Bernoulli **18**(2), 2012, 455–475
DOI: 10.3150/10-BEJ343

A multivariate piecing-together approach with an application to operational loss data

STEFAN AULBACH[1,*], VERENA BAYER[2] and MICHAEL FALK[1,**]

[1]*Institute of Mathematics, Emil-Fischer-Str. 30, D-97074 Würzburg, Germany. E-mail:* [*]*stefan.aulbach@uni-wuerzburg.de;* [**]*michael.falk@uni-wuerzburg.de*
[2]*Chair of Econometrics, Sanderring 2, D-97070 Würzburg, Germany. E-mail:* *verena.bayer@uni-wuerzburg.de*

The univariate piecing-together approach (PT) fits a univariate generalized Pareto distribution (GPD) to the upper tail of a given distribution function in a continuous manner. We propose a multivariate extension. First it is shown that an arbitrary copula is in the domain of attraction of a multivariate extreme value distribution if and only if its upper tail can be approximated by the upper tail of a multivariate GPD with uniform margins.

The multivariate PT then consists of two steps: The upper tail of a given copula C is cut off and substituted by a multivariate GPD copula in a continuous manner. The result is again a copula. The other step consists of the transformation of each margin of this new copula by a given univariate distribution function.

This provides, altogether, a multivariate distribution function with prescribed margins whose copula coincides in its central part with C and in its upper tail with a GPD copula.

When applied to data, this approach also enables the evaluation of a wide range of rational scenarios for the upper tail of the underlying distribution function in the multivariate case. We apply this approach to operational loss data in order to evaluate the range of operational risk.

Keywords: copula; domain of multivariate attraction; GPD copula; multivariate extreme value distribution; multivariate generalized Pareto distribution; operational loss; peaks over threshold; piecing together

1. Introduction

The peaks over threshold approach (POT) shows that the upper tail of a univariate distribution function F can reasonably be approximated only by that of a generalized Pareto distribution (GPD). This result goes back to Balkema and de Haan [2] and Pickands [30]. A univariate GPD W is derived from an extreme value distribution (EVD) G by the equality

$$W(x) = 1 + \log(G(x)), \qquad 1/\mathrm{e} \le G(x),$$

where, with a shape parameter $\alpha > 0$, the family of standardized EVD is given by

$$\begin{aligned} G_{1,\alpha}(x) &= \exp(-x^{-\alpha}), \qquad x > 0, \\ G_{2,\alpha}(x) &= \exp(-(-x)^{\alpha}), \qquad x \le 0, \\ G_3(x) &= \exp(-\mathrm{e}^{-x}), \qquad x \in \mathbb{R}, \end{aligned} \tag{1}$$

being the Fréchet, (reverse) Weibull and Gumbel case of an EVD.

The family of univariate standardized GPD is, consequently, given by

$$\begin{aligned} W_{1,\alpha}(x) &= 1 - x^{-\alpha}, \qquad x \geq 1, \\ W_{2,\alpha}(x) &= 1 - (-x)^{\alpha}, \qquad -1 \leq x \leq 0, \\ W_3(x) &= 1 - \exp(-x), \qquad x \geq 0, \end{aligned}$$

being the Pareto, beta and exponential case of a GPD.

If X is a univariate random variable with distribution function F, then the distribution function $F^{[x_0]}$ of X, conditional on the event $X > x_0$, is given by

$$\begin{aligned} F^{[x_0]}(x) &= P(X \leq x \mid X > x_0) \\ &= \frac{F(x) - F(x_0)}{1 - F(x_0)}, \qquad x \geq x_0, \end{aligned}$$

where we require $F(x_0) < 1$. The POT approach shows that $F^{[x_0]}$ can reasonably be approximated only by a GPD with appropriate shape, location and scale parameter $W_{\gamma,\mu,\sigma}$. Note that

$$\begin{aligned} F(x) &= \big(1 - F(x_0)\big) F^{[x_0]}(x) + F(x_0) \\ &\approx \big(1 - F(x_0)\big) W_{\gamma,\mu,\sigma}(x) + F(x_0), \qquad x \geq x_0. \end{aligned}$$

The piecing together approach (PT) now consists in replacing the distribution function F by

$$F^*_{x_0}(x) = \begin{cases} F(x), & x < x_0, \\ \big(1 - F(x_0)\big) W_{\gamma,\mu,\sigma}(x) + F(x_0), & x \geq x_0, \end{cases} \tag{2}$$

where the shape, location and scale parameters γ, μ, σ of the GPD are typically estimated from given data. This modification aims at a more precise investigation of the upper end of the data.

Replacing F in (2) by the empirical distribution function $\hat{F}_n$ of n independent copies of X offers in particular a semi-parametric approach to the estimation of high quantiles $F^{-1}(q) = \inf\{t \in \mathbb{R}\colon F(t) \geq q\}$ outside the range of given data; see, for example, Section 2.3 of Reiss and Thomas [33].

In this paper we propose an extension of the PT in (2) to higher dimensions. When applied to data, this approach also enables the evaluation of a wide range of rational scenarios for the upper tail of the underlying distribution function in the multivariate case. This will be exemplified in Section 4 for operational loss data, where we simulate different scenarios for risk parameters such as the value at risk or the expected shortfall. In Section 2 we provide the basic mathematics for our PT approach. We will show that an arbitrary copula can reasonably be approximated in its upper tail only by a GPD with uniform margins.

The multivariate PT approach, which will be established in Section 3, now consists of two steps:

(i) The upper tail of a given m-dimensional copula C is cut off and substituted by the upper tail of multivariate GPD copula in a continuous manner such that the result is again a copula.

(ii) The other step consists of the transformation of each margin of this new copula by a given univariate distribution function F_i^*, $1 \le i \le m$.

This provides, altogether, a multivariate distribution function with prescribed margins F_i^*, whose copula coincides in its central part with C and in its upper tail with a GPD copula. In Section 4.2 we will simulate the effects that the combination of univariate and multivariate PT has on quantile functions and mean excess functions or, in terms of risk analysis, on value at risk and expected shortfall. It turns out that in our specific model, which will be specified in Section 4, the application of the multivariate PT approach leads to a rising expected shortfall while the value at risk keeps (up to a level of 99.9%) nearly unchanged.

Instead of fitting a GPD to the upper tail of a distribution, estimation of rare events in the multivariate case can also be based on the fact that the exponent measure pertaining to a multivariate GPD is homogeneous; see de Haan and Sinha [11] and de Haan and Ronde [9] for details.

For recent accounts of basic and advanced topics of extreme value theory and statistics, see the monographs by Reiss and Thomas [33], de Haan and Ferreira [10] and Resnick [34].

2. Multivariate GPD

In this section we provide the mathematics underlying our PT approach, which will be established in Section 3.

Let F be an arbitrary m-dimensional distribution function that is in the domain of attraction of an m-dimensional EVD G; that is, there exist norming constants $\mathbf{a}_n > \mathbf{0}$, $\mathbf{b}_n \in \mathbb{R}^m$ such that

$$F^n(\mathbf{a}_n\mathbf{x} + \mathbf{b}_n) \underset{n\to\infty}{\to} G(\mathbf{x}), \qquad \mathbf{x} \in \mathbb{R}^m, \tag{3}$$

where all operations on vectors are meant componentwise. The distribution function G is max stable; that is, there exist norming constants $\mathbf{c}_n > \mathbf{0}$, $\mathbf{d}_n \in \mathbb{R}^m$ with

$$G^n(\mathbf{c}_n\mathbf{x} + \mathbf{d}_n) = G(\mathbf{x}), \qquad \mathbf{x} \in \mathbb{R}^m.$$

The one-dimensional margins G_i of G are up to scale and location parameters univariate EVD in (1).

