Dependence uncertainty for aggregate risk: examples and simple bounds

Paul Embrechts and Edgars Jakobsons

Abstract Over the recent years, numerous results have been derived in order to assess the properties of regulatory risk measures (in particular VaR and ES) under dependence uncertainty. In this paper we complement this mainly methodological research by providing several numerical examples for both homogeneous as well as inhomogeneous portfolios. In particular, we investigate under which circumstances the so-called worst-case VaR can be well approximated by the worst-case (i.e. comonotonic) ES. We also study best-case values and simple lower bounds.
Keywords: Value-at-Risk, Expected Shortfall, risk aggregation, dependence uncertainty, inhomogeneous portfolio, Rearrangement Algorithm.
MSC2010 Subject Classification: 91B30, 65C50.

1 Introduction

Recent regulatory discussions in the realm of banking and insurance have brought the following quantitative aspects very much to the forefront:

i) The choice of risk measure for the calculation of regulatory capital; examples include Value-at-Risk (VaR) and Expected Shortfall (ES).
ii) The properties of statistical estimators of such risk measures, especially when based on extreme tail observations.

Paul Embrechts
Senior SFI Professor,
RiskLab, Department of Mathematics, ETH Zurich,
8092 Zurich, Switzerland
e-mail: embrechts@math.ethz.ch

Edgars Jakobsons
RiskLab, Department of Mathematics, ETH Zurich,
8092 Zurich, Switzerland
e-mail: edgars.jakobsons@math.ethz.ch

iii) The issue of Model Uncertainty (MU). The latter can be interpreted in many ways: statistical, numerical, functional, ...

In this paper we will mainly concentrate on a combination of i) and iii), and discuss the (functional) model uncertainty of risk measures VaR and ES for linear portfolios in both the homogeneous as well as inhomogeneous case. The interpretation of MU concerns portfolios with known marginal distributions but unknown interdependence. The latter is often referred to as Dependence Uncertainty (DU). It plays a crucial role in various examples throughout the banking and insurance literature. Going forward, several relevant examples will be given. One general purpose reference is Embrechts et al. (2014).

The considered setting is a static one-period model, however, the results are also applicable to the dynamic case, using marginal distributions conditioned on the available information, e.g. conditional mean and volatility. While conditioning on the volatility may lead to light-tailed marginal distributions, it is not always the case, as shown in McNeil and Frey (2000), where methods from extreme value theory are applied to GARCH-filtered (and still heavy-tailed) one-dimensional time series in order to obtain point estimates of VaR and ES. Moreover, Christoffersen and Diebold (2000) argue that the volatility of financial time series is forecastable only up to 10 days ahead (relevant for market risk) but not for longer time scales (such as 1 year, relevant for insurance). The effect of tail-heaviness of the marginal distributions on the portfolio risk will therefore be analyzed. Furthermore, the conditional dependence between the margins may change over time, as illustrated by Dias and Embrechts (2009), where a parametric copula approach is used to model and detect changes in the dependence structure. This demonstrates some of the difficulties in estimating the dependence. The mainly numerical results presented in our paper aim at understanding better the estimation of the best and worst possible risk measure values under DU. The paper only offers a first step on the way to this goal, and numerous alternative approaches as well as results are available or can be obtained. We very much hope that this paper incites other researchers to look at these and related problems of Quantitative Risk Management.

2 Homogeneous portfolios

We consider random variables $X_i \sim F_i$, $i = 1, \dots, d$ and $S_d = X_1 + \cdots + X_d$. For the purpose of this paper we assume the marginal distribution functions (dfs) F_i to be known. If $F_i = F$, $i = 1, \dots, d$, we refer to the *homogeneous* case, otherwise we refer to the *inhomogeneous* case. For most of the paper we will concentrate on the homogeneous case, though for many of the results information on the inhomogeneous case can be obtained; see Section 3 for a brief discussion. The joint distribution of $(X_1, \dots, X_d)$ is not specified, and thus the aggregate value-at-risk $\mathrm{VaR}_\alpha(S_d)$ and expected shortfall $\mathrm{ES}_\alpha(S_d)$ are not uniquely determined. In the following we consider the range of possible values these aggregate risk measures can take, under fixed marginal distributions, but unspecified joint model. This framework in risk management is referred to as *dependence uncertainty*; see Embrechts et al. (2015). In this section we focus on the homogeneous case and assume throughout that the support of F is bounded below. Define the generalized inverse

$$F^{-1}(p) = \inf\{x : F(x) \geq p\}, \quad p \in (0,1],$$

and $F^{-1}(0) = \inf\{x : F(x) > 0\}$ (the left endpoint of the support); for properties of generalized inverses, see Embrechts and Hofert (2013). Since ES and VaR are translation equivariant, for simplicity also assume $F^{-1}(0) = 0$ (via translation). To give the main results in the literature on DU bounds on risk measures in the following sections, a definition from Wang and Wang (2011) will be useful.

Definition 1 (Wang and Wang (2011)). For $d \geq 1$, a distribution function F is called *d-completely mixable* (d-CM) if there exist rvs $X_1, \dots, X_d \sim F$ and a constant k such that $X_1 + \cdots + X_d = dk$ a.s.

Examples of distributions that are completely mixable for $d \geq 2$ include uniform, Gaussian, Cauchy and other unimodal symmetric continuous distributions; for higher values of d also distributions with monotone or concave densities on a bounded support; see Puccetti et al. (2012) for further examples.

Some notation from Bernard et al. (2014a) will also be needed. With respect to a df G (to be specified), we introduce functions $H(c)$ and $D(c)$, $c \in [0,1]$ and a constant c_d:

$$\begin{aligned} H(c) &= (d-1)G^{-1}((d-1)c/d) + G^{-1}(1-c/d), \\ D(c) &= \frac{1}{1-c}\int_c^1 H(t)\mathrm{d}t = \frac{d}{1-c}\int_{(d-1)c/d}^{1-c/d} G^{-1}(t)\mathrm{d}t, \\ c_d &= \inf\{c \in [0,1] : H(c) \leq D(c)\}, \end{aligned} \tag{1}$$

where for $c = 1$ we set $D(1) := D(1-) = H(1) = dG^{-1}(1-1/d)$. Also, introduce two conditions

(A) H is non-increasing on $[0, c_d]$,
(B) The conditional distribution of G on $[(d-1)c_d/d, 1-c_d/d]$ is d-completely mixable.

These conditions will imply validity and sharpness of some of the bounds stated in the following sections. A special case in which these conditions are satisfied is given in the following lemma.

Lemma 1 (Bernard et al. (2014a)). *If the df G admits a non-increasing density on its support, then conditions (A) and (B) hold.*

Furthermore, Bernard et al. (2014a) motivate numerically that these conditions are satisfied also in some other cases, using examples with Lognormal and Gamma dfs.

2.1 Upper bound on VaR

For a random variable $X \sim F_X$ representing a loss, VaR at confidence level $\alpha \in (0,1)$ is defined as the α-quantile,

$$\mathrm{VaR}_\alpha(X) = F_X^{-1}(\alpha).$$

In turn, the upper bound on $\mathrm{VaR}_\alpha(S_d)$ over all joint models is defined as

$$\overline{\mathrm{VaR}}_\alpha(S_d) := \sup\{\mathrm{VaR}_\alpha(X_1 + \cdots + X_d) : X_i \sim F_i, i = 1,\ldots,d\}.$$

In the homogeneous case $F_i = F$, $i = 1,\ldots,d$, the upper bound $\overline{\mathrm{VaR}}_\alpha(S_d)$ can be obtained by solving the integral inequality for c_d in (1) with respect to the conditional distribution in the tail (defined below).

Proposition 1 (Embrechts and Puccetti (2006); Puccetti and Rüschendorf (2013); Wang et al. (2013)). *Define H, D and c_d in* (1) *with respect to F_α, the conditional distribution of F on $[F^{-1}(\alpha),\infty)$. If F_α admits a non-increasing density on its support, then*

$$\overline{\mathrm{VaR}}_\alpha(S_d) = D(c_d).$$

Remark 1. A non-increasing density above the α-quantile is a natural assumption for high values of α, which holds for essentially all distributions used in practice. The proof of Proposition 1 uses Lemma 1 to verify conditions (A) and (B), which are sufficient for the bound in Proposition 1 to be sharp.

Before giving the next proposition, we recall the definition of expected shortfall,

$$\mathrm{ES}_\alpha(X) = \frac{1}{1-\alpha}\int_\alpha^1 \mathrm{VaR}_q(X)\mathrm{d}q.$$

Let $\mathrm{VaR}_\alpha^+(S_d)$ and $\mathrm{ES}_\alpha^+(S_d)$ denote $\mathrm{VaR}_\alpha(S_d)$ and $\mathrm{ES}_\alpha(S_d)$ respectively, when the X_i are comonotonic. It is well-known that the *comonotonic* dependence structure gives the worst-case expected shortfall; see Dhaene et al. (2002) and p. 251 in McNeil et al. (2005).

$$\overline{\mathrm{ES}}_\alpha(S_d) := \sup\{\mathrm{ES}_\alpha(X_1+\cdots+X_d) : X_i \sim F_i, i = 1,\ldots,d\} = \sum_{i=1}^{d} \mathrm{ES}_\alpha(X_i) =: \mathrm{ES}_\alpha^+(S_d).$$

Moreover, as d increases, $\overline{\mathrm{VaR}}_\alpha(S_d)$ is asymptotically equivalent to $\mathrm{ES}_\alpha^+(S_d)$. A first result in this direction is:

Proposition 2 (Puccetti and Rüschendorf (2014)). *Suppose $(X_d)_{d\geq 1}$ is a sequence of rvs identically distributed as F, where F is integrable and has a decreasing density on $[F^{-1}(\alpha),\infty)$. Then*

$$\lim_{d\to\infty} \frac{\overline{\mathrm{VaR}}_\alpha(X_1+\cdots+X_d)}{\mathrm{ES}_\alpha^+(X_1+\cdots+X_d)} = 1.$$

Table 1 Thresholds for the number of margins d for which $\overline{\mathrm{VaR}}_\alpha(S_d)$ is within 10% of $\mathrm{ES}^+_\alpha(S_d)$. The levels $\alpha \in \{0.95, 0.99, 0.999\}$ are listed for comparison. The parameter σ for $\mathrm{LogN}(0,\sigma^2)$ is chosen to match the ratio $\mathrm{VaR}_\alpha(X_1)/\mathrm{ES}_\alpha(X_1)$ with that of Pareto(a), i.e. $P[X_1 > x] = (x+1)^{-a}$, $x \geq 0$.

