DE GRUYTER OPEN Depend. Model. 2017; 5:221–245

Research Article Open Access

K. Müller and W.-D. Richter*

Exact distributions of order statistics from $l_{n,p}$-symmetric sample distributions

https://doi.org/10.1515/demo-2017-0013
Received May 2, 2017; accepted July 9, 2017

Abstract: We derive the exact distributions of order statistics from a finite number of, in general, dependent random variables following a joint $l_{n,p}$-symmetric distribution. To this end, we first review the special cases of order statistics from spherical as well as from p-generalized Gaussian sample distributions from the literature. To study the case of general $l_{n,p}$-dependence, we use both single-out and cone decompositions of the events in the sample space that correspond to the cumulative distribution function of the kth order statistic if they are measured by the $l_{n,p}$-symmetric probability measure. We show that in each case distributions of the order statistics from $l_{n,p}$-symmetric sample distribution can be represented as mixtures of skewed $l_{n-\nu,p}$-symmetric distributions, $\nu \in \{1, \ldots, n-1\}$.

Keywords: density generator, extreme value statistics, $l_{n,p}$-dependence, measure-of-cone representation, skewed $l_{n,p}$-symmetric distribution

MSC: 60E05, 62E15, 62H10

1 Introduction

Order statistics are a useful tool in numerous scientific areas with a wide-ranging applicability. In order to mention just a few areas where order statistics are needed, we refer to applications related to health data, multiple decision rules, multiple comparison problems, tests of hypotheses and digital image processing which are presented in [15], [18], and [28], respectively. More recent applications to goodness-of-fit testing, prediction in financial markets and outlier detection techniques for EEG signals are considered in [14], [36], and [13], respectively. Many other examples such as the life time of k-out-of-n systems or linear estimation based on order statistics can be found in standard references like [6], [7], and [11].

It is outlined in [37] that a preliminary analysis of independent realistic simulations of a string signal being of interest in radio interferometry allows to show that the random process from which a certain string signal arises is well modeled by generalized Gaussian distributions in wavelet space. The exponent parameters can be considered as a measure of compressibility of the underlying distribution. Values of these parameters close to zero yield very compressible distributions being of particular interest for compressed sensing imaging techniques in radio interferometry where astrophysical signals are probed through incomplete and noisy Fourier measurements. Similarly, generalized Gaussian distributions are used to model signal gradient in compressed sensing reconstruction of a string signal from interferometric observations of the cosmic microwave background in [38]. Moreover, the convergence of the so called peeling algorithm for wavelet denoising is proved in [23] under the usual assumption of signal processing that the wavelet coefficient of a signal from an independent and identically distributed (iid) family of generalized Gaussian variables. Gener-

***Corresponding Author: W.-D. Richter:** University of Rostock, Institute of Mathematics, Ulmenstraße 69, Haus 3, 18057 Rostock, Germany, E-mail: wolf-dieter.richter@uni-rostock.de
K. Müller: University of Rostock, Institute of Mathematics, Ulmenstraße 69, Haus 3, 18057 Rostock, Germany, E-mail: klaus.mueller@uni-rostock.de

alized Gaussian distributed variables are also used as balanced multiwavelet coefficients in digital image watermarking in [21], and in scalar quantization of transform coefficients obtained by modulated lapped transform or modified cosine transform filter bank in [22]. Heavy tailed distributions, examples of which include long memory processes being appropriate for financial time series or telecommunication traffic flows, can be modeled, according to [27], as a generalized Gaussian distribution having suitable variance and shape parameters. Due to the need of continuous monitoring of the respiratory mechanics during the acute exacerbations of chronic obstructive pulmonary disease, respiratory signal modeling under non-invasive ventilation is of interest in biomedical signal processing. Modeling of the measurement noise in the respiratory system has been done in [35] with the help of the generalized Gaussian distribution. This approach appeared there to be flexible and robust enough to deal with departs of the measurements noise from the Gaussian noise to the sub-Gaussian area.

Note that within the general study of order statistics from a finite number of, in general, dependent random variables special emphasis is in the literature on the particular case of extreme values and their distributions. In the background of related studies, the event in the sample space $\mathbb{R}^n$, $A_n^n(t) = \left\{(x_1, \ldots, x_n)^{\mathrm{T}} \in \mathbb{R}^n \colon x_i < t, i = 1, \ldots, n\right\}$, $t \in \mathbb{R}$, has been considered in different ways. While part of authors make (directly or indirectly) use of the single-out decomposition

$$A_n^n(t) = \bigcup_{i=1}^{n} \left\{(x_1, \ldots, x_n)^{\mathrm{T}} \in \mathbb{R}^n \colon x_{n:n} = x_i,\ x_i < t\right\} \tag{1}$$

where $x_{n:n}$ denotes the maximum of $x_1, \ldots, x_n$, the event $A_n^n(t)$ is represented in [26] for every $\nu \in \{1, \ldots, n-1\}$ as a union of $\nu + 1$ cones such that each of them, which is an intersection of n half spaces from $\mathbb{R}^n$, contains the origin in the boundary of ν of its intersecting half spaces. Invariance properties of $l_{n,p}$-symmetric measures allow then to represent the probability of $A_n^n(t)$ as the $(\nu + 1)$ multiple of the $l_{n,p}$-symmetric measure of an arbitrary element of a class $\mathcal{C}_{n-\nu,\nu}\left(E_1^{(\nu,n)}; t1_{n-\nu}\right)$ of cones being parameterized by the matrix $E_1^{(\nu,n)} \in \mathbb{R}^{\nu\times(n-\nu)}$ whose first column is $1_\nu = (1, \ldots, 1)^{\mathrm{T}} \in \mathbb{R}^\nu$ and whose remaining columns are equal the zero vector. Each such representation can be interpreted as a measure-of-cone representation of a certain skewed $l_{n-\nu,p}$-symmetric distribution generalizing the results for two- and multivariate skewed elliptically contoured distributions in [16] and [34], respectively.

Let us emphasize at this point that, on the one hand, the sets $\left\{(x_1, \ldots, x_n)^{\mathrm{T}} \in \mathbb{R}^n \colon x_{n:n} = x_i,\ x_i < t\right\}$ are not cones. On the other hand, studies making use of decomposition (1) nevertheless commonly also end up with representations of the distribution of the maximum of the components of the considered random vector as skewed distributions. However, these representations are different to that just discussed and they are only comparable if one considers there the particular case $\nu = n - 1$. Moreover, we remark that according to [5] a skewed $l_{n-\nu,p}$-symmetric distribution can be dealt with as the conditional distribution of $X^{(1)} = (X_1, \ldots, X_{n-\nu})^{\mathrm{T}}$ given the linear random selection condition $X^{(2)} < E_1^{(\nu,n)} X^{(1)}$ is satisfied where $X^{(2)} = (X_{n-\nu+1}, \ldots, X_n)^{\mathrm{T}}$, the sign of inequality is to be read componentwise, and the random variables $X_1, \ldots, X_n$ being generally not independent follow a joint $l_{n,p}$-symmetric distribution.

The sketched way of deriving exact extreme value distributions will be extended in the present paper to the case of arbitrary order statistics if a sample vector is $l_{n,p}$-dependent, i.e. it follows a joint $l_{n,p}$-symmetric distribution. Different methods of deriving exact distributions in this case may be distinguished with respect to how to (possibly indirectly) decompose the set to be considered in the sample space $\mathbb{R}^n$,

$$A_k^n(t) = \left\{(x_1, \ldots, x_n)^{\mathrm{T}} \in \mathbb{R}^n \colon x_i < t \text{ for at least } k \text{ values } i \in \{1, \ldots, n\}\right\},$$

when studying the distribution of the kth order statistic.

In [2, 3], the cumulative distribution function (cdf) and the probability density function (pdf) of linear combinations of order statistics of a finite number of arbitrary absolutely continuous dependent random variables are determined in terms of a product of marginal pdfs and conditional distributions where particularly the special cases of arbitrary exchangeable, of elliptically contoured, and of exchangeable elliptically contoured sample distributions are examined. Using exclusively skewed distributions to represent their results, [19] present the exact distribution of order statistics and of linear combinations of order statistics of a finite number of jointly elliptically contoured distributed random variables. For some more detailed information on this work see Section 3.1.1. The consideration of order statistics in $l_{n,p}$-symmetrically distributed populations started in [17] by obtaining the joint pdf of the vector of order statistics. A versatile study of different methods for deriving exact extreme value distributions in different ways can be found in [10] and [24–26]. More generally, [8] considers multivariate order statistics from dependent and nonidentically distributed random variables.

The present paper extends the results from [26] to the case of an arbitrary order statistic of n jointly $l_{n,p}$-symmetrically distributed random variables, and at the same time extends certain results in [2, 3] and [19] from spherical dependence to the case of $l_{n,p}$-dependence. To pursue this aim, on the one hand, we follow a single-out decomposition of the random event $A_k^n(t)$ as it is applied in [19] and several references cited therein. On the other hand, we follow a cone decomposition of $A_k^n(t)$ being analogously to that of $A_n^n(t)$ mentioned above, and employ an advanced geometric method by using measure-of-cone representations of skewed $l_{n,p}$-symmetric distributions derived as well as applied in [26]. One might ask why it is useful to have different representations of one and the same cdf. A part of the answer is that sometimes different representations have different numerical and mathematical properties. For another part of the answer, we refer to [16] where it is shown that four differently motivated approaches to the skewed normal distribution can be unified from a geometric point of view. Now, given this unified approach, it has the potential property of helping to identify otherwise motivated applications of skewed distributions.

The rest of the paper is organized as follows. In Section 2, we introduce the families of $l_{n,p}$-symmetric and skewed $l_{n,p}$-symmetric distributions. In Sections 3.1.1 and 3.1.2, we review known results on the distribution of order statistics for spherical and power exponential sample distribution, respectively. In Sections 3.2 and 3.3, we present our main results on exact distributions of order statistics from continuous $l_{n,p}$-symmetric sample distributions based upon the two different approaches of single-out and cone decompositions of the event $A_k^n(t)$, $t \in \mathbb{R}$, to be measured when deriving the cdf of the kth order statistic. Additionally, some visualizations and an application are shown in Section 3.4. In order to strengthen the presentation of our main results, their proofs and some discussions are given separately in Section 4 and the derivation of the pdf concerning the cdf resulting from the second approach and its proof is outsourced to Appendices A.1 and A.2. In particular, in Section 4.5, we discuss a second way of achieving our main results based upon the single-out decompositions. In Section 5, some conclusions are drawn from the present paper and further scopes are outlined.

2 The classes of $l_{n,p}$-symmetric and skewed $l_{n,p}$-symmetric distributions

Let p be a positive real number and denote $|\cdot|_p$ the p-functional in the n-dimensional Euclidean space $\mathbb{R}^n$ defined by $|x|_p = \left(\sum_{i=1}^{n} |x_i|^p\right)^{\frac{1}{p}}$, $x = (x_1, \ldots, x_n)^{\mathrm{T}} \in \mathbb{R}^n$. Additionally, let us denote the $l_{n,p}$-generalized surface content of the $l_{n,p}$-unit sphere $S_{n,p} = \{x \in \mathbb{R}^n\colon |x|_p = 1\}$ by $\omega_{n,p}$. For the local definition and calculation of the value $\omega_{n,p} = \frac{\left(2\Gamma\left(\frac{1}{p}\right)\right)^n}{p^{n-1}\Gamma\left(\frac{n}{p}\right)}$ see [29], and for a global or differential geometric approach to this notion see [30]. We remark that extensions of this basic notion to ellipsoids, norm and antinorm spheres as well as to more general star spheres are given in [31–33], respectively.

A function $g\colon (0,\infty) \to (0,\infty)$ satisfying $0 < I_n(g) < \infty$ is called a density generating function (dgf) of an n-variate distribution where $I_n(g) = \int_0^\infty r^{n-1} g(r)\, dr$. For achieving uniqueness of this notion we assume that $I_n(g) = \frac{1}{\omega_{n,p}}$. Such dgf is called a density generator (dg) and denoted by $g^{(n)}$.

An n-dimensional random vector $X\colon \Omega \to \mathbb{R}^n$ defined on a probability space $(\Omega, \mathfrak{A}, P)$ and having the pdf

$$\varphi_{g^{(n)},p}(x) = g^{(n)}\left(|x|_p\right), \quad x \in \mathbb{R}^n, \tag{2}$$

is said to follow a continuous $l_{n,p}$-symmetric distribution with dg $g^{(n)}$ and shape parameter $p > 0$ (also tail parameter). Particularly, note that the class of $l_{n,2}$-symmetric distributions, $n \in \mathbb{N}$, coincides with that of spherical distributions. Further, the probability law of X is denoted by $\Phi_{g^{(n)},p}$ and the components of X are uncorrelated but generally dependent. In particular, they are independent if and only if X is n-dimensional p-generalized Gaussian distributed. That is an important special case of $l_{n,p}$-symmetric distributions which is realized by choosing the dg $g^{(n)}$ as $g_{PE}^{(n)}$ where

$$g_{PE}^{(n)}(r) = \left(\frac{p^{1-\frac{1}{p}}}{2\Gamma\left(\frac{1}{p}\right)}\right)^n \exp\left\{-\frac{r^p}{p}\right\}, \quad r > 0. \tag{3}$$

Concerning these and more details on $l_{n,p}$-symmetric distributions, references to the literature and examples of dgs, we refer to [17], [30] and [25, 26], respectively, and references cited therein.

To introduce the skewed $l_{k,p}$-symmetric distributions, let I_m be the $m \times m$ unit matrix and 0_m the zero vector in $\mathbb{R}^m$. Furthermore, let $X = \left(X^{(1)^{\mathrm{T}}}, X^{(2)^{\mathrm{T}}}\right)^{\mathrm{T}}$ be a random vector following a continuous $l_{k+m,p}$-symmetric distribution with dg $g^{(k+m)}$ where $X^{(1)}: \Omega \to \mathbb{R}^k$ and $X^{(2)}: \Omega \to \mathbb{R}^m$. Recalling the minor change of notation in [32] compared to earlier publications, described for the particular case of $l_{n,p}$-symmetric distributions in [25], and taking this subsequently into account, according to [5], the dg $g_{(k+m)}^{(k)}$ of the marginal distribution of $X^{(1)}$ in $\mathbb{R}^k$ allows the representation

$$g_{(k+m)}^{(k)}(z) = \frac{\omega_{m,p}}{p} \int_{z^p}^{\infty} (y - z^p)^{\frac{m}{p}-1} g^{(k+m)}(\sqrt[p]{y})\, dy, \quad z \in (0, \infty). \tag{4}$$

Additionally, for $\Lambda \in \mathbb{R}^{m \times k}$, $\Gamma = (\Lambda, -I_m)$, and $\Sigma = \Gamma\Gamma^{\mathrm{T}} = I_m + \Lambda\Lambda^{\mathrm{T}}$, the cdf of ΓX will be denoted by $F_{m,p}^{(2)}\left(x; \Sigma, g_{(k+m)}^{(m)}\right)$, $x \in \mathbb{R}^m$. Moreover, for every $x^{(1)} \in \mathbb{R}^k$, the conditional density of $X^{(2)}$ given $X^{(1)} = x^{(1)}$ is

$$\frac{g^{(k+m)}\left(\sqrt[p]{|x^{(1)}|_p^p + |x^{(2)}|_p^p}\right)}{g_{(k+m)}^{(k)}(|x^{(1)}|_p)} = g_{[|x^{(1)}|_p]}^{(m)}(|x^{(2)}|_p), \quad x^{(2)} \in \mathbb{R}^m, \tag{5}$$

and the corresponding distribution law is $\Phi_{g_{[|x^{(1)}|_p]}^{(m)},p}$. Let Y be a random vector following this distribution, $Y \sim \Phi_{g_{[|x^{(1)}|_p]}^{(m)},p}$, then its cdf satisfies

$$F_{m,p}^{(1)}\left(x; g_{[|x^{(1)}|_p]}^{(m)}\right) = \int_{\mathbb{R}_+^m} g_{[|x^{(1)}|_p]}^{(m)}\left(|x-u|_p\right) du = \int_{v<x} g_{[|x^{(1)}|_p]}^{(m)}(|v|_p)\, dv, \quad x \in \mathbb{R}^m.$$

A k-dimensional random vector Z having a pdf of the form

$$f_Z(z) = \frac{1}{F_{m,p}^{(2)}(0_m; \Sigma, g_{(k+m)}^{(m)})} g_{(k+m)}^{(k)}\left(|z|_p\right) F_{m,p}^{(1)}\left(\Lambda z; g_{[|z|_p]}^{(m)}\right), \quad z \in \mathbb{R}^k, \tag{6}$$

is said to follow the skewed $l_{k,p}$-symmetric distribution $SS_{k,m,p}\left(\Lambda, g^{(k+m)}\right)$ with dimensionality parameter m, dg $g^{(k+m)}$ and skewness/ shape matrix-parameter Λ. Further, the parameter k is called the co-dimensionality parameter and the cdf of Z is denoted by $F_{k,m,p}\left(\cdot; \Lambda, g^{(k+m)}\right)$. If Σ is a diagonal matrix, the normalizing constant in (6) is $F_{m,p}^{(2)}\left(0_m; \Sigma, g_{(k+m)}^{(m)}\right) = 2^{-m}$. In the present framework, however, it will not be necessary to determine this constant explicitly even for rather arbitrary matrix Σ.

