$ is compact in $\mathbb{R}^d$. Since the mapping $C(\mathcal{K}^d) \ni M \mapsto \delta(M)$ is continuous with respect to the Hausdorff metric, another appeal to the continuous mapping theorem yields

$$c_n^{-1} D_n = \delta(conv\{c_n^{-1} X_1, \ldots, c_n^{-1} X_n\}) \xrightarrow{\mathcal{D}} \delta(conv\{\mathcal{P}_k : k \geq 1\}),$$

as was to be shown. $\square$

Writing $\sim$ for equality in distribution, we have

$$(2.6) \qquad \delta(conv\{\mathcal{P}_k : k \geq 1\}) \;\sim\; \sup_{1 \leq i < j < \infty} \|Y_i U_i - Y_j U_j\|,$$

where $Y_1, Y_2, \ldots$ is a sequence of random variables such that, for fixed $k \geq 1$, $(Y_1, \ldots, Y_k)$ is a k-dimensional Φ_α-extremal variate with density

$$(2.7) \qquad h_\alpha^{(k)}(t_1, \ldots, t_k) = \alpha^k \exp\left(-t_k^{-\alpha}\right) \cdot \prod_{i=1}^{k} t_i^{-(\alpha+1)}, \quad t_1 \geq t_2 \geq \ldots \geq t_k > 0$$

(see, e.g. Theorem 4.2.8 and Example 4.2.9 of Embrechts et al. [3], and $U_1, U_2, \ldots$ is a sequence of i.i.d. copies of U which is independent of $Y_1, Y_2, \ldots$. In fact, Y_j may be considered as the j-th largest of $\|\mathcal{P}_1\|, \|\mathcal{P}_2\|, \ldots$ and U_j as the directional part of the respective point. Notice that U_j has the same distribution as $X_j/\|X_j\|$.

Remark 2.2 One could prove the weak convergence of $c_n^{-1}D_n$ to the right hand side of (2.6) along the lines of Henze [5] by showing weak convergence of the random sequence $\mathcal{Z}_n = (Z_{n,k})_{k\geq 1}$, where $Z_{n,k} = c_n^{-1}R_{k,n}U_{[k,n]}$ if $k \leq n$ and $Z_{n,k} = 0$, otherwise, in the separable Banach space $c_0^d = \{x = (x_j)_{j\geq 1} \in [\mathbb{R}^d]^\infty : \lim_{j\to\infty} x_j = 0\}$, equipped with the supremum norm $\|x\|_\infty := \sup_{j\geq 1}\|x_j\|$. Here, $R_{1,n} \geq \ldots \geq R_{n,n}$ denote the order statistics of $R_1, \ldots, R_n$, and $U_{[1,n]}, \ldots, U_{[n,n]}$ are the concomitants of the order statistics $R_{1,n}, \ldots, R_{n,n}$, i.e. we have $U_{[k,n]} = U_j$ if $R_{k,n} = R_j$. However, such an approach would hide an important structural feature of the results.

Corollary 2.3 Under the standing assumptions, suppose that the function L figuring in (2.1) satisfies

(2.8) $\lim_{s\uparrow\infty} L(s) = \gamma$

for some positive finite γ. We then have

$$(\gamma n)^{-1/\alpha} D_n \xrightarrow{\mathcal{D}} \delta(conv\{\mathcal{P}_k : k \geq 1\}),$$

where $\{\mathcal{P}_k : k \geq 1\}$ is given in Theorem 1.

Condition (2.8) covers the case of Pareto-like distributions for R, like the Pareto distribution, the Cauchy distribution, the Burr distribution and stable distributions with exponent $\alpha < 2$.

3 Bounds for the limit law and examples

Writing D for the right-hand side of (2.6), the triangle inequality and the inequalities $Y_1 \geq Y_2 \geq \ldots$ yield

(3.1) $D \leq Y_1 + Y_2.$

On the other hand, the fact that $Y_j \to 0$ almost surely as $j \to \infty$ implies

$$(3.2) \qquad D \geq Y_1$$

almost surely. This means that the limit distribution of the largest interpoint distance is stochastically bounded by the first extremal variate and the sum of the first and second extremal variates. In particular, the distribution of D is always non-degenerate whatever choice of the directional distribution.

