ICIR International Center for Insurance Regulation
Goethe University Frankfurt

ICIR Working Paper Series No. 33/2019

Edited by Helmut Gründl and Manfred Wandt

Sensitivity-implied tail-correlation matrices*

Joachim Paulusch† Sebastian Schlütter‡

This version: 7th August 2021

Abstract

Tail-correlation matrices are an important tool for aggregating risk measurements across risk categories, asset classes and/or business segments. This paper demonstrates that traditional tail-correlation matrices—which are conventionally assumed to have ones on the diagonal—can lead to substantial biases of the aggregate risk measurement's sensitivities with respect to risk exposures. Due to these biases, decision-makers receive an odd view of the effects of portfolio changes and may be unable to identify the optimal portfolio from a risk-return perspective. To overcome these issues, we introduce the "sensitivity-implied tail-correlation matrix". The proposed tail-correlation matrix allows for a simple deterministic risk aggregation approach which reasonably approximates the true aggregate risk measurement according to the complete multivariate risk distribution. Numerical examples demonstrate that our approach is a better basis for portfolio optimization than the Value-at-Risk implied tail-correlation matrix, especially if the calibration portfolio (or current portfolio) deviates from the optimal portfolio.

JEL classification: G22, G28, G32.

Keywords: Risk aggregation, Capital allocation, Portfolio optimization.

---

*This paper represents the authors' personal opinions, not necessarily those of their employers. A previous version of this paper was titled "Making the square-root formula compatible with capital allocation". We wish to thank Bruce Auchinleck, Roland Eisen, Gregor Fels, Helmut Gründl, Mike Hoy and Jianxi Su as well as the participants of the 23rd International Congress on Insurance: Mathematics and Economics, the 46th Annual Seminar of the European Group of Risk & Insurance Economists, the 10th CEQURA Conference on Advances in Financial and Insurance Risk Management, the 2021 Annual Meeting of the American Risk and Insurance Association and the research seminar at the Technische Universität Kaiserslautern for worthwhile discussions. All errors are ours.

†R+V Lebensversicherung AG, Raiffeisenplatz 2, 65389 Wiesbaden; email: joachim.paulusch@ruv.de

‡Mainz University of Applied Sciences, School of Business, Lucy-Hillebrand-Str. 2, 55128 Mainz, Germany; Fellow of the International Center for Insurance Regulation, Goethe University Frankfurt; email: sebastian.schluetter@hs-mainz.de

1 Introduction

Tail-correlation matrices offer an approach to aggregate risks in a simple deterministic manner. The correlation-based risk aggregation is employed in various contexts, including the calculation of regulatory capital requirements or a firm's economic capital. Regarding a n-risks-portfolio, the approach starts from n univariate risk measurements, which are collected in a vector $x \in \mathbb{R}^n$. Then, a $n \times n$-matrix R of correlation parameters is used to calculate the aggregate risk measurement as

$$\sqrt{x^{\mathrm{T}} R\, x} \tag{1}$$

The approach in line (1) is employed in the Solvency II standard formula, which is used to determine the regulatory capital requirement for most insurance companies in the European Union (EU). Apart from the EU, the approach (1) is used in insurance regulation in the United States ("Risk-Based Capital"), China ("C-ROSS") and the International Capital Standard. In the banking industry, the approach is referred to as the variance-covariance approach and is popular in banks' internal risk assessments (cf. Mathur, 2015, pp. 272-274; Li et al., 2015). Moreover, the approach can be used for investment portfolio optimization (cf. Mittnik 2014). Structurally, the calculation of the portfolio risk using (1) mimics the calculation of the standard deviation of portfolio risk. Hence, portfolio selection problems in connection with the risk measurement in (1) can be studied analogously to those of the mean-variance framework of Markowitz (1952).[1]

[1] Apart from investment portfolio optimization, the mean-variance framework has been employed in an insurance context. For example, Eckert & Gatzert (2018) identify an insurer's optimal risk-return combination against the background of policyholders' willingness to pay depending on the insurer's solvency level. Braun et al. (2017) investigate insurers' asset allocations in a mean-variance framework when they face a regulatory capital requirement determined by the Solvency II standard formula. Braun et al. (2017) find that the standard formula tends to promote inefficient portfolios over efficient ones.

Using correlation parameters derived from the covariance matrix, approach (1) can guarantee an exact aggregation of risk measurements only if risks follow a multivariate elliptical risk distribution (cf. McNeil et al. 2015, pp. 295 ff.). If risks exhibit heavy tails or non-linear dependencies,[2] the aggregate risk measurement based on (1) can substantially differ from the "true" result in accordance with the complete multivariate risk distribution (Pfeifer & Strassburger 2008, Li et al. 2015).[3] To eliminate this bias and in connection with the risk measure Value-at-Risk (VaR), so-called VaR-implied tail-correlations have been proposed (Campbell et al. 2002, Mittnik 2014). According to EIOPA (2014, p. 9), the risk aggregation in the Solvency II standard formula has been calibrated based on VaR-implied tail-correlations.