Pareto(a),	$a =$	10.0	5.0	3.0	2.5	2.0	1.7	1.6	1.5
$\overline{\mathrm{VaR}}_{0.95}$		3	3	4	4	6	12	17	29
$\mathrm{LogN}(0,\sigma^2)$,	$\sigma =$	0.75	0.90	1.13	1.26	1.49	1.71	1.81	1.94
$\overline{\mathrm{VaR}}_{0.95}$		3	3	3	4	5	6	7	9
Pareto(a),	$a =$	10.0	5.0	3.0	2.5	2.0	1.7	1.6	1.5
$\overline{\mathrm{VaR}}_{0.99}$		2	3	3	4	6	11	16	27
$\mathrm{LogN}(0,\sigma^2)$,	$\sigma =$	0.70	0.89	1.20	1.37	1.66	1.95	2.08	2.23
$\overline{\mathrm{VaR}}_{0.99}$		2	3	3	4	5	6	7	9
Pareto(a),	$a =$	10.0	5.0	3.0	2.5	2.0	1.7	1.6	1.5
$\overline{\mathrm{VaR}}_{0.999}$		2	3	3	4	6	11	15	27
$\mathrm{LogN}(0,\sigma^2)$,	$\sigma =$	0.67	0.92	1.33	1.56	1.93	2.28	2.44	2.63
$\overline{\mathrm{VaR}}_{0.999}$		2	2	3	4	5	7	8	11

Remark 2. Proposition 2 holds under much more general conditions, also in the heterogeneous case; see Embrechts et al. (2015). Results of this type are relevant for regulatory practice for both the banking world (the so-called Basel framework) as well as insurance (Solvency 2); see Embrechts et al. (2015).

In Table 1 the smallest number of margins d is given for which $\overline{\mathrm{VaR}}_\alpha(S_d)$ is within 10% of $\mathrm{ES}^+_\alpha(S_d)$, i.e. after which the asymptotically equivalent sequence gives a reasonable approximation. Notice that in all but the most heavy-tailed cases, $d \approx 10$ was sufficient. Hence, if the number of margins is greater than that, we may use the easily calculated $\mathrm{ES}^+_\alpha(S_d)$ as a reasonable estimate of the conservative $\overline{\mathrm{VaR}}_\alpha(S_d)$.

The convergence rate in Proposition 2 is also known.

Proposition 3 (Embrechts et al. (2015)). *If* $\mathbb{E}[|X_1 - \mathbb{E}[X_1]|^k]$ *is finite for some* $k > 1$ *and* $\mathrm{ES}_\alpha(X_1) > 0$*, then, as* $d \to \infty$*,*

$$\frac{\overline{\mathrm{VaR}}_\alpha(S_d)}{\mathrm{ES}^+_\alpha(S_d)} = 1 - O(d^{-1+1/k}).$$

So in particular, if all moments are finite, then the convergence rate is $O(d^{-1})$, and if the distribution has a regularly varying tail (see Section 7.3 in McNeil et al. (2005)) with index $-\rho$, for example $X_1 \sim \mathrm{Pareto}(\rho)$, then the convergence is slower, since $k < \rho < \infty$. In Figure 1 the differences $1 - \overline{\mathrm{VaR}}_\alpha(S_d)/\mathrm{ES}^+_\alpha(S_d)$ as d increases are plotted on a logarithmic scale for different Pareto and Lognormal distributions. We observe that, although initially the rate of convergence is slower, for large d it seems faster than the theoretical one. Moreover, for small d the rates are not very different between Pareto and Lognormal. Tail-heaviness is often measured using the notion of regular variation, and also Proposition 3 determines the convergence rate according to the highest existing moment. However, regular variation concerns quantiles asymptotically approaching 1, and Proposition 3 holds for d tending to infinity. In

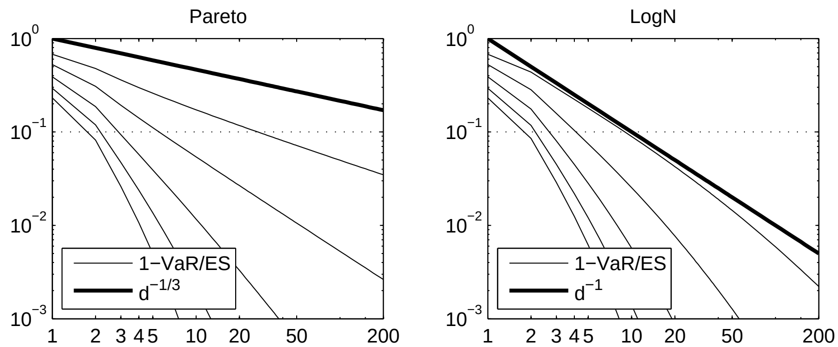

Fig. 1 Relative differences $1-\overline{\mathrm{VaR}}_\alpha(S_d)/\mathrm{ES}^+_\alpha(S_d)$ for $\alpha=0.99$ on the vertical axis versus d on the horizontal, on a log-log scale. Below the dotted line the relative difference is smaller than 10%. The left panel contains Pareto(a) distributions, $a=1.5,2,3,5,10$ from top down. The bold line shows the theoretical convergence rate $O(d^{-1/3})$ for Pareto(1.5) according to Proposition 3. In the right panel the Lognormal $\mathrm{LogN}(0,\sigma^2)$ case is plotted, $\sigma=2.23,1.66,1.20,0.89,0.70$ from top down, chosen to match the ratio $\mathrm{VaR}_\alpha(X_1)/\mathrm{ES}_\alpha(X_1)$ with that of Pareto. The bold line shows the theoretical convergence rate $O(d^{-1})$ for distributions with all moments finite.

order to analyze $\overline{\mathrm{VaR}}_\alpha(S_d)$ at a fixed level, say, $\alpha=0.99$, and for small d, we need a measure of tail-heaviness that captures the behavior for moderate α and d. We refrain from using skewness, kurtosis and higher moments, as they may not exist for power tails. Instead, we will define a different measure of tail-heaviness in the following section.

2.1.1 The normalized mean-median ratio

For $X\sim F$ define the *normalized mean-median ratio* as

$$\mathrm{M}_\alpha(F):=\frac{\mathrm{ES}_\alpha(X)-\mathrm{VaR}_\alpha(X)}{\mathrm{MS}_\alpha(X)-\mathrm{VaR}_\alpha(X)},$$

where MS is the *median shortfall*, $\mathrm{MS}_{1-p}(X):=\mathrm{VaR}_{1-p/2}(X)$; see Kou et al. (2013). $\mathrm{M}_\alpha(F)$ is a measure of tail-heaviness defined for regularly varying, subexponential, as well as exponential tails, and only requires the first moment to exist. $\mathrm{M}_\alpha(F)$ is invariant under scaling and translation, since the risk measures used in the definition are positively homogeneous and translation equivariant. In Figure 2 M_α is plotted for Gamma, Lognormal and Pareto distributions, as a function of the shape parameter. For all three families M_α diverges to infinity as the tail becomes heavier, and converges to a finite value for light tails.

For the Gamma distribution $\Gamma(k,1)$ the values are computed numerically; to reach values in the typical range for Lognormal and Pareto, a very small parameter k is required (e.g. $k=0.0027$ for $\mathrm{M}_\alpha(\Gamma)=3$). For $k=1$ we recover the exponential

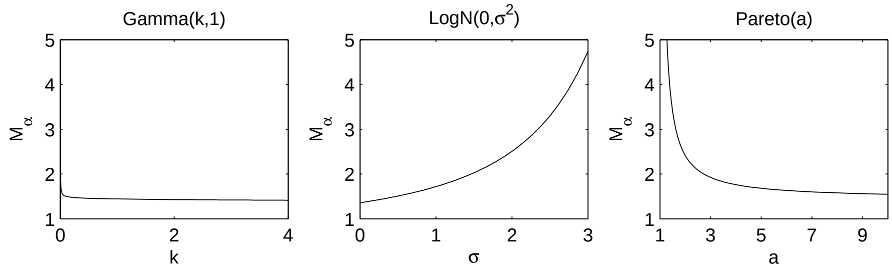

Fig. 2 $\mathrm{M}_{0.99}$ for $\Gamma(k,1)$, $\mathrm{LogN}(0,\sigma^2)$ and Pareto(a) distribution, depending on the shape parameters k,σ,a, respectively.

distribution $\mathscr{E}(1)$, for which

$$\mathrm{M}_\alpha(\mathscr{E}) = \frac{-\log(1-\alpha)+1+\log(1-\alpha)}{-\log((1-\alpha)/2)+\log(1-\alpha)} = \frac{1}{\log(2)} \approx 1.4427, \tag{2}$$

independent of α (and scale). And as $k\to\infty$, $\Gamma(k,1)$ converges (under scaling and translation) to the standard Normal distribution Φ, so

$$\lim_{k\to\infty} \mathrm{M}_\alpha(\Gamma(k,1)) = \frac{\varphi(\Phi^{-1}(\alpha))/(1-\alpha)-\Phi^{-1}(\alpha)}{\Phi^{-1}(1-(1-\alpha)/2)-\Phi^{-1}(\alpha)} \quad (\approx 1.3583 \text{ for } \alpha=0.99), \tag{3}$$

using the analytic formula for ES of the Normal distribution; see p. 45 in McNeil et al. (2005).