The lower bound (3.2) is attained if the distribution of U is the Dirac measure centred on some point a such that $\|a\| = 1$. We conjecture that the upper bound (3.1) is not attained regardless of the distribution of U and that, for any given radial distribution, the distribution of D is stochastically largest if U has a uniform distribution over $\mathcal{S}^{d-1}$, i.e., X has a spherically symmetric distribution. If the distribution of U is a symmetric mixture $(\delta_a + \delta_{-a})/2$ of Dirac measures at opposite points on the sphere surface, the distribution of D is an infinite mixture of the distributions of $Y_1 + Y_j$, $j \geq 2$. More precisely, we have

$$(3.3) \qquad D \sim \sum_{j=1}^{\infty} 2^{-j} P^{Y_1 + Y_{j+1}},$$

where, generally, P^Z stands for the distribution of a random variable Z. To prove (3.3), suppose without loss of generality that $U_1 = a$. Then, because of $Y_1 \geq Y_2 \geq \ldots$, we have $D = Y_1 + Y_j$, where $j = \min\{m \geq 2 : U_m = -a\}$. Since $U_2, U_3, \ldots$ are independent, and $P(U_m = a) = P(U_m = -a) = 1/2$, (3.3) follows.

The following are some examples with simulation results.

Example 3.1 Suppose R has an (absolute) Cauchy distribution so that $F(s) = \frac{2}{\pi} \arctan(s)$, $s > 0$. Then (2.1) holds with $\alpha = 1$ and $L(s) = s \cdot (1 - \frac{2}{\pi} \arctan(s))$. Since $\lim_{s \uparrow \infty} L(s) = 2/\pi$, it follows from Corollary 2.3 that

$$\left(\frac{2n}{\pi}\right)^{-1} D_n \xrightarrow{\mathcal{D}} \sup_{1 \leq i < j < \infty} \|Y_i U_i - Y_j U_j\|,$$

where $Y_1, Y_2, \ldots$ and $U_1, U_2, \ldots$ are given in (2.6), and the sequence $(Y_j)_{j \geq 1}$ has finite-dimensional distributions (2.7) with $\alpha = 1$.

Example 3.2 If R has the Pareto distribution function $F(s) = 1 - \left(\frac{\kappa}{\kappa + s}\right)^{\alpha}$ for some $\alpha > 0$, $\kappa > 0$, we have $1 - F(s) = s^{-\alpha} \cdot L(s)$ with $L(s) = \left(\frac{s\kappa}{\kappa + s}\right)^{\alpha}$. Since $L(s) \to \kappa^{\alpha}$ as $s \uparrow \infty$, Corollary 2.3 yields

$$\kappa^{-1} n^{-1/\alpha} \cdot D_n \xrightarrow{\mathcal{D}} \sup_{1 \leq i < j < \infty} \|Y_i U_i - Y_j U_j\|,$$

where $Y_1, Y_2, \ldots$ and $U_1, U_2, \ldots$ are given in (2.6).

Figure 1 (left) shows a simulation of the empirical distribution function (edf) of $(2n/\pi)^{-1} \cdot D_n$ if R has an absolute Cauchy distribution. Figure 1 (right) exhibits the edf of $n^{-1/4} \cdot D_n$, where R has the Pareto distribution with $\kappa = 1$ and $\alpha = 4$. In both cases, the distribution of U is uniform over the sphere surface, the number of points is 10^6, and the number of replications is 10^4. The upper and lower curves in both figures are the corresponding distribution functions of Y_1 and $Y_1 + Y_2$, respectively.