It is well known that (3) is equivalent with convergence of the univariate margins *together* with convergence of the *copulas*

$$\lim_{n\to\infty} C_F^n(\mathbf{u}^{1/n}) = C_G(\mathbf{u}) = G(G_1^{-1}(u_1), \dots, G_m^{-1}(u_m)), \qquad \mathbf{u} \in (0,1)^m, \tag{4}$$

(Deheuvels [12,13], Galambos [19]). For a recent account on copulas, see Nelsen [29]. Some elementary computations as in Falk [16], Section 6, or de Haan and Ronde [9], Section 4.2, entail that convergence (4) is equivalent with

$$\lim_{t\downarrow 0} \frac{1}{t}\big(1 - C_F(\mathbf{1} + t\mathbf{x})\big) = l_G(\mathbf{x}) := -\log(C_G(\exp(\mathbf{x}))), \qquad \mathbf{x} \le \mathbf{0}, \tag{5}$$

where l_G is known as the *stable tail dependence function* introduced by Huang [22]. For a detailed discussion of the stable tail dependence function, see Beirlant *et al.* [3]. The stable tail dependence function is homogeneous $tl_G(\mathbf{x}) = l_G(t\mathbf{x})$, $t \geq 0$, and, thus, (5) becomes

$$\frac{1 - C_F(\mathbf{1} + t\mathbf{x}) - l_G(t\mathbf{x})}{t} \to_{t \downarrow 0} 0.$$

Observe that $l_G(\mathbf{x}) = 1 - H(\mathbf{x})$, $\mathbf{x} \leq \mathbf{0}$, where H is a multivariate GP *function* with uniform margins $H_i(x) = 1 + x$, $x \leq 0$, $i \leq m$; that is,

$$H(\mathbf{x}) = 1 + \log(\widetilde{G}(\mathbf{x})), \qquad \mathbf{x} \leq \mathbf{0},$$

and $\widetilde{G}$ is a multivariate EVD with negative exponential margins $\widetilde{G}_i(x) = \exp(x)$, $x \leq 0$, $i \leq m$.

We call in general an m-dimensional *distribution function* W a multivariate GPD if its upper tail coincides with a GP function; that is, there exist a multivariate EVD G and a vector $\mathbf{x}_0 \in \mathbb{R}^m$ with $G(\mathbf{x}_0) < 1$ such that

$$W(\mathbf{x}) = 1 + \log(G(\mathbf{x})), \qquad \mathbf{x} \geq \mathbf{x}_0. \tag{6}$$

Note that $H(\mathbf{x}) = 1 + \log(G(\mathbf{x}))$, $G(\mathbf{x}) \geq 1/\mathrm{e}$, does not define a distribution function unless $m \in \{1, 2\}$; see Michel [27], Theorem 6. We, therefore, call H a GP function. It is, actually, a quasi-copula (Alsina *et al.* [1], Genest *et al.* [20] and Section 5.1 in Falk *et al.* [17]). Lemma 5.1.5 in Falk *et al.* [17] implies, on the other hand, that for any GP function there exists a distribution function W satisfying (6).

The preceding considerations together with elementary computations entail now the following characterization of domains of attraction in terms of a GPD. By $\|\cdot\|$ we denote an arbitrary norm on $\mathbb{R}^m$.

Theorem 2.1. *An arbitrary distribution function F is in the domain of attraction of a multivariate EVD G if and only if this is true for the univariate margins and if there exists a GPD W with ultimately uniform margins $W_i(x) = 1 + x$, $x_0 \leq x \leq 0$, $i \leq m$, such that*

$$C_F(\mathbf{y}) = W(\mathbf{y} - \mathbf{1}) + \mathrm{o}(\|\mathbf{y} - \mathbf{1}\|)$$

uniformly for $\mathbf{y} \in [0, 1]^m$.

We have the following equivalences for an arbitrary copula C to lie in the domain of attraction of an EVD.

Corollary 2.2. *C is in the domain of attraction of an EVD G*

$\iff$ *There exists a GPD W with ultimately uniform margins such that*

$$C(\mathbf{y}) = W(\mathbf{y} - \mathbf{1}) + \mathrm{o}(\|\mathbf{y} - \mathbf{1}\|),$$

uniformly for $\mathbf{y} \in [0, 1]^m$. *In this case* $W(\mathbf{x}) = 1 + \log(G(\mathbf{x}))$, $\mathbf{x}_0 \leq \mathbf{x} \leq \mathbf{0} \in \mathbb{R}^m$.

$\iff$ *There exists a norm* $\|\cdot\|_D$ *on* $\mathbb{R}^m$ *such that*

$$C(\mathbf{y}) = 1 - \|\mathbf{y} - \mathbf{1}\|_D + \mathrm{o}(\|\mathbf{y} - \mathbf{1}\|_D),$$

uniformly for $\mathbf{y} \in [0,1]^m$. *In this case* $G(\mathbf{x}) = \exp(-\|\mathbf{x}\|_D)$, $\mathbf{x} \le \mathbf{0}$.

Recall that all norms on $\mathbb{R}^m$ are equivalent and, thus, $\mathrm{o}(\|\mathbf{y} - \mathbf{1}\|_D)$ in the second equivalence above can be substituted by $\mathrm{o}(\|\mathbf{y} - \mathbf{1}\|)$ with an arbitrary norm $\|\cdot\|$ on $\mathbb{R}^m$.

The preceding results show that the upper tail of the copula C_F of a distribution function F can reasonably be approximated only by that of a GPD W with ultimately uniform margins. To the best of our knowledge, this provides new insight into the significance of multivariate GPD. But it is in accordance with Rootzén and Tajvidi [35], who showed that, in the multivariate case, modelling exceedances of a random variable over a high threshold can rationally be done only by a multivariate GPD.

Proof of Corollary 2.2. It is well known that a GPD W with ultimately uniform margins can be written as

$$W(\mathbf{x}) = 1 - \|\mathbf{x}\|_D, \qquad \mathbf{x}_0 \le \mathbf{x} \le \mathbf{0},$$

where $\|\cdot\|_D$ is a norm on $\mathbb{R}^m$ with particular properties, called a D-norm; see Section 4.4 in Falk *et al.* [17]. In particular, $G(\mathbf{x}) = \exp(-\|\mathbf{x}\|_D)$, $\mathbf{x} \le \mathbf{0}$, defines an EVD on $\mathbb{R}^m$. If $C(\mathbf{y}) = W(\mathbf{y} - \mathbf{1}) + \mathrm{o}(\|\mathbf{y} - \mathbf{1}\|)$, $\mathbf{y} \in [0,1]^m$, for some norm $\|\cdot\|$ on $\mathbb{R}^m$, then

$$\begin{aligned} C^n\left(\mathbf{1} + \frac{y}{n}\right) &= \left(1 - \frac{1}{n}\|\mathbf{y}\|_D + \mathrm{o}\left(\frac{1}{n}\|\mathbf{y}\|\right)\right)^n \\ &\underset{n\to\infty}{\to} \exp(-\|\mathbf{y}\|_D) = G(\mathbf{y}), \qquad \mathbf{y} \le \mathbf{0}. \end{aligned}$$

Together with Theorem 2.1 this implies Corollary 2.2. □

In the final equivalence of Corollary 2.2, the norm can obviously be computed as

$$\|\mathbf{x}\|_D = \lim_{t \downarrow 0} \frac{1 - C(\mathbf{1} + t\mathbf{x})}{t} = l(\mathbf{x}), \qquad \mathbf{x} \le \mathbf{0};$$

that is, it is the stable tail dependence function. It turns out that any stable tail dependence function is actually a norm. This explains why it is a convex function and homogeneous of order one.

Example 2.3. Take an arbitrary Archimedean copula

$$C_\varphi(\mathbf{u}) = \varphi^{-1}\big(\varphi(u_1) + \cdots + \varphi(u_m)\big),$$

where the generator $\varphi : (0,\infty) \to [0,\infty)$ is a continuous function that is strictly decreasing on $(0,1]$, $\varphi(1) = 0$, $\lim_{x\downarrow 0} \varphi(x) = \infty$ and $\varphi^{-1}(t) = \inf\{x > 0 \colon \varphi(x) \le t\}$, $t \ge 0$.

Note that C_φ is not automatically a copula for each function φ: $(0,\infty) \to [0,\infty)$ as above. While, in the bivariate case $m = 2$, convexity of φ^{-1} is a necessary and sufficient condition, this is no longer true in higher dimension $m \geq 3$. Instead, C_φ is, for general dimension $m \geq 2$, a copula if and only if φ^{-1} is differentiable up to order $m-2$, the derivatives satisfy $(-1)^k(\varphi^{-1})^{(k)}(x) \geq 0$, $k = 0, \dots, m-2$, $x \in (0,\infty)$ and further if $(-1)^{m-2}(\varphi^{-1})^{(m-2)}$ is non-increasing and convex in $(0,\infty)$; see McNeil and Nešlehová [24], Theorem 2.2.