For the Lognormal distribution $\mathrm{LogN}(\mu,\sigma^2)$,

$$\mathrm{M}_\alpha(\mathrm{LogN}(\mu,\sigma^2)) = \frac{e^{\sigma^2/2}(1-\Phi(\Phi^{-1}(\alpha)-\sigma))/(1-\alpha)-e^{\sigma\Phi^{-1}(\alpha)}}{e^{\sigma\Phi^{-1}(1-(1-\alpha)/2)}-e^{\sigma\Phi^{-1}(\alpha)}},$$

independent of the scale e^μ. Since for $Z\sim\Phi$ we have $e^{\sigma Z}\sim\mathrm{LogN}(0,\sigma^2)$, as $\sigma\to 0$, using the Taylor series of e^x about $x=0$ we recover the same limit 1.3583 as in (3).

For the Pareto distribution $\mathscr{P}(a)$, $a>1$,

$$\mathrm{M}_\alpha(\mathscr{P}(a)) = \frac{\frac{a}{a-1}(1-\alpha)^{-1/a}-(1-\alpha)^{-1/a}}{((1-\alpha)/2)^{-1/a}-(1-\alpha)^{-1/a}} = \frac{1}{(a-1)(2^{-1/a}-1)},$$

independent of the level α. As $a\to\infty$, we obtain $1/\log(2)$ as the limit, the same value as in (2). This is due the fact that if $X\sim\mathscr{E}(1)$, then $e^{X/a}-1\sim\mathscr{P}(a)$, and we can again use the Taylor series of e^x about $x=0$.

We shall use M_α as a measure of tail-heaviness in order to compare the tails of different families of distributions, and from that deduce how close $\overline{\mathrm{VaR}}_\alpha(S_d)$ is to $\mathrm{ES}^+_\alpha(S_d)$ for small values of d. Notice that $\mathrm{M}_\alpha(F)$ exactly determines the worst-case VaR when $d=2$ and the density is decreasing beyond $F^{-1}(\alpha)=\mathrm{VaR}_\alpha(X_1)$. This is because for $d=2$ the worst-case dependence structure for VaR_α is *counter-*

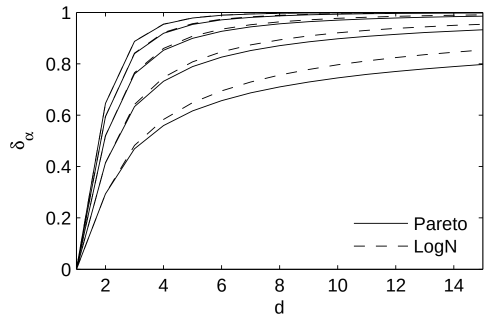

Fig. 3 $\delta_\alpha(S_d)$ plotted in dependence on d. Pareto(a) with $a = 10, 5, 3, 2, 1.5$ from top down, with tail-heaviness $\mathrm{M}_\alpha = 1.55, 1.68, 1.92, 2.41, 3.40$, respectively, and LogN(μ, σ^2) with the parameters chosen to match the M_α, $\alpha = 0.99$.

monotonicity above the α-quantile of X_1 and X_2, viz. $X_2 = F^{-1}(\alpha + (1 - F(X_1)))$ when $F(X_1) \geq \alpha$; see Makarov (1982) and Embrechts et al. (2005). Furthermore, when the density is decreasing, the minimal sum $X_1 + X_2$ in the tail is attained when $X_1 = X_2 = \mathrm{MS}_\alpha(X_1)$. Hence $\overline{\mathrm{VaR}}(S_2) = 2\,\mathrm{MS}_\alpha(X_1)$. See also (4) and Figure 3 in the next section.

2.1.2 The normalized VaR bound

Since VaR is a positively homogeneous and translation equivariant risk measure, it is easy to recalculate VaR under these operations. The upper bound on VaR depends only on the conditional distribution F on the interval $[F^{-1}(\alpha), \infty)$. Moreover, it is easy to see that

$$\overline{\mathrm{VaR}}_\alpha(S_d) \in [\mathrm{VaR}^+_\alpha(S_d), \mathrm{ES}^+_\alpha(S_d)].$$

So, in order to focus on the issue of specifying *where* in this interval $\overline{\mathrm{VaR}}_\alpha$ lies, we define the *normalized VaR bound*

$$\delta_\alpha(S_d) := \frac{\overline{\mathrm{VaR}}_\alpha(S_d) - \mathrm{VaR}^+_\alpha(S_d)}{\mathrm{ES}^+_\alpha(S_d) - \mathrm{VaR}^+_\alpha(S_d)} \in [0, 1].$$

Note that $\delta_\alpha(S_d)$ does not depend on the scale and location of F. In Figure 3 the values of $\delta_\alpha(S_d)$ are plotted in dependence on d. The parameters for Pareto and Lognormal distributions are chosen so that the M_α values match, thus, as explained in the previous section,

$$\delta_\alpha(S_2) = \frac{2\,\mathrm{MS}_\alpha(X_1) - 2\,\mathrm{VaR}_\alpha(X_1)}{2\,\mathrm{ES}_\alpha(X_1) - 2\,\mathrm{VaR}_\alpha(X_1)} = 1/\mathrm{M}_\alpha(F), \tag{4}$$

and the upper VaR bounds agree for $d = 2$. For $d > 2$ the bounds for the matched Lognormal and Pareto dfs converge at a similar rate, especially for lighter tails, when convergence to 1 is rather fast. The difference $1 - \delta_\alpha$ seems to decrease as a power of d, so in Figure 4 we plot this difference versus the dimension d on a logarithmic scale. We observe that, apart from the "kink" at $d = 2$, the dependence is approximately linear on the log-log scale. Based on this numerical evidence, and

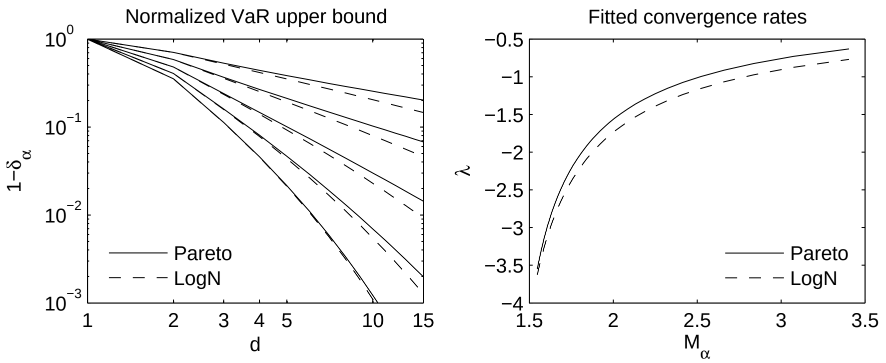

Fig. 4 Left panel: normalized VaR bound δ_α, $\alpha = 0.99$, depending on the dimension d. Solid lines correspond to Pareto(a) distributions with $a = 1.5, 2, 3, 5, 10$ and $\mathrm{M}_\alpha = 3.40, 2.41, 1.92, 1.68, 1.55$ from top down. Dashed lines correspond to Lognormal distributions with parameters selected to match the M_α values. Right panel: the fitted values of $\lambda(\mathrm{M}_\alpha)$ obtained by linear regression for different tail-weight levels M_α.

fixing the value at $d = 2$ according to (4), we consider an approximate model

$$1 - \delta_\alpha \approx (1 - 1/\mathrm{M}_\alpha) \left(\frac{d}{2}\right)^{\lambda(\mathrm{M}_\alpha)}. \tag{5}$$

Taking logs from both sides we obtain a linear dependence, and using linear regression for $d \in \{2 \ldots 15\}$ we fit the slope parameter λ (the intercept at $d = 2$ is fixed) for each considered M_α. The obtained values of λ are plotted on the right panel of Figure 4. They show that the convergence rate is indeed determined by the tail-heaviness M_α, and the values of $\lambda(M_\alpha)$ for the Pareto and Lognormal families are similar. Note also that e.g. for Pareto(1.5) the convergence rate according to Proposition 3 is $O(d^{-1/3})$ as $d \to \infty$, whereas from Figure 4 we read that for small d the difference $1 - \delta_\alpha$ decreases approximately as $d^{-2/3}$ for the corresponding $M_\alpha = 3.4$. Finally, the fitted curves for δ_α using λ values from the linear regression are plotted in Figure 5. The close match shows that the model (5) works well for small dimensions.