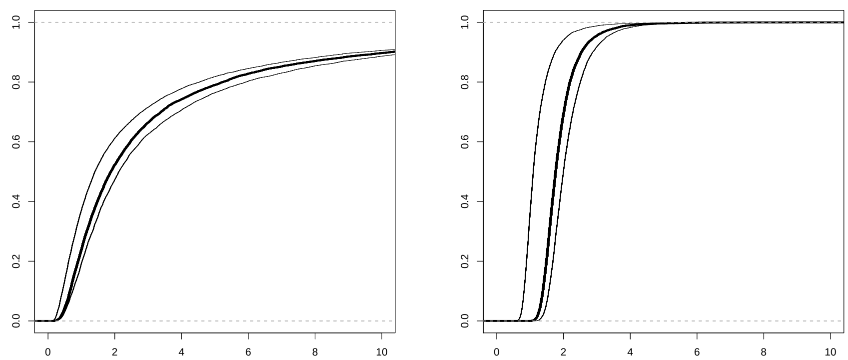

Figure 1: Edf of $(2n/\pi)^{-1} \cdot D_n$ (left) and $n^{-1/4} \cdot D_n$ (right), corresponding to Examples 1 and 2. The upper (lower) curve is the distribution function of Y_1 and $Y_1 + Y_2$, respectively.

4 Generalizations

4.1 Affine Transformations, elliptically symmetric distributions

Suppose that $\tilde{X} = AX$, where the distribution of $X = RU$ satisfies the standing assumptions, and A is a regular $d \times d$-matrix. Notice that this assumption covers the case of elliptical symmetric distributions (arising if the distribution of U is uniform) that exhibit a power law decay of the distribution function of $\|X\|$. Suppose that $\tilde{X}_1, \tilde{X}_2, \ldots$ is an i.i.d. sequence of copies of $\tilde{X}$. Writing

$$\tilde{D}_n = \max_{1\le i<j\le n} \|\tilde{X}_i - \tilde{X}_j\|$$

for the maximum interpoint distance of $\tilde{X}_1, \ldots, \tilde{X}_n$, we have the following result.

Theorem 4.1 *We have*

$$c_n^{-1}\tilde{D}_n \xrightarrow{\mathcal{D}} \delta(conv\{A\mathcal{P}_j : j \ge 1\}),$$

where c_n and $\{\mathcal{P}_j : j \ge 1\}$ are given in the statement of Theorem 1.

Proof. Using (2.5) and the fact that the mapping $C(\mathcal{K}^d) \ni K \mapsto AK$ is continuous with respect to the Hausdorff metric (notice that $\rho(AH, AK) \le \|A\| \cdot \rho(H, K)$, where $\|A\|$ is the spectral norm of A), we have

$$\begin{aligned}
c_n^{-1}\tilde{D}_n &= c_n^{-1}\delta(conv\{AX_1, \ldots, AX_n\}) \\
&= \delta(A\ conv\{c_n^{-1}X_1, \ldots, c_n^{-1}X_n\}) \\
&\xrightarrow{\mathcal{D}} \delta(A\ conv\{\mathcal{P}_j : j \ge 1\}) \\
&= \delta(conv\{A\mathcal{P}_j : j \ge 1\}).
\end{aligned}$$

□

As an example, let $\tilde{X} = AX$ with a regular $d\times d$-matrix A, and X having the spherically symmetric density $f(x) = \Gamma(a)/(\Gamma(a-d/2)(\pi\kappa)^{d/2})(1+\|x\|^2)/\kappa)^{-a}$, $x \in \mathbb{R}^d$, where $\kappa > 0$ and $a > d/2$. Putting $\Delta = AA^\top$, where $\top$ denotes transpose of matrices and vectors, $\tilde{X}$ has the density

$$\tilde{f}(x) = \frac{\Gamma(a)}{\Gamma(a-d/2)(\pi\kappa)^{d/2}|\Delta|^{1/2}} \cdot \left(1 + \frac{x^\top\Delta^{-1}x}{\kappa}\right)^{-a},$$