If φ is differentiable from the left in $x = 1$ with left derivative $\varphi'(1-) \neq 0$, then

$$\lim_{t\downarrow 0} \frac{1 - C_\varphi(\mathbf{1} + t\mathbf{x})}{t} = \sum_{i\leq m} |x_i| = \|\mathbf{x}\|_1, \qquad \mathbf{x} \leq \mathbf{0};$$

that is, each Archimedean copula with a generator φ as above is in the domain of attraction of the EVD $G(\mathbf{x}) = \exp(-\|\mathbf{x}\|_1)$, $\mathbf{x} \leq \mathbf{0}$, with independent margins. The margins of C_φ are, therefore, tail independent; that is, the tail dependence parameters vanish:

$$\chi(i,j) := \lim_{x\uparrow 1} P(U_i > x \mid U_j > x) = 0, \qquad 1 \leq i \neq j \leq m,$$

where the random vector $(U_1, \dots, U_m)$ follows the distribution function C_φ. For a discussion of the tail dependence parameter and further literature, see Section 6.1 in Falk *et al.* [17].

The preceding considerations concern, for example, the Clayton and the Frank copula, which have generators $\varphi_C(t) = \vartheta^{-1}(t^{-\vartheta} - 1)$ and $\varphi_F(t) = -\log((\exp(-\vartheta t) - 1)/(\exp(-\vartheta) - 1))$, $\vartheta > 0$, but not the Gumbel copula with parameter $\lambda > 1$, which has generator $\varphi_G(t) = -(\log(t))^\lambda$, $\lambda \geq 1$, $0 < t \leq 1$.

Any multivariate EVD G has univariate EVD margins and any multivariate GPD W has univariate GPD margins in its upper tail. We can transform an arbitrary multivariate EVD to an EVD with negative exponential margins by just transforming the margins. Equally, we can transform an arbitrary W to a GPD with uniform margins by just transforming the margins. This transformation can also be done backwards; see Section 5.6 in Falk *et al.* [17]. We will, therefore, consider in what follows multivariate GPD derived from an EVD G with negative exponential margins. For a recent account on multivariate GPD, see Michel [25].

From the de Haan–Resnick–Pickands representation of a multivariate EVD, it is well known that a function G on $(-\infty, 0]^m$ is the distribution function of an EVD with negative standard exponential margins $G_i(x) = \exp(x)$, $x \leq 0$, $i \leq m$, if and only if it can be represented as

$$G(\mathbf{x}) = \exp\left(\int_{S_m} \min_{i\leq m}(x_i t_i)\mu(\mathrm{d}\mathbf{t})\right), \qquad \mathbf{x} \leq \mathbf{0},$$

where μ is a finite measure on $S_m := \{\mathbf{t} \geq \mathbf{0}\colon \sum_{i\leq m} t_i = 1\}$, called an *angular measure*, with the characteristic property $\int_{S_m} t_i \mu(\mathrm{d}\mathbf{t}) = 1$, $i \leq m$; see Section 4.2 in Falk *et al.* [17]. Note that this integrability condition on μ implies that $\mu(S_m) = \int_{S_m} 1\,\mathrm{d}\mu = \int_{S_m} \sum_{i\leq m} t_i \mu(\mathrm{d}\mathbf{t}) = \sum_{i\leq m} \int_{S_m} t_i \mu(\mathrm{d}\mathbf{t}) = m$.

As a consequence we obtain that a multivariate GPD W with standard uniform margins $1 - W_i(x) = x$, $i \le m$, in a left neighborhood of $\mathbf{0} \in \mathbb{R}^m$ can be represented as

$$\begin{aligned} W(\mathbf{x}) &= 1 + \left(\sum_{j \le m} x_j\right) \int_{S_m} \max_{i \le m}(\tilde{x}_i t_i)\mu(d\mathbf{t}) \\ &=: 1 + \left(\sum_{j \le m} x_j\right) D(\tilde{x}_1, \ldots, \tilde{x}_{m-1}) \end{aligned} \tag{7}$$

for $\mathbf{x}_0 \le \mathbf{x} \le \mathbf{0}$, where μ is as above, $\tilde{x}_i = x_i / \sum_{j \le m} x_j$ and D: $\{\mathbf{u} \in [0,1]^{m-1}: \sum_{j \le m-1} u_j \le 1\} \to [1/m, 1]$ is a *Pickands dependence function* (Section 4.3 in Falk *et al.* [17]).

The following result characterizes a GPD with uniform margins in terms of random variables. It provides an easy way to generate a multivariate GPD, thus extending the bivariate approach proposed by Buishand *et al.* [6] to an arbitrary dimension. Recall that an arbitrary multivariate GPD can be obtained from a GPD with ultimately uniform margins by just transforming the margins. For a recent account on simulation techniques of multivariate GPD, see Michel [26].

Proposition 2.4.

(i) *Let W be a multivariate GPD with standard uniform margins in a left neighborhood of $\mathbf{0} \in \mathbb{R}^m$. Then there is a random vector $\mathbf{Z} = (Z_1, \ldots, Z_m)$ with $Z_i \in [0, m]$ and $E(Z_i) = 1$, $i \le m$, and a vector $(-1/m, \ldots, -1/m) \le \mathbf{x}_0 < \mathbf{0}$ such that*

$$W(\mathbf{x}) = P\left(-U\left(\frac{1}{Z_1}, \ldots, \frac{1}{Z_m}\right) \le \mathbf{x}\right), \qquad \mathbf{x}_0 \le \mathbf{x} \le \mathbf{0},$$

where the random variable U is uniformly distributed on $(0, 1)$ *and independent of* $\mathbf{Z}$.

(ii) *The random vector $-U(1/Z_1, \ldots, 1/Z_m)$ follows a GPD with standard uniform margins in a left neighborhood of $\mathbf{0} \in \mathbb{R}^m$ if U is independent of $\mathbf{Z} = (Z_1, \ldots, Z_m)$ and $0 \le Z_i \le c_i$ a.s. with $E(Z_i) = 1$, $i \le m$, for some $c_1, \ldots, c_m \ge 1$.*

Note that the case of a GPD W with arbitrary uniform margins $W_i(x) = 1 - a_i x$ in a left neighborhood of $\mathbf{0}$ with arbitrary scaling factors $a_i > 0$, $i \le m$, immediately follows from the preceding result by substituting Z_i by $a_i Z_i$.

Proof of Proposition 2.4. First we establish part (i). From representation (7) we obtain that for $\mathbf{x}$ in a left neighborhood of $\mathbf{0} \in \mathbb{R}^m$

$$W(\mathbf{x}) = 1 + \left(\sum_{j \le m} x_j\right) \int_{S_m} \max_{i \le m}(\tilde{x}_i t_i)\mu(d\mathbf{t})$$

with some measure μ on S_m such that $\mu(S_m) = m$ and $\int_{S_m} t_i \mu(d\mathbf{t}) = 1$, $i \le m$.

Now $\tilde{\mu}(\cdot) = \mu(\cdot)/m$ defines a probability measure on S_m. Let $\mathbf{T} = (T_1, \ldots, T_m)$ be a random vector with values in S_m that has distribution $\tilde{\mu}$ and put $\mathbf{Z} := m\mathbf{T}$. Then $\mathbf{Z} \in [0, m]^m$ and $E(Z_i) =$

$\int_{S_m} t_i \mu(d\mathbf{t}) = 1$, $i \le m$. We have, further, for $\mathbf{x} \le \mathbf{0} \in \mathbb{R}^m$ with $x_j \ge -1/m$, $j \le m$,

$$\begin{aligned}
&P\left(-U\left(\frac{1}{Z_1},\dots,\frac{1}{Z_m}\right) \le \mathbf{x}\right)\\
&= P\left(-U\left(\frac{1}{T_1},\dots,\frac{1}{T_m}\right) \le m\mathbf{x}\right)\\
&= \int_{S_m} P\left(-U\left(\frac{1}{t_1},\dots,\frac{1}{t_m}\right) \le m\mathbf{x} \mid \mathbf{T} = \mathbf{t}\right)(P * \mathbf{T})(d\mathbf{t})\\
&= \int_{S_m} P\left(-U\left(\frac{1}{t_1},\dots,\frac{1}{t_m}\right) \le m\mathbf{x}\right)\tilde{\mu}(d\mathbf{t})\\
&= \frac{1}{m}\int_{S_m} P\left(-U\left(\frac{1}{t_1},\dots,\frac{1}{t_m}\right) \le m\mathbf{x}\right)\mu(d\mathbf{t})\\
&= \frac{1}{m}\int_{S_m} P\left(U \ge m \max_{i\le m}(-x_i t_i)\right)\mu(d\mathbf{t})\\
&= \frac{1}{m}\int_{S_m} P\left(U \ge -m\left(\sum_{j\le m} x_j\right)\max_{i\le m}(\tilde{x}_i t_i)\right)\mu(d\mathbf{t})\\
&= \frac{1}{m}\int_{S_m} 1 + m\left(\sum_{j\le m} x_j\right)\max_{i\le m}(\tilde{x}_i t_i)\mu(d\mathbf{t})\\
&= 1 + \left(\sum_{j\le m} x_j\right)\int_{S_m} \max_{i\le m}(\tilde{x}_i t_i)\mu(d\mathbf{t}).
\end{aligned}$$

This implies part (i) of the proposition.