2.2 *Lower bound on VaR*

The lower bound on $\mathrm{VaR}_\alpha(S_d)$ over all joint models is defined as

$$\underline{\mathrm{VaR}}_\alpha(S_d) := \inf\{\mathrm{VaR}_\alpha(X_1 + \cdots + X_d) : X_i \sim F_i, i = 1, \ldots, d\}.$$

Before providing the basic result for the lower bound on VaR in the homogeneous case, we define the *left-tail expected shortfall*

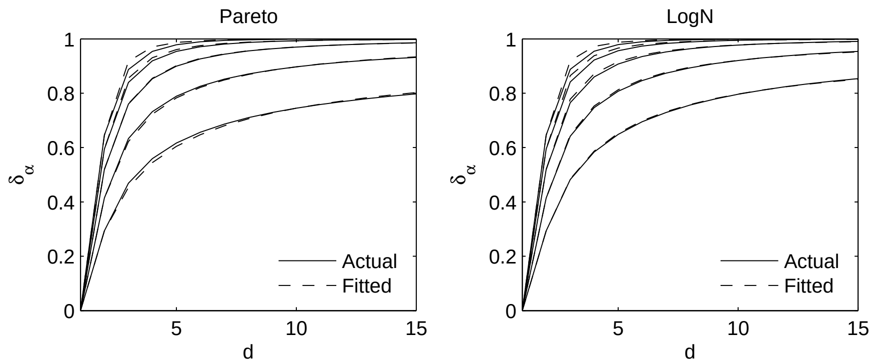

Fig. 5 Normalized VaR bound δ_α, $\alpha = 0.99$ (solid lines) compared to the fitted values (dashed) according to the model (5), depending on the dimension d. Left panel: Pareto(a) distributions with $a = 1.5, 2, 3, 5, 10$ and $\mathrm{M}_\alpha = 3.40, 2.41, 1.92, 1.68, 1.55$ from bottom up. Right panel: Lognormal distributions with parameters selected to match the M_α values.

$$\mathrm{LES}_\alpha(X) = \frac{1}{\alpha}\int_0^\alpha \mathrm{VaR}_p(X)\,\mathrm{d}p,$$

and the comonotonic $\mathrm{LES}^+_\alpha(S_d) = \sum_{i=1}^d \mathrm{LES}_\alpha(X_i) = d\,\mathrm{LES}_\alpha(X_1)$ (since $X_i \sim F$, $\forall i$).

Proposition 4 (Bernard et al. (2014a)). *If the support of F is bounded below, then*

$$\underline{\mathrm{VaR}}_\alpha(X_1+\cdots+X_d) \geq \max\{(d-1)F^{-1}(0)+F^{-1}(\alpha)\ ,\ d\,\mathrm{LES}_\alpha(X_1)\}. \tag{6}$$

Furthermore, define H and D as in (1) *with respect to F^α, the conditional distribution of F on $[F^{-1}(0), F^{-1}(\alpha)]$. If conditions (A) and (B) with respect to F^α are satisfied, then* (6) *holds with equality, i.e.*

$$\underline{\mathrm{VaR}}_\alpha(S_d) = \max\{H(0), D(0)\}.$$

Since we assume that the essential support of F is the positive half-axis, (6) reduces to $\underline{\mathrm{VaR}}_\alpha(X_1+\cdots+X_d) \geq \max\{F^{-1}(\alpha), d\,\mathrm{LES}_\alpha(X_1)\}$. In contrast to Proposition 1, in Proposition 4 we include the conditions (A) and (B) explicitly. Recall that previously these were implied by the non-increasing density assumption. While it is reasonable to assume this for the upper tail, in the lower tail this assumption holds for some distributions (e.g. Pareto) but fails for other (e.g. Lognormal). In such cases, one may resort to testing (A) numerically and testing for complete mixability in (B) using the method in Puccetti and Wang (2014). Note, however, that if $F^{-1}(\alpha) > d\,\mathrm{LES}_\alpha(X_1)$ then the bound can be sharp also without the d-CM condition. Typically, however, letting $d^* = F^{-1}(\alpha)/\mathrm{LES}_\alpha(X_1)$, we may use as a good approximation

$$\begin{aligned} &\underline{\mathrm{VaR}}_\alpha(S_d) \approx F^{-1}(\alpha) && \text{for } d < d^*,\\ &\underline{\mathrm{VaR}}_\alpha(S_d) \approx d\,\mathrm{LES}_\alpha(X_1) && \text{for } d \geq d^*. \end{aligned} \tag{7}$$

Table 2 Thresholds for the number of margins d for which $\underline{\mathrm{VaR}}_\alpha(S_d)/\mathrm{LES}^+_\alpha(S_d)$ has converged to 1. Parameters for the distributions are chosen to match the ratio $d^* = \mathrm{VaR}_\alpha(X_1)/\mathrm{LES}_\alpha(X_1)$.

$\alpha = 0.99$		d^*	2.36	3.42	5.59	6.55	8.19	11.00	14.91
	Pareto(a),	$a =$	-	-	10	5	3	2	1.5
		d	-	-	6	7	9	11	15
	LogN$(0,\sigma^2)$,	$\sigma =$	0.40	0.59	0.88	0.98	1.13	1.35	1.60
		d	3	4	6	7	9	11	15
	Gamma$(k,1)$,	$k =$	5.00	2.00	0.77	0.58	0.40	0.25	0.16
		d	3	4	6	7	9	11	15

An alternative numerical approach for computing an approximation of the sharp bounds on VaR under DU is given by the Rearrangement Algorithm (RA); see Embrechts et al. (2013) or Puccetti and Rüschendorf (2012, 2013) for earlier formulations.[1] In Table 2 the values for Pareto, Lognormal and Gamma distributed variables are listed for which the ratio $\underline{\mathrm{VaR}}_\alpha(S_d)/\mathrm{LES}^+_\alpha(S_d)$ has reached the limit 1. Due to a non-increasing density, in the Pareto case (6) is always sharp and the limit is reached at $d = \lceil d^* \rceil$. As a side note, for the Pareto(a) distribution at level $p = \alpha$, as $a \to \infty$,

$$d^* = \frac{(1-p)^{-1/a} - 1}{(1-(1-p)^{1-1/a})/((1-1/a)/p) - 1} \longrightarrow \frac{-\log(1-p)}{1 + ((1-p)/p)\log(1-p)},$$

so for $p = 0.99$ we have $d^* > 4.8298$, i.e. no parameter a gives a lower value of d^*, hence they are not listed in Table 2. For the Lognormal and Gamma dfs, the values were obtained by the RA, using as discretization parameter $N = 10^5$ and the stopping condition $\varepsilon = 10^{-4}$. In Figure 6 the normalized lower VaR bounds are plotted. The approximate bounds from (7) are the same for the three families. These are sharp for the Pareto, but differ from the sharp bounds (here, computed using the RA) for the Lognormal and Gamma families in the lighter-tailed cases, since the densities are not decreasing and the conditions in Proposition 4 are not satisfied. In particular, for small d it is the condition (A) that matters, and is not satisfied e.g. for the LogN$(0, 0.59^2)$ df with $d = 3$; see Figure 7. With parameter $k \leq 1$, however, $\Gamma(k,1)$ has a decreasing density, so by Lemma 1 satisfies the conditions and the approximate bound (6) is sharp. Overall, (6) gives an easily calculated lower bound on aggregate VaR, which is close to the sharp bound.

2.3 Lower bound on ES

The lower bound on $\mathrm{ES}_\alpha(S_d)$ under dependence uncertainty is defined as

[1] A website set up by Giovanni Puccetti with the title “The Rearrangement Algorithm project” provides full details and recent developments on the RA; see https://sites.google.com/site/rearrangementalgorithm/.

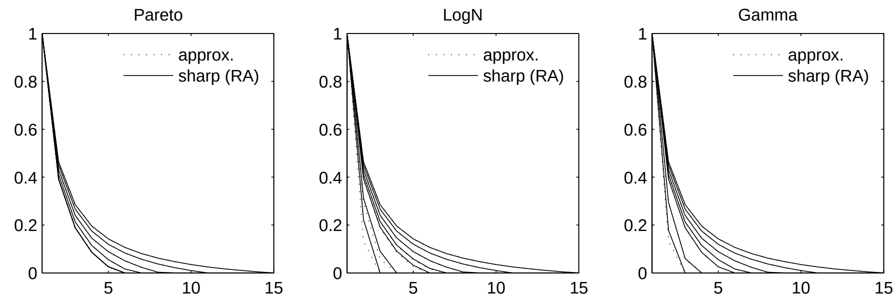

Fig. 6 $(\underline{\mathrm{VaR}}_\alpha(S_d) - \mathrm{LES}^+_\alpha(S_d))/(\mathrm{VaR}^+_\alpha(S_d) - \mathrm{LES}^+_\alpha(S_d))$ plotted in dependence on d for Pareto(a), LogN$(0,\sigma^2)$ and $\Gamma(k,1)$ distributions with the parameters chosen to match the d^* (see Table 2). For Pareto, the approximate bounds (7) are sharp, whereas for LogN and Gamma there are small deviations for the lighter-tailed cases in comparison to the sharp bounds given by the RA.

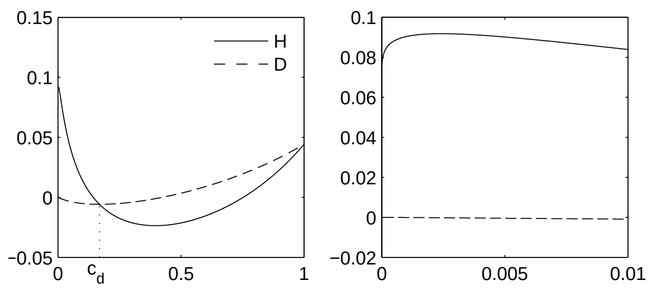

Fig. 7 Functions H and D with respect to F^α, the conditional distribution of $F = \mathrm{LogN}(0, 0.59^2)$ on $[0, F^{-1}(\alpha)]$, $d = 3$, $\alpha = 0.99$. The vertical scale is normalized to show the interval $[\mathrm{LES}^+_\alpha(S_d), \mathrm{VaR}^+_\alpha(S_d)]$ as $[0,1]$. Left panel: full range $c \in [0,1]$. Right panel: close-up of the range $[0, 0.01]$ shows that the function H is increasing initially. $H(0) = 0.058$, while the sharp lower bound (from the RA) is 0.091.

$$\underline{\mathrm{ES}}_\alpha(S_d) := \inf\{\mathrm{ES}_\alpha(X_1 + \cdots + X_d) : X_i \sim F_i, i = 1, \ldots, d\}.$$

In the literature, few results on $\underline{\mathrm{ES}}_\alpha$ are available. The two main references are Puccetti (2013) (using the RA) and Bernard et al. (2014a). The latter provides the following result in the homogeneous case.