As skewed $l_{k,p}$-symmetric distributions are constructed via selection mechanisms from $l_{k+m,p}$-symmetric distributions in [5], using the above introduced notations,

$$\mathfrak{L}\left(X^{(1)} \,\middle|\, X^{(2)} < \Lambda X^{(1)}\right) = SS_{k,m,p}\left(\Lambda, g^{(k+m)}\right) \tag{7}$$

where $\mathfrak{L}(Y)$ denotes the distribution law of the random vector Y.

Let us denote the set of all $n \times n$ permutations matrices by Π_n. Note that, according to [26], if $Z \sim SS_{k,m,p}\left(\Lambda, g^{(k+m)}\right)$ and $M_1 \in \Pi_n$, then $M_1 Z \sim SS_{k,m,p}\left(\Lambda M_1^{\mathrm{T}}, g^{(k+m)}\right)$ where ΛM_1^{T} arises from Λ by interchanging columns. Moreover, if $M_2 \in \Pi_m$, then, for $z \in \mathbb{R}^k$, $F_{k,m,p}\left(z; M_2\Lambda, g^{(k+m)}\right) = F_{k,m,p}\left(z; \Lambda, g^{(k+m)}\right)$, i.e. $SS_{k,m,p}\left(M_2\Lambda, g^{(k+m)}\right) = SS_{k,m,p}\left(\Lambda, g^{(k+m)}\right)$ where $M_2\Lambda$ arises from Λ by interchanging rows.

3 Distributions of order statistics from continuous $l_{n,p}$-symmetric sample distributions

3.1 Review of results from the literature

3.1.1 Spherically distributed sample distribution

In this section, we concisely review specific results on exact distributions of order statistics from the literature especially concentrating on spherical sample distributions and on the work [2, 3] and [19]. As it is mentioned in Section 2, the class of $l_{n,2}$-symmetric distributions, $n \in \mathbb{N}$, coincides with that of spherical distributions. Hence, our present task to determine the exact distributions of order statistics from $l_{n,p}$-symmetric sample distributions is already done for the special case $p = 2$ in the three papers mentioned above. Moreover, spherical distributions are centered elliptically contoured distributions with unit matrices as dispersion matrices, i.e. exchangeable elliptically contoured distributions with parameters $\mu = 0$, $\sigma = 1$ and $\rho = 0$ in the notation of the previous works. Let us denote the pdf and the cdf of an n-dimensional spherical distribution with density generator $h^{(n)}$ by $f_S(\cdot; h^{(n)})$ and $F_S(\cdot; h^{(n)})$, respectively. The pdf $f^{\circ}_{k:n}$ of the kth order statistic from an n-dimensional spherical distribution with functional parameter $h^{(n)}$ according to [2, page 1889], [3] and [19, Remark 5] (with results of both papers being specialized to the case $\mu = 0$, $\sigma = 1$ and $\rho = 0$) is

$$f^{\circ}_{k:n}(t) = \binom{n}{k} k\, f_S(t; h^{(1)})\, F_S(J_{n-1}t; h^{(n-1)}_{t^2}), \quad t \in \mathbb{R}, \tag{8}$$

where $J_{n-1} = \left(1^{\mathrm{T}}_{k-1}, -1^{\mathrm{T}}_{n-k}\right)^{\mathrm{T}} \in \mathbb{R}^{n-1}$ with $1_m = (1, \dots, 1)^{\mathrm{T}} \in \mathbb{R}^m$,

$$h^{(1)}(u) = \frac{\pi^{\frac{n-1}{2}}}{\Gamma\left(\frac{n-1}{2}\right)} \int_0^\infty u^{\frac{n-1}{2}-1} h^{(n)}(u+v)\, dv, \quad u > 0,$$

is an univariate marginal density generator and

$$h^{(n-1)}_{t^2}(u) = \frac{h^{(n)}(u+t^2)}{h^{(1)}(t^2)}, \quad u > 0,$$

an $(n-1)$-dimensional conditional density generator. Note that $h^{(1)}$ and $h^{(n-1)}_{t^2}$ are special cases of the marginal and conditional dgs (4) and (5), respectively, given in Section 2 having in mind the minor change of notation in [32].

Arellano-Valle and Genton achieve the above formula (8) by certain steps of specialization. Their most general results deals with the pdf of linear combinations of order statistics from arbitrary absolutely continuous sample distributions which is represented as a sum of products whose factors each are a specific marginal pdf multiplied with a specific conditional distribution. Note that these conditional distributions are different from that used in the definitions of skewed distributions in [19] and Section 2, respectively, and that their determination may be nontrivial. Next, on the one hand, this result is specified to the case of arbitrary exchangeable absolutely continuous sample distributions where the connection of distributions of linear combinations of order statistics from such sample distributions to fundamental skew distributions, see [1], is stated. On the other hand, the special case of elliptically contoured sample distributions is considered there for which the conditional distributions are calculated explicitly yielding specific elliptically contoured distributions with conditional functional parameters. Note that the corresponding pdf of a linear combination of order statistics from elliptically contoured distribution was corrected by [19, Remark 10]. Finally, both specializations are simultaneously considered resulting in a representation of the pdf of a linear combination of order statistics from exchangeable elliptically contoured sample distribution having parameters $\mu \in \mathbb{R}$, σ^2 with $\sigma > 0$ and $\rho \in [0, 1)$ and a certain functional parameter. The latter representation yields (8) by specializing the linear combination to a single order statistic and considering $\mu = 0$, $\sigma = 1$ and $\rho = 0$.

After considering various special cases in their previous work, [19] deal with distributions of order statistics and linear combination of order statistics from elliptically contoured sample distributions in term of mixtures skewed distributions. To be concrete, these are unified multivariate skew-elliptically contoured distributions being defined in this paper as specific conditional distributions and being special cases of multivariate unified skew-elliptically contoured distributions, see [4]. As it is indicated before, these conditional distributions differ from those considered in [2, 3]. With the help of these skewed distributions, on the one hand, it is shown in [19] that the distribution of an arbitrary order statistic from an elliptically contoured sample distribution is unified univariate skew-elliptically contoured distributed where the corresponding parameters are given. This result is specialized to exchangeable elliptically contoured sample distributions yielding (8) if $\mu = 0$, $\sigma = 1$, and $\rho = 0$ as well as to the cases of multivariate normal and t sample distribution, respectively. On the other hand, the distribution of an linear combination of order statistics from an elliptically contoured sample distribution is established to be a mixture of unified multivariate skew-elliptically contoured distributions having specific parameters. Note that this result, combined with Remark 10 in [19], coincides with that in [2] if the skewed distributions are dissolved. Finally, the three special cases of exchangeable elliptically contoured, multivariate normal and multivariate t sample distributions are discussed in [19].

3.1.2 Power exponential sample distribution

In this section, let the random vector $\xi = (\xi_1, \dots, \xi_n)^{\mathrm{T}}$ follow the n-dimensional p-generalized Gaussian distribution, $\xi \sim \Phi_{g_{PE}^{(n)},p}$. It is already mentioned in Section 2 that the components of an continuous $l_{n,p}$-symmetrically distributed random vector having dg $g^{(n)}$ are independent if $g^{(n)} = g_{PE}^{(n)}$ and only in that case. Furthermore, they are identically distributed due to their $l_{1,p}$-symmetrically distributed univariate marginals with dg $g_{(n)}^{(1)}$ according to [5] and to permutation invariance of $\Phi_{g^{(n)},p}$ according to [26].

The cdf $F_{k:n}^{\star}$ and the pdf $f_{k:n}^{\star}$ of the kth order statistic of $X_1^{\star}, \dots, X_n^{\star}$ being i.i.d. random variables having cdf F and pdf f, respectively, are

$$F_{k:n}^{\star}(t) = \binom{n}{k} k \int_{-\infty}^{t} f(y) \left(F(y)\right)^{k-1} \left(1 - F(y)\right)^{n-k} \, dy \tag{9}$$

$$= \sum_{j=k}^{n} \binom{n}{j} \left(F(t)\right)^{j} \left(1 - F(t)\right)^{n-j}, \quad t \in \mathbb{R}, \tag{10}$$

and

$$f_{k:n}^{\star}(t) = \binom{n}{k} k f(t) \left(F(t)\right)^{k-1} \left(1 - F(t)\right)^{n-k}, \quad t \in \mathbb{R}, \tag{11}$$

see [11]. We note that there are two representations of $F_{k:n}^{\star}$ being structured differently.

Let us denote the pdf and the cdf of the univariate distribution of ξ_1 by $\varphi_p(t) = g_{PE}^{(1)}(|t|)$ and $\Phi_p(t) = \int_{-\infty}^{t} \varphi_p(s)\, ds$, respectively. Then, because of continuity and symmetry with respect to the point $\left(0, \frac{1}{2}\right)$, we have $1 - \Phi_p(t) = \Phi_p(-t)$ and the pdf of the kth order statistic of the components of ξ follows from (11) with $F = \Phi_p$

and $f = \varphi_p$,

$$f^{\star}_{k:n}(t) = \binom{n}{k} k\, \varphi_p(t) \left(\Phi_p(t)\right)^{k-1} \left(\Phi_p(-t)\right)^{n-k}, \quad t \in \mathbb{R}. \tag{12}$$

For visualizations of $f^{\star}_{4:4}$ and $f^{\star}_{3:4}$ for different values of $p > 0$ we refer to Figures 5 and 6. Additionally, analog representations can be achieved for the cdf of the kth order statistic of the components of ξ according to (9) and (10), respectively.

In the next Sections 3.2 and 3.3, we provide our main results (Theorem 1 and Theorem 2).

3.2 The general case of $l_{n,p}$-dependence making use of single-out decompositions of $A^n_k(t)$

For the rest of the paper, let the random vector $X = (X_1, \dots, X_n)^{\mathrm{T}}$ follow the $l_{n,p}$-symmetric distribution with an arbitrary dg $g^{(n)}$, $X \sim \Phi_{g^{(n)},p}$, and denote the cdf and the pdf of the kth order statistic of $X_1, \dots, X_n$, $X_{k:n}$, by $F_{k:n}$ and $f_{k:n}$, respectively, $k = 1, \dots, n$. Moreover, we put $E^{(n-1)} = 1_{n-1} 1^{\mathrm{T}}_{n-1}$ where $1_{n-1} = (1, \dots, 1)^{\mathrm{T}} \in \mathbb{R}^{n-1}$.

Theorem 1. *For $k \in \{1, \dots, n\}$ and every set $J \subseteq \{1, \dots, n-1\}$ consisting of $k-1$ (different) elements, $F_{k:n}$ allows the representation*

$$F_{k:n}(t) = \binom{n}{k} k\, F^{(2)}_{n-1,p}\left(0_{n-1}; I_{n-1} + S^{(n-1)}_J E^{(n-1)} S^{(n-1)}_J, g^{(n-1)}_{(n)}\right) \cdot F_{1,n-1,p}\left(t; S^{(n-1)}_J 1_{n-1}, g^{(n)}\right), \quad t \in \mathbb{R},$$

where $S^{(n-1)}_J = \mathrm{diag}(s_1, \dots, s_{n-1})$ is a diagonal matrix with $s_j = \begin{cases} 1 & , j \in J \\ -1 & , \text{otherwise} \end{cases}$.

The following corollary is an immediate consequence of Theorem 1.

Corollary 1. *For $k \in \{1, \dots, n\}$ and every set $J \subseteq \{1, \dots, n-1\}$ having $k-1$ elements, $f_{k:n}$ allows the representation*

$$f_{k:n}(t) = \binom{n}{k} k\, g^{(1)}_{(n)}(|t|)\, F^{(1)}_{n-1,p}\left(S^{(n-1)}_J 1_{n-1} t; g^{(n-1)}_{[|t|]}\right), \quad t \in \mathbb{R}. \tag{13}$$

The above representation of the distribution of the kth order statistic conforms well with earlier results in the literature. Note that the representation of $f_{2:2}$ is derived first in [10]. Moreover, the general results of Theorem 1 and Corollary 1 coincide in the case $k = n$ with those derived in [26] if one puts there $\nu = n-1$. For the particular case $p = 2$ and the particular choice of diagonal matrix $S^{(n-1)}_J$ such that $S^{(n-1)}_J 1_{n-1} = \left(1^{\mathrm{T}}_{k-1}, -1^{\mathrm{T}}_{n-k}\right)^{\mathrm{T}} = J_{n-1}$, the result of Corollary 1 coincides with (8) and thus is covered by [2, 3] and [19], respectively. Aside from that [19] also proved Theorem 1 in the special case of $p = 2$ but allowing more general moments.

Notice that the result of Theorem 1 can be reformulated without referring to skewed distributions as

$$F_{k:n}(t) = \binom{n}{k} k \int_{-\infty}^{t} g^{(1)}_{(n)}\left(|y|_p\right) F^{(1)}_{n-1,p}\left(S^{(n-1)}_J 1_{n-1} y; g^{(n-1)}_{[|y|_p]}\right) dy, \quad t \in \mathbb{R}, \tag{14}$$

where $g^{(1)}_{(n)}$ is the one-dimensional marginal dg of any component of X and $F^{(1)}_{n-1,p}$ is the $(n-1)$-variate cdf defined in Section 2. In some sense, the mathematical structure behind the representations (14) and (13) is thus close to that behind (9) and (11), respectively.

3.3 General representations in case of $l_{n,p}$-dependence based upon cone decompositions of $A_k^n(t)$

It was discussed in Section 1 that different decompositions of the event $A_n^n(t)$ lead to different representations of $F_{n:n}(t)$, $t \in \mathbb{R}$. Similarly, $F_{k:n}(t)$, $t \in \mathbb{R}$, allows additional representations if one uses cone decompositions of $A_k^n(t)$ instead of single-out ones. To this end, for $\nu \in \{1, \dots, n-1\}$ recalling the notation $E_1^{(\nu,n)}$ of the $\nu \times (n-\nu)$ matrix whose first column is 1_ν and whose remaining elements are equal to zero, see Section 1, for $\nu_1 \in \{1, \dots, n-k-1\}$ and $\nu_2 \in \{1, \dots, k-1\}$ with $\nu_1 + \nu_2 > 0$, let

$$E_{1,1}^{(\nu_1,\nu_2,n,k)} = \begin{pmatrix} E_1^{(\nu_1,n-k)} & 0 \\ 0 & E_1^{(\nu_2,k)} \end{pmatrix} \in \mathbb{R}^{(\nu_1+\nu_2)\times(n-\nu_1-\nu_2)}$$

be a block matrix with 0 denoting each time a suitably sized zero matrix. If $\nu_1 = 0$ or $\nu_2 = 0$, then the first or second block row in $E_{1,1}^{(\nu_1,\nu_2,n,k)}$ is omitted, respectively. Moreover, $E^{(\nu_1,\nu_2)} = \begin{pmatrix} E^{(\nu_1)} & 0 \\ 0^{\mathrm{T}} & E^{(\nu_2)} \end{pmatrix} \in \mathbb{R}^{(\nu_1+\nu_2)\times(\nu_1+\nu_2)}$ denotes a block (diagonal) matrix. Additionally, in the sequel, we define the sum $\sum_{j=n}^{n-1}$ to be zero.

