Multi-normex distributions for the sum of random vectors.
Rates of convergence

Marie Kratz\textsuperscript{1} · Evgeny Prokopenko\textsuperscript{1,2}

Received: 20 October 2021 / Revised: 19 December 2022 / Accepted: 21 December 2022 / Published online: 13 January 2023
© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2023

Abstract
We build a sharp approximation of the whole distribution of the sum of iid heavy-tailed random vectors, combining mean and extreme behaviors. It extends the so-called 'normex' approach from a univariate to a multivariate framework. We propose two possible multi-normex distributions, named $d$-Normex and MRV-Normex. Both rely on the Gaussian distribution for describing the mean behavior, via the CLT, while the difference between the two versions comes from using the exact distribution or the EV theorem for the maximum. The main theorems provide the rate of convergence for each version of the multi-normex distributions towards the distribution of the sum, assuming second order regular variation property for the norm of the parent random vector when considering the MRV-normex case. Numerical illustrations and comparisons are proposed with various dependence structures on the parent random vector, using QQ-plots based on geometrical quantiles.

Keywords Aggregation · Central limit theorem · Dependence · Extreme value theorem · Geometrical quantiles · Multivariate extremes · Multivariate regular variation · (Multivariate) Pareto distribution · Ordered statistics · QQ-plots · Rate of convergence · Second order regular variation · Sum of random vectors

AMS 2000 Subject Classifications 60E05 · 60F05 · 60G70 · 60F15 · 62G30 · 41A25

\textsuperscript{1} CREAR, ESSEC Business School Paris, Paris, France
\textsuperscript{2} Novosibirsk State University, Novosibirsk, Russia

\textsuperscript{*} Marie Kratz
kratz@essec.edu

Evgeny Prokopenko
prokopenko@math.nsc.ru
1 Introduction

Motivation Looking for the most accurate possible evaluation of the distribution of the sum of random variables, or vectors, or processes, with unknown distributions, has always been a classical problem in the probabilistic and statistical literature, with various answers depending on the given framework and on the specific application in view. On one hand, (uni- or multivariate) Central Limit Theorems (CLT) or Functional ones prove, under finite variance for the sum components or and additional conditions, the asymptotic Gaussian behavior of the sum with some rate of convergence, focusing on the 'body' of the distribution. When considering heavy-tailed marginal distributions, Generalized CLT with the convergence to stable distributions, handle the case of infinite variance (see e.g. Samorodnitsky and Taqqu (1994), Petrov (1995), and references therein), while, in the case of finite variance, an alternative way is to consider trimmed sums, removing extremes from the sample, to improve the rate of convergence: see e.g. Mori (1984), Hahn (1991) and references therein. When interested in tail distributions, CLTs may give poor results, especially when considering heavy tails. That is why different approaches have been developed, among which large deviation theorems (see e.g. Petrov (1975), Borovkov (2020) for light tails and Mikosch and Nagaev (1998), Foss et al. (2013), Lehtomaa (2017) for heavy tails, and references therein), extreme value theorems (EVT) focusing on the tail only (see e.g. Embrechts et al. (1997), de Haan and Ferreira (2006), Resnick (2007)), and hybrid distributions combining (asymptotic) distributions for both the main and extreme behaviors when considering independent random variables (see e.g. Csörgő et al. (1988), Zaliapin et al. (2005), Kratz (2014), Müller (2019), we use the name given in Kratz (2014) for this type of hybrid distribution/method/approach, namely Normex distribution/method/approach. Recall briefly the idea of Normex (for 'Norm(al)-Ex(tremes)') method. It consists of rewriting the sum of random variables as the sum of their ordered statistics, and splitting it into two main parts, a trimmed sum removing the extremes, and the extremes. Using that the trimmed sum of the first \( n-k-1 \) ordered statistics is conditionnally independent of the \( k \) largest order statistics, given the \( (n-k) \)-th order statistics, we can express the distribution of the sum, integrating w.r.t. to the \( (n-k) \)-th order statistics and using a CLT for the conditional trimmed sum, and an EVT one for the \( k \) largest order statistics. Note that a benefit of Normex approach is that it does not require any condition on the existence of moments, as the CLT applies on truncated random variables.

The following example (as developed in Kratz (2014)), simulating identically distributed and independent (iid) Pareto(\( \alpha \)) (such that \( F(x) = x^{-\alpha} \) for \( x > 1 \)) random variables (rv), with \( \alpha = 2.3 \) (finite variance, but no third moment), illustrates perfectly the adding value of combining main and extreme behaviors.

Indeed, on the QQ-plots given in Fig. 1, we can compare the fit of the distribution of the empirical sample with the following three distributions: the Gaussian one (CLT approach), the Fréchet one (EVT approach), the hybrid one (i.e. Normex, combining the CLT and the distribution of the maximum). For the hybrid Normex distribution, we may consider either the exact distribution of the
maximum, or its asymptotic approximation, the Fréchet distribution. Since both provide the same plot, we display it in the third plot on the right. Note that we choose a rather small number of components in the sum, \( n = 52 \) (corresponding to the aggregation of data on 1 year (or 52 weeks) as in financial applications), also to illustrate the speed of convergence when using asymptotic theorems. We observe that the CLT approach does not provide a sharp evaluation, even in the body of the distribution, due to this choice of \( n \), which cannot compensate yet the fact that the distribution of the rv is asymmetric and skewed; increasing \( n \) will of course improve the fit in the body of the distribution. Given the fact that the Pareto distribution belongs to the Fréchet maximum domain of attraction, using the Fréchet distribution for the distribution of the sum of Pareto rv’s gives a very sharp approximation in the tail (from the 93% quantile), but not for the average behavior, as expected. Finally, a perfect match between empirical quantiles and Normex ones is observed for the whole distribution in the right plot, even for a small number of summands.

**Goal of the study** It is natural to extend the normex approach to a multivariate framework. With this goal of proposing a multi-normex method and distribution, we consider iid random vectors \( X_1, \ldots, X_n \), with parent random vector \( X \) having a heavy-tailed \( d \)-dimensional distribution \( F_X \) and density \( f_X \) (when existing). Note that there are different ways to define multivariate extremes (see e.g. chap. 8 in Beirlant et al. (2004)). The chosen way in this paper is w.r.t. the norm \( \| \cdot \| \) in \( \mathbb{R}^d \), meaning that the ordered (w.r.t. the norm) vector of \( (X_1, \ldots, X_n) \), denoted by \( (X_{(1)}, \ldots, X_{(n)}) \), satisfies

\[
\|X_{(1)}\| \leq \|X_{(2)}\| \leq \cdots \leq \|X_{(n)}\|.
\]  

(1.1)

So, assuming \( F_X \) heavy-tailed means that \( \|X\| \) is a regularly varying rv with \( \alpha > 0 \), denoted by \( \|X\| \in \mathcal{R\nu}_-\alpha \), i.e. such that \( \lim_{t \to \infty} \mathbb{P}(\|X\| > tx)/\mathbb{P}(\|X\| > t) = x^{-\alpha} \), for \( x > 0 \).

We propose two versions of multi-normex. The first one, named \( d \)-Normex, is a natural extension to any dimension \( d \) of the univariate (\( d = 1 \)) normex method as developed in Kratz (2014): We approximate the distribution of the trimmed sum via the CLT
and consider the distribution of the maximum $X_{(n)}$. This latter distribution is approximated via the Extreme Value (EV) theorem in the second multi-normex version, named MRV-Normex.

Aiming at proving the benefit of using a multi-normex distribution for a better fit of the whole (unknown) heavy-tailed distribution $F_X$, we focus analytically on the case $\alpha \in (2, 3)$ (when $\|X\|$ has a finite second moment, but no third moment) to compare the rates of convergence when using the CLT and the multi-normex approach, respectively. Note our focus on heavy tailed distributions (i.e. distributions belonging to the max domain of attraction of Fréchet), where the impact of using Normex distribution will be much stronger than in the light tail case (because of the one big jump principle), in particular for risk analysis and management. We prove that the normex approach leads, as expected, to a better speed of convergence for evaluating the distribution of the sum than the CLT does, for such type of heavy-tailed distributions. When varying the fatness of the tail measured by $\alpha > 0$, we draw this comparison numerically, using geometrical multivariate quantiles (see e.g. Chaudhuri (1996), Dhar et al. (2014) or a brief description in Kratz and Prokopenko (2021)).

Structure of the paper In Section 2, besides general notation, we recall the normex approach and the generalized Berry-Esseen inequality. Then we give a specific result on conditional distributions of order statistics, which will be needed for the construction of multi-normex distributions. The two next sections develop the two multi-normex versions, $d$-Normex in Section 3 and MRV-Normex in Section 4. These sections have the same structure: we first define the multi-normex distribution, then we study analytically its rate of convergence, before ending with some examples. In Section 5, we consider those examples to study numerically the two versions of multi-normex distribution, comparing them with the empirical distribution of the sum (obtained via simulation) as well as, if relevant, with the Gaussian approximation when applying the CLT. Geometrical multivariate quantiles are computed to this aim and represented on QQ-plots. Section 6 concludes. The proofs of all analytical results are developed in the Appendix. More discussions on the existing literature with respect to our new results, with survey and additional examples or illustrations, can be found in Kratz and Prokopenko (2021).

2 Framework

2.1 Context and notation

Normex idea The Normex method clearly adapts to a multivariate framework. Using this approach, we split the maximum from the rest of the sum:

$$S_n = \sum_{k=1}^{n-1} X_{(k)} + X_{(n)},$$

(2.1)
taking into account the ‘principle of one big jump’, namely that the asymptotic tail behavior of the sum of heavy-tailed random vectors is led by that of the maximum. Indeed, this principle extends to dimension \( d > 1 \) by borrowing the multivariate subexponential distribution definition of Samorodnitsky and Sun (2016) (see Section 4 therein):

\[
\mathbb{P} \left( S_n > t A \right) \sim \mathbb{P} \left( X_{(n)} > t A \right),
\]

where \( A \subset \mathbb{R}^d \) is open, increasing, such that \( A^c \) convex and \( 0 \not\in \tilde{A} \). Note also earlier works on that notion by Cline and Resnick (1992) and Omey (2006).

Combining the CLT for the trimmed sum given the maximum, and the distribution of the maximum or its asymptotic distribution, leads to a multivariate version of Normex distribution. We name this multi-normex distribution as \( d \)-Normex, when using the distribution of the maximum, and as MRV-Normex when considering its asymptotic distribution (when rescaled).

**General notation** Before defining explicitly both versions of multi-normex distribution, let us introduce some general notation.

Let \( f_{(i)} \) denote the \( d \)-dimensional density, when existing, of the ordered vectors \( X_{(i)} \), for \( i = 1, \ldots, n \). The cumulative distribution function (cdf) of the norm \( \|X\| \) is denoted as \( F_{\|X\|}(\cdot) \), which we assume throughout the paper to be absolutely continuous, and its probability density function (pdf) as \( f_{\|X\|}(\cdot) \). Nevertheless, some of our results will be stated under the slightly stronger condition \( F_X \) absolutely continuous.

As we will work with truncated multidimensional distributions or vectors, let us introduce the following notions.

For any \( y > 0 \), we define the truncated (via the norm) multidimensional distribution \( F_{X/\|X\|}(\cdot \mid y) \) of \( X \) on \( \mathbb{R}^d \) as

\[
F_{X/\|X\|}(B \mid y) := \mathbb{P}(X \in B \mid \|X\| \leq y)
\]

for any event \( B \) of the Borel sigma-field \( \mathcal{B}(\mathbb{R}^d) \). We denote by \( f_{X/\|X\|}(\cdot) \) its pdf, when existing:

\[
f_{X/\|X\|}(x \mid y) = \frac{f_X(x) \mathbb{1}_{\{\|x\| \leq y\}}}{F_{\|X\|}(y)}.
\]

Let \( \overset{\circ}{X}_y \in \mathbb{R}^d \) denote the random vector with distribution \( F_{\overset{\circ}{X}_y} \) on the \((d - 1)\)-sphere \( S_y = \{x \in \mathbb{R}^d : \|x\| = y\} \) for \( y > 0 \), defined by

\[
F_{\overset{\circ}{X}_y}(B_y) = \mathbb{P} \left( \overset{\circ}{X}_y \in B_y \right) := \mathbb{P}(X \in B_y \mid \|X\| = y),
\]

where \( B_y \) belongs to the trace Borel \( \sigma \)-algebra on \( S_y \). With this definition, for any \( A \subseteq \mathbb{R}^d \), we can write
thus
\[ P(X \in A) = \int_0^\infty P(\|X\| = y) f_{\|X\|}(y) \, dy, \]

\[ P(X \in A) = \int_{\mathbb{R}_+} P(\overset{\circ}{X}_y \in A_y) f_{\|X\|}(y) \, dy, \quad \text{where} \quad A_y := A \cap S_y. \quad (2.5) \]

**Generalized Berry-Esseen inequality** As we are going to compare the rates of convergence when using, respectively, the CLT and the Normex method, let us recall the rate of convergence in the CLT provided by the Generalized Berry-Esseen inequality (see e.g. Corollary 18.3 of Bhattacharya and Rao (2010)), when assuming a 'moderate' heavy tail.