Proposition 5 (Bernard et al. (2014a)). *Define H, D and c_d as in* (1) *with respect to F. If condition (A) is satisfied, then*

$$\underline{\mathrm{ES}}_\alpha(S_d) \geq \begin{cases} \frac{1}{1-\alpha}\left[\mathbb{E}[S_d] - \alpha D(c_d)\right] & \text{if} \quad \alpha \leq 1 - c_d, \\ \frac{1}{1-\alpha}\int_0^{1-\alpha} H(t)\mathrm{d}t & \text{if} \quad \alpha > 1 - c_d. \end{cases} \tag{8}$$

Furthermore, if condition (B) is satisfied, then (8) *holds with equality.*

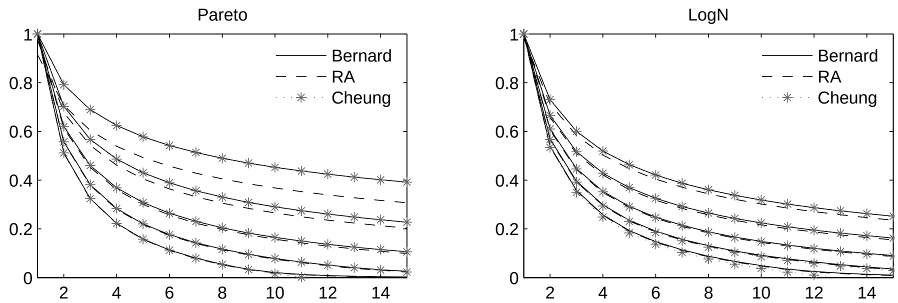

Fig. 8 $(\underline{\mathrm{ES}}_\alpha(S_d)-\mathbb{E}[S_d])/(\mathrm{ES}^+_\alpha(S_d)-\mathbb{E}[S_d])$ plotted in dependence on d. Solid lines are the bound from Proposition 8 (Bernard et al. (2014a)), dashed lines correspond to the RA (Puccetti (2013)) and dotted lines to Proposition 6 (Cheung and Lo (2013)). Left panel: Pareto(a) distributions with $a=1.5,2,3,5,10$ from top down. Right panel: LogN(μ,σ^2) distributions with the parameters chosen to match $\mathbb{E}[X_1]$ and $\mathrm{ES}_\alpha(X_1)$ with those of Pareto.

Remark 3. Due to Lemma 1, Proposition 5 is most useful in the case of a decreasing density, e.g. for Pareto or Exponential df. In other cases the conditions (A) and (B) would need to be verified numerically. Alternatively, one may apply the RA. The latter, however, uses a discretization of the marginal distributions, which may underestimate the integral over the infinite tail of the aggregate risk.

A simple lower bound on the aggregate ES can be calculated as follows.

Proposition 6 (Cheung and Lo (2013)). *If* $F^{-1}(0)\geq 0$*, then*

$$\underline{\mathrm{ES}}_{1-p}(S_d)\geq \mathrm{ES}_{1-p/d}(X_1). \tag{9}$$

Furthermore, if $F(0)\geq (d-1)/d$*, then* (9) *holds with equality.*

For this bound to be sharp, a sufficient mass at 0 (at least $p(d-1)/d$) is required for the marginal distributions, in which case Proposition 5 gives the same bound. In Figure 8 all three bounds are plotted for different Pareto and Lognormal dfs. First, notice that the convergence to the limit ($\mathbb{E}[S_d]$ in this case) is much slower than for the VaR bounds. For the Pareto df, bounds from Proposition 5 are sharp, while the error in the RA bound due to discretization is clearly visible in the heavy-tailed cases (discretization parameter $N=10^5$ and stopping condition $\varepsilon=10^{-4}$ were used). The non-sharp bound from Proposition 6, however, is very close, because the Pareto distribution has a large mass near 0. For the Lognormal dfs, the density is small near 0, hence there is a visible approximation error for the lighter tails. The bound from the RA again underestimates the tail-expectation in the heavy-tailed cases due to discretization. Overall, Proposition 5 seems to give the best bound (due to infinite tails, condition (A) is likely to hold).

3 Inhomogeneous portfolios

In this section suppose $X_i \sim F_i$, $i = 1, \ldots, d$, where the F_i's are not necessarily the same, and hence $S_d = X_1 + \cdots + X_d$ is the aggregate loss of a possibly inhomogeneous portfolio. In order to investigate the key determinants of the DU bounds for VaR and ES of an inhomogeneous portfolio, we sample at random different families and parameters for the marginal distributions. In particular, we choose F_i independently for $i = 1, \ldots, d$ from each of the following families with probability $1/3$:

- Gamma, $\Gamma(k, \mu/k)$ with $k \sim 1 + \text{Poisson}(1)$, $\mu \sim \mathscr{E}(1/5)$,
- Lognormal, $\text{LogN}(\mu, \sigma^2)$ with $\mu \sim \mathscr{N}(0.2, 0.4^2)$, $\sigma \sim \Gamma(8, 0.1)$,
- Generalized Pareto, $\text{GPD}(k, \nu, 0)$ with $k \sim 1/(1.5 + \mathscr{E}(1/5))$, $\nu \sim \Gamma(4, 0.3)$.

For the latter Pareto-type df from Extreme Value Theory, see Embrechts et al. (1997). In this manner we sample the marginal distributions of 100 portfolios for each $d = 2, \ldots, 15$; in total 1400 portfolios. Then, using only basic properties of these randomly selected marginal distributions and their respective risk quantities, such as $\text{VaR}_\alpha(X_i)$, $\text{MS}_\alpha(X_i)$ and $\text{ES}_\alpha(X_i)$, we will approximate the dependence uncertainty bounds for the aggregated portfolio risk measures.

3.1 Upper bound on VaR

In this section we focus on approximating the upper bound for $\text{VaR}_\alpha(S_d)$. As the two key drivers, similarly to the homogeneous case, we look at the dimension and the mean-median ratio, properly adapted. Since the marginal distributions can be different, the scale and shape of each margin affects the aggregate VaR. In Figure 9, the left panel, we observe that the nominal portfolio dimension d is a poor predictor of $\delta_\alpha(S_d)$. The reason is that the different scales of the marginals have the numerical effect of a dimension reduction for the underlying portfolio; for example, if one marginal has a much larger scale then the others, then the VaR of this marginal gives a good estimate of the aggregate VaR. In this case the dependence uncertainty is small and δ_α is close to 0 (similar to $d = 1$). If the marginals are of a similar scale, then the dependence uncertainty is greater (similar to the homogeneous case d) and δ_α is closer to 1.

Let $\text{VaR}^+(S_d)$, $\text{MS}^+(S_d)$ and $\text{ES}^+(S_d)$ denote $\text{VaR}(S_d)$, $\text{MS}(S_d)$ and $\text{ES}(S_d)$ respectively, when the X_i are comonotonic. Define the *effective dimension* as

$$\widetilde{d}(S_d) := \frac{\text{ES}^+_\alpha(S_d) - \text{VaR}^+_\alpha(S_d)}{\max_{i=1,\ldots,d}\{\text{ES}_\alpha(X_i) - \text{VaR}_\alpha(X_i)\}}.$$

On the right panel of Figure 9 we observe that the effective dimension $\widetilde{d}$ determines the normalized VaR-bound δ_α better. In order to take into account also the tail-heaviness of the marginal distributions, define the *average mean-median ratio* as

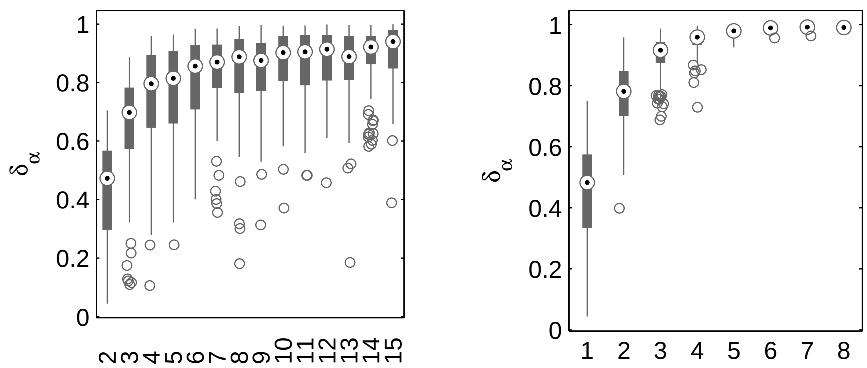

Fig. 9 Boxplots of $\delta_{0.99}(S_d)$ for portfolios with randomly sampled marginal distributions in dependence on the nominal d on the left panel, and in dependence on the effective $\widetilde{d}$ (rounded) on the right panel.