Theorem 2. *For $k \in \{1, \dots, n\}$, every $\nu_{j,1} \in \{0, \dots, n-j-1\}$ and $\nu_{j,2} \in \{0, \dots, j-1\}$ with $\nu_j = \nu_{j,1}+\nu_{j,2} > 0$ for $j = k, \dots, n-1$, and every $\nu_n \in \{1, \dots, n-1\}$, the cdf $F_{k:n}$ allows the representation*

$$\begin{aligned} F_{k:n}(t) = \sum_{j=k}^{n-1} \binom{n}{j} (\nu_{j,1}+1)(\nu_{j,2}+1) F^{(2)}_{\nu_j,p}\left(0_{\nu_j}; I_{\nu_j} + E^{(\nu_{j,1},\nu_{j,2})}, g^{(\nu_j)}_{(n)}\right) \\ \cdot F_{n-\nu_j,\nu_j,p}\left(\begin{pmatrix} -t1_{n-j-\nu_{j,1}} \\ t1_{j-\nu_{j,2}} \end{pmatrix}; E_{1,1}^{(\nu_{j,1},\nu_{j,2},n,j)}, g^{(n)}\right) \\ + (\nu_n+1) F^{(2)}_{\nu_n,p}\left(0_{\nu_n}; I_{\nu_n} + E^{(\nu_n)}, g^{(\nu_n)}_{(n)}\right) F_{n-\nu_n,\nu_n,p}\left(t1_{n-\nu_n}; E_1^{(\nu_n,n)}, g^{(n)}\right), \quad t \in \mathbb{R}. \end{aligned}$$

If $k = n$, then Theorems 1 and 2 lead to the same results if one puts $\nu_n = n-1$ in Theorem 2. Additionally, note that the last summand in the representation of $F_{k:n}(t)$ in Theorem 2 is already well known from [26] to be the maximum cdf in $l_{n,p}$-symmetrically distributed populations. We recall that the representation of $F_{k:n}$ given in Theorem 1 makes use of a skewed distribution with dimensionality parameter $n-1$ and co-dimensionality parameter 1. The result of Theorem 2 means that $F_{k:n}$ can also be represented as a mixture of $n-k+1$ skewed $l_{n-\nu_j,p}$-symmetric distribution functions with dimensionality parameters ν_j each being smaller than $n-1$ for $j = k, \dots, n-1$ and ν_n (and co-dimensionality parameters $n-\nu_j$ as well as matrix parameter Λ_j for $j = k, \dots, n$, respectively). Note that $\Lambda_j = E_{1,1}^{(\nu_{j,1},\nu_{j,2},n,j)}$ if $j \in \{k, \dots, n-1\}$ and $\Lambda_n = E_1^{(\nu_n,n)}$. Wherever, vice versa, one meets elsewhere some quantity expressed as such a mixture of skewed distributions one thus may think there about a comprehensive representation in terms of a single skewed distribution similar to the one given in Theorem 1.

Furthermore, notice that the mathematical structure of the representation given in Theorem 2 is in a way linked to that of (10), since the ith summand of both formulae represents the probability that exactly $i+k-1$ of the random variables $X_1, \dots, X_n$ and $X_1^\star, \dots, X_n^\star$ are less than t, $i = 1, \dots, n-k+1$, respectively.

Because of symmetry with respect to the origin and continuity of the probability distribution of X, $\Phi_{g^{(n)},p}$, for $k = 1, \dots, n$ and every $t \in \mathbb{R}$, the well-known relation

$$F_{n-k+1:n}(t) = 1 - F_{k:n}(-t), \quad t \in \mathbb{R}, \tag{15}$$

is fulfilled. This relation may be used to reduce the number of summands when applying the representation of Theorem 2. More precisely, denoting the floor and the ceiling function by $\lfloor\cdot\rfloor$ and $\lceil\cdot\rceil$, respectively, if $1 \le k \le \lfloor \frac{n+1}{2} \rfloor$, then, using first relation (15) and afterwards Theorem 2, the cdf consists of $n-(n-k+1)+1 = k \le \lfloor \frac{n+1}{2} \rfloor$ summands whereas the cdf from the straightforward use of Theorem 2 consists of $n-k+1 \ge n+1-\lfloor \frac{n+1}{2} \rfloor \ge \lfloor \frac{n+1}{2} \rfloor$ summands. In order to reduce both the computational expense and time, this fact may be succesfully used when implementing $F_{k:n}$ as given in Theorem 2.

Moreover, notice that, according to the representation of the cdf $F_{k:n}$ from Theorem 2, the pdf of the kth order statistic $X_{k:n}$, $f_{k:n}(t)$, $t \in \mathbb{R}$, is considered for $k = 1, \dots, n$ in Appendix A.

At the end of this section, in view of what was said in Section 1 on the potential usefulness of a variety of representations of one and the same numerical quantity, we present further representations of $F_{k:n}$ according to the approach via cone decompositions of the event $A_k^n(t)$, $t \in \mathbb{R}$. In order to do this, extending the notation of the matrix $E_1^{(\nu,n)}$, for $i = 1, \dots, n-\nu$, let $E_i^{(\nu,n)}$ be the $\nu \times (n-\nu)$ matrix whose ith column is 1_ν and whose remaining columns are 0_ν. Analogously, for $i_1 = 1, \dots, n-k-\nu_1$ and $i_2 = 1, \dots, k-\nu_2$, the block matrices

$$E_{i_1,i_2}^{(\nu_1,\nu_2,n,k)} = \begin{pmatrix} E_{i_1}^{(\nu_1,n-k)} & 0 \\ 0 & E_{i_2}^{(\nu_2,k)} \end{pmatrix} \in \mathbb{R}^{(\nu_1+\nu_2)\times(n-\nu_1-\nu_2)}$$

extend the abvoe notation of $E_{1,1}^{(\nu_1,\nu_2,n,k)}$. Recalling the notation Π_n of the set of all $n \times n$ permutation matrices, Theorem 2 is a special case of the following theorem.

Theorem 3. *For $k \in \{1, \dots, n\}$, every $\nu_{j,1} \in \{0, \dots, n-j-1\}$ and $\nu_{j,2} \in \{0, \dots, j-1\}$ such that $\nu_j = \nu_{j,1} + \nu_{j,2} > 0$, every $i_{j,1} \in \{1, \dots, n-j-\nu_{j,1}\}$ and $i_{j,2} \in \{1, \dots, j-\nu_{j,2}\}$, every $M_{j,1} \in \Pi_{n-\nu_j}$ and $M_{j,2} \in \Pi_{\nu_j}$ for $j = 1, \dots, n-1$, and every $\nu_n \in \{1, \dots, n-1\}$ and $i_n \in \{1, \dots, n-\nu_n\}$, the cdf $F_{k:n}$ allows the representation*

$$\begin{aligned} F_{k:n}(t) = \sum_{j=k}^{n-1} \binom{n}{j} (\nu_{j,1}+1)(\nu_{j,2}+1) F^{(2)}_{\nu_j,p}\left(0_{\nu_j}; I_{\nu_j} + E^{(\nu_{j,1},\nu_{j,2})}, g^{(\nu_j)}_{(n)}\right) \\ \cdot F_{n-\nu_j,\nu_j,p}\left(M_{j,1}\begin{pmatrix} -t1_{n-j-\nu_{j,1}} \\ t1_{j-\nu_{j,2}} \end{pmatrix}; M_{j,2}E^{(\nu_{j,1},\nu_{j,2},n,j)}_{i_{j,1},i_{j,2}} M^{\mathrm{T}}_{j,1}, g^{(n)}\right) \\ + (\nu_n+1) F^{(2)}_{\nu_n,p}\left(0_{\nu_n}; I_{\nu_n} + E^{(\nu_n)}, g^{(\nu_n)}_{(n)}\right) F_{n-\nu_n,\nu_n,p}\left(t1_{n-\nu_n}; E^{(\nu_n,n)}_{i_n}, g^{(n)}\right), \quad t \in \mathbb{R}. \end{aligned}$$

3.4 Visualization and an application

In the present section, on the one hand, we visualize the pdf of the maximum statistic and the third order statistic from $l_{4,3}$-symmetric Kotz type and Pearson Type VII sample distributions, respectively, see Figures 1-4. For the definitions of the corresponding dgs $g^{(4)}_{Kt;M,\beta,\gamma}$ and $g^{(4)}_{PT7;M,\nu}$ of these subclasses of continuous $l_{4,p}$-symmetric distributions, we refer to [26]. Note that the choice of parameters M, β, γ in Figures 1 and 2 and M, ν in Figures 3 and 4 coincides with that in Figures 2d and 3d in [25] (where the median is considered for sample size three) as well as with that in Figures 4d and 5d in [26] (where the maximum is considered for sample size three), respectively. We recall that p-generalized Gaussian distributions are particular cases of the class of light tailed Kotz type distributions and Student-t as well as Cauchy distributions belong to the family of heavy tailed Pearson Type VII distributions. It turns out that the corresponding graphs of $f_{3:4}$ and $f_{4:4}$ seem to be more similar to each other if the sample vector distribution is heavy tailed than if it is light tailed. For a study of certain domain quantiles of $l_{n,p}$-symmetric distributions, $n \in \{1, 2, 3\}$, we refer to [26].

On the other hand, assume data suggest that, as appropriate after centering and standardization, the largest wavelet coefficient of each of n signals, $n \in \{2, 3, \dots\}$, follows one and the same symmetric distribution having a pdf $Dg(|x|)$, $x \in \mathbb{R}$, for some $D > 0$. Moreover, assume that an additional visual inspection of the level sets of their joint density indicates a joint pdf

$$C_n h\left(|x|_p\right) = C_n h\left(\left(\sum_{i=1}^n |x_i|^p\right)^{1/p}\right), \quad x = (x_1, \dots, x_n)^{\mathrm{T}} \in \mathbb{R}^n,$$

could be true for some $p > 0$, a function $h\colon [0,\infty) \to [0,\infty)$ satisfying $0 < \int_0^\infty r^{n-1}h(r)\,dr < \infty$, and a suitable normalizing constant $C_n > 0$. Then, an independence assumption with respect to the n signals is supported by the data if and only if h can be chosen as $h(r) = e^{-\frac{r^p}{p}}$ and in consequence $g(r) = e^{-\frac{r^p}{p}}$. Let us consider the $n = 4$ largest wavelet coefficients in an independence model for different values of p. Figures 5 and 6 then show the density $f^{\star}_{4:4}$ of the maximum and the density $f^{\star}_{3:4}$ of the second largest order statistic, respectively. For comparison, in the special case of $p = 3$ we refer again to the specific visualizations of $f_{4:4}$ and $f_{3:4}$ in light tailed dependence models of Kotz type in Figures 1 and 2 and in heavy tailed dependence models of Pearson Type VII in Figures 3 and 4.

Summarizing this section, numerous comparisons regarding these visualizations can be made. First, Figures 1–4 show the effects of different choices of distributional parameters on the pdf of the maximum statistic and the second largest order statistic in $l_{4,3}$-symmetric Kotz type and Pearson Type VII modelings, respec-

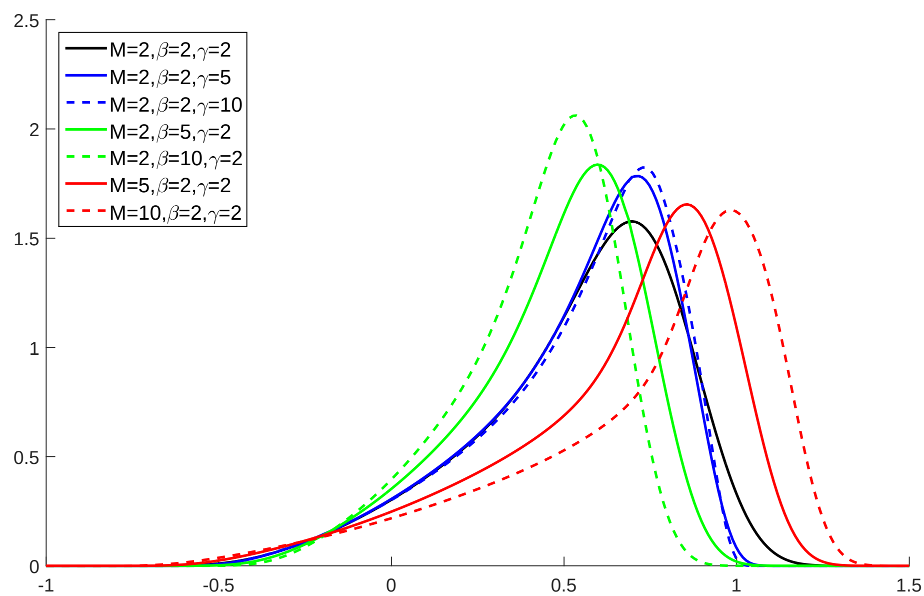

Figure 1: Maximum pdf from $l_{4,3}$-symmetric sample vector distribution, $f_{4:4}$, if the dg is $g^{(4)}_{Kt;M,\beta,\gamma}$.

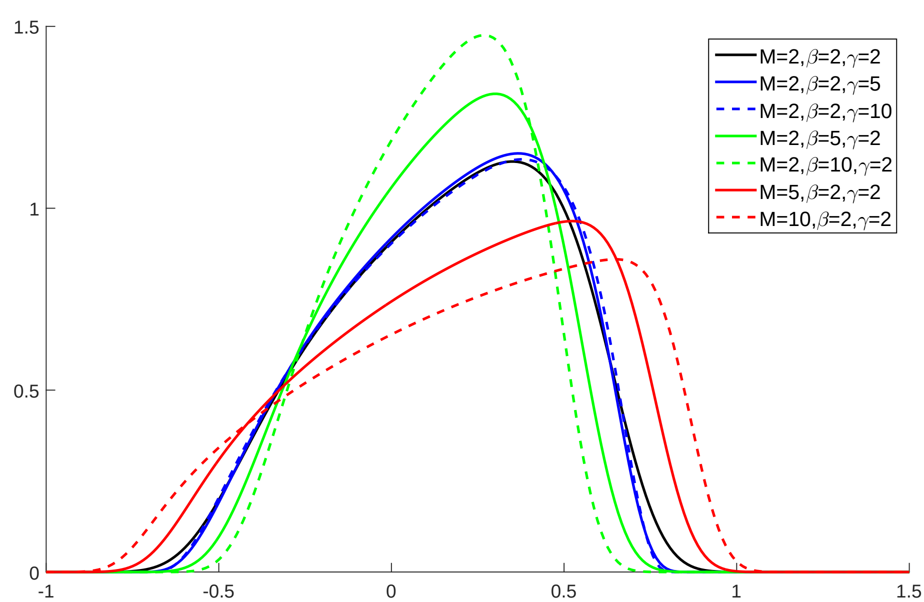

Figure 2: Pdf of the third order statistic from $l_{4,3}$-symmetric sample vector distribution, $f_{3:4}$, if the dg is $g^{(4)}_{Kt;M,\beta,\gamma}$.

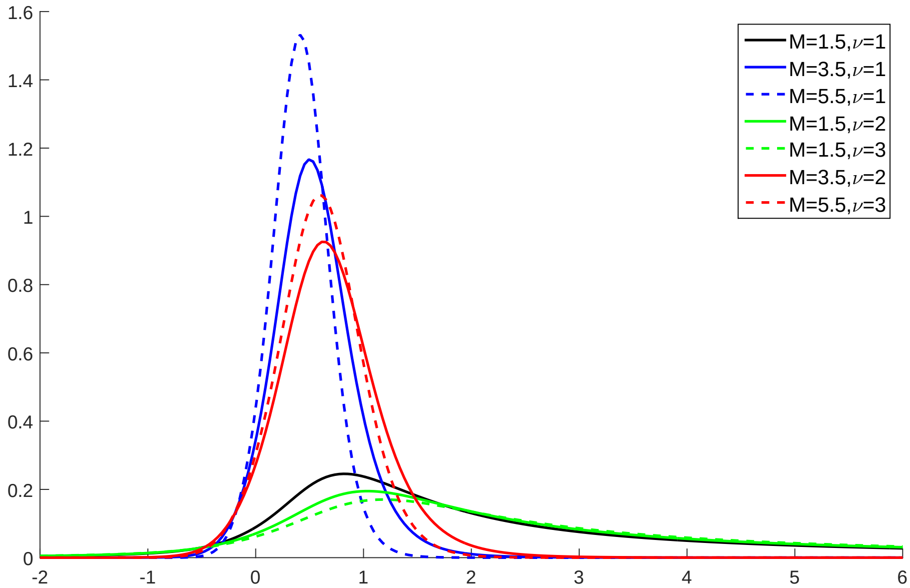

Figure 3: Maximum pdf from $l_{4,3}$-symmetric sample vector distribution, $f_{4:4}$, if the dg is $g^{(4)}_{PT7;M,\nu}$.