On the other hand, we have for $\mathbf{x} \le \mathbf{0}$ and large $s > 0$

$$\begin{aligned}
&P\left(-U\left(\frac{1}{Z_1},\dots,\frac{1}{Z_m}\right) \le \frac{1}{s}\mathbf{x}\right)^s\\
&= \left(\int_{[\mathbf{0},\mathbf{c}]} P\left(U \ge \frac{1}{s}\max_{i\le m}(-x_i z_i)\right)(P * \mathbf{Z})(d\mathbf{z})\right)^s\\
&= \left(1 - \frac{1}{s}\int_{[\mathbf{0},\mathbf{c}]} \max_{i\le m}(-x_i z_i)(P * \mathbf{Z})(d\mathbf{z})\right)^s\\
&\to_{s\to\infty} \exp\left(-\int_{[\mathbf{0},\mathbf{c}]} \max_{i\le m}(-x_i z_i)(P * \mathbf{Z})(d\mathbf{z})\right)\\
&=: G(\mathbf{x})
\end{aligned}$$

with $\mathbf{c} = (c_1,\dots,c_m)$.

Lemma 7.2.1 in Reiss [32] now implies that G is a distribution function that is obviously max stable: $G^s(s^{-1}\mathbf{x}) = G(\mathbf{x})$, $s > 0$; that is, G is a multivariate EVD and has negative standard exponential margins $G_i(x) = \exp(xE(Z_i)) = \exp(x)$, $x \le 0$. As a consequence, $1 + \log(G(\mathbf{x}))$ is a GP function with

$$\begin{aligned} 1 + \log(G(\mathbf{x})) &= 1 - \int_{[\mathbf{0},\mathbf{c}]} \max_{i \le m}(-x_i z_i)(P * \mathbf{Z})(\mathrm{d}\mathbf{z}) \\ &= P\left(-U\left(\frac{1}{Z_1}, \dots, \frac{1}{Z_m}\right) \le \mathbf{x}\right) \end{aligned}$$

for $\mathbf{x}_0 \le \mathbf{x} \le \mathbf{0}$ and some $\mathbf{x}_0 < \mathbf{0}$. □

Let, for instance, C be an arbitrary m-dimensional *copula*; that is, C is the distribution function of a random vector $\mathbf{S}$ with uniform margins $P(S_i \le s) = s$, $s \in (0, 1)$, $i \le m$, (Nelsen [29]). Then $\mathbf{Z} := 2\mathbf{S}$ is a proper choice in part (ii) of Proposition 2.4. Proposition 2.4, therefore, maps the set of copulas in a natural way to the set of multivariate GPDs, thus opening a wide range of possible scenarios.

According to Theorem 2.1, we call a copula C_W a *GPD copula on* $[0, 1]^m$ or simply a *GPD copula* if there exists $\mathbf{y}_0 < \mathbf{1}$ such that

$$C_W(\mathbf{y}) = W(\mathbf{y} - \mathbf{1}), \qquad \mathbf{y}_0 \le \mathbf{y} \le \mathbf{1},$$

where W is a GPD with standard uniform margins in a left neighborhood of zero.

For mathematical convenience we temporarily shift a copula to the interval $[-1, 0]^m$ by shifting each univariate margin by -1. Thus we obtain a distribution function $\widetilde{C}_W$ from a GPD copula C_W, whose marginal distribution functions are the uniform distribution on $[-1, 0]$, and $\widetilde{C}_W$ coincides close to zero with a GPD W as in equation (7); that is, there exists $\mathbf{x}_0 < \mathbf{0}$ such that

$$\begin{aligned} \widetilde{C}_W(\mathbf{x}) &= W(x_1, \dots, x_m) \\ &= 1 + \left(\sum_{j \le m} x_j\right) \int_{S_m} \max_{i \le m}\left(t_i \frac{x_i}{\sum_{j \le m} x_j}\right) \mu(\mathrm{d}\mathbf{t}), \qquad \mathbf{x} \in [\mathbf{x}_0, \mathbf{0}]. \end{aligned}$$

Because $\widetilde{C}_W$ inherits its properties from the original GPD copula C_W, we call $\widetilde{C}_W$ a *GPD copula on* $[-1, 0]^m$.

For later purposes we remark that a random vector $\mathbf{V} \in [-1, 0]^m$ following a GPD copula on $[-1, 0]^m$ can easily be generated as follows, using Proposition 2.4. Let U be uniformly distributed on $(0, 1)$ and independent of the vector $\mathbf{S} = (S_1, \dots, S_m)$, which follows an arbitrary copula on $[0, 1]^m$. Then we have for $i \le m$

$$\begin{aligned} P\left(-U \frac{1}{2S_i} \le x\right) &= \begin{cases} 1 + x, & \text{if } -\dfrac{1}{2} \le x \le 0, \\ \dfrac{1}{4|x|}, & \text{if } x < -\dfrac{1}{2}, \end{cases} \\ &=: H(x), \qquad x \le 0, \end{aligned}$$

and, consequently,

$$\mathbf{V} := \left(H\left(-\frac{U}{2S_1}\right) - 1, \dots, H\left(-\frac{U}{2S_m}\right) - 1 \right) = (V_1, \dots, V_m)$$

with

$$V_i = \begin{cases} -\dfrac{U}{2S_i}, & \text{if } U \le S_i, \\ \dfrac{S_i}{2U} - 1, & \text{if } U > S_i, \end{cases} \tag{8}$$

follows by Proposition 2.4 a GPD copula on $[-1, 0]^m$.

3. Multivariate piecing together

The multivariate PT approach consists of two steps. In a first step, the upper tail of a given m-dimensional copula C is cut off and substituted by the upper tail of multivariate GPD copula in a continuous manner. The result is again a copula, that is, an m-dimensional distribution with uniform margins. The other step consists of the transformation of each margin of this copula by a given univariate distribution function F_i^*, $1 \le i \le m$. This provides, altogether, a multivariate distribution function with prescribed margins F_i^* whose copula coincides in its central part with C and in its upper tail with a GPD copula.

We start with fitting a GPD copula to the upper tail of a given copula C on $[-1, 0]^m$. Recall that for mathematical convenience we shift any copula $\tilde{C}(\mathbf{u})$, $\mathbf{u} \in [0, 1]^m$, to a copula on $[-1, 0]^m$ by setting $C(\mathbf{v}) = \tilde{C}(\mathbf{1} + \mathbf{v})$, $\mathbf{v} \in [-1, 0]^m$.

Let $\mathbf{V} = (V_1, \dots, V_m)$ follow a GPD copula on $[-1, 0]^m$; that is, $P(V_i \le x) = 1 + x$, $-1 \le x \le 0$, is for each $i \le m$ the uniform distribution on $[-1, 0]$, and there exists $\mathbf{x}_0 = (x_0^{(1)}, \dots, x_0^{(m)}) < \mathbf{0}$ such that for each $\mathbf{x} = (x_1, \dots, x_m) \in [\mathbf{x}_0, \mathbf{0}]$

$$P(\mathbf{V} \le \mathbf{x}) = 1 + \left(\sum_{i \le m} x_i \right) D\left(\frac{x_1}{\sum_{i \le m} x_i}, \dots, \frac{x_{m-1}}{\sum_{i \le m} x_i} \right),$$

where D is a Pickands dependence function.

Let $\mathbf{Y} = (Y_1, \dots, Y_m)$ follow an arbitrary copula C on $[-1, 0]^m$ and suppose that $\mathbf{Y}$ is independent of $\mathbf{V}$. Choose a threshold $\mathbf{y} = (y_1, \dots, y_m) \in [-1, 0]^m$ and put

$$Q_i := Y_i 1_{(Y_i \le y_i)} - y_i V_i 1_{(Y_i > y_i)}, \qquad i \le m. \tag{9}$$

The random vector $\mathbf{Q}$ then follows a GPD copula on $[-1, 0]^m$, which coincides with C on $\times_{i \le m} [-1, y_i]$. This is the content of the main result of this section.