**Proposition 2.1** (Generalized Berry-Esseen inequality)

Let \( X_1, \ldots, X_n \) be i.i.d. centered random vectors with parent random vector \( X \) with values in \( \mathbb{R}^d \), with positive-definite covariance matrix \( \Sigma \). If
\[ \mathbb{E}\|X\|^\alpha < \infty, \quad \text{for some} \quad \alpha \in [2, 3], \]

then
\[ \sup_{\mathcal{B} \in \mathcal{C}} \left| P\left( S_n \in \mathcal{B} \right) - \Phi_0,\Sigma(\mathcal{B}) \right| \leq c \mathbb{E}\left( \|\Sigma^{-1/2}X\|^\alpha \right) n^{-(\alpha-2)/2}, \quad (2.6) \]
where \( \mathcal{C} \) is the class of all Borel-measurable convex subsets of \( \mathbb{R}^d \), \( c \) is a positive universal constant, and \( \Phi_0,\Sigma \) is the cdf of the centered normal multivariate distribution with covariance matrix \( \Sigma \).

Note that the (non-generalized) Berry-Esseen inequality, which holds for any \( \alpha \geq 3 \), corresponds to (2.6) when taking \( \alpha = 3 \).

### 2.2 A preliminary result on order statistics

We present a simple but elegant result on order statistics, needed for the proofs of the theorems, and also of interest in itself, completing the vast literature on order statistics. Its proof is given in Appendix 1.

**Lemma 2.1** The distribution of the order statistics \( X_{(1)}, \ldots, X_{(n-1)}, X_{(n)} \), conditionally on the event \( \|X_{(n)}\| = y \), is the distribution of the \( n-1 \) ordered statistics from the truncated distribution \( F_{X_{(1)}}(\cdot \mid y) \), and of an independent random vector \( X_y \) defined on the \((d-1)\)-sphere \( S_y \), for \( y > 0 \):
Multi-normex distributions for the sum of random vectors,…

\[ \mathcal{L}(X_1, \ldots, X_{n-1}, X_n \mid \|X_n\| = y) = \mathcal{L}(Y_1, \ldots, Y_{(n-1)}) \times \mathcal{L}(\hat{X}_y), \quad (2.7) \]

where \( Y_1, \ldots, Y_{n-1} \) are i.i.d. random vectors with multidimensional truncated distribution \( F_{X_1\mid\|X\|}(\cdot \mid y) \) defined in (2.2), and the random vector \( X_y \) has the distribution \( F_{X_y}(\cdot) \) defined in (2.4).

In particular, we have

\[ \mathcal{L}(X_1, \ldots, X_{n-1} \mid \|X_n\| = y) = \mathcal{L}(Y_1, \ldots, Y_{(n-1)}). \quad (2.8) \]

3 A first multi-normex version: \( d \)-Normex

We start building a first multi-normex version, using the Normex approach (2.1), then approximating the distribution of the trimmed sum via the CLT and keeping the distribution of the maximum \( X_{(n)} \). It is a natural extension to any dimension \( d \) of the univariate \((d = 1)\) Normex distribution as developed in Kratz (2014). Note that, when turning to data, the distribution of \( X_{(n)} \) may be approximated e.g. via simulations or, as will be done in the MRV-Normex, using another asymptotic theorem, the Extreme Value (EV) one.

3.1 Definition and Rate of Convergence

**Definition 3.1** The so-called \( d \)-Normex distribution function is defined, for \( B \subset \mathbb{R}^d \), as:

\[ G_n(B) = \mathbb{P}(\hat{Z} + X_{(n)} \in B) \]

where \( Z \) is, conditionally on the event \( (\|X_{(n)}\| = y) \), a Gaussian random vector with mean \( (n-1)\mu(y) \) and covariance matrix \( (n-1)\Sigma(y) \), the functions \( \mu(\cdot) \) and \( \Sigma(\cdot) \) are, respectively, the mean vector and covariance matrix of the truncated distribution \( F_{X_1\mid\|X\|}(\cdot) \) defined in (2.2).

Another way to formulate Definition 3.1 is the following:

\[ G_n(B) = \int_{\mathbb{R}^d} f_{(n)}(x) \Phi_{(n-1)\mu(\|x\|), (n-1)\Sigma(\|x\|)}(B - x) \, dx, \quad (3.1) \]

where \( \Phi_{m, \Gamma} \) denotes the cdf of the Gaussian vector with mean \( m \) and covariance matrix \( \Gamma \).

Let us turn to the evaluation of the Normex distribution for approximating the distribution of the sum of iid random vectors, studying its rate of convergence. We do it analytically. Although the multi-normex construction holds for any \( \alpha > 0 \), we state the result when assuming the same condition on moments as in the generalized Berry-Esseen inequality, namely \( \alpha \in (2, 3] \), to be able to compare the
results and show explicitly the benefit of using Normex approximation. Then we show numerically the general (for any $\alpha > 0$) good fit of $d$-Normex in an example (see Section 3.2).

The analytical result given in Theorem 3.1 shows that applying Normex method rather than the multivariate CLT improves, as expected, the accuracy of the evaluation of the (tail) distribution of the sum of heavy tailed vectors, with a better rate of convergence than the one of the CLT whenever the shape parameter $\alpha \in (2, 3]$.

**Theorem 3.1** Let $X_1, \ldots, X_n$ be i.i.d. random vectors with parent random vector $X$ with values in $\mathbb{R}^d$ such that:

(C1) For all $y > 0$, the truncated (w.r.t. the norm) distribution $F_{X, ||X||}(\cdot | y)$ defined in (2.2) is nondegenerate (i.e. for all $y > 0$ there is no hyperplane $H \subset \mathbb{R}^d$ such that $\mathbb{P}(X \in H \mid ||X|| \leq y) = 1$).

(C2) The distribution of the r.v $||X||$ is absolutely continuous and regularly varying at infinity: $||X|| \in RV_{-\alpha}$, with $\alpha > 0$.

Then, for any $\alpha \in (2, 3]$, there exists a slowly varying function $L(\cdot)$ such that

$$\sup_{B \in C} |\mathbb{P}(S_n \in B) - G_n(B)| \leq L(n) n^{-\frac{1}{2} + \frac{3-\alpha}{\alpha}},$$

where $C$ is the class of all Borel-measurable convex subsets of $\mathbb{R}^d$.

Let us briefly indicate how to reach the upper bound of this main result; for more details, see the proof developed in Appendix 2. First, we use the law of total probability conditioning by $X(n)$. Second, we apply the (non-generalized) Berry-Esseen inequality for the truncated r.v. $\{Y_i\}_{i \leq n-1}$ given the event $(||X(n)|| = y)$. The right-hand side of the inequality is of the order of $\frac{1}{\sqrt{n}} \mathbb{E}||Y||^3$, which is equivalent to $\frac{1}{\sqrt{n}} \mathbb{E}||X(n)||^{3-\alpha}$ whenever $\alpha \leq 3$. Finally, to derive the upper bound of the main result, we use that $||X(n)||$ is of the order $n^{1/\alpha}$ under (C2).

**Remark 3.1**

(i) We consider the case $\alpha \in (2, 3]$ since it is the condition under which the Generalized Berry-Esseen holds. For $\alpha > 3$, the bound given in Theorem 2.1 is the same as that of Berry-Esseen inequality, making the analytical comparison useless. Indeed, in such a case, the bound $\frac{1}{\sqrt{n}} \mathbb{E}||Y||^3$ reduces simply to the order $\frac{1}{\sqrt{n}}$ (see (8.6) in the proof), giving back the same rate as for the CLT. It means to look for an alternative way if we want to study analytically the Normex rate of convergence. We might use Edgeworth expansions, but it evolves too heavy computations (as we could experience for r.v (case $d = 1$), conditioning on...
X\((\alpha)\). This is why we show numerically the benefit of using Normex distribution, as illustrated in Section 5.

(ii) The rate of convergence given in Theorem 3.1 is better than the one provided in the generalized Berry-Esseen inequality (Proposition 2.1), whenever \(\alpha \in (2, 3)\) (and whatever \(n\)), as \(\frac{\alpha-2}{2} < \frac{1}{2} - \frac{3-\alpha}{4}\). Note also that, in the case \(\alpha = 3\) and \(\mathbb{E}\|X\|^3 = \infty\), the inequality in Theorem 3.1 is slightly sharper than the inequality that can be obtained by the Berry-Esseen theorem (replacing \(n^{-1/2+\varepsilon}, \varepsilon > 0\), by \(L(n)n^{-1/2}\)).

(iii) One can apply the Normex method with any norm on \(\mathbb{R}^d\), for instance the \(L^1\) norm defined, for \(x = (x_1, \ldots, x_d) \in \mathbb{R}^d\), by \(\|x\|_1 := \sum_{i=1}^d |x_i|\). In such a case, for positive random variables, Condition (C2) translates into the assumption

\[(C2^\ast) \quad S_d := \sum_{i=1}^d X^{(i)} \in \mathcal{RV}_{-\alpha},\]

where \(X^{(i)}\), for \(i = 1, \ldots, d\), denote the components of \(X\). We may want to relate this \(\mathcal{RV}\) property on the sum, with conditions on the random vector itself. This topic has already been investigated in the literature; see e.g. Basrak et al. (2002), Barbe et al. (2006), Mainik and Embrechts (2013), Cuberos et al. (2015). For instance, assuming \(X\) multivariate regularly varying, \(X \in \mathcal{MRV}_{-\alpha}(b, \nu)\) implies that the sum \(S_d \in \mathcal{RV}_{-\alpha}(b)\). We will come back on the MRV notion in Section 4.

### 3.2 Example of the Multivariate Pareto-Lomax distribution

Let us consider a \(d\)–dimensional random vector \(X = (X^{(1)}, \ldots, X^{(d)})\) having a multivariate Pareto-Lomax(\(\alpha\)) distribution, with \(\alpha > 0\), i.e. with survival distribution function defined, for any non-negative real numbers \(x_1, \ldots, x_d\), by

\[F_X(x_1, \ldots, x_d) := \mathbb{P}\left(X^{(i)} > x_i, i = 1, \ldots, d\right) = \left(1 + \sum_{i=1}^d x_i\right)^{-\alpha},\]

from which marginal distributions, expectation and covariance matrix of a multivariate Pareto-Lomax vector follow via straightforward computations. We consider the case \(\alpha \in (2, 3)\), as in Theorem 3.1 (even if the construction holds for any \(\alpha > 0\)). As \(d\)-Normex can be applied for any norm on \(\mathbb{R}^d\), we choose the example of the \(L_1\) norm

\[\|x\|_1 := \sum_{i=1}^d x_i\]

to simplify the computations. We also take the example of \(d = 3\) for illustration.

We can express the cdf of the rv \(\|X\|\), for \(y > 0\), as
Now, let us compute the moments for the \(d\)-dimensional truncated Pareto-Lomax random vector, denoted by \(Y\), having cdf \(F_{X\mid\|X\|}(\cdot\mid y)\) (see Definition 2.2), expectation \(\mu(y)\) and covariance matrix \(\Sigma(y)\). We have, for any \(j \in \{1, 2\}\), if \(\alpha \neq 1\) (which is our case), and, if \(\alpha \neq 1, 2\) (also our case), from which can be deduced the covariance matrix \(\Sigma(y) = (\Sigma_{ij}(y))_{i,j}\).

\[
\begin{align*}
\mathbb{E}[Y_j] & : = \mathbb{E}[Y_j] = \mathbb{E}[X_j \mid \|X\| \leq y] \\
& = \frac{(y + 1)^{-\alpha-2} - (\alpha + 2)y(1 + y)^{\alpha+1}(\alpha y + 3)/6 - 1}{(\alpha - 1)F_{\|X\|}(y)},
\end{align*}
\]

and, if \(\alpha \neq 1, 2\) (also our case),

\[
\begin{align*}
\mathbb{E}[Y_1 Y_2] & = \mathbb{E}[X_1 X_2 \mid \|X\| \leq y] = \frac{1}{F_{\|X\|}(y)} \left( 1 - (y + 1)^{-\alpha-2} \right) \\
& \times \left( (\alpha + 2)y((\alpha + 1)y(\alpha y - 1)/24 + 1/6) + 1/2 + 1) - 1 \right),
\end{align*}
\]

\[
\begin{align*}
\mathbb{E}[Y_1^2] & = \mathbb{E}[X_1^2 \mid \|X\| \leq y] = \frac{2}{F_{\|X\|}(y)} \left( 1 - (y + 1)^{-\alpha-2} \right) \\
& \times \left[ (\alpha + 2)y((\alpha y(\alpha y(\alpha y - 1)/24 + 1/6) + 1/2) + 1) + 1 \right],
\end{align*}
\]

from which can be deduced the covariance matrix \(\Sigma(y) = (\Sigma_{ij}(y))_{i,j}\).