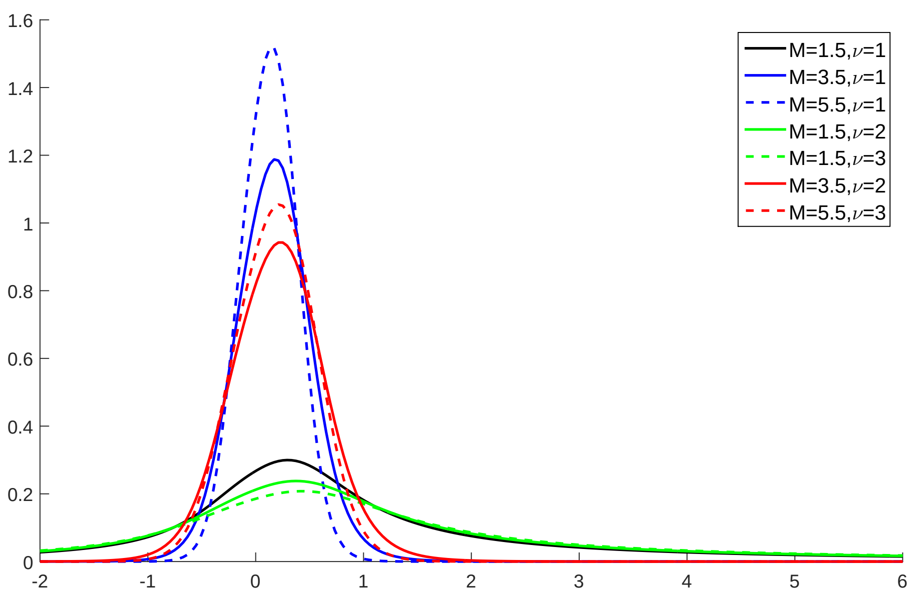

Figure 4: Pdf of the third order statistic from $l_{4,3}$-symmetric sample vector distribution, $f_{3:4}$, if the dg is $g^{(4)}_{PT7;M,\nu}$.

tively. Moreover, in these figures, one can compare $f_{4:4}$ and $f_{3:4}$ for a particular choice of the type of dependence model and of the corresponding distributional parameters. Second, the impact of the parameter $p > 0$ on the graphs of $f^{\star}_{4:4}$ and $f^{\star}_{3:4}$ with $F = \Phi_p$ and $f = \varphi_p$, see (12), is shown in Figures 5 and 6. Third, as it is stated before, one can draw comparisons between independence modeling and light as well as heavy tailed dependence modeling in Figures 1, 3 and 5 and in Figures 2, 4 and 6 if $p = 3$.

Additionally, because of (30), we remark that the visualizations of $f_{k:4}$ for $k \in \{1, 2\}$ can be achieved by reflecting the graph of $f_{k:4}$ for $k \in \{3, 4\}$, respectively, with respect to the ordinate axis.

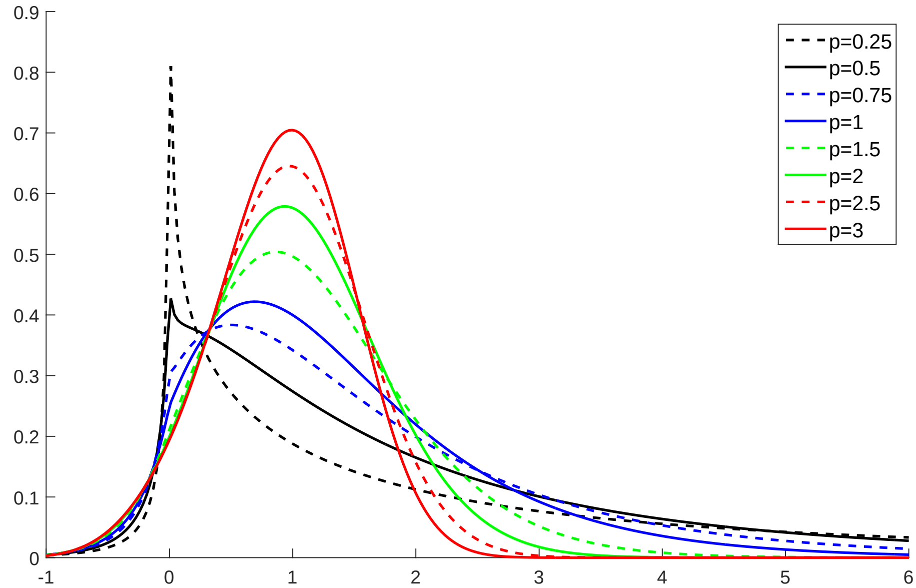

Figure 5: Maximum pdf from 4-dimensional p-generalized Gaussian sample distribution, $f^{\star}_{4:4}$, for $p \in \left\{\frac{1}{4}, \frac{1}{2}, \frac{3}{4}, 1, \frac{3}{2}, 2, \frac{5}{2}, 3\right\}$.

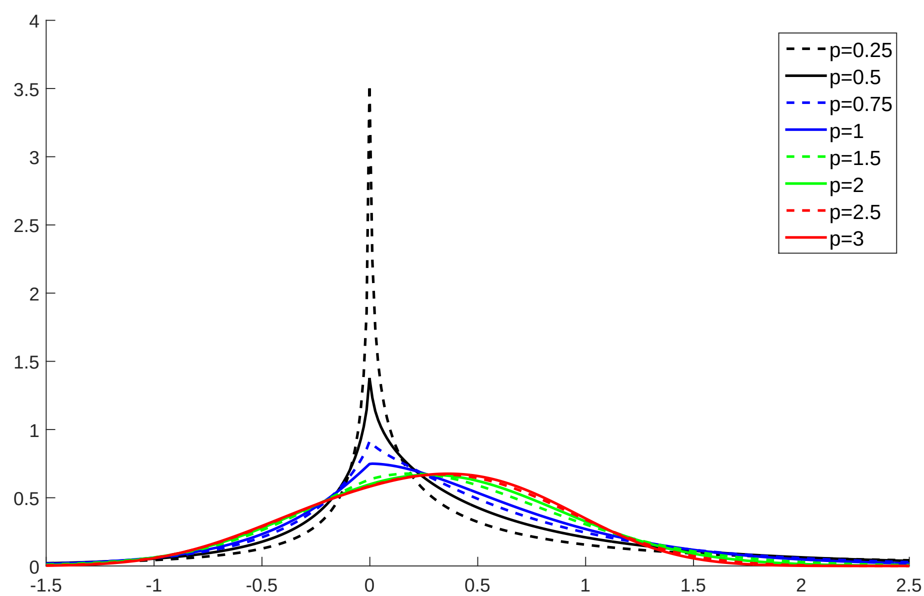

Figure 6: Pdf of the third order statistic from 4-dimensional p-generalized Gaussian sample distribution, $f^{\star}_{3:4}$, for $p \in \left\{\frac{1}{4}, \frac{1}{2}, \frac{3}{4}, 1, \frac{3}{2}, 2, \frac{5}{2}, 3\right\}$.

4 Proofs and discussions

4.1 Proof of Theorem 1

Due to the single-out decomposition of $A_k^n(t)$,

$$A_k^n(t) = \bigcup_{i=1}^{n} \left\{ (x_1, \ldots, x_n)^{\mathrm{T}} \in \mathbb{R}^n : x_i < t,\ x_{k:n} = x_i \right\},$$

we have

$$F_{k:n}(t) = P(X_{k:n} < t) = \sum_{i=1}^{n} P(X_i < t,\ \ X_i = X_{k:n}), \quad t \in \mathbb{R}. \tag{16}$$

The vector remaining after eliminating the ith component of X will be denoted by $X[i]$, and the matrix after eliminating both the ith column and the ith row of the quadratic matrix M by $M[i]$. Furthermore, $Y \overset{d}{=} Z$ means that the random vectors Y and Z follow the same distribution law. We recall that $S_J^{(n)}[i]X[i] \overset{d}{=} X[i]$ for $i = 1, \ldots, n$ and $J \subseteq \{1, \ldots, n\} \setminus \{i\}$, thus

$$\begin{aligned} P\big(X_i < t, X_i = X_{k:n}\big) &= \sum_{\substack{J_1 \subseteq \{1,\ldots,n\}\setminus\{i\} \\ |J_1|=k-1}} P\Big(X_i < t,\ \max\big\{X_j : j \in J_1\big\} < X_i,\ X_i < \min\big\{X_l : l \in J_1^C \setminus \{i\}\big\}\Big) \\ &= \sum_{\substack{J_1 \subseteq \{1,\ldots,n\}\setminus\{i\} \\ |J_1|=k-1}} P\Big(X_i < t,\ X[i] < S_{J_1}^{(n)}[i] 1_{n-1} X_i\Big) \\ &= \sum_{\substack{J_1 \subseteq \{1,\ldots,n\}\setminus\{i\} \\ |J_1|=k-1}} P\Big(X[i] < S_{J_1}^{(n)}[i] 1_{n-1} X_i\Big)\, P\Big(X_i < t \bigm| X[i] < S_{J_1}^{(n)}[i] 1_{n-1} X_i\Big) \end{aligned}$$

where $J_1^C = \{1, \ldots, n\} \setminus J_1$ is the complementary set of J_1 with respect to $\{1, \ldots, n\}$. From the stochastic representation of the skewed $l_{\varkappa+m,p}$-symmetric distribution, see (7) in Section 2, with $\varkappa = 1$, $m = n-1$, $X^{(1)} = X_i$, $X^{(2)} = X[i]$ and $\Lambda = S_{J_1}^{(n)}[i]1_{n-1}$ for $i = 1, \ldots, n$, it follows that

$$\mathfrak{L}\Big(X_i \bigm| X[i] < S_{J_1}^{(n)}[i] 1_{n-1} X_i\Big) = SS_{1,n-1,p}\left(S_{J_1}^{(n)}[i] 1_{n-1}; g^{(n)}\right).$$

Hence,

$$\begin{aligned} P\Big(X_i < t,\ X_i = X_{k:n}\Big) = \sum_{\substack{J_1 \subseteq \{1,\ldots,n\}\setminus\{i\} \\ |J_1|=k-1}} & F_{n-1,p}^{(2)}\left(0_{n-1}; I_{n-1} + \Gamma_{i,J_1}^{(n-1)}, g_{(n)}^{(n-1)}\right) \\ & \cdot F_{1,n-1,p}\left(t; S_{J_1}^{(n)}[i] 1_{n-1}, g^{(n)}\right) \end{aligned}$$

where $\Gamma_{i,J_1}^{(n-1)} = \left(S_{J_1}^{(n)}[i]1_{n-1}\right)\left(S_{J_1}^{(n)}[i]1_{n-1}\right)^{\mathrm{T}} = S_{J_1}^{(n)}[i]E^{(n-1)}S_{J_1}^{(n)}[i]$. For $i \in \{1, \ldots, n\}$ and $J_1, J_2 \subseteq \{1, \ldots, n\} \setminus \{i\}$ with $|J_1| = k-1 = |J_2|$ and $n \times n$ matrix $M_{J_1,J_2}^{(n)}$ being the permutation matrix describing the permutation $\sigma_{J_1,J_2}^{(n)}$ with $\sigma_{J_1,J_2}^{(n)}(J_1) = J_2$ it follows $SS_{1,n-1,p}\left(S_{J_2}^{(n)}[i]1_{n-1}; g^{(n)}\right) = SS_{1,n-1,p}\left(M_{J_1,J_2}^{(n)} \cdot S_{J_1}^{(n)}[i]1_{n-1}; g^{(n)}\right)$. Here, $|J|$ denote the number of elements of the set J. For $i \in \{1, \ldots, n\}$, there are $\binom{n-1}{k-1}$ possible choices of index sets $J_{(i)} \subseteq \{1, \ldots, n\} \setminus \{i\}$ with $|J_{(i)}| = k-1$, thus

$$P\left(X_i < t,\ X_i = X_{k:n}\right) = \binom{n-1}{k-1} F_{n-1,p}^{(2)}\left(0_{n-1}; I_{n-1} + \Gamma_{i,J_{(i)}}^{(n-1)}, g_{(n)}^{(n-1)}\right) F_{1,n-1,p}\left(t; S_{J_{(i)}}^{(n)}[i]1_{n-1}, g^{(n)}\right).$$

Because $P(X_i < t, X_i = X_{k:n}) = P(X_j < t, X_j = X_{k:n})$ for all $i, j \in \{1, \ldots, n\}$, it follows

$$F_{k:n}(t) = n\, P\left(X_n < t,\ X_n = X_{k:n}\right)$$

$$= n\binom{n-1}{k-1} F^{(2)}_{n-1,p}\left(0_{n-1}; I_{n-1} + S^{(n)}_{J_{(n)}}[n]E^{(n-1)}S^{(n)}_{J_{(n)}}[n], g^{(n-1)}_{(n)}\right)$$
$$\cdot F_{1,n-1,p}\left(t; S^{(n)}_{J_{(n)}}[n]1_{n-1}, g^{(n)}\right), \quad t \in \mathbb{R},$$

where $J_{(n)}$ is an arbitrary subset of $\{1, \ldots, n-1\}$ having $k-1$ elements. On using $n\binom{n-1}{k-1} = \binom{n}{k}k$ and $S^{(n)}_{J_{(n)}}[n] = S^{(n-1)}_{J_{(n)}}$, the assertion follows.

4.2 Measure-of-cone representations of skewed $l_{n,p}$-symmetric distributions

Because the derivation of the results on Section 3.3 makes basically use of so called measure-of-cone representations of skewed $l_{n,p}$-symmetric distributions, the present section deals with results of this type which may be of independent interest, too. Measure-of-cone representations of elliptically contoured distributions are derived in [16] and [34] for dimension two and for arbitrary finite dimension, respectively. A representation of more general type is proved in [26] for $l_{n,p}$-symmetric distributions.

For $\varkappa \in \mathbb{N}$ and $m \in \mathbb{N}$, let $\Lambda = (\Lambda_{j,i})_{\substack{j=1,\ldots,m \\ i=1,\ldots,\varkappa}}$ be a matrix from $\mathbb{R}^{m\times\varkappa}$, and $I = \{i_1, \ldots, i_\varkappa\} \subset \{1, \ldots, \varkappa + m\}$ with $|I| = \varkappa$. Furthermore, let us denote the set of all $(\varkappa + m) \times (\varkappa + m)$ signs matrices by $\mathcal{S}_{\varkappa+m}$ and assume that D is an element of this set. Moreover, for $z = (z_1, \ldots, z_\varkappa) \in \mathbb{R}^\varkappa$, we put $V_{I,D}(\Lambda; z) = D \cdot (v_1, \ldots, v_{\varkappa+m})^{\mathrm{T}}$ where $v_{i_l} = z_l$ for $l = 1, \ldots, \varkappa$ and $v_{j_\nu} = (\Lambda z)_\nu = \sum_{l=1}^{\varkappa} \Lambda_{\nu,l} z_l$ for $j_\nu \in \{1, \ldots, \varkappa + m\} \setminus I$ and $\nu = 1, \ldots, m$. It turns out that the set

$$C_{I,D}(\Lambda; z) = \{Dx \in \mathbb{R}^{\varkappa+m} \colon x_{i_l} < z_l, l = 1, \ldots, \varkappa,$$
$$x_{j_\nu} < \sum_{l=1}^{\varkappa} \Lambda_{\nu,l} x_{i_l}, j_\nu \in \{1, \ldots, \varkappa + m\} \setminus I, \nu = 1, \ldots, m\}$$

is a cone with vertex in $V_{I,D}(\Lambda; z)$, for every $z \in \mathbb{R}^\varkappa$. The class of all such cones will be denoted by $\mathcal{C}_{\varkappa,m}(\Lambda; z)$, i.e.

$$\mathcal{C}_{\varkappa,m}(\Lambda; z) = \{C_{I,D}(\Lambda; z) \colon I \subset \{1, \ldots, \varkappa + m\} \text{ with } |I| = \varkappa, D \in \mathcal{S}_{\varkappa+m}\}.$$

Recalling the notation $\Pi_{\varkappa+m}$ for the set of all $(\varkappa + m) \times (\varkappa + m)$ permutation matrices from Section 2 and denoting the particular cone $C_{\{1,\ldots,\varkappa\},I_{\varkappa+m}}(\Lambda; z)$ with vertex in $V_{\{1,\ldots,\varkappa\},I_{\varkappa+m}}(\Lambda; z) = \left(z^{\mathrm{T}}, (\Lambda z)^{\mathrm{T}}\right)^{\mathrm{T}}$ by $A_0(z)$, it follows that $\mathcal{C}_{\varkappa,m}(\Lambda; z) = \{MDA_0(z) \colon M \in \Pi_{\varkappa+m}, D \in \mathcal{S}_{\varkappa+m}\}$.