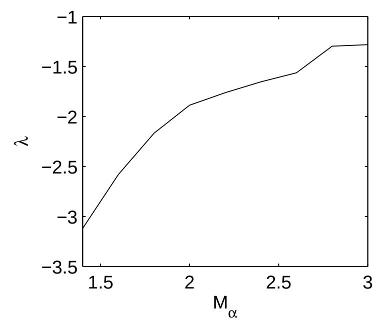

Fig. 10 The fitted values of $\lambda(\widetilde{\mathrm{M}}_\alpha)$ obtained by linear regression for different tail-weight levels $\widetilde{\mathrm{M}}_\alpha$.

$$\widetilde{\mathrm{M}}_\alpha(S_d) := \frac{\mathrm{ES}^+_\alpha(S_d) - \mathrm{VaR}^+_\alpha(S_d)}{\mathrm{MS}^+_\alpha(S_d) - \mathrm{VaR}^+_\alpha(S_d)}.$$

We speculate that, similar to the homogeneous case, the rate of convergence depends on the tail-heaviness $\widetilde{\mathrm{M}}_\alpha(S_d)$, and consider an approximate model

$$\delta_\alpha \approx 1 - \widetilde{d}^{\,\lambda(\widetilde{\mathrm{M}}_\alpha)}.$$

Rearranging and taking log, we obtain a linear dependence $\log(1-\delta_\alpha) \approx \lambda \log(\widetilde{d})$ for which we estimate λ using linear regression. The estimated parameters are plotted in Figure 10. As expected, for light tails the convergence is fast, $O(\widetilde{d}^{\,-3})$, and slower for heavy tails. The fitted lines for δ_α for three levels of the average tail-weight $\widetilde{\mathrm{M}}_\alpha$, as well as the true VaR bound for portfolios with similar tail-weight, are plotted in Figure 11. We observe that the two key determinants $\widetilde{d}$ and $\widetilde{\mathrm{M}}_\alpha$ provide a reasonable approximation of $\delta_\alpha(S_d)$ and hence of $\overline{\mathrm{VaR}}_\alpha(S_d)$. Of course, these

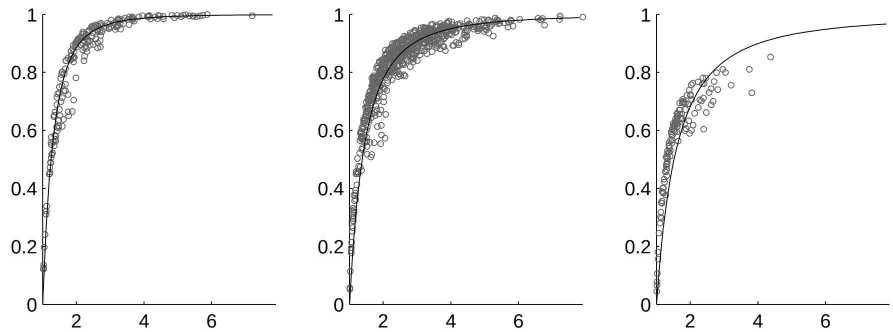

Fig. 11 Fitted curves $\delta_{.99} = 1 - \widetilde{d}^{\lambda}$ for tail-heaviness levels $\widetilde{\mathrm{M}}_{\alpha} = 1.4, 1.8, 2.4$, superimposed on scatterplots of the true δ_{99} of those sampled portfolios with tail-heaviness within 10% of these levels.

conclusions are to be understood in the specific context of the example. An out-of-sample test would be required to determine whether this is a good fit in general.

3.2 Lower bound on VaR

Similar to the homogeneous case (6), an easily calculated but non-sharp bound is

$$\underline{\mathrm{VaR}}_{\alpha}(S_d) \geq \max\{\mathrm{VaR}_{\alpha}(X_i), i = 1, \ldots, d\} \vee \mathrm{LES}^{+}_{\alpha}(S_d). \tag{10}$$

In Figure 12 the error of this approximation is shown, relative to the length of the possible interval $\underline{\mathrm{VaR}}_{\alpha}(S_d) \in [\mathrm{LES}^{+}_{\alpha}(S_d), \mathrm{VaR}^{+}_{\alpha}(S_d)]$. We see that of the 1400 portfolios, only 4 have a relative error above 5%. The marginal densities for one of these portfolios is plotted in the right panel of Figure 12. Notice the light left tails of the Lognormal margins (especially for $\sigma < 0.5$); these do not have enough mass to compensate for the heaviest right tail, and hence the approximate bound (10) is not sharp. This is in agreement with the results in the homogeneous case; see Figure 6. Keeping this observation in mind, we can in most cases use (10) as a good approximation for the most favorable aggregate VaR.

3.3 Lower bound on ES

In the inhomogeneous case a lower bound on the ES is given in Jakobsons et al. (2015), based on a similar construction to that of (1). In this case, however, due to the lack of symmetry, one cannot combine the tails of the margins at the same rate; a

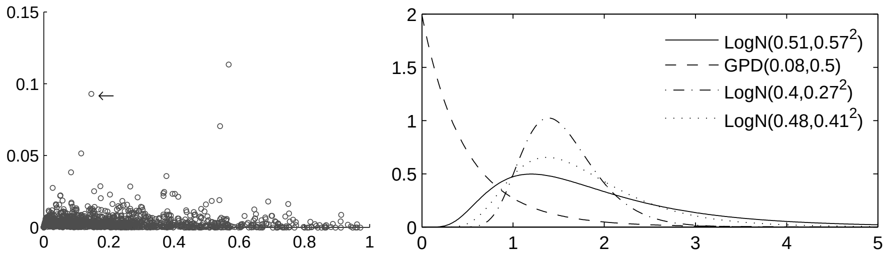

Fig. 12 Left panel: True $\underline{\mathrm{VaR}}_\alpha(S_d)$ on the horizontal axis (computed using the RA) versus the error of the approximate bound (10) on the vertical axis. For each portfolio the values in $[\mathrm{LES}^+_\alpha(S_d), \mathrm{VaR}^+_\alpha(S_d)]$ are normalized to lie within $[0,1]$. The arrow indicates an example portfolio, for which the marginal densities are plotted on the Right panel. Notice the light left tail of $\mathrm{LogN}(0.4, 0.27^2)$. Here $\underline{\mathrm{VaR}}_\alpha(S_4) = 7.18$, while $\max\{\mathrm{VaR}_\alpha(X_i), i = 1,\dots,4\} = 6.22$. The possible range is $[\mathrm{LES}^+(S_4), \mathrm{VaR}^+(S_4)] = [5.67, 16.01]$, hence the error is 9% of this range.

dynamical weighting needs to be found by solving a system of functional equations. Also, to check the validity and sharpness of the obtained bounds, conditions similar to (A) and (B) would need to be verified. Instead, based on the good performance of the approximate bounds in Section 2.3, we consider the following generalization, based on Theorem 4.1 in Cheung and Lo (2013).

Theorem 1. *Let* $\zeta = \min\{x : \sum_{i=1}^d F_i(x) \geq d - (1-\alpha)\}$. *Then*

$$\underline{\mathrm{ES}}_\alpha(S_d) \geq \frac{1}{1-\alpha} \sum_{i=1}^d \mathbb{E}[X_i \mathrm{I}_{\{X_i > \zeta\}}]. \tag{11}$$

Furthermore, if $\sum_{i=1}^d F_i(0) \geq d-1$, *then* (11) *holds with equality.*

Proof. Let $S'_d = \sum_{i=1}^d X_i \mathrm{I}_{\{X_i > \zeta\}}$. Clearly $S_d \geq S'_d$, so by monotonicity of ES,

$$\mathrm{ES}_\alpha(S_d) \geq \mathrm{ES}_\alpha(S'_d) \geq \mathrm{ES}_\alpha(S_i^M) = \frac{1}{1-\alpha} \sum_{i=1}^d \mathbb{E}[X_i \mathrm{I}_{\{X_i > \zeta\}}],$$

where $S_d^M = \sum_{i=1}^d Y_i$ for $Y_i \stackrel{\mathrm{d}}{=} X_i \mathrm{I}_{\{X_i > \zeta\}}$ such that $Y_1,\dots,Y_d$ are *mutually exclusive* and hence $S_d^M \leq_{sl} S'_d$ (stop-loss order); see Dhaene and Denuit (1999). If $\sum_{i=1}^d F_i(0) \geq d-1$, then there exist mutually exclusive $X_i \sim F_i$, $i = 1,\dots,d$ and (11) is an equality. □

Also in the inhomogeneous case the RA (Puccetti (2013)) can be applied, with the caveat that the obtained bound may underestimate the expectation in the infinite tail. We computed both bounds at the level $\alpha = 0.99$ for the sampled portfolios, the results are plotted in Figure 13. First, notice that the bound for higher-dimensional portfolios is closer to $\mathbb{E}[S_d]$. However, for a fixed d, the RA underestimates the bound (due to discretization) when $\underline{\mathrm{ES}}_\alpha(S_d)$ is closer to $\mathrm{ES}^+_\alpha(S_d)$, since this is often

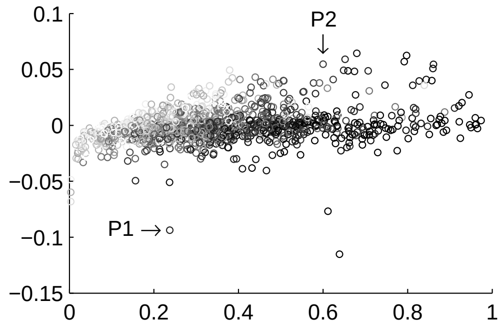

Fig. 13 On the horizontal axis: $\underline{\mathrm{ES}}_\alpha(S_d)$ of the sampled portfolios as approximated by the RA (with discretization parameter $N = 10^5$ and stopping condition $\varepsilon = 10^{-4}$). Vertical axis: the difference between the non-sharp bound (11) and the RA bound. For portfolio labeled P1 the error is negative, so (11) is away from sharpness, while for P2 the error is positive, so the RA bound underestimates the tail integral. The marginal densities for these portfolios are plotted in Figure 14. The color corresponds to the dimension from $d = 2$ (dark) to $d = 15$ (light). For each portfolio the values in $[\mathbb{E}[S_d], \mathrm{ES}^+_\alpha(S_d)]$ are normalized to lie within $[0,1]$.

due to one or more heavy-tailed margins (e.g. the GPD with parameter $k > 0.5$, which has infinite variance). On the other hand, the RA gives a sharper bound than (11) in the presence of marginals with light left tails (e.g. $\mathrm{LogN}(\mu, \sigma^2)$ with $\sigma < 0.5$). Marginal densities for example portfolios in each of these two cases are given in Figure 14. Out of the 1400 portfolios, in only 7 cases the RA was worse than (11) by more than 5% points, and in only 6 cases the RA was better by more than 5% points (relative to the possible interval $\underline{\mathrm{ES}}_\alpha(S_d) \in [\mathbb{E}[S_d], \mathrm{ES}^+_\alpha(S_d)]$). Hence, depending on the portfolio, either of the two bounds may give a value closer to the sharp lower bound on ES, but in most cases the difference is small (and both of the bounds are close to sharpness).