Therefore, the Gaussian cdf \(\Phi_{(n-1)\mu(\|X\|), (n-1)\Sigma(\|X\|)}\) introduced in Definition 3.1 is explicitly determined, and so is the \(d\)-Normex distribution \(G_n\) defined in (3.1).

### 4 MRV-Normex

Here we investigate a more universal version of multi-normex, named MRV-Normex, using an asymptotic theorem for the maximum, namely the Extreme Value (EV) one. Given our focus on the sum of iid heavy-tailed (w.r.t. the norm) random vectors, we consider the standard extreme value theory (EVT) framework of multivariate regularly varying (MRV), a natural extension of the regular variation in a multivariate framework. In fact, to obtain the rate of convergence of this multi-normex approximation, we assume slightly stronger assumption than MRV, asking for a uniform asymptotic independence of the polar coordinates of the random vector, as made explicit in Condition \((M_\delta)\) of Theorem 4.1.
4.1 Rate of convergence in the EV Theorem: Discussion of its assumptions

In order to obtain the rate of convergence for the MRV-Normex approximation of the sum, we first need to discuss the rate of convergence in the Extreme Value Theorem to control the difference between the norm of the maximum $\|X_{(n)}\|$ and the limit Fréchet distribution.

Let $\{X_n, n \geq 1\}$ be i.i.d. random variables with c.d.f. $F_X$. By the EV Theorem, $F_X$ belongs to the maximum domain of attraction (MDA) of an extreme-value distribution $G_\gamma$, with $\gamma \in \mathbb{R}$, i.e. there exist normalizing constants $a_n > 0$ and $b_n \in \mathbb{R}$ such that

$$\mathbb{P}\left( \max_{1 \leq i \leq n} X_i \leq a_n x + b_n \right) = F_X^n(a_n x + b_n) \xrightarrow{n \to \infty} G_\gamma(x).$$  \hspace{1cm} (4.1)

Note that the limit in (4.1) remains unchanged when replacing $(a_n)$ and $(b_n)$ with $(\tilde{a}_n)$ and $(\tilde{b}_n)$, as long as

$$\lim_{n \to \infty} \frac{\tilde{a}_n}{a_n} = 1 \quad \text{and} \quad \lim_{n \to \infty} \frac{b_n - \tilde{b}_n}{a_n} = 0. \hspace{1cm} (4.2)$$

Further discussion on the choice of $(a_n)$ and $(b_n)$ can be found in Kratz and Prokopenko (2021). Let us introduce the real function $g$ defined on $\mathbb{R}^+$ by:

$$g : = \left( \frac{1}{-\log F_X} \right)^{\leftarrow} \hspace{1cm} (4.3)$$

($\leftarrow$ denoting the left-continuous inverse function).

It is straightforward to show that the convergence (4.1) is equivalent to

$$\lim_{t \to \infty} \frac{g(tx) - g(t)}{a(t)} = \frac{x^\gamma - 1}{\gamma}, \quad \forall x > 0, \hspace{1cm} (4.4)$$

for some $\gamma \in \mathbb{R}$ and auxiliary positive function $a$ defined on $\mathbb{R}^+$. (The function $g$ is said to be of extended regular variation, $g \in ERV_\gamma(a)$).

To describe the rate of convergence in the EV theorem, we refer to two studies developed for $F_X$ belonging to any MDA, under slightly different assumptions (discussed in Kratz and Prokopenko (2021)), namely Falk and Marohn (1993) with a direct condition on the derivative of the distribution $F_X$, and de Haan and Resnick (1996) assuming a second-order von Mises condition on $g$ defined in (4.3). Focusing here on the case of $F_X \in MDA(\text{Fréchet})$, we look for a condition on $F_X$ involving $RV$ properties to replace the second-order von Mises condition on $g$ and to retrieve the exact rate of convergence described in de Haan and Resnick (1996). This is presented in Proposition 4.1.

**Proposition 4.1** Suppose $\bar{F}_X \in RV_{-\alpha}, \alpha > 0$, and $F_X$ is twice differentiable with pdf $f_X$. The rate of convergence for the EV Theorem as given in de Haan and Resnick (1996), Theorem 4.1, holds when replacing the condition $g \in 2-von~Mises(-\alpha,-\rho)$ ($g$ being defined in (4.3)), where $\rho > 0$, by the condition
Namely, there exists a constant $C > 0$ (that is defined explicitly) such that

\[
\lim_{n \to \infty} \frac{\sup_{A \in B(\mathbb{R})} \mathbb{P}\left( a_n^{-1} \left( \max_{1 \leq i \leq n} X_i - b_n \right) \in A \right) - G_r(A)}{|A(n)|} = C,
\]

where $a_n = ng'(n)$ and $b_n = g(n)$.

Under these assumptions, the function $A(t) := tg''(t) - \gamma + 1$ belongs to the $\mathcal{RV}_{-\beta}$ class with $\rho := -\min \left\{ 1, \frac{\beta}{\alpha} \right\}$.

Indeed, since $F_X \in \text{MDA}(\text{Fréchet})$, the Potter bounds can be directly established from a $2\mathcal{RV}$ condition on $g'$ via Proposition 4 in Hua and Joe (2011), which in turn follows from a $2\mathcal{RV}$ condition on $f_X$ by Lemma 4.1 (which proof is provided in the Appendix 3). This latter condition is equivalent to our assumption (4.5), as stated in Lemma 3 from Hua and Joe (2011). Once we have the Potter bounds, we can replicate exactly the proof of Theorem 4.1 as given in de Haan and Resnick (1996), obtaining the same rate of convergence. In fact, Hua and Joe (2011) proved their results, Lemma 3 and Proposition 4, in the case $\alpha < 0$, as they used them for $g(t) = \bar{F}(t)$. But their proof can be repeated line by line for an arbitrary $\alpha \in \mathbb{R}$.

**Lemma 4.1** If $f_X \in 2\mathcal{RV}_{-\alpha-1,-\beta}$, with $\alpha > 0$ and $\beta > 0$, then the derivative $g'$ of $g$ defined in (4.3), satisfies $g' \in 2\mathcal{RV}_{\frac{1}{\alpha}-1,\rho}$ where $\rho := -\min \left\{ 1, \frac{\beta}{\alpha} \right\} (< 0)$.

### 4.2 Rate of convergence for MRV-Normex

First, let us recall the MRV definition based on the pseudo-polar representation.

**Definition 4.1** The random vector $X \in \mathcal{MRV}_{-\alpha}$, with $\alpha > 0$, if there exists a $d$-dimensional random vector $\Theta$ with values in the unit sphere $S_1$ in $\mathbb{R}^d$ w.r.t. the norm $\| \cdot \|$, such that, for $t > 0$,

\[
\frac{\mathbb{P}(\|X\| > tu, X/\|X\| \in \cdot)}{\mathbb{P}(\|X\| > u)} \to t^{-\alpha} \mathbb{P}(\Theta \in \cdot) \quad \text{as} \quad u \to \infty.
\]

Using this definition of MRV, in particular the random vector $\Theta$, we can define the MRV-Normex distribution as follows.

**Definition 4.2** The so-called MRV-Normex distribution function is defined, for $B \subset \mathbb{R}^d$, by:
Multi-normex distributions for the sum of random vectors,…

\[ GM_n(\mathbf{B}) := \mathbb{P}(H_{\alpha,n} \Theta + Z \in \mathbf{B}). \] (4.8)

with \( H_{\alpha,n} := a_n H + b_n \) where the random variable \( H_\alpha \) (with \( \alpha > 0 \)) is Fréchet distributed (i.e. \( \mathbb{P}(H_\alpha \leq x) = e^{-x^{-\alpha}}, \) for \( x > 0 \)) and independent of the random vector \( \Theta \) introduced in (4.7), the normalizing sequences satisfy the standard conditions of the EV theorem, namely \( a_n = c_n^{1/\alpha} \) with \( c_\alpha := \lim_{y \to \infty} y^{\alpha} \hat{F}(\|X\|)(y), \) and \( b_n = 0. \) The \( d \)-dimensional random vector \( Z \), also assumed to be independent of \( \Theta \), is, conditionally to the event \((H_{\alpha,n} = y), \) with \( y > 0 \), normally \( \mathcal{N}(n^{-1} y, (n^{-1})\Sigma(y)) \)-distributed, the mean vector and covariance matrix being those of the truncated distribution \( \hat{F}_{X_\|X\|}(.|y) \) defined in (2.2).

The MRV-Normex cdf can be rewritten as

\[ GM_n(\mathbf{B}) := \int_{0}^{\infty} f_{H_{\alpha,n}}(y) \mathbb{P}(y \Theta + Z_y \in \mathbf{B}) dy, \] (4.9)

where \( Z_y \) is \( \mathcal{N}(n^{-1} \mu(y), (n^{-1})\Sigma(y)) \)-distributed and independent of \( \Theta \).

**Rate of convergence for the MRV-Normex approximation** We have the following result, which proof can be found in Appendix 3.

**Theorem 4.1** Let \( X_1, \ldots, X_n \) be i.i.d. random vectors with parent random vector \( X \) having values in \( \mathbb{R}^d \). Assume the following conditions:

(C1) given in Theorem 3.1 (namely, \( F_{X_\|X\|}(.|y) \) non degenerate \( \forall y > 0 \));

(M\( _{\|\|} \)) The distribution of the rv \( \|X\| \) is absolutely continuous and its pdf \( f_{\|X\|} \in 2\mathcal{RV}_{\alpha-1,-\rho} \) with \( \alpha > 0, \beta > 0 \);

(M\( _{\Theta} \)) There exists a function \( A \) such that \( A(t) \to 0, A(t) \in \mathcal{RV}_{-\rho} \) with \( \rho > 0 \), and

\[ \sup_{\mathbf{B} \in \mathcal{S}_1} \left| \mathbb{P}\left(\frac{X}{\|X\|} \in \mathbf{B} \mid \|X\| > t \right) - \mathbb{P}(\Theta \in \mathbf{B}) \right| \sim_{t \to \infty} A(t), \]

where the supremum is taken over all measurable subsets of \( \mathcal{S}_1 \).

Then, for any \( \alpha \in (2, 3), \beta > 0 \) and \( \rho > 0 \), there exists a slowly varying function \( L(\cdot) \) such that

\[ \sup_{\mathbf{B} \in \mathcal{C}} |\mathbb{P}(S_n \in \mathbf{B}) - GM_n(\mathbf{B})| \leq \left( n^{-\frac{1}{2} + \frac{1-\alpha}{\alpha}} + n^{-\frac{\beta}{\beta-\rho}} + n^{-\min(1, \frac{\rho}{\alpha})} \right) L(n), \] (4.10)

where \( \mathcal{C} \) is the class of all Borel-measurable convex subsets of \( \mathbb{R}^d \) and \( GM_n \) is defined in (4.9).
Remark 4.1

1. Compared with $d$-Normex (Theorem 3.1), there are two additional error terms in (4.10), price to pay for using the approximation for the maximum $\|X_{(n)}\|$. Nevertheless, if $\rho > \alpha$ and $\beta > \alpha$, then (4.10) rewrites as

$$\sup_{B \in \mathcal{C}} |\mathbb{P}(S_n \in B) - GM_n(B)| \leq n^{-\frac{1}{2} + \frac{\alpha}{\alpha - \beta} - L(n)},$$

providing the same rate of convergence given in Theorem 3.1.

2. If the supremum considered in $(M_\Theta)$ has a faster speed of convergence than that of $\mathcal{R}V$, then we can set $\rho = \infty$ and exclude the term $n^{-\frac{\alpha}{\alpha - \beta}}$ from inequality (4.10).

3. Discussion on Condition $(M_\Theta)$ (see Kratz and Prokopenko (2021) for further details and proofs of the following statements):
   - Assuming $\|X\| \in \mathcal{R}V$, Condition $(M_\Theta)$, which requires uniform convergence, is closely related to the MRV definition (4.7). Replacing this technical condition $(M_\Theta)$ with (4.7) might be investigated further.
   - If the norm $\|X\|$ and the direction $X/\|X\|$ of $X$ are independent, then $(M_\Theta)$ is satisfied.
   - Assuming that the density $f_X(x)$ depends only on the norm $\|x\|$, does not guarantee that the distribution of $X/\|X\|$ is uniform on the unit sphere $S_1$. It is uniform on $S_1$ for $L^p$- norms (or their weighted versions) if and only if $p = 1, 2, \infty$ (as a measure on the unit sphere is not proportional to a measure on the unit ball for $p \neq 1, 2, \infty$).