Making use of this notations, in [26], the following measure-of-cone representations of skewed $l_{\varkappa,p}$-symmetric distributions are derived.

Proposition 1. *Let $Z \sim SS_{\varkappa,m,p}\left(\Lambda, g^{(\varkappa+m)}\right)$. For every element $C_{I,D}(\Lambda; z)$ of $\mathcal{C}_{\varkappa,m}(\Lambda; z)$, the cdf of Z satisfies the representation*

$$F_{\varkappa,m,p}\left(z; \Lambda, g^{(\varkappa+m)}\right) = \frac{1}{F^{(2)}_{m,p}(0_m; I_m + \Lambda\Lambda^\top, g^{(m)}_{(\varkappa+m)})} \Phi_{g^{(\varkappa+m)},p}\left(C_{I,D}(\Lambda; z)\right), \quad z \in \mathbb{R}^\varkappa.$$

4.3 Proof of Theorem 2

The event in the sample space $\mathbb{R}^n$ which is measured by the value of the cdf $F_{k:n}(t)$ can be represented for each $k \in \{1, \ldots, n\}$ as follows:

$$A^n_k(t) = \bigcup_{\substack{J\subseteq\{1,\ldots,n\} \\ |J|=k}} \{x \in \mathbb{R}^n \colon x_j < t \;\forall j \in J\} \tag{17}$$

$$= \bigcup_{\substack{J\subseteq\{1,\ldots,n\} \\ |J|\geq k}} \{x \in \mathbb{R}^n \colon x_j < t \;\forall j \in J,\; x_l \geq t \;\forall l \in \{1, \ldots, n\} \setminus J\}. \tag{18}$$

Note that these decompositions make use of intersections of half spaces being pairwise disjoint in (18) but not in (17). For a visualization of the set $A_k^n(t)$ for some low-dimensional cases, see [24, 25]. Furthermore, $A_k^n(t)$ is an open cone with vertex in $t1_n$.

For $\nu \in \{1, \dots, n-1\}$, $t \in \mathbb{R}$ and every $i \in \{1, \dots, n-\nu\}$, let us denote $B_i^{(n,\nu)}(t) = \tilde{B}_{i,0}^{(n,\nu)}(t)$ and

$$\begin{aligned}\tilde{B}_{i,j}^{(n,\nu)}(t) = \{x \in \mathbb{R}^n \colon x_{i+j} < t,\ & x_{l_1} \le x_{i+j}, \forall\, l_1 \in \{i, \dots, i+j\} \setminus \{i+j\},\\ & x_{l_2} < x_{i+j}, \forall\, l_2 \in \{i+j, \dots, i+\nu\} \setminus \{i+j\},\\ & x_{l_3} < t, \forall\, l_3 \in \{1, \dots, n\} \setminus \{i, \dots, i+\nu\}\}\end{aligned}$$

for $j = 0, 1, \dots, \nu$ which are cones with vertex in $t1_n$ being neither open nor closed. According to the proof of Theorem 1 in [26], for $t \in \mathbb{R}$ and every $\nu \in \{1, \dots, n-1\}$, the cone $A_n^n(t)$ can be decomposed into $\nu+1$ disjoint cones such that each of them, which is an intersection of half spaces of $\mathbb{R}^n$, contains the origin in the boundary of ν of its n intersecting half spaces,

$$A_n^n(t) = B_i^{(n,\nu)}(t) \cup \left(\bigcup_{j=1}^{\nu} \tilde{B}_{i,j}^{(n,\nu)}(t) \right) \tag{19}$$

where the parameter i can be chosen arbitrarily from $\{1, \dots, n-\nu\}$. Hence, the parameter $\nu \in \{1, \dots, n-1\}$ describes a certain degree of decomposition. Moreover, according to [26], it follows that

$$\Phi_{g^{(n)},p}(A_n^n(t)) = (\nu+1)\, \Phi_{g^{(n)},p}\left(B_i^{(\nu)}(t)\right), \quad t \in \mathbb{R}. \tag{20}$$

Unless for the just considered case of the maximum statistic, in Lemma 1, we provide a recursive result on the cdf $F_{k:n}$ of the kth order statistic $X_{k:n}$ of the $l_{n,p}$-symmetrically distributed random vector X with the help of the cdf $F_{k+1:n}$ of $X_{k+1:n}$. To this end, we define

$$D^{(n,k)}(t) = F_{k:n}(t) - F_{k+1:n}(t), \quad t \in \mathbb{R},\ \text{for } k \in \{1, \dots, n-1\}\,.$$

The following lemma will form the basis of the proof of Theorem 2.

Lemma 1. *For $k \in \{1, \dots, n-1\}$ and every $\nu_1 \in \{0, \dots, n-k-1\}$ and $\nu_2 \in \{0, \dots, k-1\}$ such that $\nu = \nu_1 + \nu_2 > 0$, the function $D^{(n,k)}(t)$, $t \in \mathbb{R}$, allows the representation*

$$\frac{D^{(n,k)}(t)}{\binom{n}{k}(\nu_1+1)(\nu_2+1)} = F_{\nu,p}^{(2)}\left(0_\nu; I_\nu + E^{(\nu_1,\nu_2)}, g_{(n)}^{(\nu)}\right) F_{n-\nu,\nu,p}\left(\left(\begin{array}{c} -t1_{n-k-\nu_1} \\ t1_{k-\nu_2} \end{array}\right); E_{1,1}^{(\nu_1,\nu_2,n,k)}, g^{(n)}\right).$$

Proof of Lemma 1. For $k = 1, \dots, n-1$, according to the disjoint decomposition (18), $A_k^n(t)$ can basically be partitioned into the following disjoint recursive union

$$A_k^n(t) = A_{k+1}^n(t) \cup \bigcup_{\substack{J_1 \subset \{1,\dots,n\} \\ |J_1| = k}} \left\{x \in \mathbb{R}^n \colon x_j < t\ \forall j \in J_1,\ x_l \ge t\ \forall l \in \{1, \dots, n\} \setminus J_1\right\}, \quad t \in \mathbb{R}.$$

Hence, with $F_{k:n}(t) = P\left(X \in A_k^n(t)\right)$, we have that

$$D^{(n,k)}(t) = \sum_{\substack{J_1 \subset \{1,\dots,n\} \\ |J_1| = k}} \Phi_{g^{(n)},p}\left(\left\{x \in \mathbb{R}^n \colon x_j < t\ \forall j \in J_1,\ x_l \ge t\ \forall l \in \{1, \dots, n\} \setminus J_1\right\}\right).$$

For two sets $J_1, J_2 \subseteq \{1, \dots, n\}$ with $|J_1| = |J_2|$, let $\sigma_{J_1,J_2}^{(n)}$ be the permutation such that $\sigma_{J_1,J_2}^{(n)}(J_1) = J_2$, and let $M_{J_1,J_2}^{(n)}$ denote the corresponding permutation matrix. Furthermore, we recall the definition of the sign matrix $S_{J_1}^{(n)}$ from Theorem 1. For the rest of this proof, let $J = \{n-k+1, \dots, n\}$ be fixed with $|J| = k$. Because of the permutation and sign invariance of the $l_{n,p}$-symmetric probability measure, see [26], and its continuity, it follows that

$$D^{(n,k)}(t) = \sum_{\substack{J_1 \subset \{1,\dots,n\} \\ |J_1| = k}} \Phi_{g^{(n)},p}\left(M_{J_1,J}^{(n)}\left\{x \in \mathbb{R}^n \colon x_j < t\ \forall j \in J_1,\ x_l \ge t\ \forall l \in \{1, \dots, n\} \setminus J_1\right\}\right)$$

$$= \binom{n}{k} \Phi_{g^{(n)},p}\left(\{x \in \mathbb{R}^n : x_{n-k+1} < t, \ldots, x_n < t, x_1 > t, \ldots, x_{n-k} > t\}\right)$$
$$= \binom{n}{k} \Phi_{g^{(n)},p}\left(S_J^{(n)}\{x \in \mathbb{R}^n : x_{n-k+1} < t, \ldots, x_n < t, x_1 < -t, \ldots, x_{n-k} < -t\}\right)$$
$$= \binom{n}{k} \Phi_{g^{(n)},p}\left(\{x \in \mathbb{R}^n : (x_1, \ldots, x_{n-k})^{\mathrm{T}} \in A_{n-k}^{n-k}(-t),\ (x_{n-k+1}, \ldots, x_n)^{\mathrm{T}} \in A_k^k(t)\}\right).$$

In the following, at least one of the cones $A_{n-k}^{n-k}(-t)$ and $A_k^k(t)$ will be decomposed. To this end, according to (19), for $\nu_1 \in \{0, \ldots, n-k-1\}$ and $\nu_2 \in \{0, \ldots, k-1\}$ with $\nu_1 + \nu_2 > 0$, we choose

$$A_{n-k}^{n-k}(-t) = B_1^{(n-k,\nu_1)}(-t) \cup \left(\bigcup_{j=1}^{\nu_1} \tilde{B}_{1,j}^{(n-k,\nu_1)}(-t)\right) \tag{21}$$

and

$$A_k^k(t) = B_1^{(k,\nu_2)}(t) \cup \left(\bigcup_{j=1}^{\nu_2} \tilde{B}_{1,j}^{(k,\nu_2)}(t)\right). \tag{22}$$

Here, $\nu_1 = 0$ and $\nu_2 = 0$ means that $A_{n-k}^{n-k}(-t)$ and $A_k^k(t)$ are not decomposed, respectively. Hence, (20) implies

$$\frac{D^{(n,k)}(t)}{\binom{n}{k}(\nu_1+1)(\nu_2+1)} = \Phi_{g^{(n)},p}\Big(\{x \in \mathbb{R}^n : (x_1, \ldots, x_{n-k})^{\mathrm{T}} \in B_1^{(n-k,\nu_1)}(-t),$$
$$(x_{n-k+1}, \ldots, x_n)^{\mathrm{T}} \in B_1^{(k,\nu_2)}(t)\}\Big). \tag{23}$$

Note that the set measured by $\Phi_{g^{(n)},p}$ in (23) is an element of the class of cones

$$\mathcal{C}_{n-\nu_1-\nu_2,\nu_1+\nu_2}\left(E_{1,1}^{(\nu_1,\nu_2,n,k)}; \begin{pmatrix} -t1_{n-k-\nu_1} \\ t1_{k-\nu_2} \end{pmatrix}\right). \tag{24}$$

Because of the identity $E_{1,1}^{(\nu_1,\nu_2,n,k)} E_{1,1}^{(\nu_1,\nu_2,n,k)^{\mathrm{T}}} = E^{(\nu_1,\nu_2)}$ and $\nu = \nu_1 + \nu_2$, the measure-of-cone representation from Proposition 1 yields

$$\Phi_{g^{(n)},p}\left(\{x \in \mathbb{R}^n : (x_1, \ldots, x_{n-k})^{\mathrm{T}} \in B_1^{(n-k,\nu_1)}(-t), (x_{n-k+1}, \ldots, x_n)^{\mathrm{T}} \in B_1^{(k,\nu_2)}(t)\}\right)$$
$$= F_{\nu,p}^{(2)}\left(0_\nu; I_\nu + E^{(\nu_1,\nu_2)}, g_{(n)}^{(\nu)}\right) F_{n-\nu,\nu,p}\left(\begin{pmatrix} -t1_{n-k-\nu_1} \\ t1_{k-\nu_2} \end{pmatrix}; E_{1,1}^{(\nu_1,\nu_2,n,k)}, g^{(n)}\right)$$

and the assertion follows from (23). □

In order to proof Theorem 2, we start from the representation

$$F_{k:n}(t) = \left(\sum_{j=k}^{n-1} D^{(n,j)}(t)\right) + F_{n:n}(t), \quad t \in \mathbb{R}, \tag{25}$$

being a telescoping sum. Applying Lemma 1 to each summand $D^{(n,j)}(t)$, $j = k, \ldots, n-1$, in each step the two parameters $\nu_{j,1} \in \{0, \ldots, n-j-1\}$ and $\nu_{j,2} \in \{0, \ldots, j-1\}$ satisfying $\nu_j = \nu_{j,1} + \nu_{j,2} > 0$ can be chosen differently. The assertion of Theorem 2 follows by inserting the cdf $F_{n:n}$ of the maximum statistic derived in [26] and based upon the cone decomposition (19) as well.

4.4 Proof of Theorem 3

Looking through the proof of Theorem 1 in [24] once again, there appear representations of the cdf $F_{n:n}$ of the maximum statistic $X_{n:n}$ in $l_{n,p}$-symmetrically distributed populations not used so far, see Lemma 2.

Lemma 2. *For every $\nu \in \{1, \dots, n-1\}$ and $i \in \{1, \dots, n-\nu\}$, the cdf $F_{n:n}$ allows the representation*

$$F_{n:n}(t) = (\nu+1) F^{(2)}_{\nu,p}\left(0_\nu; \Sigma, g^{(\nu)}_{(n)}\right) F_{n-\nu,\nu,p}\left(t1_{n-\nu}; E^{(\nu,n)}_i, g^{(n)}\right), \quad t \in \mathbb{R},$$

where $\Sigma = I_\nu + E^{(\nu,n)}_i E^{(\nu,n)^\top}_i = I_\nu + E^{(\nu)}$.

As stated in [24], because of the exchangeability of the components of $l_{n,p}$-symmetrically distributed random vectors, the additional parameter i has no numerical impact when evaluating the cdf $F_{n:n}$. Nevertheless, these alternative representations may be helpful in identifying a maximum distribution if particular skewed $l_{n,p}$-symmetric distributions occur in other contexts.

Recall that the cones $A^{n-k}_{n-k}(-t)$ and $A^k_k(t)$ are decomposed in relations (21) and (22) with the help of (19), choosing in both cases particularly $i = 1$. Using now the entire variety of decompositions in (19), for every $i_1 \in \{1, \dots, n-k-\nu_1\}$ and $i_2 \in \{1, \dots, k-\nu_2\}$,

$$A^{n-k}_{n-k}(-t) = B^{(n-k,\nu_1)}_{i_1}(-t) \cup \left(\bigcup_{j=1}^{\nu_1} \tilde{B}^{(n-k,\nu_1)}_{i_1,j}(-t)\right) \tag{26}$$

und

$$A^k_k(t) = B^{(k,\nu_2)}_{i_2}(t) \cup \left(\bigcup_{j=1}^{\nu_2} \tilde{B}^{(k,\nu_2)}_{i_2,j}(t)\right). \tag{27}$$

Following the rest of the proof of Lemma 1, but using now decompositions (26) and (27) instead of (21) and (22), leads to the following result.