4 Conclusions and further research

Model uncertainty in general and dependence uncertainty in particular have taken center stage of the regulatory discourse of financial and insurance regulation. In the present paper we have mainly looked at some recent results from the latter, i.e. properties of extremal regulatory risk measures, where we assume full knowledge of the marginal risk distributions, but we do not have, or want, to make assumptions on their interdependence. Whereas numerous analytic results have recently appeared on this issue, there is an increasing need for a better understanding of the underlying numerical problems. This paper provides a first (small) step in this direction. Comparing simple approximate DU bounds for VaR and ES with more advanced ones, we observe that the former often give reasonable estimates of the sharp bounds.

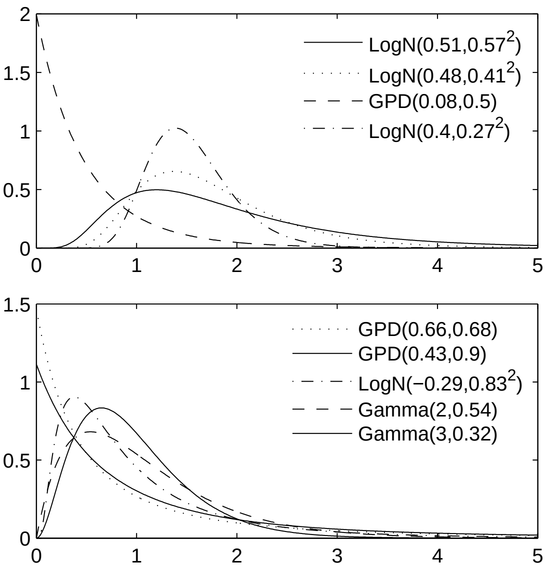

Fig. 14 Top panel: marginal densities of portfolio P1 from Figure 13. Notice the light left tails of the Lognormal margins. RA yields $\underline{\mathrm{ES}}_\alpha(S_4) \geq 8.98$ while (11) $\underline{\mathrm{ES}}_\alpha(S_4) \geq 7.73$. The possible range is $[\mathbb{E}[S_4], \mathrm{ES}^+_\alpha(S_4)] = [5.81, 19.20]$, hence the difference is -9% of this range. Bottom panel: marginal densities of portfolio P2. Notice the heavy right tails of the GPD margins. RA yields $\underline{\mathrm{ES}}_\alpha(S_5) \geq 64.63$ while (11) $\underline{\mathrm{ES}}_\alpha(S_5) \geq 69.90$. The possible range is $[6.73, 103.19]$, hence the relative difference is 5%.

Moreover, we identify the cases when the approximate bounds are far from the sharp ones. We also note that the convergence rate of $\overline{\mathrm{VaR}}$ as the portfolio dimension increases is different for small dimensions compared to the theoretical asymptotic rate. The above examples should serve as illustrations for the recent analytic results in the literature, build intuition and motivate further research in the area of dependence uncertainty.

An important direction of research is investigating what partial information on the dependence structure helps to obtain narrower bounds, which under complete dependence uncertainty tend to be very wide. An early approach considers the case when some multivariate marginals are known, and it is found that it can lead to strongly improved bounds; see Rüschendorf (1991), Embrechts and Puccetti (2010), and Embrechts et al. (2013). Another approach is taken in Bernard and Vanduffel (2014), where the full-dimensional joint density of the assets in portfolio is assumed to be known, but only on a "trusted region", which is a subset of the support. They find that the bounds quickly deteriorate as the trusted region deviates from the entire support. Bernard et al. (2014b) argue that variance of the sum of marginals (portfolio variance) is often available, and that this dependence information often leads to substantially narrower VaR bounds. Recently, dynamic factor multivariate GARCH models have been proposed for forecasting financial time series; see Alessi et al. (2009) and Santos and Moura (2014). Consequently, a new and relevant approach would be using factor models in order to reduce the DU to a smaller dimension, namely, the number of factors. Recall the observations from Section 3, which show that for small-dimensional portfolios the bounds are closer to the comonotonic case (narrower). Gordy (2003) even considers the portfolio VaR in a one-factor model, eliminating DU entirely. Of course, any additional assumptions on the structure should be justified, so it is equally important to understand what additional information on the dependence is typically available in practice. An illuminating case study in the applied setting of economic capital computation for a bank is Aas and Puccetti (2014), where the DU bounds are calculated, compared with the value corresponding to a t-copula, and other practical issues discussed.

Acknowledgements The authors would like to thank the Swiss Finance Institute for financial support. They also thank the referee and the editor for valuable comments, and Ruodu Wang and Giovanni Puccetti for numerous discussions on issues related to the broader research area of dependence uncertainty within Quantitative Risk Management.

References

Aas, K. and Puccetti, G. (2014). Bounds on total economic capital: the DNB case study. *Extremes*, 17(4):693–715.

Alessi, L., Barigozzi, M., and Capasso, M. (2009). Estimation and forecasting in large datasets with conditionally heteroskedastic dynamic common factors. ECB Working Paper No. 1115.

Bernard, C., Jiang, X., and Wang, R. (2014a). Risk aggregation with dependence uncertainty. *Insurance: Mathematics and Economics*, 54:93–108.

Bernard, C., Rüschendorf, L., and Vanduffel, S. (2014b). Value-at-Risk bounds with variance constraints. Available at SSRN 2342068.

Bernard, C. and Vanduffel, S. (2014). A new approach to assessing model risk in high dimensions. Available at SSRN 2393054.

Cheung, K. C. and Lo, A. (2013). General lower bounds on convex functionals of aggregate sums. *Insurance: Mathematics and Economics*, 53(3):884–896.

Christoffersen, P. F. and Diebold, F. X. (2000). How relevant is volatility forecasting for financial risk management? *Review of Economics and Statistics*, 82(1):12–22.

Dhaene, J. and Denuit, M. (1999). The safest dependence structure among risks. *Insurance: Mathematics and Economics*, 25(1):11–21.

Dhaene, J., Denuit, M., Goovaerts, M. J., Kaas, R., and Vyncke, D. (2002). The concept of comonotonicity in actuarial science and finance: Theory. *Insurance: Mathematics and Economics*, 31(1):3–33.

Dias, A. and Embrechts, P. (2009). Testing for structural changes in exchange rates' dependence beyond linear correlation. *The European Journal of Finance*, 15(7-8):619–637.

Embrechts, P. and Hofert, M. (2013). A note on generalized inverses. *Mathematical Methods of Operations Research*, 77(3):423–432.

Embrechts, P., Höing, A., and Puccetti, G. (2005). Worst VaR scenarios. *Insurance: Mathematics and Economics*, 37(1):115–134.

Embrechts, P., Klüppelberg, C., and Mikosch, T. (1997). *Modelling Extremal Events for Insurance and Finance*, volume 33. Springer.

Embrechts, P. and Puccetti, G. (2006). Bounds for functions of dependent risks. *Finance and Stochastics*, 10(3):341–352.

Embrechts, P. and Puccetti, G. (2010). Bounds for the sum of dependent risks having overlapping marginals. *Journal of Multivariate Analysis*, 101(1):177–190.

Embrechts, P., Puccetti, G., and Rüschendorf, L. (2013). Model uncertainty and VaR aggregation. *Journal of Banking & Finance*, 37(8):2750–2764.

Embrechts, P., Puccetti, G., Rüschendorf, L., Wang, R., and Beleraj, A. (2014). An academic response to Basel 3.5. *Risks*, 2(1):25–48.

Embrechts, P., Wang, B., and Wang, R. (2015). Aggregation-robustness and model uncertainty of regulatory risk measures. *Finance and Stochastics*, to appear.

Gordy, M. B. (2003). A risk-factor model foundation for ratings-based bank capital rules. *Journal of Financial Intermediation*, 12(3):199–232.

Jakobsons, E., Han, X., and Wang, R. (2015). General convex order on risk aggregation. *Scandinavian Actuarial Journal*, to appear.

Kou, S., Peng, X., and Heyde, C. C. (2013). External risk measures and Basel accords. *Mathematics of Operations Research*, 38(3):393–417.

Makarov, G. D. (1982). Estimates for the distribution function of a sum of two random variables when the marginal distributions are fixed. *Theory of Probability and its Applications*, 26(4):803–806.