4.3 Examples

Let us develop two examples such that the marginal distributions are Pareto-Lomax, as in Example 3.2, so that it allows comparison with $d$-Normex. We consider two cases for the parent random vector $X$, when assuming its components to be, on one hand, independent, on the other hand, related with a survival Clayton copula. These are standard examples considered in the actuarial and risk literature (see e.g. Das and Kratz (2020)), in particular in reinsurance context for the Clayton copula is (see e.g. Dacorogna et al. (2018) and references therein). This second example includes itself two cases, when the polar coordinates of the considered vector $X$ are dependent (but asymptotically independent), and when they are independent. This latter case corresponds to Example 3.2.

We check that the conditions of Theorem 4.1 are satisfied whenever $\alpha \in (2, 3]$ (recall that this constraint on $\alpha$ appears only for the analytical comparison with the generalized Berry-Esseen inequality), but apply the MRV-Normex distribution for any positive $\alpha$, as the construction via asymptotic theorems remains valid whatever the value of this parameter.
4.3.1 Independent Pareto-Lomax marginals

Assume the components of the random vector \( \mathbf{X} \) to be iid with Pareto-Lomax(\( \alpha \)) distribution ((3.2) with \( d = 1 \)). Then its (non truncated) moments remain the same as in Example 3.2 and its covariance matrix is diagonal. As we developed Example 3.2 with the \( L_1 \)-norm, let us switch here to the \( L_\infty \)-norm \( \| \cdot \|_\infty \), more convenient in terms of computations in this framework.

The cdf of the norm of the vector \( \mathbf{X} \) being, for \( \alpha > 0 \),

\[
F_{\|\mathbf{X}\|_\infty}(y) = \mathbb{P}\left( \max\{X(1), \ldots, X(d)\} \leq y \right) = (1 - (1 + y)^{-\alpha})^d, \quad \text{for } y > 0,
\]

straightforward calculations give the following expressions for the truncated moments:

\[
\mu^{(1)}(y) = \mathbb{E}\left( X(1) \mid \|\mathbf{X}\|_\infty \leq y \right) = \mathbb{E}\left( X(1) \mid X(1) \leq y \right) = \frac{1}{1 - (1 + y)^{-\alpha}} \left( \frac{1}{\alpha - 1} (1 - (1 + y)^{-\alpha + 1}) - y(1 + y)^{-\alpha} \right),
\]

\[
\mathbb{E}\left( X^{(1)}^2 \mid \|\mathbf{X}\|_\infty \leq y \right) = \frac{2 - (1 + y)^{-\alpha} (ay((\alpha - 1)y + 2) + 2)}{(\alpha - 2)(\alpha - 1)(1 - (1 + y)^{-\alpha})},
\]

and 0 for the truncated covariances. When looking for the distribution of \( \Theta \), notice that, for any \( i \neq j \), for any \( \varepsilon_i > 0 \) and \( \varepsilon_j > 0 \), we have

\[
\lim_{n \to \infty} \mathbb{P}\left( X^{(i)} > \varepsilon_i t, X^{(j)} > \varepsilon_j t \mid \|\mathbf{X}\|_\infty > t \right) = 0.
\]

Therefore, the distribution of the random vector \( \Theta \) is discrete on the unit sphere \( S_1 \), with values given by the basis vectors \( e_i = (0, \ldots, 1, \ldots, 0) \) (where 1 is for the \( i \)-th component). It is straightforward to verify that \( F_{\|\mathbf{X}\|} \in \mathcal{RV}_{-\alpha,-\alpha} \), so that \( (M_{\|\|}) \) is satisfied, and that Condition \( (M_{\Theta}) \) holds with auxiliary function

\[
A(\cdot) \in \mathcal{RV}_{-\alpha},
\]

Finally, one may chose the normalizing sequences as \( a_n = (dn)^{1/\alpha} \) and \( b_n = -1 \) (see Kratz and Prokopenko (2021) for further details). The numerical implementation of this MRV-Normex approximation is developed in Section 5, along with that of the \( d \)-Normex one, for any positive \( \alpha \), both multi-normex methods being compared to the Gaussian approximation whenever \( \alpha \geq 2 \). The QQ-plots are drawn in Fig. 5.

4.3.2 Pareto-Lomax marginal distribution with survival Clayton copula

We introduce in this example some dependence among the components of \( \mathbf{X} \), choosing a survival Clayton copula (so, with upper tail dependence). To lighten the expressions of the computed moments, we choose \( d = 2 \). We consider \( \mathbf{X} = (X_1, X_2) \) with identical Pareto-Lomax (\( \alpha, 1 \)) marginal distributions, \( \alpha > 1 \), i.e.

\[
\hat{F}_1(x) = \hat{F}_2(x) = (1 + x)^{-\alpha}, \quad \forall x > 0,
\]

and survival Clayton copula on \([0, 1]^2\), with parameter \( \theta > 0 \), defined by
\[ P(X_1 > x_1, X_2 > x_2) = \left[ (1 + x_1)^{\alpha \theta} + (1 + x_2)^{\alpha \theta} - 1 \right]^{-1/\theta}, \]

with pdf

\[ f(x_1, x_2) = \frac{\alpha^2(1 + \theta)(1 + x_1)^{\alpha \theta - 1}(1 + x_2)^{\alpha \theta - 1}}{\left( (1 + x_1)^{\alpha \theta} + (1 + x_2)^{\alpha \theta} - 1 \right)^{1/\theta + 2}}. \]

Considering the \( L^\infty \)-norm, \( \| \cdot \|_\infty \), the survival cdf of the norm of the vector is:

\[ \overline{F}_{\|X\|_\infty}(t) = P\left( \max \{ X_1, X_2 \} > t \right) = 2(1 + t)^{-\alpha} - (2(1 + t)\alpha^{\theta} - 1)^{-1/\theta}. \]

We computed the truncated moments (of order 1 and 2), using an integral calculator (based on Maxima, a computer algebra system developed by W. Schelter, MIT), providing explicit but long expressions (displayed in Kratz and Prokopenko (2021)). For positive \( u_1 \) and \( u_2 \) such that \( \max \{ u_1, u_2 \} \geq 1 \), we can write

\[ \frac{f_X(tu)}{\overline{F}_{\|X\|_\infty}(t)} = \frac{\alpha^2(1 + \theta)(u_1 + \frac{1}{\theta})(u_2 + \frac{1}{\theta})^{\alpha \theta - 1}}{(2 - 2^{-1/\theta})(1 + O(\frac{1}{\theta}))\left( (u_1 + \frac{1}{\theta})^{\alpha \theta} + (u_2 + \frac{1}{\theta})^{\alpha \theta} \right)^{1/\theta + 2}}, \]

from which we deduce the limit as \( t \to \infty \), namely

\[ \lim_{t \to \infty} \frac{f_X(tu)}{\overline{F}_{\|X\|_\infty}(t)} = \frac{\alpha^2(1 + \theta)u_1^{\alpha \theta - 1}u_2^{\alpha \theta - 1}}{(2 - 2^{-1/\theta})(u_1^{\alpha \theta} + u_2^{\alpha \theta})^{1/\theta + 2}} = : f_{\Theta}(u). \]

Note that, for \( \| \cdot \| = \| \cdot \|_\infty \), the function \( \frac{f_X(tu)}{\overline{F}_{\|X\|_\infty}(t)} \) depends on \( t \), therefore the rv \( \|X\| \) and random vector \( X/\|X\| \) are not independent. They will be independent when replacing the \( L^\infty \)-norm with the \( L^1 \)-norm (\( \| \cdot \| = \| \cdot _1 \| \)) and choosing \( \alpha \theta = 1 \); this corresponds to the Pareto-Lomax Example 3.2. Turning to the conditions of Theorem 4.1, it is straightforward to check that \( \overline{F}_{\|X\|} \in 2\mathcal{R}V_{\max, \infty}(\alpha \theta, 1) \), so that \( (M_{\| \cdot \|}) \) is satisfied. Some computations are required for Condition \( (M_{\Theta}) \). We keep the maximum norm, i.e. \( \| \cdot \| = \| \cdot \|_\infty \), so that we exhibit an example with dependence between the polar coordinates (but with asymptotic independence), but consider the case \( \alpha \theta = 1 \) to simplify the computations. We obtain:

\[
\sup_{B} \left| P\left( \frac{X}{\|X\|} \in B \mid \|X\| \geq t \right) - P(\Theta \in B) \right| = \frac{1}{2} \int_{\|u\| \geq 1} \left| \frac{f_X(tu)}{\overline{F}_{\|X\|_\infty}(t)} - f_{\Theta}(u) \right| \, du
\]

\[
= \frac{\alpha(\alpha + 1)}{2t} \int_{\|u\| \geq 1} \left( \frac{\alpha^2(\alpha + 1)^2}{\overline{F}(t)(\alpha + 1)^{\alpha + 1}} - \frac{c_{\alpha}u^{\alpha + 1}}{\overline{F}(t)(\alpha + 1)^{\alpha + 1}c_{\alpha}u^{\alpha + 1}} \right) \, du,
\]

\[(4.11)\]
where \( c_a = (2 - 2^{-a}) \) and \(|\mathbf{u}| = u_1 + u_2\).

We can easily find an upper bound of type \( c / |\mathbf{u}|^{a+2} \) for the integrand (4.11), and, noticing that this integrand converges, as \( t \to \infty \), to \( \frac{|c_a| |\mathbf{u}|+c_a(\alpha+2)}{c^2|\mathbf{u}|^{a+3}} \) with \( c_a := 2^{-a-1} - 2 \), we can conclude, via the dominated convergence theorem, that the last integral in (4.11) converges to

\[
\int_{|\mathbf{u}| \geq 1} \frac{|c_a| |\mathbf{u}| + c_a(\alpha + 2)}{c^2|\mathbf{u}|^{a+3}} d\mathbf{u} < \infty.
\]  

(4.12)

Combining (4.11) and (4.12) provides that Condition \((M_\Theta)\) holds with \( A(t) = C_a / t \) for some constant \( C_a \in (0, \infty) \). As in the previous example (case of independent components), one may choose the normalizing sequences \( a_n = (c_a n)^{1/a} \) and \( b_n = -1 \). We refer to the next section for the numerical implementation of this example; see Fig. 6a and b for the QQ-plots (see Kratz and Prokopenko (2021) for additional illustrations).

5 QQ-plots of the various examples, illustrating both versions of the multi-normex method

5.1 Construction of the QQ-plots

Considering the examples given so far, we illustrate the benefit of the multi-normex method on \( d \)-dimensional QQ-plots based on geometrical quantiles. We refer mainly to Dhar et al. (2014) for definitions and detailed explanations; for a brief overview, see Kratz and Prokopenko (2021). Here are a few first key ideas of these objects, to help interpret the plots displayed in this section. The geometrical quantile, as given in Definition 5.1, is a generalization to higher dimension of the 1-dimensional quantile that can be defined as the solution of some optimization problem. So, one can formulate the same optimization problem in the \( d \)-dimensional case, for which the solution will be named geometrical quantile.

Definition 5.1 (Geometrical quantile, see Chaudhuri (1996)) For a random vector \( \mathbf{X} \) with a probability distribution \( F \) on \( \mathbb{R}^d \), the \( d \)-dimensional spatial quantile or geometrical quantile \( Q_F(\mathbf{u}) = (Q_{F,1}(\mathbf{u}), \ldots, Q_{F,d}(\mathbf{u})) \) is defined as

\[
Q_F(\mathbf{u}) = \arg \min_{\mathbf{Q} \in \mathbb{R}^d} \mathbb{E}\{\|\mathbf{X} - \mathbf{Q}\| - \|\mathbf{X}\| - \langle \mathbf{u}, \mathbf{Q} \rangle\},
\]

(4.13)

with \( \mathbf{u} \in B^d := \{\mathbf{v} \in \mathbb{R}^d, \|\mathbf{v}\| < 1\} \) and \( \langle \cdot, \cdot \rangle \) denoting the inner product.

In this way, the geometrical quantile is a point of \( \mathbb{R}^d \). While 1-dimensional quantiles are parameterized by the interval \((0, 1)\), multidimensional ones are parameterized by a \( d \)-dimensional unit ball \( B^d \). In this context, we refer to vectors of the unit ball as levels. Although geometrical quantiles reflect the structure of a \( d \)-dimensional distribution, they are abstract objects and do not have a nice interpretation as
1-dimensional quantiles do. Only the median has a geometrical sense: given a random vector, the median is a point of $\mathbb{R}^d$ such that the overall sum of the distances from this point to all values of the random vector is minimum (note that distances are multiplied by the 'probabilities' that the vector takes the considered values, respectively). Moreover, if the vector has a finite second moment, then its extreme quantiles will share the same speed of convergence towards infinity as any vector having the same covariance matrix (see Girard and Stupfler (2015, 2017)). Nevertheless, we can construct QQ-plots with these geometrical quantiles, with the aim at comparing $d-$dimensional distributions by comparing the plots between each other. The construction of QQ-plots is similar as in the 1-dimensional case. Considering two distributions on $\mathbb{R}^d$, we can solve the optimization problem for a fixed number of levels, say $N$, and obtain two sets of $N$ geometrical quantiles: $\{q_1, \ldots, q_N\}$ for the first distribution, and $\{q'_1, \ldots, q'_N\}$ for the second. Then we draw a 2-dimensional QQ-plot for each component of the $\mathbb{R}^d$-geometrical quantiles to get a visualization, obtaining $d$ QQ-plots: $\{(q_{ij}, q'_{ij}) : j = 1, \ldots, N\}$, for $i = 1, \ldots, d$. In Dhar et al. (2014), Theorem 2.2, the authors showed that for all $i = 1, \ldots, d$, any points (pairs of quantiles) in the $i$-th 2-dimensional plot will lie close to a straight line with slope 1 and intercept 0 if and only if the two distributions are equal. We are going to apply this result in our case, to check how the Normex distribution approximates the distribution of the sum.