$x \in \mathbb{R}^d$, where $|\Delta|$ is the determinant of Δ. Notice that, apart from an additional location parameter μ (which has no influence on distances between points), $\tilde{f}$ is the density of the (elliptically) symmetric multivariate Pearson type VII distribution, denoted by $\tilde{X} \sim MPVII_d(\kappa, a, \mu, \Delta)$ (see, e.g. Section 3.3 of Fang et al. [4]. Straightforward calculations yield

$$P(\|X\| > t) = \frac{1}{B(d/2, a-d/2)} \cdot \int_0^{\kappa/(\kappa+t^2)} y^{a-d/2-1}(1-y)^{d/2-1}\, dy,$$

where $B(\alpha,\beta) = \Gamma(\alpha)\Gamma(\beta)/\Gamma(\alpha+\beta)$ is the Beta function. It follows that $P(\|X\| > t) = L(t)t^{-(2a-d)}$, $t > 0$, where L is a slowly varying function satisfying

$$\lim_{t\to\infty} L(t) = \frac{\kappa^{a-d/2}}{B(d/2, a-d/2)(a-d/2)}. \tag{4.1}$$

From Corollary 2.3 and Theorem 4.1 we thus have the following result.

Corollary 4.2 If $\tilde{X}_1, \tilde{X}_2, \ldots$ are i.i.d. points from the symmetric multivariate Pearson type VII distribution $MPVII_d(\kappa, a, \mu, \Delta)$, we have

$$(\gamma n)^{-1/(2a-d)} \tilde{D}_n \xrightarrow{\mathcal{D}} \delta(conv\{A\mathcal{P}_j : j \geq 1\}) \qquad \text{as } n \to \infty,$$

where γ is given by the right-hand side of (4.1). Notice that the Poisson process $\{\mathcal{P}_j : j \geq 1\}$ is isotropic, and that the limit distribution depends on the matrix A only via $\Delta = AA^\top$.

4.2 Random sample sizes

Under the conditions of Theorem 2.1, suppose that $(Z(t))_{t\geq 0}$ is a process of integer-valued random variables, which is independent of $(X_j)_{j\geq 1}$ and satisfies

$$\frac{Z(t)}{t} \xrightarrow{\text{P}} \lambda$$

as $t \to \infty$ for some $\lambda \in (0, \infty)$. For example, $(Z(t))_{t\geq 0}$ could be a homogeneous Poisson Process on $[0, \infty)$ with intensity λ. Putting

$$D_n^* := \max_{1\leq i<j\leq Z(n)} \|X_i - X_j\|,$$

we have the following result.

Theorem 4.3 *Under the standing assumptions on the distribution of X, we have*

$$c_n^{-1} D_n^* \xrightarrow{\mathcal{D}} \delta(conv\{\mathcal{Q}_k : k \geq 1\}),$$

where $\{\mathcal{Q}_k : k \geq 1\}$ are the points of a Poisson process on $(E, \mathcal{E})$ with intensity measure $\lambda\nu$, with ν given in (2.3).

Proof. Define point processes on $(E, \mathcal{E})$ by

$$\xi_n = \sum_{j=1}^{Z(n)} \epsilon_{c_n^{-1} X_j}, \qquad \xi = \sum_{j=1}^{\infty} \epsilon_{\mathcal{Q}_j},$$

and write $\psi_{\xi_n}(f) = E\exp(-\xi_n f)$, $\psi_\xi(f) = E\exp(-\xi f)$ for the Laplace functionals of ξ_n and ξ, where f is a continuous nonnegative function on E with compact support. By conditioning on $Z(n)$ and using the independence of $Z(n)$ and $X_1, \ldots, X_n$, we obtain

$$\begin{aligned}
\psi_{\xi_n}(f) &= E\left[E\left(\exp\left\{-\sum_{j=1}^{Z(n)} f(c_n^{-1}X_j)\right\}\middle|Z(n)\right)\right] \\
&= E\left[\left(\left\{E\exp(-f(c_n^{-1}X_1)\right\}^n\right)^{Z(n)/n}\right] \\
&\to \exp\left\{-\int_E \left(1-e^{-f(x)}\right)\lambda\nu(dx)\right\} \quad \text{as } n\to\infty \\
&= \psi_\xi(f),
\end{aligned}$$

where the convergence follows from uniform integrability and the fact that the process N_n defined in (2.4) converges to a Poisson process with intensity measure ν, having Laplace functional $\exp(-\int_E(1-e^{-f(x)})\nu(dx)$. Thus, $\xi_n \xrightarrow{\mathcal{D}} \xi$ as $n\to\infty$, and the rest of the argument follows the lines of the proof of Theorem 1. □