McNeil, A. J. and Frey, R. (2000). Estimation of tail-related risk measures for heteroscedastic financial time series: an extreme value approach. *Journal of Empirical Finance*, 7(3):271–300.

McNeil, A. J., Frey, R., and Embrechts, P. (2005). *Quantitative Risk Management: Concepts, Techniques, Tools*. Princeton University Press.

Puccetti, G. (2013). Sharp bounds on the expected shortfall for a sum of dependent random variables. *Statistics & Probability Letters*, 83(4):1227–1232.

Puccetti, G. and Rüschendorf, L. (2012). Computation of sharp bounds on the distribution of a function of dependent risks. *Journal of Computational and Applied Mathematics*, 236(7):1833–1840.

Puccetti, G. and Rüschendorf, L. (2013). Sharp bounds for sums of dependent risks. *Journal of Applied Probability*, 50(1):42–53.

Puccetti, G. and Rüschendorf, L. (2014). Asymptotic equivalence of conservative value-at-risk-and expected shortfall-based capital charges. *Journal of Risk*, 16(3):3–22.

Puccetti, G., Wang, B., Wang, R., et al. (2012). Advances in complete mixability. *Journal of Applied Probability*, 49(2):430–440.

Puccetti, G. and Wang, R. (2014). Detecting complete and joint mixability. Available at SSRN 2476802.

Rüschendorf, L. (1991). Bounds for distributions with multivariate marginals. In Mosler, K. and M, S., editors, *Stochastic Orders and Decisions*, volume 19, pages 285–310. JSTOR. IMS Lecture Notes.

Santos, A. A. P. and Moura, G. V. (2014). Dynamic factor multivariate GARCH model. *Computational Statistics & Data Analysis*, 76:606–617.

Wang, B. and Wang, R. (2011). The complete mixability and convex minimization problems with monotone marginal densities. *Journal of Multivariate Analysis*, 102(10):1344–1360.

Wang, R., Peng, L., and Yang, J. (2013). Bounds for the sum of dependent risks and worst Value-at-Risk with monotone marginal densities. *Finance and Stochastics*, 17(2):395–417.

Appendix

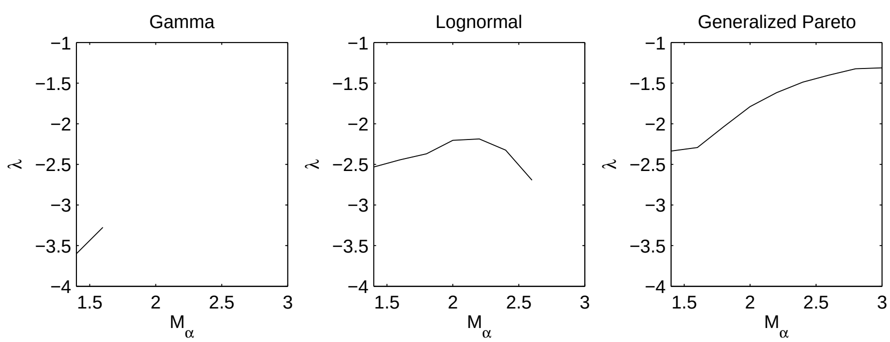

Fig. 15 The fitted values of $\lambda(\widetilde{\mathrm{M}}_\alpha)$, $\alpha = 0.99$ obtained by regression for different tail-weight levels $\widetilde{\mathrm{M}}_\alpha$, for three collections of portfolios: either only Gamma, Lognormal, or Generalized Pareto marginal distributions.

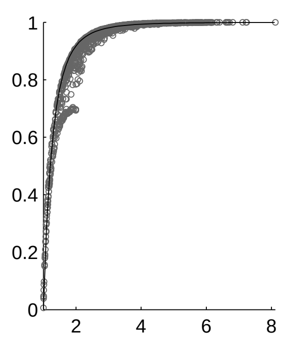

Fig. 16 Fitted curve $\delta_\alpha = 1 - \widetilde{d}^{\lambda}$, $\alpha = 0.99$ for tail-heaviness level $\widetilde{\mathrm{M}}_\alpha = 1.4$, superimposed on scatterplots of the true $\delta_\alpha(S_d)$ of the sampled portfolios. Portfolios with only Gamma marginal distributions.

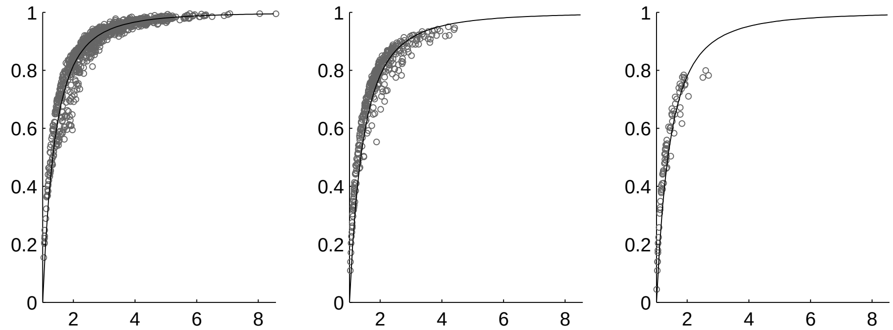

Fig. 17 Fitted curves $\delta_\alpha = 1 - \widetilde{d}^{\lambda}$, $\alpha = 0.99$ for tail-heaviness levels $\widetilde{\mathrm{M}}_\alpha = 1.6, 2.0, 2.2$, superimposed on scatterplots of the true $\delta_\alpha(S_d)$ of those sampled portfolios with tail-heaviness within 10% of these levels. Portfolios with only Lognormal marginal distributions.

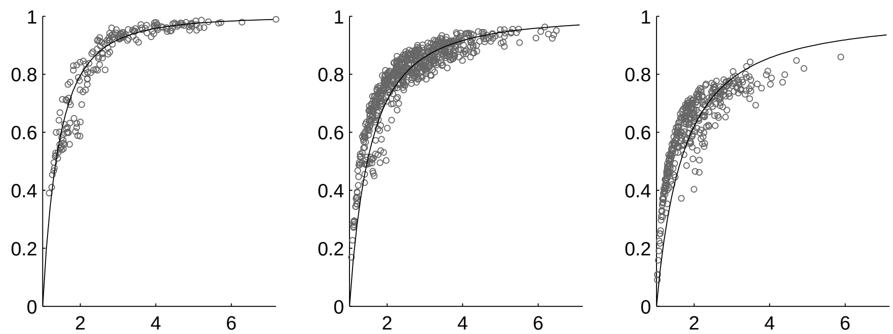

Fig. 18 Fitted curves $\delta_\alpha = 1 - \widetilde{d}^{\lambda}$, $\alpha = 0.99$ for tail-heaviness levels $\widetilde{\mathrm{M}}_\alpha = 1.6, 2.0, 2.6$, superimposed on scatterplots of the true $\delta_\alpha(S_d)$ of those sampled portfolios with tail-heaviness within 10% of these levels. Portfolios with only Generalized Pareto marginal distributions.

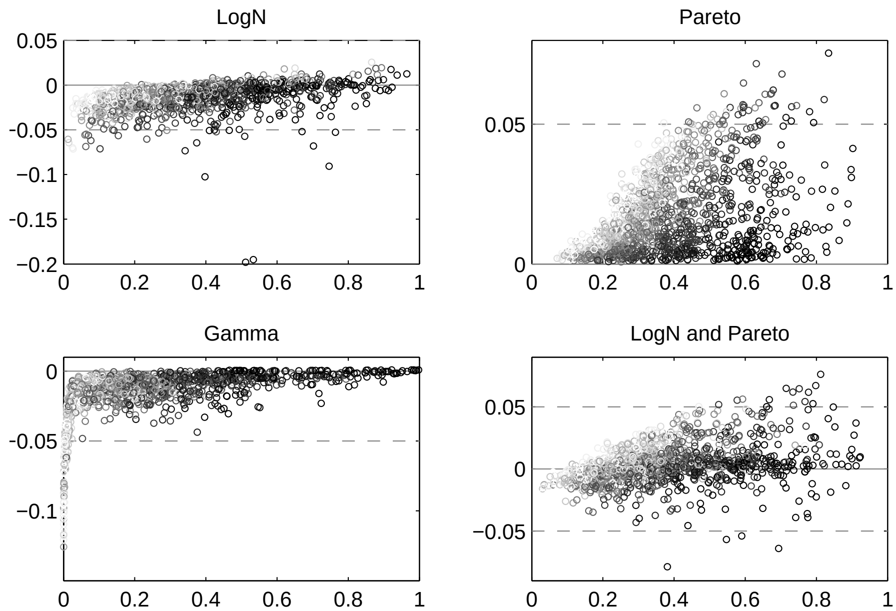

Fig. 19 On the horizontal axis: $\underline{\mathrm{ES}}_\alpha(S_d)$ of the sampled portfolios as approximated by the RA (with discretization parameter $N = 10^5$ and stopping condition $\varepsilon = 10^{-4}$). Vertical axis: the difference between the non-sharp bound (11) and the RA bound. For each portfolio the values in $[\mathbb{E}[S_d], \mathrm{ES}^+_\alpha(S_d)]$ are normalized to lie within $[0,1]$. The color corresponds to the dimension from $d = 2$ (dark) to $d = 15$ (light). The marginal dfs are sampled from the families given above each plot. For LogN and Gamma distributions, light lower tails are possible, so (11) is away from sharpness and gives a worse bound than the RA. For Pareto, (11) always gives a better bound due to heavy upper tails. For LogN and Pareto mixed, either of the two bounds may be sharper.