Turning to our previous examples, we consider multidimensional distributions with Pareto-Lomax($\alpha$) marginals, varying their heaviness through the parameter $\alpha \in \{1.5, 2.3, 3.5\}$, with the different structures of dependence given in Examples 3.2 & 4.3. We choose $d \in \{2, 3\}$ and $n = 52$ (a rather small number to better illustrate the fast convergence of Normex). For each case, we evaluate both multinormex distributions and the normal distribution (except for $\alpha = 1.5$) obtained by the CLT; those distributions are evaluated empirically, via simulations, each sample being of size $10^7$. Then we proceed to their comparisons via the QQ-plots. We construct the QQ-plots in 4 main steps:

1. Simulate all the distributions to be compared: the distribution of the sum $S_n$, the normal distribution from CLT, the $d$-Normex distribution $G_n$, and the MRV-Normex distribution $MG_n$. Namely,

   (a) To obtain a simulated sample of size $10^7$ for the sum $S_n$, we simulate $n \times 10^7$ random vectors $X$ from the considered distribution.

   (b) To obtain a simulated sample from the $d$-Normex distribution defined in (3.1): First, we build $n = 52$ samples (of size $10^7$) from the $d$-dimensional multivariate Pareto distribution, from which we deduce a sample (of size $10^7$) for the ($d$-dimensional) maximum $X_{(n)}$ (see (1.1)). Second, for each element of the latter sample, we calculate its norm $y = ||X_{(n)}||$ and simulate a normal vector with mean $(n-1)\mu(y)$ and covariance matrix $(n-1)\Sigma(y)$ (described in Definition 3.1), collecting then $10^7$ Gaussian vectors. Finally, we sum maximum and normal vectors to produce a sample (of size $10^7$) from the $d$-Normex distribution.
(c) In order to simulate the MRV-Normex distribution defined in (4.8), we start simulating a sample (of size $10^7$) for the vector $\Theta$ (representing the direction) and an independent sample for the Fréchet distributed rv $H_{a,n}$ introduced in (4.9). Next, we collect the $10^7$ normal vectors with mean $(n-1)\mu(y)$ and covariance matrix $(n-1)\Sigma(y)$, where, now, $y = H_{a,n}$. Finally, we aggregate all the constructed samples according to (4.8).

2. Fix the set of levels $\mathcal{L} \subset \{ v \in \mathbb{R}^d, \|v\| < 1 \}$ with different lengths and directions. For $d = 3$, we choose 10 lengths, $\|v\| \in \{0, 0.2, 0.4, 0.6, 0.8, 0.9, 0.9225, 0.945, 0.9675, 0.99\}$ (half for the body of the distribution and half for its tail), and all the directions with angle (or factor) $\frac{\pi}{4}$ (to uniformly cover possible directions). It represents a total of 235 vectors. For $d = 2$, we choose 19 lengths and 8 directions with factor $\frac{\pi}{4}$, meaning 145 vectors.

3. Calculate the geometrical quantiles for the simulated samples of all considered distributions. It means to solve the numerically optimization problem (5.1) for each empirical distribution and for all levels $v$ from the set $\mathcal{L}$.

4. Draw a QQ-plot for the three pairs (or two, when the CLT cannot be applied): (sum, CLT), (sum, $d$-Normex) and (sum, MRV-Normex).

Note that the numerical implementation has been performed with Python (SciPy library). The scipy.optimize.minimize function based on the quasi-Newton method of Broyden, Fletcher, Goldfarb, and Shanno (see p.136 in Nocedal and Wright (2006)) has helped solve the numerical optimization problem (5.1) for empirical distributions. The gradient for the function has been calculated analytically. The computation time of one geometrical quantile was, on average, 53 seconds (on the computer i7 2GHz, 16 GB RAM).

5.2 Multivariate Pareto-Lomax distribution - Example 3.2

We first show the QQ-plots for the multivariate Pareto distribution, developed in Section 3.2, considering the dimension $d = 3$, the number of summands $n = 52$, and varying $\alpha$. We choose a small number of summands to highlight the performance of multi-normex methods, even in this case. The QQ-plots are given on a same row for each of the three components, and for each given approximation method (CLT, $d$-Normex and MRV-Normex). As expected, the plots are similar componentwise. A zoom of the center of the graph is given in the upper left corner of each plot, since it is the only region where the quantiles are so concentrated, making it more difficult to judge if they form a straight line. It should be noted that the center corresponds not only to levels with length less than or equal to 0.9 (marked in blue), but also to extreme levels (in red) with length greater than 0.9.

Case $\alpha \in (2, 3]$ Figure 3 exhibits the QQ-plots in the case $\alpha = 2.3$, i.e. in the framework of the multi-normex theorems, when using the Gaussian approximation (via the CLT) and both multi-normex approximations. We observe that all the points (pairs of quantiles) for $d$-Normex and MRV-Normex lie much closer to the line with slope
1 and intercept 0 than for the Gaussian distribution. Thus, we can see numerically that both multi-normex approximations better describe the distribution of the vectorial sum $\mathbf{S}_n$, as proved analytically in Theorems 3.1 and 4.1. While the QQ-plots look quite similar for the two versions of multi-normex approximations, when zooming, the fit is slightly better when using the $d$-Normex distribution than the MRV-Normex one that uses the Fréchet distribution for the distribution of the rescaled maximum.

**Arbitrary $\alpha > 0$** Here we consider other examples of heavy tails than the case $\alpha \in (2, 3]$, to illustrate the benefit of using Normex distributions rather than applying the generalized CLT with a Gaussian distribution (for finite variance) or a stable one (when the variance is infinite). We do it numerically as the upper bound of the Berry-Esseen inequality does not allow us to compare analytically the rates of convergence in terms of the number of summands $n$. Nevertheless, the constant given in terms of $\alpha$ and the number $k$ of largest order statistics when trimming the sum, may make a difference, as noticed in the 1-dimensional case (see Kratz (2014)). We give two examples, when $\alpha = 1.5$ (Fig. 2), case where the summands have no variance, and when $\alpha = 3.5$ (Fig. 4), i.e. beyond the frame of the generalized Berry-Esseen inequality. In Fig. 2, we only have two rows as the CLT does not apply and we did not build the QQ-plot for the stable distribution. The fit looks very good for both multi-normex methods, with barely no difference of fit between the two, as shown in the zoomed

**Multivariate Pareto ($\alpha = 1.5$)**

![3-dimensional QQ-plots for the empirical distribution (sample size = $10^7$) of the sum of 52 iid trivariate Pareto ($\alpha = 1.5$) random vectors, with two different approximations of the sum distribution: $d$-Normex (first row) and MRV-Normex (second row). Each column corresponds to a component (from the 1st to the 3rd). The red points on the plots correspond to extreme geometric quantiles (when the norm of the parameterized vectors is greater than 0.9).](Fig. 2 3-dimensional QQ-plots for the empirical distribution (sample size = $10^7$) of the sum of 52 iid trivariate Pareto ($\alpha = 1.5$) random vectors, with two different approximations of the sum distribution: $d$-Normex (first row) and MRV-Normex (second row). Each column corresponds to a component (from the 1st to the 3rd). The red points on the plots correspond to extreme geometric quantiles (when the norm of the parameterized vectors is greater than 0.9).)
Multi-normex distributions for the sum of random vectors...

Multivariate Pareto ($\alpha = 2.3$)

Fig. 3 3-dimensional QQ-plots for the empirical distribution (sample size = $10^7$) of the sum of 52 iid trivariate Pareto ($\alpha = 2.3$) random vectors, with three different approximations of the sum distribution: CLT (first row), $d$-Normex (second row) and MRV-Normex (third row). Each column corresponds to a component (from the 1st to the 3rd). The red points on the plots correspond to extreme geometric quantiles (when the norm of the parameterized vectors is greater than 0.9)

part. Whenever $\alpha = 3.5$, we again clearly observe in Fig. 4 an overall fit that gives the advantage to the multi-normex distributions, but with more difference, when zooming, between $d$-normex and MRV-normex than in the previous cases $\alpha = 1.5$ and $\alpha = 2.3$.

5.3 Pareto-Lomax marginal distributions with various dependence structures - Examples 4.3

Here, we consider the examples developed in Section 4.3, with Pareto-Lomax($\alpha$) marginal distributions, various dependence structures and the norm $\| \cdot \|_\infty$ instead of $\| \cdot \|_1$ as in the previous subsection. Figure 5 displays QQ-plots for independent
components of the vector $\mathbf{X}$, taking $\alpha = 2.3$ to be in the half-closed interval $(2,3]$ considered in our theorems. Here also, the good fit of the multi-normex distributions appears clearly, in particular when comparing with the Gaussian approximation. Nevertheless, the difference between the two multi-normex distributions is more pronounced in the center, as can be observed in the zoomed part.

When the dependence structure is given via a survival Clayton copula with parameter $\theta$, we provide the QQ-plots in Fig. 6a assuming $\alpha \theta = 0.5$. The previous
Independent Pareto-Lomax components ($\alpha = 2.3$)

Fig. 5 3-dimensional QQ-plots for the empirical distribution (sample size $= 10^7$) of the sum of 52 iid random vectors with independent Pareto-Lomax components of the vectors ($\alpha = 2.3$), with three different approximations of the sum distribution: CLT (first row), $d$-Normex (second row) and MRV-Normex (third row). Each column corresponds to a component (from the 1st to the 3rd). The red points on the plots correspond to extreme geometric quantiles (when the norm of the parameterized vectors is greater than 0.9).

observations hold too, with an increasing difference in the center between the two multi-normex distributions. The case $\alpha \theta = 1$ is illustrated in Fig. 6b. Note that this choice of $\theta = 1/\alpha$ is a standard example in the literature as it makes analytical computations much more tractable. It would also correspond to Example 3.2 if choosing the $L_1$-norm $\| \cdot \|_1$, allowing then a comparison of the results obtained respectively with the two norms. Recall that the analytical results (Theorems 3.1 and 4.1) are independent of the norm; it can also be observed numerically on (ranked) scatter plots (see Kratz and Prokopenko (2021)).
Conclusion

The purpose of this study was to build a sharp approximation of the whole distribution of the sum of iid random vectors under the presence of heavy tails. It has been done by extending the normex approach from a univariate to a multivariate framework, combining mean and extreme behaviors. We proposed two possible multi-normex distributions, named $d$-Normex and MRV-Normex. Both rely on the Gaussian distribution for describing the mean behavior, via the CLT, while the difference between the two versions comes from using the EV theorem or the exact distribution for the maximum. The main theorems establish the rate of convergence of each version of the multi-normex distributions towards the distribution of the sum. This is done analytically whenever the shape parameter $\alpha$ of the tail of the marginal distribution belongs to the interval $(2, 3]$, making the comparison with the generalized Berry-Esseen inequality relevant. For the
MRV-Normex, second order regular variation conditions are needed to obtain the main theorem. Numerical comparisons are developed for any value of \( \alpha \), for both multi-normex distributions, considering examples with different structure of dependence for the random vectors. Illustrations are made through multidimensional QQ-plots based on geometrical quantiles.

We focused on the case of heavy tailed random vectors, as it is of most interest in the risk literature. Nevertheless, this method could be extended to light tails vectors (whenever \( 1/\alpha = 0 \)) as the rate of convergence for the EV theorem is also known in such a case. It would then require to introduce a specific metric to evaluate the error for the whole distribution taking into account the impact of extremes. Moreover, the MRV-Normex approach has been developed conditioning on the norm of the maximum. It could be done conditioning on the maximum itself (a vector). To widen the applicability of the multi-normex methods, simple approximations of truncated moments could also be suggested (as numerical ones or evaluating them using e.g. a Pareto approximation). Finally, generalization of multi-normex distributions will be studied when introducing dependence between random vectors, then considering random processes. We intend to explore such topics in the near future. In the meantime, we are developing the statistical side of multi-normex, including building a statistical package.