Proposition 3.1. *Suppose that* $P(\mathbf{Y} > \mathbf{y}) > 0$. *Each* Q_i *defined in* (9) *follows the uniform distribution on* $[-1, 0]$. *The random vector* $\mathbf{Q} = (Q_1, \dots, Q_m)$ *follows a GPD copula on* $[-1, 0]^m$, *which coincides with* C *on* $\times_{i \le m} [-1, y_i]$; *that is,*

$$P(\mathbf{Q} \le \mathbf{x}) = C(\mathbf{x}), \qquad \mathbf{x} \le \mathbf{y}.$$

We have, moreover, with $x_i \in [\max(y_i, x_0^{(i)}), 0]$, $i \le m$, *for any non-empty subset* K *of* $\{1, \dots, m\}$

$$P(Q_i \ge x_i, i \in K) = P(V_i \ge b_{i,K} x_i, i \in K),$$

where

$$b_{i,K} := \frac{P(Y_j > y_j, j \in K)}{-y_i} = \frac{P(Y_j > y_j, j \in K)}{P(Y_i > y_i)} \in (0, 1], \qquad i \in K.$$

Proof. First we show that each Q_i follows the uniform distribution on $[-1, 0]$. We have for $-1 \le x \le y_i$

$$\begin{aligned} P(Q_i \le x) &= P(Q_i \le x, Y_i \le y_i) + P(Q_i \le x, Y_i > y_i) \\ &= P(Y_i \le x) \\ &= 1 + x, \end{aligned}$$

whereas for $y_i < x \le 0$ we obtain

$$\begin{aligned} P(Q_i \le x) &= P(Y_i \le y_i) + P(-y_i V_i \le x) P(Y_i > y_i) \\ &= 1 + y_i + P\left(V_i \le -\frac{x}{y_i}\right)(-y_i) \\ &= 1 + y_i + \left(1 - \frac{x}{y_i}\right)(-y_i) \\ &= 1 + x. \end{aligned}$$

The random vector $\mathbf{Q}$, thus, follows a copula on $[-1, 0]^m$. We have, further, for $\mathbf{x} \le \mathbf{y}$

$$\begin{aligned} P(\mathbf{Q} \le \mathbf{x}) &= P(\mathbf{Q} \le \mathbf{x}, \mathbf{Y} \le \mathbf{y}) + P(\mathbf{Q} \le \mathbf{x}, \mathbf{Y} \not\le \mathbf{y}) \\ &= P(\mathbf{Y} \le \mathbf{x}) \\ &= C(\mathbf{x}). \end{aligned}$$

By Proposition 2.1 in Falk and Michel [18] we have with $x_i \in [\max(y_i, \omega_i), 0]$, $i \le m$, $t \in [0, 1]$ and an arbitrary subset $K \subset \{1, \dots, m\}$

$$\begin{aligned} P(Q_j > t x_j, j \in K) &= P(Q_j > t x_j, Y_j > y_j, j \in K) \\ &= P(-y_j V_j > t x_j, j \in K) P(Y_j > y_j, j \in K) \\ &= t P(-y_j V_j > x_j, j \in K) P(Y_j > y_j, j \in K) \\ &= t P(Q_j > x_j, j \in K), \end{aligned}$$

which, again by Proposition 2.1 in Falk and Michel [18], implies that $\mathbf{Q}$ follows a GPD.

We have, moreover, with $x_i \in [\max(y_i, x_0^{(i)}), 0]$, $i \le m$,

$$\begin{aligned}
&P(Q_i \ge x_i, i \in K) \\
&\quad = P(Q_i \ge x_i, Y_i > y_i, i \in K) + P(Q_i \ge x_i, i \in K, Y_j \le y_j \text{ for some } j \in K) \\
&\quad = P(-y_i V_i \ge x_i, i \in K) P(Y_i > y_i, i \in K) \\
&\quad = P\left(V_i \ge -\frac{x_i}{y_i}, i \in K\right) P(Y_i > y_i, i \in K) \\
&\quad = P(V_i \ge b_{i,K} x_i, i \in K). \qquad \square
\end{aligned}$$

The above approach provides an easy way to generate a random vector $\mathbf{X} \in \mathbb{R}^m$ with prescribed margins F_i^*, $i \le m$, such that $\mathbf{X}$ has a given copula in the central part of the data, whereas in the upper tail it has a GPD copula.

Take $\mathbf{Q} = (Q_1, \dots, Q_m)$ as in (9) and put $\widetilde{\mathbf{Q}} := (Q_1 + 1, \dots, Q_m + 1)$. Then each component $\widetilde{Q}_i$ of $\widetilde{\mathbf{Q}}$ is uniformly distributed on $(0, 1)$ and thus

$$\mathbf{X} := (X_1, \dots, X_m) := (F_1^{*-1}(\widetilde{Q}_1), \dots, F_m^{*-1}(\widetilde{Q}_m)) \tag{10}$$

has the desired properties.

Combining the univariate *and* the multivariate PT approach now consists in choosing a threshold $u(i) \in \mathbb{R}$ for each dimension $i \le m$ and a univariate distribution function F_i together with an arbitrary univariate GPD $W_{\gamma_i,\mu_i\sigma_i}$, and putting for $i \le m$

$$F_i^*(x) := \begin{cases} F_i(x), & \text{if } x \le u(i) \\ \big(1 - F_i(u(i))\big) W_{\gamma_i,\mu_i\sigma_i}(x) + F_i(u(i)), & \text{if } x > u(i) \end{cases}. \tag{11}$$

This is typically done in a way such that F_i^* is a continuous function.

4. An application to operational loss data

In this section we apply our multivariate PT approach to operational loss data. For an excellent introduction to operational risk and insurance analytics, see Chapter 10 of McNeil *et al.* [23] and the literature cited there. In the sequel we give a brief summary.

According to the New Basel Capital Accord (Basel II), banks are required to determine the regulatory capital charge for operational risk, defined as the risk of losses resulting from inadequate or failed internal processes, people and systems or from external events. The Basel Committee on Banking Supervision encourages the use and further development of advanced modelling techniques to quantify operational risk. The most risk-sensitive methodology is the *loss distribution approach* using bank internal data to estimate probability distribution functions for each business line/event type category. To provide a greater consistency of loss data collection within and between banks, operational losses are classified in eight business lines and seven event types. To calculate the capital charge for each business line/event type combination, a risk measure such as value at risk to the 99.9% confidence level over a one-year holding period is chosen. A conservative way to assess a bank's total capital requirement is to sum up the capital charges across business line/event type classes assuming perfect dependence and disregarding diversification effects in operational risk. Actually, the fact that all severe losses occur in the same year is rather dubious. Therefore, the dependence structure among losses of different business line/event type categories needs to be modelled explicitly. For simplicity, we consider in what follows only business lines and not event types.

The frequency of a loss event for business line i over a one-year time horizon will be denoted by $N(i)$. The random loss associated with the kth loss event for business line i will be denoted by $\zeta_k(i)$.

The random loss $L(i)$ over one year for business line i is, therefore, modelled as

$$L(i) = \sum_{k=1}^{N(i)} \zeta_k(i),$$

where $\zeta_1(i), \zeta_2(i), \dots$ are assumed to be i.i.d. with distribution function F_i and they are independent of their total number $N(i)$.

The goal is to model the total loss distribution for operational risk; that is, the distribution of

$$L := \sum_{i=1}^{m} L(i),$$

or parameters of it such as the value at risk $\mathrm{VAR}(\alpha)$ at the probability level α satisfying

$$P\big(L \geq \mathrm{VAR}(\alpha)\big) = 1 - \alpha$$

or the expected shortfall at the probability level α

$$\mathrm{ES}(\alpha) := E\big(L \mid L \geq \mathrm{VAR}(\alpha)\big).$$

In order to assess the total capital charge, the traditional models for measuring operational risk determine $\mathrm{VAR}(\alpha)$ and $\mathrm{ES}(\alpha)$ for each of the m business lines separately and then simply sum up the corresponding capital charges.

In contrast, Di Clemente and Romano [14] suggest modelling the dependence structure among $L(1), \dots, L(m)$ by a copula function, precisely, by the copula corresponding to the m-dimensional t-distribution with ν degrees of freedom. For a closer look at the issue of modelling the dependence among components of a random vector of financial risk factors using the concept of a copula, see Chapter 5 of McNeil *et al.* [23].