Lemma 3. *For $k \in \{1, \dots, n-1\}$, every $\nu_1 \in \{0, \dots, n-k-1\}$ and $\nu_2 \in \{0, \dots, k-1\}$ such that $\nu = \nu_1 + \nu_2 > 0$, and every $i_1 \in \{1, \dots, n-k-\nu_1\}$ and $i_2 \in \{1, \dots, k-\nu_2\}$, the function $D^{(n,k)}(t)$, $t \in \mathbb{R}$, satisfies the representation*

$$\frac{D^{(n,k)}(t)}{\binom{n}{k}(\nu_1+1)(\nu_2+1)} = F^{(2)}_{\nu,p}\left(0_\nu; I_\nu + E^{(\nu_1,\nu_2)}, g^{(\nu)}_{(n)}\right) F_{n-\nu,\nu,p}\left(\begin{pmatrix} -t1_{n-k-\nu_1} \\ t1_{k-\nu_2} \end{pmatrix}; E^{(\nu_1,\nu_2,n,k)}_{i_1,i_2}, g^{(n)}\right).$$

In the same sense as we commented the role of the parameter i in the representation of $F_{n:n}$ in Lemma 2, the additional parameters i_1 and i_2 in Lemma 3 do not have any numerical impact on the evaluation of $D^{(n,k)}$. Subsequently, a further extension of the Lemmas 1 and 3 is considered. Both, after the first third of the proof of Lemma 1 and of the complete proof of Lemma 3, respectively, using the permutation invariance of $\Phi_{g^{(n)},p}$, an arbitrary subset of $\{1, \dots, n\}$ with k elements is specialized to $J = \{n-k+1, \dots, n\}$. This choice influences the positioning of both the rows of matrix parameter $E^{(\nu_1,\nu_2,n,k)}_{i_1,i_2}$ and the negative components of $(-t1^\top_{n-k-\nu_1}, t1^\top_{k-\nu_2})^\top$ of the cone

$$\mathcal{C}_{n-\nu_1-\nu_2,\nu_1+\nu_2}\left(E^{(\nu_1,\nu_2,n,k)}_{i_1,i_2}; \begin{pmatrix} -t1_{n-k-\nu_1} \\ t1_{k-\nu_2} \end{pmatrix}\right).$$

By the application of the measure-of-cone representations of skewed $l_{n,p}$-symmetric distributions, see Proposition 1, this impact is transmitted to the resulting difference function $D^{(n,k)}$ in Lemma 3. In Corollary 2, this restriction is subsequently removed using the properties of skewed $l_{n,p}$-symmetric distributions concerning the permutation of the columns and the rows, respectively, in the matrix parameter described in the last paragraph of Section 2.

Corollary 2. *For $k \in \{1, \dots, n-1\}$, every $\nu_1 \in \{0, \dots, n-k-1\}$ and $\nu_2 \in \{0, \dots, k-1\}$ such that $\nu = \nu_1 + \nu_2 > 0$, every $i_1 \in \{1, \dots, n-k-\nu_1\}$ and $i_2 \in \{1, \dots, k-\nu_2\}$, every $M_1 \in \Pi_{n-\nu}$ and $M_2 \in \Pi_\nu$, the function $D^{(n,k)}$ satisfies the representation*

$$D^{(n,k)}(t) = \binom{n}{k}(\nu_1+1)(\nu_2+1)\, F^{(2)}_{\nu,p}\left(0_\nu; I_\nu + E^{(\nu_1,\nu_2)}, g^{(\nu)}_{(n)}\right)$$

$$\cdot F_{n-\nu,\nu,p}\left(M_1\begin{pmatrix}-t1_{n-k-\nu_1}\\ t1_{k-\nu_2}\end{pmatrix}; M_2E_{i_1,i_2}^{(\nu_1,\nu_2,n,k)}M_1^{\mathrm{T}}, g^{(n)}\right), \quad t\in\mathbb{R}.$$

Note that a possible inclusion of permutation matrices in the representations of $F_{n:n}$ in Lemma 2 would not yield any new results since permuting the rows of $1_{n-\nu}$ and $E_i^{(\nu,n)}$, respectively, would not yield any numerical changes, and the effect of eventually permuting the columns of $E_i^{(\nu,n)}$ would be covered by the effect of varying the additional parameter $i\in\{1,\dots,n-\nu\}$. Using Lemma 2 and Corollary 2 instead of Theorem 1 from [26] and Lemma 1 in the proof of Theorem 2, see (25), we achieve the results of Theorem 3 on the cdf $F_{k:n}$ of the kth order statistic in $l_{n,p}$-symmetrically distributed populations.

4.5 Discussion of combining Proposition 1 and the single-out decomposition of $A_k^n(t)$

In this section, the consideration of the approach via single-out decomposition of the event $A_k^n(t)$ is extended concerning the usage of measure-of-cone representations of skewed $l_{n,p}$-symmetric distributions.

As it is established in [20] for the case of order statistics of two jointly $l_{2,p}$-symmetrically distributed random variables, the approach in [19] and others which is based upon the single-out decomposition of $A_2^2(t)$, see Section 3.2, yields the same result as applying the geometric measure representation in [30] directly, see [24], and transforming the results into terms of skewed distributions, see [10]. In higher dimensional cases, analogous relations are not yet investigated. However, applying the advanced geometric method of proof using measure-of-cone representations of skewed $l_{n,p}$-symmetric distributions yields more representations of the cdf of an arbitrary order statistic than the direct application of the geometric measure representation.

According to Section 4.1, and further using the P-almost sure uniqueness of order statistics from the components of absolutely continuous random vectors defined on a probability space $(\Omega,\mathfrak{A},P)$, we have

$$\begin{aligned}A_k^n(t)&=\bigcup_{i=1}^n\left\{x\in\mathbb{R}^n: x_i<t,\ x_i=x_{k:n}\right\}\\&=\bigcup_{i=1}^n\Big\{x\in\mathbb{R}^n: x_i<t,\ \text{exactly } k-1 \text{ components of } x \text{ are less than } x_i\\&\qquad\qquad\text{and exactly } n-k \text{ are greater than } x_i\Big\}\\&=\bigcup_{i=1}^n\ \bigcup_{\substack{J_1\subset\{1,\dots,n\}\setminus\{i\}\\|J_1|=k-1}}\left\{x\in\mathbb{R}^n: x_i<t,\ x_j<x_i\ \forall j\in J_1,\ x_i<x_l\ \forall l\in\{1,\dots,n\}\setminus(J_1\cup\{i\})\right\}.\end{aligned}\tag{28}$$

Note that (28) is a decomposition of $A_k^n(t)$ into cones such that each of them contains the origin in the boundary of $n-1$ of its intersecting half spaces. Further, representation (28) does not follow from the decompositions in (17) and (18), and vice versa. Moreover, for the particular case $k=n$, (28) is covered by (19) for $\nu=n-1$, i.e.

$$\bigcup_{i=1}^n\left\{x\in\mathbb{R}^n: x_i<t,\ x_j<x_i\ \forall j\in[1,n]\setminus\{i\}\right\}=B_1^{(n,n-1)}(t)\cup\left(\bigcup_{j=1}^{n-1}\tilde{B}_{1,j}^{(n,n-1)}(t)\right)\tag{29}$$

where $i=1$ is implicitly given by $\nu=n-1$.

Remark 1. In view of (28), Theorem 1 can also be achieved applying the measure-of-cone representations from Proposition 1.

Proof of Remark 1. For $k\in\{1,\dots,n\}$ and for an arbitrary but fixed index set $J\subseteq\{1,\dots,n-1\}$ with $k-1$ elements, the application of (28) and of permutation and sign invariance property of $\Phi_{g^{(n)},p}$ yields

$$F_{k:n}(t)=\binom{n}{k}k\,\Phi_{g^{(n)},p}\left(\left\{x\in\mathbb{R}^n: x_n<t,\ x_j<x_n\ \forall j\in J,\ x_l<-x_n\ \forall l\in\{1,\dots,n-1\}\setminus J\right\}\right).$$

Since the above argument of $\Phi_{g^{(n)},p}$ is a member of the class $\mathcal{C}_{1,n-1}\left(S_J^{(n-1)}1_{n-1};t\right)$ of cones, the result of Theorem 1 is achieved by the application of Proposition 1. □

5 Conclusion and outlook

Exact distributions of order statistics of a finite number of independent and identically distributed random variables are studied extensively. For the cases of dependent random variables, on the one hand, [2, 3] determine the distribution of order statistics of the components of arbitrary absolutely continuous random vectors in terms of conditional distribution and, particularly, emphasize the case of elliptically contoured distributions. On the other hand, in [19], considering elliptically contoured distributed populations as well, the distributions of order statistics are represented as mixtures of unified skew-elliptical distributions. In the present paper, extending our studies from [26] and references cited therein, we examine the distributions of order statistics from $l_{n,p}$-symmetric distributions in terms of mixtures of skewed $l_{\varkappa,p}$-symmetric distributions, thus generalizing part of results from [2, 3] and [19] to the case of $l_{n,p}$-dependence.

Just to give a short outlook, in the future, one could study exact distributions of multivariate order statistics as considered under another type of model assumptions in [12] as well as [9]. Furthermore, an extension of our approaches to uni- or multivariate order statistics from dependent and nonidentically distributed random variables is conceivable where studies on such model assumptions are made in [8].

Acknowledgement: The authors are grateful to the reviewers for their valuable hints leading to various improvements of exposition and making the paper more accessible to a broader readership.

A Additional remarks

A.1 Pdf $f_{k:n}$ according to $F_{k:n}$ from Theorem 2

In this section, the pdf of the kth order statistic $X_{k:n}$, $f_{k:n}(t)$, $t\in\mathbb{R}$, is determined for $k=1,\dots,n$ according to the representation of the cdf $F_{k:n}$ from Theorem 2. In order to represent this density in a concise manner, for $n\in\mathbb{N}$, a dg $g^{(n)}$, parameters $k\in\{1,\dots,n-1\}$, $\nu_1\in\{0,\dots,n-k-1\}$ and $\nu_2\in\{0,\dots,k-1\}$ or parameters $k=n$ and $(\nu_1,\nu_2)\in\bigcup\limits_{l=1}^{n-1}\{(l,0),(0,l)\}$ and with $\nu=\nu_1+\nu_2>0$, we put

$$\frac{G_j^{(\nu_1,\nu_2,n,k)}(x;t)}{\binom{n}{k}(\nu_1+1)(\nu_2+1)}=\int\limits_{-\infty}^{x}g_{(n)}^{(n-\nu)}\left(\sqrt[p]{|z|_p^p+|t|^p}\right)F_{\nu,p}^{(1)}\left(a_j^{(\nu_1,\nu_2,n,k)}(z;t);g^{(\nu)}_{\left[\sqrt[p]{|z|_p^p+|t|^p}\right]}\right)dz,$$

for $t\in\mathbb{R}$ and $x\in\mathbb{R}^{n-\nu-1}$, $j\in\{1,2,3\}$. Here, for $z=(z_1,\dots,z_{n-\nu-1})^{\mathrm{T}}\in\mathbb{R}^{n-\nu-1}$, $a_1^{(\nu_1,\nu_2,n,k)}(z;t)=\begin{pmatrix}z_1 1_{\nu_1}\\ z_{n-\nu-1}1_{\nu_2}\end{pmatrix}$, $a_2^{(\nu_1,\nu_2,n,k)}(z;t)=\begin{pmatrix}-t1_{\nu_1}\\ z_{n-\nu-1}1_{\nu_2}\end{pmatrix}$, and $a_3^{(\nu_1,\nu_2,n,k)}(z;t)=\begin{pmatrix}z_1 1_{\nu_1}\\ t1_{\nu_2}\end{pmatrix}$. For the particular case of $k=n$ and $(\nu_1,\nu_2)\in\{(n-1,0),(0,n-1)\}$, we put

$$G_3^{(0,\nu,n,n)}(t1_{n-\nu-1};t)=n\,g_{(n)}^{(1)}(|t|)F_{n-1,p}^{(1)}\left(t1_{n-1};g_{[|t|]}^{(n-1)}\right),\quad t\in\mathbb{R},$$

and assume that $G_1^{(\nu,0,n,n)}(t1_{n-\nu-1};t)$ is finite.

Corollary 3. *For $k\in\{1,\dots,n\}$, every $\nu_{j,1}\in\{0,\dots,n-j-1\}$ and $\nu_{j,2}\in\{0,\dots,j-1\}$ such that $\nu_j=\nu_{j,1}+\nu_{j,2}>0$, $j=k,\dots,n-1$, and every $\nu_n\in\{1,\dots,n-1\}$, the pdf $f_{k:n}$ allows the representation*

$$f_{k:n}(t)=\sum_{j=k}^{n-1}\left[(k-\nu_{j,2}-1)\,G_1^{(\nu_{j,1},\nu_{j,2},n,k)}\left(\begin{pmatrix}-t1_{n-k-\nu_{j,1}}\\ t1_{k-\nu_{j,2}-1}\end{pmatrix};t\right)\right.$$

$$- (n-k-\nu_{j,1}-1)\, G_1^{(\nu_{j,1},\nu_{j,2},n,k)}\left(\begin{pmatrix} -t1_{n-k-\nu_{j,1}-1} \\ t1_{k-\nu_{j,2}} \end{pmatrix}; t\right)$$
$$\left. - G_2^{(\nu_{j,1},\nu_{j,2},n,k)}\left(\begin{pmatrix} -t1_{n-k-\nu_{j,1}-1} \\ t1_{k-\nu_{j,2}} \end{pmatrix}; t\right) + G_3^{(\nu_{j,1},\nu_{j,2},n,k)}\left(\begin{pmatrix} -t1_{n-k-\nu_{j,1}} \\ t1_{k-\nu_{j,2}-1} \end{pmatrix}; t\right)\right]$$
$$+ \left\langle (n-\nu_n-1)G_1^{(\nu_n,0,n,n)}(t1_{n-\nu_n-1}; t) + G_3^{(0,\nu_n,n,n)}(t1_{n-\nu_n-1}; t)\right\rangle, \quad t \in \mathbb{R}.$$

Note that the sum within the brackets $\langle\cdot\rangle$ is equal to the maximum pdf $f_{n:n}$ derived in [26]. Moreover, choosing $k = n$ and $\nu_n = n-1$ yields the representation from Corollary 1.

In order to implement $f_{k:n}$ according to Corollary 3, as in the case of $F_{k:n}$, the number of summands of $f_{k:n}$ can be reduced making use of the general relation

$$f_{n-k+1:n}(t) = f_{k:n}(-t), \quad t \in \mathbb{R}, \tag{30}$$

if $1 \le k \le \left\lfloor \frac{n+1}{2} \right\rfloor$ in advance of applying Corollary 3.

Considering the particular case of order statistics from an n-dimensional p-generalized Gaussian distribution, i.e. the particular dg $g^{(n)} = g^{(n)}_{PE}$, see (3), for every $k \in \{1,\dots,n\}$ and any choice of parameters $\nu_{j,1} \in \{0,\dots,n-j-1\}$ and $\nu_{j,2} \in \{0,\dots,j-1\}$ such that $\nu_j = \nu_{j,1}+\nu_{j,2} > 0$ for $j = k,\dots,n-1$ and $\nu_n \in \{1,\dots,n-1\}$, the representation of $f_{k:n}$ from Corollary 3 admits the form (11). Recalling the notations $\Phi_p(t)$ and $\varphi_p(t)$ of the cdf and the pdf of the one-dimensional p-generalized Gaussian distribution, respectively, we get the following result.

Remark 2. For $g^{(n)} = g^{(n)}_{PE}$ and $k \in \{1,\dots,n\}$, the pdf $f_{k:n}$ from Corollary 3 attains the representation in (12) being nothing else then $f^{\star}_{k:n}$ for $F = \Phi_p$ and $f = \varphi_p$.

A.2 Proofs of Corollary 3 and Remark 2

As before, for $a^{(n)} \in \mathbb{R}^n$ and $i \in \{1,\dots,n\}$, let $a^{(n)}_i$ denote the ith component of $a^{(n)}$ and $a^{(n)}[i]$ the $(n-1)$-dimensional vector after eliminating $a^{(n)}_i$ out of $a^{(n)}$. The following remark is useful for the derivation of the pdf $f_{k:n}$.