### Appendix 1. Proof of Lemma 2.1

Let us first prove (2.8) with a very standard way (see e.g. David and Nagaraja (2004)). For this we prove the following result.

\[
\text{Let } f_{X_1, \ldots, X_{n-1} | ||X_n||}(x_1, \ldots, x_{n-1} | y) \text{ be the conditional pdf of } (X_1, \ldots, X_{n-1}) \text{ given the event } (||X_n|| = y). \text{ For any } x_1, \ldots, x_{n-1} \in \mathbb{R}^d, y \geq 0, \text{ such that } ||x_1|| \leq \cdots \leq ||x_{n-1}|| \leq y, \text{ we have }
\]

\[
f_{X_1, \ldots, X_{n-1} | ||X_n||}(x_1, \ldots, x_{n-1} | y) = (n-1)! \prod_{i=1}^{n-1} f_{X_i | ||X||}(x_i | y), \tag{7.1}
\]

where \( f_{X_i | ||X||}(\cdot | y) \) is defined in (2.3).

**Proof** Let \([y, y + \delta y]\) be a ‘small’ interval from \(\mathbb{R}^+\) and, for \(i = 1, \ldots, n-1\), let \([x_i, \delta x_i]\) denote a ‘small’ cube in \(\mathbb{R}^d\) with initial vertex \(x_i\) and measure \(\delta x_i\). As \(X_1, \ldots, X_{n}\) are the ordered (by norm) statistics (see (1.1)), there are \(n!\) permutations of the vector \((X_1, \ldots, X_n)\) to be \((X_{(1)}, \ldots, X_{(n)})\), hence we can write

\[
P\left( X_{(1)} \in [x_1, \delta x_1), \ldots, X_{(n-1)} \in [x_{n-1}, \delta x_{n-1}), ||X_{(n)}|| \in [y, y + \delta y) \right)
\]

\[
= n! \prod_{i=1}^{n-1} \left( f(x_i) \delta x_i \right) f_{||X||}(y) \delta y + O((\delta y)^2).
\]

Also, as \(||X_{(n)}||\) is the maximum of the iid rv \(||X_1||, \ldots, ||X_n||\), we have

\[
\odot \text{ Springer}
\]
Since \( h \) is continuous, \( \epsilon \rightarrow y \) and \( y \rightarrow y \), hence \( \epsilon \rightarrow y \). From which the results \((7.1) \) and \((7.8) \) follow.

Turning to \((7.7) \), it can be formalized as follows, using characteristic functions: \( \forall x_1, \ldots, x_n \in \mathbb{R}^d \) and any real \( y > 0 \), we have

\[
\mathbb{E} \left[ e^{\sum_{i=1}^{n} i \langle x_i, X_i \rangle} \| X_n \| = y \right] = \mathbb{E} \left[ e^{\sum_{i=1}^{n} i \langle x_i, Y_i \rangle} \right] \mathbb{E} \left[ e^{i \langle x_n, X_n \rangle} \| X_n \| = y \right]. \tag{7.3}
\]

**Proof** To simplify the notation, we denote by \( R_1, \ldots, R_n \) the norms \( \| X_1 \|, \ldots, \| X_n \| \). Since \( (R_1, \ldots, R_n) \) are iid, for all \( x_1, \ldots, x_n \in \mathbb{R}^d, y > 0 \), we have

\[
\mathbb{E} \left[ e^{\sum_{i=1}^{n} i \langle x_i, X_i \rangle} \mid R_1 = y \right] = n! \mathbb{E} \left[ e^{\sum_{i=1}^{n-1} i \langle x_i, X_i \rangle \| R_1\leq \cdots \leq R_{n-1}\leq R_n \} R_n = y \right] = (n - 1)! \mathbb{E} \left[ \left( e^{\sum_{i=1}^{n-1} i \langle x_i, X_i \rangle \| R_1\leq \cdots \leq R_{n-1}\leq R_n \} e^{n i \langle x_n, X_n \rangle \| R_n = y \} \right) \mid R_1 = y \right].
\]

The random variables in the two big brackets being independent, since one depends on \( X_1, \ldots, X_{n-1} \) and the other on \( X_n \), the last expression equals

\[
\mathbb{E} \left[ (n - 1)! e^{\sum_{i=1}^{n-1} i \langle x_i, X_i \rangle \| R_1\leq \cdots \leq R_{n-1}\leq y \} \right] \mathbb{E} \left[ e^{i \langle x_n, X_n \rangle \| R_n = y \} \right]. \tag{7.4}
\]

Note that, by \((7.8) \), the first expectation in \((7.4) \) equals to \( \mathbb{E} \left[ e^{\sum_{i=1}^{n-1} i \langle x_1, Y_i \rangle} \right] \).

Let us calculate the second expectation in \((7.4) \). Considering the \( \sigma \)-algebra \( \sigma(R) \) generated by the iid rv \( (R_1, \ldots, R_n) \) (with parent rv \( R \)), we can write, for any rv \( \xi \),

\[
\mathbb{E} [\xi \mid R_1] = \mathbb{E} [\mathbb{E} [\xi \mid \sigma(R)] \mid R_1] \quad \text{and} \quad \mathbb{E} [e^{i \langle x_n, X_n \rangle} \mid \sigma(R)] = \mathbb{E} [e^{i \langle x_n, X_n \rangle} \mid R_n] = h_{x_n} (R_n).
\]

We deduce that, for all \( y > 0 \),

\[
\mathbb{E} [n \{ R_n = y \} e^{i \langle x_n, X_n \rangle} \mid R_1] = \mathbb{E} [n \{ R_n = y \} h_{x_n} (R_n) \mid R_1] = h_{x_n} (y) \mathbb{E} [n \{ R_n = y \} \mid R_1] = h_{x_n} (y).
\]

Combining the results obtained for the two expectations in \((7.4) \) provides \((7.3) \), hence \((7.7) \).
Appendix 2. Proof of Theorem 3.1

Preliminary result  To prove Theorem 3.1, we need to show that the inverse matrix of the covariance matrix $\Sigma(y)$ of the truncated random vector $Y$ is bounded, namely:

**Lemma 8.1** Under the settings of Theorem 3.1, for some $C > 0$, there exists $\delta > 0$ such that $\mathbb{P}(\|X\| \leq \delta) < 1$ and

$$\|\Sigma(y)^{-1/2}\| \leq C \quad \text{for all } y \geq \delta.$$

**Proof** From the definition of the truncated distribution $F_X(\cdot | \|X\|)$ in (2.2) and continuity of $F_{\|X\|}(\cdot)$, one can show that, for any sequence $(y_n)$ converging to $y_0 \in \mathbb{R}^+$, with $\mathbb{P}(\|X\| \leq y_0) > 0$, we have

$$\Sigma(y_n) \to \Sigma(y_0) \text{ as } n \to \infty,$$

where, for convenience, we suppose $\Sigma(\infty) := \Sigma$, covariance matrix of the initial random vector $X$ with c.d.f. $F$. Condition $(C2)$ in Theorem 3.1 and the definition of $y_0$ imply that $\Sigma(y_0)$ is positive definite. Consequently, the square root of its inverse matrix exists, with finite norm:

$$\|\Sigma(y_0)^{-1/2}\| < \infty. \quad (8.1)$$

Now, assume that the statement of Lemma 8.1 is false. Then, for all $C > 0$ and $\delta > 0$ such that $\mathbb{P}(\|X\| \leq \delta) < 1$, there exists $y \geq \delta$ such that $\|\Sigma(y)^{-1/2}\| > C$.

Therefore, there exists a sequence $(y_n)_n \geq \delta$ such that

$$\|\Sigma(y_n)^{-1/2}\| \to \infty \text{ as } n \to \infty. \quad (8.2)$$

But we can choose a subsequence $(y_{n_k})_{k \geq 1}$ such that, as $k \to \infty$, either $y_{n_k} \to \infty$, or $y_{n_k} \to y_0 \geq \delta$ for some point $y_0 \geq \delta$. In both cases, (8.2) contradicts (8.1). We conclude that the statement of Lemma 8.1 is true.

**Proof of Theorem 3.1** Let $C$ denote a positive constant that may vary from line to line, all along the proof. Let $Y_1, \ldots, Y_{n-1}$ be i.i.d. random vectors, with parent random vector $Y$, having the truncated distribution $F_X(\cdot | \|X\|)$ defined in (2.2). We have, using conditional probabilities and applying (2.7) in Lemma 2.1,

$$\mathbb{P}(S_n \in B) = \mathbb{P}(X_{(1)} + \cdots + X_{(n)} \in B) \quad \text{for any } B \in \mathcal{C},$$

$$= \int_{\mathbb{R}^d} f(y_n)(x) \mathbb{P}(X_{(1)} + \cdots + X_{(n-1)} \in B - x \mid X_{(n)} = x) \, dx \quad \text{for any } B \in \mathcal{C}.$$

Using Definition 3.1 of the $d$-Normex distribution, we can write, for any $B \in \mathcal{C}$,
\[
\left| \mathbb{P}(S_n \in B) - G_n(B) \right| = \int_{\mathbb{R}^d} f_{(\alpha)}(x) \left| \mathbb{P}(T_{n-1} \in B - x) - \Phi_{(n-1)\mu(||x||),(n-1)\Sigma(||x||)}(B - x) \right| dx. \tag{8.3}
\]

Further, under (C2), the pdf of the rv $||X||$ is regularly varying, $f_{||X||}(\cdot) \in \mathcal{RY}_{-\alpha-1}$, i.e. (see e.g., Bingham et al. (1987)) there exists a slowly varying function $L(\cdot)$ such that
\[
f_{||X||}(y) = L(y)y^{-\alpha-1}. \tag{8.4}
\]

Let us choose $\delta > 0$ that satisfies Lemma 8.1 and such that the function of $y$ defined by $\sup_{t \in [\delta, y]} L(t)$ is slowly varying (see e.g. ex.4 p.58 in Bingham et al. (1987)).

Splitting the integration domain of the integral of (8.3) into two disjoint sets \{x : ||x|| < \delta\} and \{x : ||x|| \geq \delta\}, we have, on one hand,
\[
\int_{||x|| < \delta} f_{(\alpha)}(x) \left| \mathbb{P}(T_{n-1} \in B - x) - \Phi_{(n-1)\mu(||x||),(n-1)\Sigma(||x||)}(B - x) \right| dx \leq \int_{||x|| < \delta} f_{(\alpha)}(x)dx = \mathbb{P}(||X|| < \delta) = (\mathbb{P}(||X|| < \delta))^a. \tag{8.5}
\]

On the other hand, since $Y_1, \ldots, Y_{n-1}$ are i.i.d bounded random vectors with a non degenerate distribution (via (C2)), we can use the non-generalized Berry-Esseen inquality recalled in Proposition 3.1, and obtain
\[
\int_{||x|| \geq \delta} f_{(\alpha)}(x) \left| \mathbb{P}(T_{n-1} \in B - x) - \Phi_{(n-1)\mu(||x||),(n-1)\Sigma(||x||)}(B - x) \right| dx \leq \frac{C}{\sqrt{n}} \int_{||x|| \geq \delta} f_{(\alpha)}(x) \mathbb{E} \left( ||(\Sigma(||x||))^{-1/2}(Y - \mu(||x||))||^3 \right) dx, \tag{8.6}
\]

$Y$ denoting the parent random vector of $(Y_1, \ldots, Y_{n-1})$. It is straightforward to see that the last integral term in (8.6) can be written only in terms of $||x||$, which we denote by $y$, and of the pdf of the rv $||X_{(\alpha)}||$:
\[
\int_{||x|| \geq \delta} f_{(\alpha)}(x) \mathbb{E} \left( ||(\Sigma(||x||))^{-1/2}(Y - \mu(||x||))||^3 \right) dx = \int_{y \geq \delta} f_{||X_{(\alpha)}||}(y) \mathbb{E} \left( ||(\Sigma(y))^{-1/2}(Y - \mu(y))||^3 \right) dy. \tag{8.7}
\]

For $y \geq \delta$, using Lemma 8.1, it is straightforward to see that
\[
\mathbb{E} \left( ||\Sigma^{-1/2}(y)(Y - \mu(y))||^3 \right) \leq C \mathbb{E} ||Y - \mu(y)||^3 \leq C (\mathbb{E} ||Y||^3 + ||\mu(y)||^3) \leq C ||Y||^3.
\]
Splitting the latter expectation into 2 parts according to $\|X\| \leq \delta$ or $\|X\| \in [\delta, y]$ (and recalling that $Y$ has the truncated cdf $F_X|\|X\|\cdot [\cdot \cdot |y]$), we can write

$$E\left(\|\Sigma^{-1/2}(y)(Y - \mu(y))\|^3 \right) \leq C \left( \delta^3 + E\left(\|X\|^3 \mathbb{1}_{\{\|X\|\in[\delta, y]\}}\right) \right)$$

$$\leq C \left( \delta^3 + \int_\delta^y \frac{t^3}{\delta} \ln L(t)n^{-\alpha-1} \, dt \right) \leq CL(y) y^{\alpha-\alpha},$$