In our application we analyse operational losses of the external database *SAS OpRisk Global Data*, which contains worldwide information on publicly reported operational losses over US \$100 000. Since we do not know the probability of losses lying under US \$100 000, we neglect this cut-off limit in modelling the severity and frequency of the data. We concentrate on two business lines of the financial sector, *Commercial banking* and *Retail banking*.

First we follow the copula extreme value theory approach for modelling operational loss data as outlined in Di Clemente and Romano [14], but we add the multivariate PT approach developed in Section 3.

4.1. Estimation

Before applying the multivariate PT approach, we explore the characteristics of the empirical distributions of the two business lines and estimate the model's parameters. Thereby the severity distribution of the random variable $\zeta_k(i)$ and the frequency distribution of the random variable $N(i)$, $i = 1, 2$, are treated separately. To obtain the distribution function of the total loss $L(i)$ of the business line i over a one-year time horizon we accomplish a Monte Carlo simulation combining the severity distribution with the frequency distribution.

First we analyse the empirical distributions of the loss severity. The measures skewness and kurtosis indicate that the empirical distributions of the two business lines are skewed to the right and very heavy tailed.

In the next step, parametric distributions (i.e., Weibull, gamma and lognormal distribution) are fitted to the data. The parameters are estimated by the maximum likelihood method. With the help of graphical analysis (QQ plots, theoretical versus empirical distribution function plots) and goodness-of-fit tests (Anderson–Darling test, Cramer–von Mises (CvM) test), we conclude, that none of the selected distributions provides a good fit to the complete data sets. (For a detailed presentation and discussion of goodness-of-fit techniques, see D'Agostino and Stephens [8].) However, the lognormal distribution fits the body of the data very well, while it underestimates the severity of the data in the right tail.

Therefore, to fit the tail data accurately, the univariate POT method in the model of the severity distribution of the losses $\zeta_k(i)$ is applied: The existence of a threshold $u(i)$ for each business line i is assumed such that $\zeta_k(i)$ follows a lognormal distribution function below $u(i)$, whereas above $u(i)$ it follows a univariate GPD, that is,

$$\begin{aligned} &P\big(\zeta_k(i) \le x\big) \\ &= \begin{cases} F_i(x), & x \le u(i), \\ F_i(u(i)) + \big(1 - F_i(u(i))\big)\,\mathrm{GPD}_{\beta(i),\xi(i)}\big(x - u(i)\big), & x \ge u(i), \end{cases} \end{aligned} \tag{12}$$

where the GPD is given by

$$\mathrm{GPD}_{\beta(i),\xi(i)}(z) := 1 - \left(1 + \xi(i)\frac{z}{\beta(i)}\right)^{-1/\xi(i)}, \qquad z \ge 0,$$

with shape and scale parameter $\xi(i) > 0$, $\beta(i) > 0$. Furthermore, F_i is defined as $F_i(x) := \Phi((\log(x) - \mu(i))/\sigma(i))$, where Φ is the standard normal distribution function and $\mu(i) \in \mathbb{R}$, $\sigma(i) > 0$ are location and scale parameters of the lognormal distribution.

In this case we obtain for $x \ge u(i)$

$$P\big(\zeta_k(i) > x\big) = P\big(\zeta_k(i) > u(i)\big)\left(1 + \xi(i)\frac{x - u(i)}{\beta(i)}\right)^{-1/\xi(i)}.$$

The threshold $u(i)$ is chosen with the help of mean excess plots. The shape and scale parameters of the GPD are estimated by the maximum likelihood method. For a discussion of the parameter estimation of a GPD and optimal choice of the threshold, see Section 6 of Embrechts *et al.* [15].

Table 1. Estimated model parameters

	$\alpha(i)$	$r(i)$	$\mu(i)$	$\sigma(i)$	$u(i)$	$\beta(i)$	$\xi(i)$
Commercial banking	0.74	46.10	2.19	2.23	918.02	609.84	0.82
Retail banking	0.39	162.04	0.88	2.06	69.18	99.75	1.02

To determine the frequency distribution of the random variable $N(i)$, the Poisson and negative binomial distribution are fitted to the total number of losses per year. The parameters of these distributions are estimated by the method of moments. Since the negative binomial distribution has two parameters, α and r, it is more flexible and often provides a better fit to operational loss data than the Poisson distribution; see Cruz [7], page 89. With the help of the χ^2 goodness-of-fit test, this expectation is confirmed. Therefore, the random variable $N(i)$ is modelled by the negative binomial distribution, whose probability mass function is expressed as

$$P\big(N(i) = n\big) = \binom{\alpha(i) + n - 1}{n} \left(\frac{1}{1 + r(i)}\right)^{\alpha(i)} \left(\frac{r(i)}{1 + r(i)}\right)^{n}, \qquad n \in \mathbb{N}_0,$$

with $\alpha(i) > 0$, $r(i) > 0$. The resulting estimates of the model's parameters are given in Table 1.

In the following, we model the dependence structure among $L(1)$ and $L(2)$ by a copula function. The assumption of a normal copula has been quite popular in finance for modelling the dependence between different risks, but it puts less weight on observations that are large in each component; see, for example, Rachev *et al.* [31]. The t copula is more heavily tailed and, therefore, better suited for modelling operational risk.

In our bivariate case the t copula is fitted to the total loss data over a one-year time horizon. The parameter of the correlation matrix and the degrees of freedom ν are estimated by the maximum likelihood method. For a discussion of the problem of fitting copulas to data, see Section 5.5 of McNeil *et al.* [23].

Table 2 contains the estimated correlation matrix for the t copula with $\nu = 8.64$ estimated degrees of freedom.

To evaluate the goodness of fit of the t copula, the CvM test is applied. For recent reviews of copula goodness-of-fit testing, see Berg [4] and Genest *et al.* [21]. In Table 3 the CvM test value and corresponding p-value are reported.

Since the p-value is 0.521978, the null hypothesis that the dependence structure of the data follows a t copula is not rejected.

Table 2. Estimated correlation matrix for the t copula

	Commercial	Retail
Commercial	1	0.76
Retail	0.76	1

Table 3. Goodness-of-fit test for the t copula

CvM statistic	p-value
0.02543851	0.521978

4.2. Simulation

In the previous section we estimated the parameters that specify a situation where only the univariate PT approach is applied. This model is similar to the one described in Di Clemente and Romano [14]. Now we show by simulations how popular risk measures such as value at risk or expected shortfall can be influenced by the combination of univariate and multivariate piecing together.

First we take the t copula derived in Section 4.1and add the multivariate PT approach developed in Section 3. We simulate 10^4 independent copies of $\widetilde{\mathbf{Y}} = (\widetilde{Y}_1, \widetilde{Y}_2)$ and $\mathbf{V} = (V_1, V_2)$, which follow the t copula from above and a GPD copula on $[-1, 0]^2$, respectively. The realizations of $\mathbf{Y} := \widetilde{\mathbf{Y}} - \mathbf{1}$ and $\mathbf{V}$ are then combined with those of a random vector $\mathbf{Q}$ according to definition (9). The distribution of $\mathbf{Q}$ is then a GPD copula on $[-1, 0]^2$ which coincides with the previously mentioned t copula – shifted by minus one – below some threshold vector $\mathbf{y} = (y_1, y_2)$. The last step consists in shifting the realizations of $\mathbf{Q}$ to the interval $[0, 1]^2$ and transforming the margins by F_1^*, F_2^* according to equation (10). (F_1^*, F_2^* are derived from Monte Carlo simulations as described in Section 4.1.) Thus we obtain 10^4 realizations of a random vector $\mathbf{X}$ that follows a multivariate distribution function that has the marginal distribution functions F_1^*, F_2^* and the associated copula is a GPD copula that coincides with the original t copula in its central part. These realizations of $\mathbf{X}$ are then taken to compute the empirical counterparts of the value at risk and the expected shortfall.

Before we apply these steps, we remark that there are still two remaining degrees of freedom in our model: the GPD copula that underlies $\mathbf{V}$ and the copula threshold vector $\mathbf{y}$. Our goal in this section is to give a first insight into the consequences of replacing the upper tail of a given copula with a GPD copula. (Note that this procedure is justified by Theorem 2.1 and Corollary 2.2.) For this purpose, we assume a simple model:

(i) We define the GPD copula underlying $\mathbf{V}$ indirectly by setting $\mathbf{Z} := 2\mathbf{S}$ in Proposition 2.4, where $\mathbf{S}$ follows a bivariate normal copula.