Remark 3. If $Z^{(n)}$ is an n-dimensional random vector having pdf $f_{Z^{(n)}}$, then

$$\frac{d}{dt}\left(P\left(Z^{(n)} < a^{(n)}t\right)\right) = \begin{cases} a^{(1)} f_{Z^{(1)}}(a^{(1)}t) & , \ n = 1 \\ \mathcal{D}_n(f_{Z^{(n)}}, a^{(n)}, t) & , \ n \ge 2 \end{cases}, \quad t \in \mathbb{R},$$

for $a^{(n)} \in \mathbb{R}^n$ where

$$\mathcal{D}_n(f_{Z^{(n)}}, a^{(n)}, t) = \sum_{i=1}^{n} a^{(n)}_i \int\limits_{D^{(n)}_i} f_{Z^{(n)}}(z_1,\dots,z_{i-1},a^{(n)}_i t, z_{i+1},\dots,z_n)\,dz$$
$$= \sum_{i=1}^{n} a^{(n)}_i \int\limits_{-\infty}^{a^{(n)}[i]\,t} f_{Z^{(n)}}(z_1,\dots,z_{i-1},a^{(n)}_i t, z_{i+1},\dots,z_n)\,d(z_1,\dots,z_{i-1},z_{i+1},\dots,z_n)^{\mathrm{T}}$$

and for $i = 1,\dots,n$

$$D^{(n)}_i = D^{(n)}_i(a^{(n)}, t) = \{z = (z_1,\dots,z_{i-1},z_{i+1},\dots,z_n)^{\mathrm{T}} \in \mathbb{R}^{n-1} : z < a^{(n)}[i]\,t\}.$$

Proof of Remark 3. This proof will be given by induction. If $n = 1$, the assertion follows immediately applying the chain rule. If $n = 2$, the Leibniz integral rule yields

$$\frac{d}{dt}\left(P\left(Z^{(2)} < a^{(2)}t\right)\right) = \int\limits_{-\infty}^{a^{(2)}_1 t} \frac{\partial}{\partial t}\left(\int\limits_{-\infty}^{a^{(2)}_2 t} f_{Z^{(2)}}(z_1,z_2)\,dz_1\right) dz_2 + a^{(2)}_1 \int\limits_{-\infty}^{a^{(2)}_2 t} f_{Z^{(2)}}(z_1, a^{(2)}_1 t)\,dz_1$$

$$= a_1^{(2)} \int_{-\infty}^{a_2^{(2)}t} f_{Z^{(2)}}(a_1^{(2)}t, z_2)\, dz_2 + a_2^{(2)} \int_{-\infty}^{a_1^{(2)}t} f_{Z^{(2)}}(z_1, a_2^{(2)}t)\, dz_1$$

$$= \mathcal{D}_2(f_{Z^{(2)}}, a^{(2)}, t), \quad t \in \mathbb{R}.$$

The step of induction from $n-1$ to n reads as

$$\frac{d}{dt}\left(P\left(Z^{(n)} < a^{(n)}t\right)\right) = \frac{d}{dt}\left(\int_{-\infty}^{a_{-n}^{(n)}t}\int_{-\infty}^{a_n^{(n)}t} f_{Z^{(n)}}(z_1,\ldots,z_{n-1},z_n)\, dz_n\, d(z_1,\ldots,z_{n-1})^{\mathrm{T}}\right)$$

$$= a_n^{(n)} \int_{-\infty}^{a_{-n}^{(n)}t} f_{Z^{(n)}}(z_1,\ldots,z_{n-1},a_n^{(n)}t)\, d(z_1,\ldots,z_{n-1})^{\mathrm{T}}$$

$$+ \mathcal{D}_{n-1}\left(\left(\int_{-\infty}^{a_n^{(n)}t} f_{Z^{(n)}}(\cdot,\ldots,\cdot,z_n)\, dz_n\right), a_{-n}^{(n)}, t\right)$$

$$= a_n^{(n)} \int_{-\infty}^{a_{-n}^{(n)}t} f_{Z^{(n)}}(z_1,\ldots,z_{n-1},a_n^{(n)}t)\, d(z_1,\ldots,z_{n-1})^{\mathrm{T}}$$

$$+ \sum_{i=1}^{n-1}\left(a_{-n}^{(n)}\right)_i \int_{-\infty}^{\left(a_{-n}^{(n)}\right)_{-i}t}\int_{-\infty}^{a_n^{(n)}t} f_{Z^{(n)}}\left(z_1,\ldots,z_{i-1},\left(a_{-n}^{(n)}\right)_i t, z_{i+1},\ldots,z_{n-1},z_n\right) dz_n\, d(z_1,\ldots,z_{i-1},z_{i+1}\ldots,z_{n-1})^{\mathrm{T}}$$

$$= \mathcal{D}_n(f_{Z^{(n)}}, a^{(n)}, t), \quad t \in \mathbb{R},$$

as $\left(a^{(n)}[n]\right)_i = a_i^{(n)}$ and $\left(\left(a^{(n)}[n][i]\right)^{\mathrm{T}}, a_n^{(n)}\right)^{\mathrm{T}} = a^{(n)}[i]$ for $i = 1,\ldots,n-1$. □

Note that the integrands $f_{Z^{(n)}}(z_1,\ldots,z_{i-1},a_i^{(n)}t,z_{i+1},\ldots,z_n)$, $i = 1,\ldots,n$, appearing in the above integral representation of $\mathcal{D}_n(f_{Z^{(n)}}, a^{(n)}, t)$ simplify to $f_{Z^{(n)}}(a_1^{(n)}t, z_2,\ldots,z_n)$ and $f_{Z^{(n)}}(z_1,\ldots,z_{n-1},a_n^{(n)}t)$ if $i = 1$ and $i = n$, respectively. Furthermore, the particular case of $a^{(n)} = 1_n$ is already proven in [26].

Denoting the derivative of $D^{(n,k)}(t)$ with respect to t by $d^{(n,k)}(t)$, we have

$$d^{(n,k)}(t) = f_{k:n}(t) - f_{k+1:n}(t) \quad \text{for } k \in \{1,\ldots,n-1\}\,.$$

The following lemma provides a significant step for the derivation of the representation of the pdf $f_{k:n}$ in Corollary 3.

Lemma 4. *For $k \in \{1,\ldots,n-1\}$ and every $\nu_1 \in \{0,\ldots,n-k-1\}$ and $\nu_2 \in \{0,\ldots,k-1\}$ with $\nu = \nu_1+\nu_2 > 0$, the function $d^{(n,k)}$ allows the representation*

$$d^{(n,k)}(t) = (k-\nu_2-1)\, G_1^{(\nu_1,\nu_2,n,k)}\left(\begin{pmatrix} -t1_{n-k-\nu_1} \\ t1_{k-\nu_2-1}\end{pmatrix}; t\right)$$

$$-(n-k-\nu_1-1)\, G_1^{(\nu_1,\nu_2,n,k)}\left(\begin{pmatrix} -t1_{n-k-\nu_1-1} \\ t1_{k-\nu_2}\end{pmatrix}; t\right)$$

$$-G_2^{(\nu_1,\nu_2,n,k)}\left(\begin{pmatrix} -t1_{n-k-\nu_1-1} \\ t1_{k-\nu_2}\end{pmatrix}; t\right) + G_3^{(\nu_1,\nu_2,n,k)}\left(\begin{pmatrix} -t1_{n-k-\nu_1} \\ t1_{k-\nu_2-1}\end{pmatrix}; t\right), \quad t \in \mathbb{R}.$$

Proof of Lemma 4. Let $a^{(n-\nu)} = \begin{pmatrix} -1_{n-k-\nu_1} \\ 1_{k-\nu_2}\end{pmatrix}$, and the $(n-\nu)$-dimensional random vector $Z^{(n-\nu)}$ follow the distribution $SS_{n-\nu,\nu,p}\left(E_{1,1}^{(\nu_1,\nu_2,n,k)}, g^{(n)}\right)$ having pdf $f_{Z^{(n-\nu)}}$. According to Lemma 1 and Remark 3, for $t \in \mathbb{R}$, it follows

$$
\begin{aligned}
&\frac{1}{\binom{n}{k}(\nu_1+1)(\nu_2+1)}\, d^{(n,k)}(t)\\
&= F^{(2)}_{\nu,p}\left(0_\nu; I_\nu + E^{(\nu_1,\nu_2)}, g^{(\nu)}_{(n)}\right) \frac{d}{dt}\left(P\left(Z^{(n-\nu)} < a^{(n-\nu)} t\right)\right)\\
&= F^{(2)}_{\nu,p}\left(0_\nu; I_\nu + E^{(\nu_1,\nu_2)}, g^{(\nu)}_{(n)}\right) \mathcal{D}_{n-\nu}\left(f_{Z^{(n-\nu)}}, a^{(n-\nu)}, t\right)\\
&= \sum_{j=1}^{n-\nu} a_j^{(n-\nu)} \int_{-\infty}^{a^{(n-\nu)}[j]\, t} g^{(n-\nu)}_{(n)}\left(\sqrt[p]{|z^{(j,1)}|_p^p + |a_j^{(n-\nu)} t|^p + |z^{(j,2)}|_p^p}\right)\\
&\qquad \cdot F^{(1)}_{\nu,p}\left(E^{(\nu_1,\nu_2,n,k)}_{1,1}\left(z^{(j,1)^{\mathrm{T}}}, a_j^{(n-\nu)} t, z^{(j,2)^{\mathrm{T}}}\right)^{\mathrm{T}}; g^{(\nu)}_{\left[\sqrt[p]{|z^{(j,1)}|_p^p + |a_j^{(n-\nu)} t|^p + |z^{(j,2)}|_p^p}\right]}\right) d\begin{pmatrix} z^{(j,1)}\\ z^{(j,2)}\end{pmatrix},
\end{aligned}
$$

where $z^{(j,1)} = (z_1, \ldots, z_{j-1}) \in \mathbb{R}^{j-1}$ and $z^{(j,2)} = (z_{j+1}, \ldots, z_{n-\nu}) \in \mathbb{R}^{n-\nu-j}$ for $j = 1, \ldots, n-\nu$. With regard to the particular form of the vector $a^{(n-\nu)}$ and the matrix $E^{(\nu_1,\nu_2,n,k)}_{1,1}$, we distinguish the four cases $j = 1$, $j \in \{1, \ldots, n-k-\nu_1\} \setminus \{1\}$, $j = j^*$ and $j \in \{n-k-\nu_1+1, \ldots, n-\nu\} \setminus \{j^*\}$ where $j^* = n-k-\nu_1+1$. On the one hand, we have $a_j^{(n-\nu)} = -1$ if $j \in \{1, \ldots, n-k-\nu_1\}$ and $a_j^{(n-\nu)} = 1$ if $j \in \{n-k-\nu_1+1, \ldots, n-\nu\}$. On the other hand, the matrix vector product $E^{(\nu_1,\nu_2,n,k)}_{1,1}\left(z^{(j,1)^{\mathrm{T}}}, a_j^{(n-\nu)} t, z^{(j,2)^{\mathrm{T}}}\right)^{\mathrm{T}}$ takes the value $\left(z_1 1^{\mathrm{T}}_{\nu_1}, z_{j^*} 1^{\mathrm{T}}_{\nu_2}\right)^{\mathrm{T}}$ if $j \in \{1, \ldots, n-\nu\} \setminus \{1, j^*\}$, $\left(-t 1^{\mathrm{T}}_{\nu_1}, z_{j^*} 1^{\mathrm{T}}_{\nu_2}\right)^{\mathrm{T}}$ if $j = 1$, and $\left(z_1 1^{\mathrm{T}}_{\nu_1}, t 1^{\mathrm{T}}_{\nu_2}\right)^{\mathrm{T}}$ if $j = j^*$. Thus,

$$
\begin{aligned}
\frac{1}{\binom{n}{k}(\nu_1+1)(\nu_2+1)}\, d^{(n,k)}(t) = &-\sum_{j=2}^{n-k-\nu_1} \int_{-\infty}^{\zeta_1} g^{(n-\nu)}_{(n)}\left(\sqrt[p]{|z^{(j,1)}|_p^p + |-t|^p + |z^{(j,2)}|_p^p}\right)\\
&\qquad \cdot F^{(1)}_{\nu,p}\left(\begin{pmatrix} z_1 1_{\nu_1}\\ z_{j^*} 1_{\nu_2}\end{pmatrix}; g^{(\nu)}_{\left[\sqrt[p]{|z^{(j,1)}|_p^p + |-t|^p + |z^{(j,2)}|_p^p}\right]}\right) d\begin{pmatrix} z^{(j,1)}\\ z^{(j,2)}\end{pmatrix}\\
&-\int_{-\infty}^{\zeta_1} g^{(n-\nu)}_{(n)}\left(\sqrt[p]{|-t|^p + |z^{(1,2)}|_p^p}\right)\\
&\qquad \cdot F^{(1)}_{\nu,p}\left(\begin{pmatrix} -t 1_{\nu_1}\\ z_{j^*} 1_{\nu_2}\end{pmatrix}; g^{(\nu)}_{\left[\sqrt[p]{|-t|^p + |z^{(1,2)}|_p^p}\right]}\right) dz^{(1,2)}\\
&+\sum_{j=n-k-\nu_1+2}^{n-\nu} \int_{-\infty}^{\zeta_2} g^{(n-\nu)}_{(n)}\left(\sqrt[p]{|z^{(j,1)}|_p^p + |t|^p + |z^{(j,2)}|_p^p}\right)\\
&\qquad \cdot F^{(1)}_{\nu,p}\left(\begin{pmatrix} z_1 1_{\nu_1}\\ z_{j^*} 1_{\nu_2}\end{pmatrix}; g^{(\nu)}_{\left[\sqrt[p]{|z^{(j,1)}|_p^p + |t|^p + |z^{(j,2)}|_p^p}\right]}\right) d\begin{pmatrix} z^{(j,1)}\\ z^{(j,2)}\end{pmatrix}\\
&+\int_{-\infty}^{\zeta_2} g^{(n-\nu)}_{(n)}\left(\sqrt[p]{|z^{(j^*,1)}|_p^p + |t|^p + |z^{(j^*,2)}|_p^p}\right)\\
&\qquad \cdot F^{(1)}_{\nu,p}\left(\begin{pmatrix} z_1 1_{\nu_1}\\ t 1_{\nu_2}\end{pmatrix}; g^{(\nu)}_{\left[\sqrt[p]{|z^{(j^*,1)}|_p^p + |t|^p + |z^{(j^*,2)}|_p^p}\right]}\right) d\begin{pmatrix} z^{(j^*,1)}\\ z^{(j^*,2)}\end{pmatrix}
\end{aligned}
$$

where $\zeta_1 = \left(-t 1^{\mathrm{T}}_{n-k-\nu_1-1}, t 1^{\mathrm{T}}_{k-\nu_2}\right)^{\mathrm{T}}$ and $\zeta_2 = \left(-t 1^{\mathrm{T}}_{n-k-\nu_1}, t 1^{\mathrm{T}}_{k-\nu_2-1}\right) \mathrm{T}$. The assertion follows by renaming the variables $(z^{(j,1)^{\mathrm{T}}}, z^{(j,2)^{\mathrm{T}}})^{\mathrm{T}}$, $j = 1, \ldots, n-\nu$, to $z \in \mathbb{R}^{n-\nu-1}$, respectively, and the associated cancellation of sums. □

According to (25), the pdf $f_{k:n}$ satisfies

$$
f_{k:n}(t) = \left(\sum_{j=k}^{n-1} d^{(n,j)}(t)\right) + f_{n:n}(t), \quad t \in \mathbb{R}.
$$

By the application of Lemma 4 to each of the summands $d^{(n,j)}(t)$, $j = k, \dots, n-1$, associated with the corresponding parameter choices of $\nu_{j,1}$ and $\nu_{j,2}$ and of $f_{n:n}$ derived in [26] but represented here in terms of $G_i^{(\nu_1,\nu_2,n,k)}$, $i = 1, 2, 3$,

$$f_{n:n}(t) = (n-\nu-1) G_1^{(\nu,0,n,n)}(t1_{n-\nu-1}; t) + G_3^{(0,\nu,n,n)}(t1_{n-\nu-1}; t), \quad t \in \mathbb{R}.$$

Thus, the assertion of Corollary 3 is established.