(8.8)

where the last inequality comes from the Karamata’s integral theorem (Karamata (1933), see e.g. Proposition 1.5.8 p.26 in Bingham et al. (1987)). So, from (8.3)–(8.8) we obtain

$$\left| \mathbb{P}(S_n \in B) - G_n(B) \right| \leq \frac{C}{\sqrt{n}} \int_{y > \delta} f_{\|X_{(n)}\|}(y) y^{\alpha-\alpha} \, dy$$

$$= \frac{C}{\sqrt{n}} E\left( L(\|X_{(n)}\|) \|X_{(n)}\|^{\alpha-\alpha} \mathbb{1}_{\{\|X_{(n)}\| > \alpha/\alpha\}} \right).$$

(8.9)

For $n$ large enough such that $n^{1/\alpha} > \delta$, using (7.2) and (8.4) gives

$$E\left( L(\|X_{(n)}\|) \|X_{(n)}\|^{\alpha-\alpha} \mathbb{1}_{\{\|X_{(n)}\| > \alpha/\alpha\}} \right)$$

$$= \int_{n^{1/\alpha}}^{\infty} L(y) y^{\alpha-\alpha} n^{3\alpha} F_{\|X_{(n)}\|}(y) \, dy$$

$$\leq n \int_{n^{1/\alpha}}^{\infty} L(y) y^{\alpha-\alpha} \, dy \sim L(n^{1/\alpha})n^{3\alpha}.$$

Consequently, we have, $L$ denoting a slowly varying function at infinity that may vary from line to line,

$$E\left( L(\|X_{(n)}\|) \|X_{(n)}\|^{\alpha-\alpha} \mathbb{1}_{\{\|X_{(n)}\| > \alpha/\alpha\}} \right)$$

$$\leq E\left( L(\|X_{(n)}\|) \|X_{(n)}\|^{\alpha-\alpha} \left( \mathbb{1}_{\{\|X_{(n)}\| \in [\alpha, n^{1/\alpha}]\}} + \mathbb{1}_{\{\|X_{(n)}\| > n^{1/\alpha}\}} \right) \right)$$

(8.10)

$$\leq \sup_{y \in [\alpha, n^{1/\alpha}]^\alpha} L(y)n^{3\alpha} + L(n^{1/\alpha})n^{3\alpha} = L(n)n^{3\alpha}.$$

Combining inequalities (8.9) and (8.10) provides the result.

**Appendix 3. Proof of Lemma 4.1 and Theorem 4.1**

**Proof of Lemma 4.1** First we prove that $g$ belongs to $2\mathcal{R}_{\gamma-1,\alpha}$, with $\gamma = 1/\alpha$.

It is straightforward to check that:

$$g(t) = F_X^{-1}(e^{-\gamma}) = \bar{F}_X^{-1}(1 - e^{-\gamma}).$$

(9.1)
Notice that the function \( h(t) := 1 - e^{-\frac{1}{t}} \) belongs to \( 2RV_{-1,-1} \). Now, from the \( 2RV \) condition on \( f_X \), we obtain that \( \tilde{F}_X \in 2RV_{-\alpha,-\beta} \) with parameters \( \alpha > 0, \beta > 0 \); see e.g. Proposition 6 in Hua and Joe (2011). Then, applying Proposition 2.6 in Lv et al. (2012) and denoting by \( 2RV_\alpha(0+) \) the class of regularly varying functions at \( 0+ \), we have

\[
\tilde{F}_X \in 2RV_{-\alpha,-\beta}(0+)
\]

from which we deduce, using Proposition 2.8 in Lv et al. (2012) (for a composition of functions) and (9.1),

\[
g \in 2RV_{\frac{1}{a}-\min\left\{1, \frac{\rho}{a}\right\}}.
\]

Now let us look at the derivative \( g' \). From (9.1), we have

\[
g'(t) = \frac{e^{-\frac{1}{t}} t^{-2}}{f_X(g(t))}.
\]

Again, by Proposition 2.8 and Proposition 2.5 in Lv et al. (2012), we have

\[
\frac{1}{f_X(g(t))} \in 2RV_{\frac{1}{a}+1, -\min\left\{1, \frac{\rho}{a}\right\}} \quad \text{and} \quad e^{-\frac{1}{t}} t^{-2} \in 2RV_{-2,-1}.
\]

Hence, multiplying the \( 2RV \) functions, we obtain \( g' \in 2RV_{\frac{1}{a}-1, -\min\left\{1, \frac{\rho}{a}\right\}} \).

**Proof of Theorem 4.1** Recall that \( Y_1, \ldots, Y_{n-1} \) denote i.i.d. random vectors having the truncated distribution \( F_X(\cdot \mid y) \) defined in (2.2), while \( X_0 \) is a family of random vectors with distribution (2.4), independent of \( \{ Y_k \} \). Using conditional probabilities and applying Lemma 2.1, we have

\[
\mathbb{P}(S_n \in B) = \int_0^\infty \mathbb{P}(X_1 + \cdots + X_{n-1} + X_n \in B \mid \|X_n\| = y) \, dy
\]

\[
= \int_0^\infty \mathbb{P}(T_{n-1} + Y_n \in B) \, dy, \quad \text{where} \quad T_{n-1} := \sum_{k=1}^{n-1} Y_k.
\]

Using the definition (4.9) of the MRV-Normex cdf, then the triangle inequality, we have, for any \( B \in \mathcal{C} \),

\[
\mathbb{P}(S_n \in B) - MG_n(B) = \int_0^\infty \mathbb{P}(X_n + T_{n-1} \in B) \, dy - \int_0^\infty \mathbb{P}(y \cdot \Theta + Z_n \in B) \, dy
\]

\[
\leq \int_0^\infty \left| \mathbb{P}(X_n + T_{n-1} \in B) - \mathbb{P}(y \cdot \Theta + Z_n \in B) \right| \, dy
\]

\[
+ \int_0^\infty \mathbb{P}(X_n + T_{n-1} \in B) - \mathbb{P}(X_n + Z_n \in B) \, dy
\]

\[
+ \int_0^\infty \mathbb{P}(X_n + Z_n \in B) - \mathbb{P}(y \cdot \Theta + Z_n) \, dy.
\]
Recall that, via the Scheffé theorem (see e.g. Billingsley (1968), p. 224), if the cdf $F$ and $G$ are absolutely continuous, then

$$
\sup_{A \in B(R)} |F(A) - G(A)| = \frac{1}{2} \int_{-\infty}^{\infty} |F'(x) - G'(x)| \, dx.
$$

Therefore, for the first integral in the right-hand side of the inequality (9.2), using Proposition 4.1 and this recall, there exists a slowly-varying function at infinity $L$ such that, for all $n \geq 1$,

$$
\int_{0}^{\infty} \left| f_{\|X_{n}\|}(y) - f_{H_{n}}(y) \right| \, dy \leq L(n)n^{-\min\{1, \frac{\varphi}{a} \}}. \quad (9.3)
$$

For the second integral appearing in (9.2), following the same approach as for the proof of Theorem 3.1, we can prove that there exists a slowly varying function at infinity, $L$, such that, for all $n \geq 1$,

$$
\int_{0}^{\infty} f_{\|X_{n}\|}(y) \left| \mathbb{P}\left( \overset{\circ}{X}_{y} + T_{n-1} \in \mathbb{B} \right) - \mathbb{P}\left( \overset{\circ}{X}_{y} + Z_{y} \in \mathbb{B} \right) \right| \, dy \leq L(n)n^{-\frac{1}{2} + \frac{1-a}{a}}. \quad (9.4)
$$

Indeed, choosing the constant $\delta > 0$ as in the proof of Theorem 3.1 and splitting the integration domain of the integral in (9.4) into two disjoint sets $\{y : y < \delta\}$ and $\{y : y \geq \delta\}$, we can write

$$
\int_{0}^{\delta} f_{\|X_{n}\|}(y) \left| \mathbb{P}\left( T_{n-1} + \overset{\circ}{X}_{y} \in \mathbb{B} \right) - \mathbb{P}\left( \overset{\circ}{X}_{y} + Z_{y} \in \mathbb{B} \right) \right| \, dy \leq \int_{0}^{\delta} f_{\|X_{n}\|}(y) \, dy = (F_{\|X\|}(\delta))^{n}. \quad (9.5)
$$

Let $C$ denote a positive constant and $L$ a slowly varying function at infinity, which may vary from line to line. Since $Y_{1}, \ldots, Y_{n-1}$ are i.i.d bounded random vectors ($Y$ denoting their parent random vector) with non degenerate distribution (Condition (C2)), we can use once again the non-generalized Berry-Esseen inequality, and obtain

$$
\int_{\delta}^{\infty} f_{\|X_{n}\|}(y) \left| \mathbb{P}\left( T_{n-1} + \overset{\circ}{X}_{y} \in \mathbb{B} \right) - \mathbb{P}\left( \overset{\circ}{X}_{y} + Z_{y} \in \mathbb{B} \right) \right| \, dy
\leq \int_{\delta}^{\infty} f_{\|X_{n}\|}(y) \sup_{\mathbb{B} \in \mathcal{C}} \left| \mathbb{P}(T_{n-1} \in \mathbb{B}) - \mathbb{P}(Z_{y} \in \mathbb{B}) \right| \, dy \quad (9.6)
\leq \frac{C}{\sqrt{n}} \int_{\delta}^{\infty} f_{\|X_{n}\|}(y) \mathbb{E}\left( \left\| (\Sigma(y))^{-1/2}(Y - \mu(y)) \right\|_{2}^{3} \right) \, dy.
$$
Replicating the arguments in the proof of Theorem 3.1, we obtain (8.8), i.e.

\[ \mathbb{E}\left( \left\| \Sigma^{-1/2}(y)(Y - \mu(y)) \right\|_2^3 \right) \leq C L(y) y^{3-a}. \]

We deduce that

\[ \mathbb{E}\left( (\|X\|) L(y) y^{3-a} \right) \leq C n \frac{y^{3-a}}{\sqrt{n}} \mathbb{E}\left( \left\| X_n \right\| \left\| X_n \right\|^{3-a} \mathbb{I}\{ \|X\| \geq \delta \} \right). \]

(9.7)

We are back to the upper bound given in (8.9). Combining it with (8.10) provides the claimed result (9.4).

The last integral in (9.2) is taken care of, in the following technical lemma, proved below.

**Lemma 9.1** There exists a slowly varying function at infinity, \( L \), such that, for all \( n \geq 1 \), for any \( B \in \mathcal{C} \),

\[ \int_{y \geq \delta} f_{\|X_n\|}(y)L(y)y^{3-a} dy = \frac{C}{\sqrt{n}} \mathbb{E}\left( L(\|X_n\|)\|X_n\|^{3-a} \mathbb{I}\{ \|X\| \geq \delta \} \right). \]

Combining inequalities (9.2), (9.3), (9.4) and Lemma 9.1 provides the statement of Theorem 4.1.

Let us turn to the proof of Lemma 9.1.

**Proof of Lemma 9.1** Define the sequence \((a_n)_{n \geq 1}\) with

\[ a_n := F_{\|X\|}^{-1}\left(n^{-\ln n} \right) \rightarrow_{n \to \infty} \infty. \]

We have

\[ \mathbb{P}\left( \|X_n\| \leq a_n \right) = (F_{\|X\|}(a_n))^n = n^{-\ln n}. \]

(9.8)

Note that \( a_n = F_{\|X\|}^{-1}\left(1 - n^{-\ln n} \right) \). Since the function \( u(n) := 1 - n^{-\ln n} \in \mathcal{R}_{V_1} \) (at infinity) and \( \|X\| \in \mathcal{R}_{V_1} \),\( \|X\| \in \mathcal{R}_{V_1-\alpha} \), so \( F_{\|X\|}^{-1} \in \mathcal{R}_{V_{1-\alpha}} \) at \( 0^+ \), therefore

\[ a_n = F_{\|X\|}^{-1} u(n) \in \mathcal{R}_{V_1/\alpha} \quad \text{(at infinity)} \]

(9.9)

(for a reminder on properties of \( \mathcal{R}_{V_1} \), see e.g. Lv et al. (2012)).