(ii) The copula threshold vector is obtained by $\mathbf{y} := (F_1(u(1)), F_2(u(2))) - \mathbf{1}$; that is, the thresholds for the marginal distributions in the univariate PT approach (see equation (12)) are transformed and used for the multivariate PT approach, too.

This way, we construct a parametric model that is only dependent on the correlation matrix

$$\mathbf{\Sigma} = \begin{pmatrix} 1 & \varrho \\ \varrho & 1 \end{pmatrix}, \qquad \varrho \in [0, 1),$$

of the random vector $\mathbf{S}$. It is typically very difficult, particularly in higher dimensions, to find a good multivariate model that describes both marginal behavior and dependence structure effectively. The advantage of the preceding copula model is that it depends on just one parameter

$\varrho \in [0,1)$, allowing dependence and independence of the margins in a simple and continuous manner. We refer again to McNeil *et al.* [23].

The simulations as described above were now done for various values of ϱ. To attain more reliable estimates for the value at risk and the expected shortfall, we simulated not only once but 50 times and took the average. Additionally, these values were also computed for the case in which only the univariate PT approach is applied and the t copula kept unchanged. This procedure allows us to identify the effect the multivariate PT approach has on the mentioned risk measures.

In the following we state our results from the simulation series with $\varrho = 0.7$, which models the case of relative high dependence but not complete dependence between the business lines. Although a graphical analysis of the used GPD copulas suggested that the degree of dependence in the upper tail was increasing with ϱ getting larger, there was no observable trend in the estimates of the value at risk and the expected shortfall. Further research is necessary to derive criteria for the optimal choice of ϱ or, more generally, of the GPD copula underlying $\mathbf{V}$ and the copula threshold $\mathbf{y}$.

We now start with the presentation of the results. For simplicity, we identify the business lines Commercial banking and Retail banking with the cases $i = 1$ and $i = 2$, respectively. The random vector $\mathbf{X}$ from above models the combined random losses for these two business lines over one year, that is, $\mathbf{X} = (L(1), L(2))$. The value at risk and the expected shortfall of $L(1)$, $L(2)$ and $L = L(1) + L(2)$ were computed using their empirical counterparts

$$\widehat{\mathrm{VAR}}(\alpha) = \widehat{F}^{-1}(\alpha),$$

where $\widehat{F}$ is the empirical distribution function of $L(1)$, $L(2)$ or L, respectively, and

$$\widehat{\mathrm{ES}}(\alpha) = \frac{1}{n(1-\alpha)} \sum_{i=1}^{n} l_i 1_{[\widehat{\mathrm{VAR}}(\alpha),\infty)}(l_i),$$

where l_i is the ith realization of $L(1)$, $L(2)$ or L, respectively. By 1_B we denote the indicator function of a set B, that is, $1_B(x) = 1$ if $x \in B$ and $1_B(x) = 0$ otherwise.

Table 4 gives the means of 50 independent simulations resulting from the multivariate PT approach, whereas Table 5 makes use of the univariate approach only. Since the marginal distributions are the same in both cases, namely F_1^*, F_2^*, the value at risk estimates concerning $L(1), L(2)$ nearly coincide across both tables. It is apparent that the respective values for L, the total loss, are only slightly different.

Table 4. Estimated value at risk and expected shortfall, GPD copula

	$\widehat{\mathrm{VAR}}(\alpha)$				$\widehat{\mathrm{ES}}(\alpha)$			
α	95%	99%	99.5%	99.9%	95%	99%	99.5%	99.9%
$L(1)$	13 638	32 899	49 650	159 442	46 436	154 758	270 036	1 038 077
$L(2)$	12 586	45 370	84 386	390 127	93 365	381 142	702 350	2 880 672
L	26 578	75 518	127 042	533 701	135 581	512 781	930 472	3 746 889

Table 5. Estimated value at risk and expected shortfall, t copula

	$\widehat{\mathrm{VAR}}(\alpha)$				$\widehat{\mathrm{ES}}(\alpha)$			
α	95%	99%	99.5%	99.9%	95%	99%	99.5%	99.9%
$L(1)$	13 638	32 667	49 196	153 322	37 674	111 075	182 956	608 755
$L(2)$	12 590	45 601	83 414	392 673	71 288	270 821	482 058	1774 252
L	25 428	75 674	131 267	533 710	105 122	366 904	637 184	2261 144

On the other hand, the empirical expected shortfalls attain clearly higher values if the upper tail of the t copula is substituted by a GPD. This behavior is independent of the α level and holds for the losses in the single business lines as well as for the total loss in our simulation. This is remarkable since the corresponding estimates for the value at risk are in both cases nearly the same, indicating that there are some extreme high losses that occur very rarely.

Clearly, with underlying estimated GPD shape parameters $\xi(1) = 0.82$ and $\xi(2) = 1.02$, the theoretical expected shortfall exists only for $i = 1$ but not for $i = 2$. The significant increases for $\widehat{\mathrm{ES}}(\alpha)$ in line $L(1)$ and $L(2)$ in Table 4 should, therefore, only be due to the high volatility of the empirical expected shortfall, whereas the significant increase in line L should be caused by the substituted GPD copula as well. This example of real operational loss data might be considered as a warning, not to underestimate the effects of rare events that nevertheless might occur simultaneously.

In cases of no existing theoretical expected shortfall, Moscadelli [28] suggests the risk measure median shortfall (MS) that is defined regardless of the values of the shape parameter ξ. If the respective distribution function is continuous, the median shortfall has the representation

$$\mathrm{MS}(\alpha) = \mathrm{VAR}\left(\frac{1+\alpha}{2}\right),$$

see Biagini and Ulmer [5], pages 749–750. In addition to our previous results, Table 6 states the estimates for the median shortfall, which we computed as $\widehat{\mathrm{MS}}(\alpha) = \widehat{\mathrm{VAR}}((1+\alpha)/2)$.

Unlike the expected shortfall, the median shortfall as a robust measure is not as heavily influenced by extreme values in the upper tail. Therefore, the median shortfall estimates for $L(1), L(2)$ are closer to each other than the expected shortfall estimates if we compare the t copula case with the GPD copula case. However, the values for the median shortfall of L at the

Table 6. Estimated median shortfall

	$\widehat{\mathrm{MS}}(\alpha)$, t copula				$\widehat{\mathrm{MS}}(\alpha)$, GPD copula			
α	95%	99%	99.5%	99.9%	95%	99%	99.5%	99.9%
$L(1)$	19 829	49 196	78 678	243 938	19 829	49 650	80 292	313 246
$L(2)$	21 494	83 414	162 438	793 252	21 600	84 386	162 866	782 938
L	40 340	131 267	234 910	962 458	42 463	127 042	229 260	1085 283

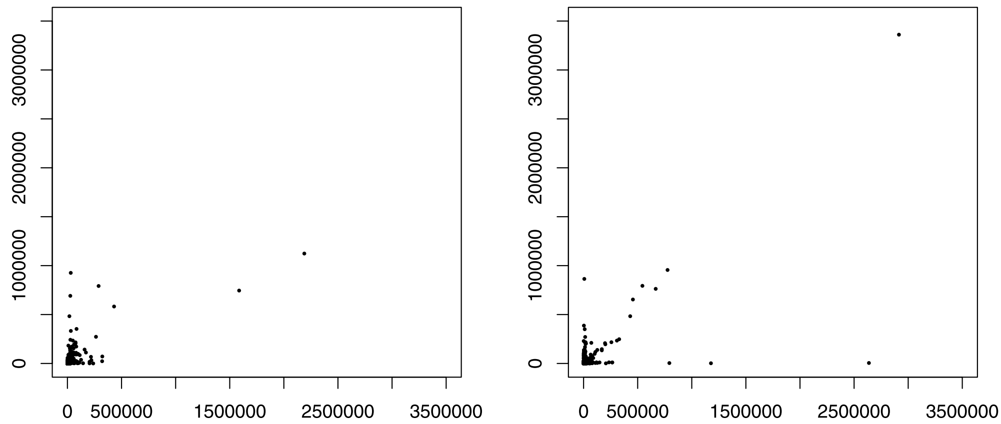

Figure 1. 10^4 random deviates of $(L(1), L(2))$ based on the original t copula (left) and on the GPD copula (right).