Remark 4. Choosing in each case the largest possible parameters $\nu_{j,1} = n-j-1$ and $\nu_{j,2} = j-1$ for $j = k, \dots, n-1$, and $\nu_n = n-1$, according to Corollary 3, the representation of $f_{k:n}$ simplifies as

$$\begin{aligned} f_{k:n}(t) = &\sum_{j=k}^{n-1} \binom{n}{j} (n-j)j \int_{-\infty}^{-t} g^{(2)}\left(\sqrt[p]{|z|^p + |t|^p}\right) F_{n-2,p}^{(1)}\left(\begin{pmatrix} z1_{n-j-1} \\ t1_{j-1} \end{pmatrix}; g^{(n-2)}_{\left[\sqrt[p]{|z|^p+|t|^p}\right]}\right) dz \\ &- \sum_{j=k}^{n-1} \binom{n}{j} (n-j)j \int_{-\infty}^{t} g^{(2,p)}\left(\sqrt[p]{|t|^p + |z|^p}\right) F_{n-2,p}^{(1)}\left(\begin{pmatrix} -t1_{n-j-1} \\ z1_{j-1} \end{pmatrix}; g^{(n-2)}_{\left[\sqrt[p]{|t|^p+|z|^p}\right]}\right) dz \\ &+ n\, g^{(1,p)}(|t|)\, F_{n-1,p}^{(1)}\left(t1_{n-1}; g^{(n-1)}_{[|t|]}\right), \quad t \in \mathbb{R}. \end{aligned}$$

In order to strengthen the proof of Remark 2, the following specialization of Lemma 4 to the case of the n-dimensional p-generalized Gaussian sample distribution is considered.

Remark 5. For the particular dg $g_{PE}^{(n)}$ and $k \in \{1, \dots, n-1\}$, Lemma 4 asserts that

$$d^{(n,k)}(t) = \binom{n}{k} \varphi_p(t) \left(\Phi_p(t)\right)^{k-1} \left(\Phi_p(-t)\right)^{n-k-1} \left(k\Phi_p(-t) - (n-k)\Phi_p(t)\right), \quad t \in \mathbb{R}.$$

Proof of Remark 5. Because of the independence of the components of X where $X \sim \Phi_{g_{PE}^{(n)},p}$, for $x_1 \in \mathbb{R}^{\varkappa_1}$ and $x_2 \in \mathbb{R}^{\varkappa_2}$, the dg is factorable such that

$$g_{PE}^{(\varkappa_1+\varkappa_2)}\left(\sqrt[p]{|x_1|_p^p + |x_2|_p^p}\right) = g_{PE}^{(\varkappa_1)}(|x_1|_p)\, g_{PE}^{(\varkappa_2)}(|x_2|_p).$$

Thus, it follows $F_{\varkappa,p}^{(1)}\left(\xi; g_{PE}{}^{(\varkappa)}_{[|y|_p]}\right) = \prod_{j=1}^{\varkappa} \Phi_p(\xi_j)$ for $\varkappa \in \{1, \dots, n-1\}$, $y \in \mathbb{R}^{n-\varkappa}$ and $\xi = (\xi_1, \dots, \xi_\varkappa)^{\mathrm{T}} \in \mathbb{R}^\varkappa$.

Using this product structure, the functions $G_j^{(\nu_1,\nu_2,n,k)}(x; t)$, $j \in \{1, 2, 3\}$, for $\nu_1 \in \{0, \dots, n-k-1\}$ and $\nu_2 \in \{0, \dots, k-1\}$ satisfying $\nu = \nu_1 + \nu_2 > 0$, $x \in \mathbb{R}^{n-\nu-1}$ and $t \in \mathbb{R}$ allow the representations

$$\begin{aligned} \frac{G_1^{(\nu_1,\nu_2,n,k)}(x; t)}{\binom{n}{k}(\nu_1+1)(\nu_2+1)} &= \int_{-\infty}^{x} \varphi_p(t) \left(\prod_{i=1}^{n-\nu-1} \varphi_p(z_i)\right) \left(\Phi_p(z_1)\right)^{\nu_1} \left(\Phi_p(z_{n-\nu-1})\right)^{\nu_2} dz \\ &= \varphi_p(t) \left(\prod_{i=2}^{n-\nu-2} \Phi_p(x_i)\right) \int_{-\infty}^{x_1} \varphi_p(s) \left(\Phi_p(s)\right)^{\nu_1} ds \int_{-\infty}^{x_{n-\nu-1}} \varphi_p(s) \left(\Phi_p(s)\right)^{\nu_2} ds \end{aligned}$$

and

$$G_2^{(\nu_1,\nu_2,n,k)}(x; t) = \binom{n}{k} (\nu_1+1)(\nu_2+1)\, \varphi_p(t) \left(\Phi_p(-t)\right)^{\nu_1} \left(\prod_{i=1}^{n-\nu-2} \Phi_p(x_i)\right) \int_{-\infty}^{x_{n-\nu-1}} \varphi_p(s) \left(\Phi_p(s)\right)^{\nu_2} ds$$

as well as

$$G_3^{(\nu_1,\nu_2,n,k)}(x; t) = \binom{n}{k} (\nu_1+1)(\nu_2+1)\, \varphi_p(t) \left(\Phi_p(t)\right)^{\nu_2} \left(\prod_{i=2}^{n-\nu-1} \Phi_p(x_i)\right) \int_{-\infty}^{x_1} \varphi_p(s) \left(\Phi_p(s)\right)^{\nu_1} ds.$$

By integration by parts, $\int_{-\infty}^{y} \varphi_p(s) \left(\Phi_p(s)\right)^m \, ds = \frac{1}{m+1} \left(\Phi_p(y)\right)^{m+1}$ for $m \in \mathbb{N}$. Hence,

$$\frac{G_1^{(\nu_1,\nu_2,n,k)}(x;t)}{\binom{n}{k}} = \varphi_p(t) \left(\Phi_p(x_1)\right)^{\nu_1+1} \left(\prod_{i=2}^{n-\nu-2} \Phi_p(x_i)\right) \left(\Phi_p(x_{n-\nu-1})\right)^{\nu_2+1},$$

$$\frac{G_2^{(\nu_1,\nu_2,n,k)}(x;t)}{\binom{n}{k}(\nu_1+1)} = \varphi_p(t) \left(\Phi_p(-t)\right)^{\nu_1} \left(\prod_{i=1}^{n-\nu-2} \Phi_p(x_i)\right) \left(\Phi_p(x_{n-\nu-1})\right)^{\nu_2+1}$$

and

$$\frac{G_3^{(\nu_1,\nu_2,n,k)}(x;t)}{\binom{n}{k}(\nu_2+1)} = \varphi_p(t) \left(\Phi_p(t)\right)^{\nu_2} \left(\Phi_p(x_1)\right)^{\nu_1+1} \left(\prod_{i=2}^{n-\nu-1} \Phi_p(x_i)\right).$$

Choosing the values of x_i with view to Lemma 4 finishes the proof. □

Proof of Remark 2. By Remark 5, relation $\Phi_p(t) + \Phi_p(-t) = 1$, $t \in \mathbb{R}$, and using properties of binomial coefficients, for all $t \in \mathbb{R}$,

$$\begin{aligned}
f_{k:n}(t) &= f_{n:n}(t) + \sum_{j=k}^{n-1} d^{(n,j)}(t) \\
&= n\,\varphi_p(t) \left(\Phi_p(t)\right)^{n-1} + \sum_{j=k}^{n-1} \binom{n}{j} \varphi_p(t) \left(\Phi_p(-t)\right)^{n-j-1} \left(\Phi_p(t)\right)^{j-1} \left(j\Phi_p(-t) - (n-j)\Phi_p(t)\right) \\
&= n\,\varphi_p(t) \left(\Phi_p(t)\right)^{n-1} + n\,\varphi_p(t) \sum_{j=k}^{n-1} \binom{n-1}{j-1} \left(\Phi_p(-t)\right)^{n-j} \left(\Phi_p(t)\right)^{j-1} \\
&\quad - n\,\varphi_p(t) \left[\left(\Phi_p(-t) + \Phi_p(t)\right)^{n-1} - \sum_{j=0}^{k-1} \binom{n-1}{j} \left(\Phi_p(-t)\right)^{n-1-j} \left(\Phi_p(t)\right)^{j} \right] \\
&= n\,\varphi_p(t) \left(\Phi_p(t)\right)^{n-1} - n\,\varphi_p(t) + n\,\varphi_p(t) \sum_{j=0}^{k-1} \binom{n-1}{j} \left(\Phi_p(-t)\right)^{n-1-j} \left(\Phi_p(t)\right)^{j} \\
&\quad + n\,\varphi_p(t) \binom{n-1}{k-1} \left(\Phi_p(-t)\right)^{n-k} \left(\Phi_p(t)\right)^{k-1} \\
&\quad + n\,\varphi_p(t) \left[-\left(\Phi_p(t)\right)^{n-1} + \sum_{i=k}^{n-1} \binom{n-1}{i} \left(\Phi_p(-t)\right)^{n-1-i} \left(\Phi_p(t)\right)^{i} \right] \\
&= n\,\varphi_p(t) \binom{n-1}{k-1} \left(\Phi_p(-t)\right)^{n-k} \left(\Phi_p(t)\right)^{k-1}.
\end{aligned}$$

□

References

[1] Arellano-Valle, R. B. and M. G. Genton (2005). On fundamental skew distributions. *J. Multivariate Anal. 96*(1), 93–116.

[2] Arellano-Valle, R. B. and M. G. Genton (2007). On the exact distribution of linear combinations of order statistics from dependent random variables. *J. Multivariate Anal. 98*(10), 1876–1894.

[3] Arellano-Valle, R. B. and M. G. Genton (2008). Corrigendum to "On the exact distribution of linear combinations of order statistics from dependent random variables": [J. Multivariate Anal. 98 (2007) 1876-1894]. *J. Multivariate Anal. 99*(5), 1013.

[4] Arellano-Valle, R. B. and M. G. Genton (2010). Multivariate unified skew-elliptical distributions. *Chil. J. Stat. 1*(1), 17–33.

[5] Arellano-Valle, R. B. and W.-D. Richter (2012). On skewed continuous $l_{n,p}$-symmetric distributions. *Chil. J. Stat. 3*(2), 193–212.

[6] Balakrishnan, N. and A. C. Cohen (1991). *Order Statistics and Inference. Estimation Methods.* Academic Press, Boston MA.

[7] Balakrishnan, N. and C. R. Rao, editors (1998). *Order Statistics: Applications*. Elsevier Science, Amsterdam.

[8] Barakat, H. (2009). Multivariate order statistics based on dependent and nonidentically distributed random variables. *J. Multivariate Anal. 100*(1), 81–90.

[9] Barakat, H. M. and E. M. Nigm (2012). On multivariate order statistics. Asymptotic theory and computing moments. *Kuwait J. Sci. Eng 39*(1A), 113–127.

[10] Batún-Cutz, J., G. González-Farías, and W.-D. Richter (2013). Maximum distributions for $l_{2,p}$-symmetric vectors are skewed $l_{1,p}$-symmetric distributions. *Stat. Probab. Lett. 83*(10), 2260–2268.

[11] David, H. A. and H. N. Nagaraja (2003). *Order Statistics*. Third edition. John Wiley & Sons, Hoboken NJ.

[12] Galambos, J. (1975). Order statistics of samples from multivariate distributions. *J. Am. Stat. Assoc. 70*(351a), 674–680.

[13] Giri, B. K., S. Sarkar, S. Mazumder, and K. Das (2015). A computationally efficient order statistics based outlier detection technique for EEG signals. In *Engineering in Medicine and Biology Society (EMBC), 2015 37th Annual International Conference of the IEEE*, pp. 4765–4768.

[14] Glen, A. G., L. M. Leemis, and D. R. Barr (2001). Order statistics in goodness-of-fit testing. *IEEE T. Reliab. 50*(2), 209–213.

[15] Greenberg, B. G. and A. E. Sarhan (1958). Applications of order statistics to health data. *Am. J. Public Health N. 48*(10), 1388–1394.

[16] Günzel, T., W.-D. Richter, S. Scheutzow, K. Schicker, and J. Venz (2012). Geometric approach to the skewed normal distribution. *J. Stat. Plann. Inference 142*(12), 3209–3224.

[17] Gupta, A. K. and D. Song (1997). l_p-norm spherical distributions. *J. Stat. Plann. Inference 60*(2), 241–260.

[18] Gupta, S. S., K. Nagel, and S. Panchapakesan (1973). On the order statistics from equally correlated normal random variables. *Biometrika 60*(2), 403–413.

[19] Jamalizadeh, A. and N. Balakrishnan (2010). Distributions of order statistics and linear combinations of order statistics from an elliptical distribution as mixtures of unified skew-elliptical distributions. *J. Multivar. Anal. 101*(6), 1412–1427.

[20] Jamalizadeh, A. and N. Balakrishnan (2013). A note on "Maximum distributions for $l_{2,p}$-symmetric vectors are skewed $l_{1,p}$-symmetric distributions" by Batún-Cutz et al. (2013). *Stat. Probab. Lett. 83*(11), 2522–2523.

[21] Kejgir, S. and M. Kokare (2009). Optimization of bit plane combination for efficient digital image watermarking. *Int. J. Comput. Sci. Inf. Secur. 4*(1 & 2).

[22] Kudryashov, B. D., A. V. Porov, and E. L. Oh (2008). Scalar quantization for audio data coding. Available at https://arxiv.org/ftp/arxiv/papers/0806/0806.4293.pdf.

[23] Lacaux, C., A. Muller-Gueudin, R. Ranta, and S. Tindel (2014). Convergence and performance of the peeling wavelet denoising algorithm. *Metrika 77*(4), 509–537.

[24] Müller, K. and W.-D. Richter (2015). Exact extreme value, product, and ratio distributions under non-standard assumptions. *AStA Adv. stat. Anal. 99*(1), 1–30.

[25] Müller, K. and W.-D. Richter (2016a). Exact distributions of order statistics of dependent random variables from $l_{n,p}$-symmetric sample distributions, $n \in \{3, 4\}$. *Depend. Model. 4*(1), 1–29.

[26] Müller, K. and W.-D. Richter (2016b). Extreme value distributions for dependent jointly $l_{n,p}$-symmetrically distributed random variables. *Depend. Model. 4*(1), 30–62.

[27] Murtagh, F., P. Contreras, and J.-L. Starck (2009). Scale-based Gaussian coverings: combining intra and inter mixture models in image segmentation. *Entropy 11*(3), 513–528.

[28] Pitas, I. and A. N. Venetsanopoulos (1992). Order statistics in digital image processing. *Proc. IEEE 80*(12), 1893–1921.

[29] Richter, W.-D. (2007). Generalized spherical and simplicial coordinates. *J. Math. Anal. Appl. 336*(2), 1187–1202.

[30] Richter, W.-D. (2009). Continuous $l_{n,p}$-symmetric distributions. *Lith. Math. J. 49*(1), 93–108.

[31] Richter, W.-D. (2013). Geometric and stochastic representations for elliptically contoured distributions. *Comm. Stat. Theory Methods 42*(4), 579–602.

[32] Richter, W.-D. (2014). Geometric disintegration and star-shaped distributions. *J. Stat. Distrib. Appl.* 1:20.

[33] Richter, W.-D. (2015). Convex and radially concave contoured distributions. *J. Probab. Stat. 2015*, Article ID 165468, 12 pages.

[34] Richter, W.-D. and J. Venz (2014). Geometric representations of multivariate skewed elliptically contoured distributions. *Chil. J. Stat. 5*(2), 71–90.

[35] Saatci, E. and A. Akan (2010). Respiratory parameter estimation in non-invasive ventilation based on generalized Gaussian noise models. *Signal Process. 90*(2), 480–489.

[36] Shafiei, S., M. Doostparast, and A. Jamalizadeh (2015). Prediction of variables via their order statistics in bivariate elliptical distributions with application in the financial markets. *Commun. Stat. Theory Methods 44*(3), 627–643.

[37] Wiaux, Y., L. Jacques, G. Puy, A. M. Scaife, and P. Vandergheynst (2009). Compressed sensing imaging techniques for radio interferometry. *Mon. Not. R. Astron. Soc. 395*(3), 1733–1742.

[38] Wiaux, Y., G. Puy, and P. Vandergheynst (2010). Compressed sensing reconstruction of a string signal from interferometric observations of the cosmic microwave background. *Mon. Not. R. Astron. Soc. 402*(4), 2626–2636.