We have

$$\delta(conv\{\mathcal{Q}_k : k\geq 1\}) \sim \sup_{1\leq i<j<\infty} \|Y_{i,\lambda}U_i - Y_{j,\lambda}U_j\|,$$

where $Y_{1,\lambda}, Y_{2,\lambda}, \ldots$ is a sequence of random variables with finite-dimensional distributions given by

$$h_{\alpha,\lambda}^{(k)}(t_1,\ldots,t_k) = (\alpha\lambda)^k \exp\left(-\lambda t_k^{-\alpha}\right)\cdot\prod_{i=1}^k t_i^{-(\alpha+1)}, \quad t_1\geq t_2\geq\ldots\geq t_k>0,$$

and $U_1, U_2, \ldots$ is a sequence of i.i.d. copies of U (figuring in (2.2)), which is independent of $Y_{1,\lambda}, Y_{2,\lambda}, \ldots$ (cf. Theorem 4.3.4 of [3]). Here, $Y_{j,\lambda}$ may be considered as the j-th largest of $\|\mathcal{Q}_1\|, \|\mathcal{Q}_2\|, \ldots$ and U_j as the directional part of the respective point.

Acknowledgements

The authors would like to thank A.A. Borovkov for helpful discussions and an anonymous referee for constructive criticism on an earlier version of the paper.

References

[1] Appel, Martin J. B. and Najim, Christopher A. and Russo, Ralph P. (2002): Limit laws for the diameter of a random point set. *Adv. Appl. Probab.*, 34, 1–10.

[2] Davis, R., Mulrow, E. and Resnick, S. (1987): The convex hull of a random sample in $\mathbb{R}^2$.

Commun. Statist. - Stoch. Models 3:1-27, 1987.

[3] Embrechts, Paul and Klüppelberg, Claudia and Mikosch, Thomas (1997): *Modelling Extremal Events*. Springer, New York.

[4] Fang, Kai Tai and Kotz, Samuel and Ng, Kai Wang (1990): *Symmetric multivariate and related distributions*. Chapman and Hall Ltd., London.

[5] Henze, Norbert (1996): Empirical-distribution-function goodness-of-fit tests for discrete models. *Canad. J. Statist.*, 24, 81–93.

[6] Henze, Norbert and Klein, Timo (1996): The limit distribution of the largest interpoint distance from a symmetric Kotz sample. *J. Multiv. Anal.*, 57(2), 228–239.

[7] Kallenberg, Olav (2002): *Foundations of modern probability*, 2^{nd} edition. Springer, New York.

[8] Lao, Wei (2006): The limit law of the Maximum Distance of points in a sphere in $\mathbb{R}^d$. *Preprint*, Universität Karlsruhe (TH).

[9] Lao, Wei and Mayer, Michael (2008): U-max-statistics. *J. Multiv. Anal.*, 99, 2039–2052.

[10] Matthews, Peter C. and Rukhin, Andrew L. (1993): Asymptotic distribution of the normal sample range. *Ann. Appl. Probab.*, 3(2), 454–466.

[11] Mayer, Michael and Molchanov, Ilya (2007): Limit theorems for the diameter of a random sample in the unit ball. *Extremes*, 10, 151–174.

[12] Reitzner, Matthias (2010): Random polytopes. In: *New perspectives in stochastic geometry, W.S. Kendall & I. Molchanov (eds.),* 45–76. Oxford University Press.

[13] Resnick, Sidney I. (1986): Point processes, regular variation and weak convergence. *Adv. Appl. Probab.*, 18, 66-138.

[14] Resnick, Sidney I. (1987): *Extreme values, regular variation, and point processes.* Springer, New York.