99.9% level do clearly differ, indicating a small number of extreme high losses in the upper tail – in accordance with our considerations on the expected shortfall, see page 472.

To give visual insight into these results, Figure 1 compares the realizations of $\mathbf{X} = (L(1), L(2))$ graphically. The graphics were taken from a single simulation that contributed to the results in Table 4 and Table 5. Although the t copula itself is already heavily tailed, the substitution of its upper part by a GPD puts even more weight on observations that are very high in both components. (Recall Corollary 2.2.) The latter type of modelling, therefore, represents a higher risk of an extraordinarily high total loss over a one-year time horizon. This can be seen on the point in the upper right corner in the right scatterplot of Figure 1, which represents a fictive total loss of 6275 000, whereas the highest total loss in the respective pure t copula scenario is 3313 000.

5. Conclusions

In the present paper we extended the well known univariate PT approach to higher dimensions. This was motivated by Theorem 2.1 and Corollary 2.2, which show that it is not sufficient to apply the univariate approach to the marginal distributions of a random vector if the upper tail of its distribution is to be modelled adequately. It is, therefore, necessary to approximate the underlying copula by a multivariate GPD.

The multivariate PT approach that was introduced in (9) offers a wide range of scenarios to be modelled because it depends basically only on some random vector whose components need to be bounded and to have expectation one; see Proposition 2.4. As a consequence we also mentioned a natural way to map the set of copulas to the set of multivariate GPDs that was useful for our simulation studies to obtain a one-parametric model.

Because the values in a simulation are random, simulations occur that produce no values that are high in both components. (This is depending on the sample size, too.) Fixing this disadvantage is subject to further research and could probably be achieved by making additional restrictions on the GPD copula random vector in (9). Furthermore, goodness-of-fit testing of the compound GPD copula is required.

Nevertheless, the multivariate PT approach is a powerful and suitable tool to adequately model multivariate distributions in their upper tails. This ensures that the probability of very rare events that occur simultaneously and have a high effect if they occur is not underestimated. The high empirical expected shortfalls in Table 4 might be considered as a warning.

Acknowledgements

The authors are grateful to the referees for their constructive comments on an earlier version of this paper. The authors are also indebted to the SAS Institute, Heidelberg, for providing access to the database *SAS OpRisk Global Data*.

The first author was supported by DFG Grant FA 262/4-1.

References

[1] Alsina, C., Nelsen, R.B. and Schweizer, B. (1993). On the characterization of a class of binary operations on distribution functions. *Statist. Probab. Lett.* **17** 85–89. MR1223530

[2] Balkema, A.A. and de Haan, L. (1974). Residual life time at great age. *Ann. Probab.* **2** 792–804. MR0359049

[3] Beirlant, J., Goegebeur, Y., Teugels, J. and Segers, J. (2004). *Statistics of Extremes: Theory and Applications*. Chichester: Wiley. MR2108013

[4] Berg, D. (2009). Copula goodness-of-fit testing: An overview and power comparison. *The European Journal of Finance* **15** 675–701.

[5] Biagini, F. and Ulmer, S. (2009). Asymptotics for operational risk quantified with expected shortfall. *ASTIN Bulletin* **39** 735–752.

[6] Buishand, T.A., de Haan, L. and Zhou, C. (2008). On spatial extremes: With application to a rainfall problem. *Ann. Appl. Statist.* **2** 624–642. MR2524349

[7] Cruz, M.G. (2002). *Modeling, Measuring and Hedging Operational Risk*. Chichester: Wiley.

[8] D'Agostino, L. and Stephens, M.A. (1986). *Goodness-of-fit Techniques*. *Statistics: Textbooks and Monographs* **68**. New York: Marcel Dekker Inc. MR0874534

[9] de Haan, L. and de Ronde, J. (1998). Sea and wind: Multivariate extremes at work. *Extremes* **1** 7–45. MR1652944

[10] de Haan, L. and Ferreira, A. (2006). *Extreme Value Theory: An Introduction*. New York: Springer. MR2234156

[11] de Haan, L. and Sinha, A.K. (1999). Estimating the probability of a rare event. *Ann. Statist.* **27** 732–759. MR1714710

[12] Deheuvels, P. (1978). Caractérisation complète des lois extrême multivariées et de la convergence des types extrêmes. *Publ. Inst. Statist. Univ. Paris* **23** 1–36.

[13] Deheuvels, P. (1984). Probabilistic aspects of multivariate extremes. In *Statistical Extremes and Applications* (*Vimeiro*, 1983). *NATO Adv. Sci. Inst. Ser. C Math. Phys. Sci.* (J. Tiago de Oliveira, ed.) **131** 117–130. Dordrecht: Reidel. MR0784817

[14] Di Clemente, A. and Romano, C. (2004). A copula-extreme value theory approach for modelling operational risk. In *Operational Risk Modelling and Analysis, Theory and Practice* (M. Cruz, ed.) 189–208. London: RISK Books.

[15] Embrechts, P., Klüppelberg, C. and Mikosch, T. (1997). *Modelling Extremal Events for Insurance and Finance*. *Applications of Mathematics* (*New York*) **33**. Berlin: Springer. MR1458613
[16] Falk, M. (2008). It was 30 years ago today when Laurens de Haan went the multivariate way. *Extremes* **11** 55–80. MR2420197
[17] Falk, M., Hüsler, J. and Reiss, R.D. (2010). *Laws of Small Numbers*: *Extremes and Rare Events*, 3rd ed. Basel: Birkhäuser.
[18] Falk, M. and Michel, R. (2009). Testing for a multivariate generalized Pareto distribution. *Extremes* **12** 33–51. MR2480722
[19] Galambos, J. (1987). *The Asymptotic Theory of Extreme Order Statistics*, 2nd ed. Melbourne, FL: Robert E. Krieger Publishing Co. Inc. MR0936631
[20] Genest, C., Quesada Molina, J.J., Rodríguez Lallena, J.A. and Sempi, C. (1999). A characterization of quasi-copulas. *J. Multivariate Anal*. **69** 193–205. MR1703371
[21] Genest, C., Rémillard, B. and Beaudoin, D. (2009). Goodness-of-fit tests for copulas: A review and a power study. *Insurance Math*. *Econom*. **44** 199–213. MR2517885
[22] Huang, X. (1992). Statistics of bivariate extreme values. Ph.D. thesis, Tinbergen Institute Research Series.
[23] McNeil, A.J., Frey, R. and Embrechts, P. (2005). *Quantitative Risk Management*: *Concepts*, *Techniques and Tools*. Princeton, NJ: Princeton Univ. Press. MR2175089
[24] McNeil, A.J. and Nešlehová, J. (2009). Multivariate Archimedean copulas, d-monotone functions and l_1-norm symmetric distributions. *Ann*. *Statist*. **37** 3059–3097. MR2541455
[25] Michel, R. (2006). Simulation and estimation in multivariate generalized Pareto models. Ph.D. thesis, Univ. Würzburg.
[26] Michel, R. (2007). Simulation of certain multivariate generalized Pareto distributions. *Extremes* **10** 83–107. MR2415635
[27] Michel, R. (2008). Some notes on multivariate generalized Pareto distributions. *J. Multivariate Anal*. **99** 1288–1301. MR2432486
[28] Moscadelli, M. (2004). The modelling of operational risk: Experience with the analysis of the data collected by the Basel Committee. *Banca D'Italia, Termini di discussione No.* **517**.
[29] Nelsen, R.B. (2006). *An Introduction to Copulas*, 2nd ed. New York: Springer. MR2197664
[30] Pickands, J. III (1975). Statistical inference using extreme order statistics. *Ann*. *Statist*. **3** 119–131. MR0423667
[31] Rachev, S.T., Stein, M. and Sun, W. (2009). Copula concepts in financial markets. Technical report, Karslruhe Institute of Technology (KIT).
[32] Reiss, R.D. (1989). *Approximate Distributions of Order Statistics*: *With Applications to Nonparametric Statistics*. New York: Springer. MR0988164
[33] Reiss, R.D. and Thomas, M. (2007). *Statistical Analysis of Extreme Values*, 3rd ed. Basel: Birkhäuser. MR1819648
[34] Resnick, S.I. (2006). *Heavy-Tail Phenomena*: *Probabilistic and Statistical Modeling*. New York: Springer. MR2271424
[35] Rootzén, H. and Tajvidi, N. (2006). Multivariate generalized Pareto distributions. *Bernoulli* **12** 917–930. MR2265668

Received September 2009 and revised August 2010