As \( F_{\|X\|}(a_n) = n \left(1 - n^{-\ln n} \right) \), there exists also a slowly varying function at infinity, \( L \), such that

\[ \mathbb{P}\left( \|X_n\| \leq a_n \right) = n^{-\ln n} \]

\[ \mathbb{P}\left( \|X_n\| \leq a_n \right) = n^{-\ln n} \]
$n\bar{F}_{\|X\|}(a_n) = L(n)$. \hfill (9.10)

Splitting the integral of Lemma 9.1 into two sets $\{y \leq a_n\}$ and $\{y > a_n\}$ and using the pdf for the maximum (7.2), we can write

$$\int_0^\infty f_{\|X\|}(y)\left|\mathbb{P}\left(\overset{\circ}{X}_y + Z_y \in B\right) - \mathbb{P}\left(y \cdot \Theta + Z_y \in B\right)\right| dy \leq \mathbb{P}\left(\|X\| \leq a_n\right)$$

$$+ \int_{a_n}^\infty nf_{\|X\|}(y)F_{\|X\|}^{n-1}(y)\left|\mathbb{P}\left(\overset{\circ}{X}_y + Z_y \in B\right) - \mathbb{P}\left(y \cdot \Theta + Z_y \in B\right)\right| dy$$

$$\leq n^{-ln} + \int_{a_n}^\infty nf_{\|X\|}(y)\left|\mathbb{P}\left(\overset{\circ}{X}_y + Z_y \in B\right) - \mathbb{P}\left(y \cdot \Theta + Z_y \in B\right)\right| dy, \text{ using (C8).}$$ \hfill (9.11)

To estimate the integral given in (9.11), we use a coupling inequality (see e.g. den Hollander (2012)), namely: For any random vector $(\xi, \eta)$ and for any measurable set $B$, we have

$$|\mathbb{P}(\xi \in B) - \mathbb{P}(\eta \in B)| \leq \mathbb{P}(\xi \neq \eta).$$

So, for any joint distribution of $\overset{\circ}{X}_y$ and $\Theta$ (or, more generally, for any joint distribution of $X$ and $\Theta$), we have

$$\left|\mathbb{P}\left(\overset{\circ}{X}_y + Z_y \in B\right) - \mathbb{P}\left(y \cdot \Theta + Z_y \in B\right)\right| \leq \mathbb{P}\left(\overset{\circ}{X}_y \neq y \cdot \Theta\right),$$

which gives the following upper bound for the integral in (9.11), using (2.4):

$$\int_{a_n}^\infty nf_{\|X\|}(y)\left\{\overset{\circ}{X}_y \neq \Theta \right\} dy = n\mathbb{P}\left(\|X\| > a_n, \|X\| \neq \Theta\right).$$

Therefore, we have

$$(C11) \leq n^{-ln} + n\mathbb{P}\left(\|X\| > a_n\right)\mathbb{P}\left(\frac{X}{\|X\|} \neq \Theta \bigg| \|X\| > a_n\right). \hfill (9.12)$$

Now, at given $n$, we define the joint distribution of $X$ and $\Theta$ such that:

(i) $\Theta$ is independent of the event $\{\|X\| > a_n\}$,

(ii) $(X / \|X\|, \Theta)$ has a joint distribution defined by Dobrushin’s theorem (see Dobrushin (1970)), which can be applied since we are in $\mathbb{R}^d$.

Hence we can write, using Condition $(M_\Theta)$ to get the asymptotic behavior,
\[
\mathbb{P}\left( \frac{X}{\|X\|} \neq \Theta \middle| \|X\| > a_n \right) = \sup_{B \subseteq \mathbb{S}_1} \mathbb{P}\left( \frac{X}{\|X\|} \in B \middle| \|X\| > a_n \right) - \mathbb{P}(\Theta \in B)
\]

\[
\sim_{n \to \infty} A(a_n),
\]

where \(A(a_n) \in \mathcal{R}_{1-\rho/\alpha} \) by combining \((M_\Theta)\) and (9.9).

Reporting this last result in (9.12) and using (9.10), we obtain

\[
(C11) \leq n^{-\ln n} + n\tilde{F}_{\|X\|}(a_n) A(a_n) = n^{-\ln n} + L(n) A(a_n),
\]

from which the result of Lemma 9.1 follows.

**Acknowledgements** Evgeny Prokopenko acknowledges the support received from the National Research Agency of the French government through the program “Investment for the future” (ANR-16-IDEX-0008 CY Initiative) during his postdoctoral fellowship at ESSEC CREAR.

**Data availability** The datasets generated during the current study are available from the corresponding author on reasonable request.

**Declarations**

**Conflict of interest** The authors declare that they have no conflict of interest.

**References**

Barbe, P., Fougères, A., Genest, C.: On the tail behavior of sums of dependent risks. ASTIN Bulletin 36(2), 361–373 (2006). https://doi.org/10.1017/S0515036100014550

Basrak, B., Davis, R., Mikosch, T.: A characterization of multivariate regular variation. Annals of Applied Probability 12(3), 908–920 (2002). https://doi.org/10.1214/aoap/1031863174

Beirlant, J., Goegebeur, Y., Segers, J., Teugels, J.L.: Statistics of Extremes: Theory and Applications. Wiley Series in Probability and Statistics. John Wiley and Sons, Ltd, New Jersey (2004). https://doi.org/10.1002/04700470012382

Bhattacharya, R.N., Rao, R.R.: Normal Approximation and Asymptotic Expansions (Classics in Applied Mathematics). Classics in applied mathematics, p. 316. SIAM, Philadelphia (2010). https://doi.org/10.1137/9780898719895

Billingsley, P.: Convergence of Probability Measures. John Wiley and Sons, New Jersey (1968)

Bingham, N.H., Goldie, C.M., Teugels, J.L.: Regular Variation. Encyclopedia of Mathematics and its Applications, vol. 27, p. 494. Cambridge University Press, Cambridge (1987). https://doi.org/10.1017/CBO9780511721434

Borovkov, A.A.: Asymptotic Analysis of Random Walks: Light-Tailed Distributions. Encyclopedia of Mathematics and its Applications, Cambridge University Press, Cambridge (2020). https://doi.org/10.1017/9781139871303

Chaudhuri, P.: On a geometric notion of quantiles for multivariate data. J. Am. Stat. Assoc. 91(434), 862–872 (1996). https://doi.org/10.1080/01621459.1996.10476954

Cline, D.B.H., Resnick, S.I.: Multivariate subexponential distributions. Stochastic Processes and their Applications 42(1), 49–72 (1992). https://doi.org/10.1016/0304-4149(92)90026-M

Csörgö, S., Haeseler, E., Mason, D.M.: A probabilistic approach to the asymptotic distribution of sums of independent, identically distributed random variables. Adv. Appl. Math. 9(3), 259–333 (1988). https://doi.org/10.1016/0196-8858(88)90016-4

Cuberos, A., Masiello, E., Maume-Deschamps, V.: High level quantile estimations of sums of risks. Dependence Modeling 3, 141–158 (2015). https://doi.org/10.1515/demo-2015-0010
Multi-normex distributions for the sum of random vectors…

Dacorogna, M., Elbahtouri, L., Kratz, M.: Model validation for aggregated risks. Annals of Actuarial Science 12(2), 433–454 (2018). https://doi.org/10.1017/S1748499517000227

Das, B., Kratz, M.: Risk concentration under second order regular variation. Extremes 23, 381–410 (2020). https://doi.org/10.1007/s10687-020-00382-3

David, H.A., Nagaraja, H.N.: Order Statistics. Wiley Series in Probability and Statistics. Wiley, New Jersey (2004). https://doi.org/10.1002/0471667196.es0023

Dobrushin, R.L.: Prescribing a system of random variables by conditional distributions. Theory of Probability and Its Applications 15(3), 458–486 (1970). https://doi.org/10.1137/1115049

Embrechts, P., Klüppelberg, C., Mikosch, T.: Modelling Extreme Events for Insurance and Finance. Stochastic Modelling and Applied Probability, p. 648. Springer, Heidelberg (1997). https://doi.org/10.1007/978-3-642-33483-2

Falk, M., Marohn, F.: Von mises conditions revisited. Ann. Probab. 21(3), 1310–1328 (1993). https://doi.org/10.1214/aop/1176989120

Foss, S., Korshunov, D., Zachary, S.: An Introduction to Heavy-Tailed and Subexponential Distributions. Springer Series in Operations Research and Financial Engineering, p. 157. Springer, New York (2013). https://doi.org/10.1007/978-1-4614-7101-1

Girard, S., Stupfler, G.: Extreme geometric quantiles in a multivariate regular variation framework. Extremes 18(4), 629–663 (2015). https://doi.org/10.1007/s10687-015-0226-0

Girard, S., Stupfler, G.: Intriguing properties of extreme geometric quantiles. Revstat - Statistical Journal 15(1), 107–139 (2017)

de Haan, L., Ferreira, A.: Extreme Value Theory: An Introduction. Springer Series in Operations Research and Financial Engineering, p. 418. Springer, New York (2006). https://doi.org/10.1007/0-387-34471-3

de Haan, L., Resnick, S.I.: Second-order regular variation and rates of convergence in extreme-value theory. Ann. Probab. 24(1), 97–124 (1996). https://doi.org/10.1214/aop/1042644709

Hahn, M.G.: Sums, Trimmed Sums and Extremes. Progress in Probability, vol. 23. Birkhäuser Boston, Boston (1991). https://doi.org/10.1007/978-1-4684-6793-2

den Hollander, F.: Probability Theory: The Coupling Method, p. 664. Springer, New York (2006). https://doi.org/10.1017/978-0-511-75518-6.003

David, H.A., Nagaraja, H.N.: Order Statistics. Wiley Series in Probability and Statistics. Wiley, New Jersey (2004). https://doi.org/10.1002/0471667196.es0023

Falk, M., Marohn, F.: Von mises conditions revisited. Ann. Probab. 21(3), 1310–1328 (1993). https://doi.org/10.1214/aop/1176989120

Foss, S., Korshunov, D., Zachary, S.: An Introduction to Heavy-Tailed and Subexponential Distributions. Springer Series in Operations Research and Financial Engineering, p. 157. Springer, New York (2013). https://doi.org/10.1007/978-1-4614-7101-1

Girard, S., Stupfler, G.: Extreme geometric quantiles in a multivariate regular variation framework. Extremes 18(4), 629–663 (2015). https://doi.org/10.1007/s10687-015-0226-0

Girard, S., Stupfler, G.: Intriguing properties of extreme geometric quantiles. Revstat - Statistical Journal 15(1), 107–139 (2017)

de Haan, L., Ferreira, A.: Extreme Value Theory: An Introduction. Springer Series in Operations Research and Financial Engineering, p. 418. Springer, New York (2006). https://doi.org/10.1007/0-387-34471-3

de Haan, L., Resnick, S.I.: Second-order regular variation and rates of convergence in extreme-value theory. Ann. Probab. 24(1), 97–124 (1996). https://doi.org/10.1214/aop/1042644709

Hahn, M.G.: Sums, Trimmed Sums and Extremes. Progress in Probability, vol. 23. Birkhäuser Boston, Boston (1991). https://doi.org/10.1007/978-1-4684-6793-2

den Hollander, F.: Probability Theory: The Coupling Method, p. 74. Leiden University, Lectures Notes-Mathematical, Leiden (2012)

Hua, L., Joe, H.: Second order regular variation and conditional tail expectation of multiple risks. Insurance Math. Econom. 49(3), 537–546 (2011). https://doi.org/10.1016/j.ismteco.2011.08.013

Karamata, J.: Sur un mode de croissance régulière, théorèmes fondamentaux. Bulletin de la Société Mathématique de France 61, 55–62 (1933). 10.24033/bsmf.1196

Kratz, M.: Normex, a new method for evaluating the distribution of aggregated heavy tailed risks. applica-

Omey, E.: Subexponential distribution functions in ℝ
d

Nocedal, J., Wright, S.: Numerical Optimization. Springer Series in Operations Research and Financial Engineering, p. 664. Springer, New York (2006). https://doi.org/10.1017/978-0-387-40065-5

Omel, E.: Subexponential distribution functions in ℝ
d

Petrov, V.V.: Limit Theorem of Probability Theory: Sequences of Independent Random Variables. Oxford Studies in Probability, vol. 4. Oxford Sciences Publications, Oxford (1995)
Petrov, V.V.: Sums of Independent Random Variables. Ergebnisse der Mathematik und ihrer Grenzgebiete, p. 348. Springer, Heidelberg (1975). https://doi.org/10.1007/978-3-642-65809-9

Resnick, S.I.: Heavy Tail Phenomena: Probabilistic and Statistical Modeling. Springer Series in Operations Research and Financial Engineering, p. 404. Springer, New York (2007). https://doi.org/10.1007/978-0-387-45024-7

Samorodnitsky, G., Sun, J.: Multivariate subexponential distributions and their applications. Extremes 19, 171–196 (2016). https://doi.org/10.1007/s10687-016-0242-8

Samorodnitsky, G., Taqqu, M.S.: Stable non-Gaussian Random Processes: Stochastic Models with Infinite Variance. Chapman and Hall, New York (1994). https://doi.org/10.1201/9780203738818

Zaliapin, I.V., Kagan, Y.Y., Schoenberg, F.P.: Approximating the distribution of pareto sums. Pure Appl. Geophys. 162, 1187–1228 (2005). https://doi.org/10.1007/s00024-004-2666-3

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.