Chen et al. (2019) empirically study the sensitivities of the correlation-based risk aggregation approach with regard to the regulatory Risk-Based Capital (RBC) for US insurance companies, which is referred to as the "square-root formula" in this context.[4] The authors find that the insurers' optimal investment policy is driven by marginal capital requirements, i.e. by sensitivities of the aggregate capital requirement with respect to the size of univariate risks. Moreover, the authors demonstrate that the square-root formula has understated the marginal capital requirement of fixed-income investments and has thereby

[2]Empirical evidence indicates that correlations between asset returns are higher during periods of (stressful) downside moves, cf. for example Longin & Solnik (2001), Campbell et al. (2002) and Ang & Chen (2002). In addition, risk types such as operational risks or non-life insurance risks follow more heavily tailed distribution, see for example Bernard et al. (2018).

[3]In addition, Christiansen et al. (2012) estimate correlation coefficients between a life insurer's different types of biometric risks. The authors find that the correlation-based risk aggregation in connection with the estimated correlation coefficients leads to a conservative assessment of the aggregate risk. Breuer et al. (2010) study banks' summation of the regulatory capitals for market risks and credit risks—which can be viewed as a special case of (1) with R including only ones. The authors find that the summation is not necessarily conservative.

[4]More precisely, Chen et al. (2019) consider property and casualty insurance companies. For these insurers, the RBC includes six risk categories, such as stock risk, underwriting risk and reserving risk. Five of the risk categories are assumed to be uncorrelated and one of them (affiliated investments) is added up on the result of the square-root formula.

incentivized insurers to increase those investments. The insurers' overall risks have thus increased.

Our paper elaborates on the observations of Chen et al. (2019) in a stylized set-up. We demonstrate that the sensitivities of approach (1) can be substantially biased if R is a traditional tail-correlation matrix with ones on the diagonal, even if the calibration of R is conducted based on the complete multivariate risk distribution.

To make the correlation-based risk aggregation approach a suitable basis for portfolio management decisions, we propose taking a different view of matrix R. We show that for elliptical distributions, the entries of R globally coincide with the second-order partial derivatives of the squared aggregate risk measurement with respect to changes in the risk measurements of the univariate risks. For general distributions, these second-order partial derivatives uniquely define a symmetric matrix. We show that approach (1) in connection with this *"sensitivity-implied tail-correlation matrix"* approximates the true aggregate risk measurement in the sense of a second-order Taylor polynomial: hence, for the calibration portfolio, it yields the aggregate risk and all first and second-order sensitivities with respect to risk exposures in line with the respective results based on the true risk distribution. The deterministic risk aggregation approach (1) thereby accurately reflects diversification effects and how diversification changes when the portfolio is changed in a neighborhood of the calibration portfolio.

Our method locally approximates the true risk measurement at the calibration portfolio, but can misstate the risk of portfolios distant from the calibration portfolio. We demonstrate that the diagonal elements of the matrix R inform about the approximation error of approach (1) for stand-alone risks. For non-elliptical distributions, the diagonal

elements of the sensitivity-implied tail-correlation matrix can substantially deviate from one, and the matrix can hence cause a misstatement of stand-alone risks.

To analyze the implications of the calibration of the matrix R for portfolio optimization and business steering, we consider an example of a multiline-insurance company whose objective is the maximization of Economic Value Added (EVA) in connection with the risk measure 99.5% VaR. As a basis of comparison, we identify the "true" EVA-optimal strategy by calculating the risk measure based on the true multivariate risk distribution. Afterwards, we derive the EVA-maximizing strategy if the correlation-based risk aggregation approach is used in connection with either a traditional tail-correlation matrix or our proposed sensitivity-implied tail-correlation matrix. We find that the use of a traditional tail-correlation matrix induces a strategy which achieves a reduced EVA and goes along with a lower safety level than the true EVA-optimal strategy. In combination with the sensitivity-implied tail-correlation matrix, these distortions are very small, even if the calibration portfolio of the sensitivity-implied tail-correlation matrix clearly differs from the true EVA-optimal portfolio.

The remainder of this paper is structured as follows. Section 1 introduces the "sensitivity-implied tail-correlation" matrix and discusses its properties and calibration. Section 3 provides an overview of traditional tail-correlation matrices. Section 4 provides numerical examples including an analysis in terms of EVA optimization. Section 5 concludes and outlines possible areas of application of the sensitivity-implied tail-correlation matrix.

2 Introducing the sensitivity-implied tail-correlation matrix

2.1 Mathematical background

The mathematical core of the sensitivity-implied tail-correlation matrix is the observation that for any function $f(u)$ which is positive homogeneous of degree one and twice continuously differentiable, the second-order Taylor polynomial for $f^2(u)$ can be presented in a simple matrix form.[5]

Theorem 1. *Let $n \in \mathbb{N}$, $U \subseteq \mathbb{R}^n$ be an open and convex cone, and the function $f : U \to \mathbb{R}$ be twice continuously differentiable in a neighborhood of $u_0 \in U$. Assume that $f(u)$ is positive homogeneous of degree one on U, i.e.*

$$f(\lambda u) = \lambda \cdot f(u) \textit{ for all } \lambda > 0 \textit{ and all } u \in U$$

Then

$$f^2 : U \to \mathbb{R}, u \mapsto (f(u))^2$$

is positive homogeneous of degree two. Let $Df^2(u_0)$ and $D^2 f^2(u_0)$ denote the gradient and Hesse matrix of $f^2(u)$ with respect to u evaluated at u_0. Then, the second-order Taylor polynomial for f^2 at u_0,

$$P_{f^2}(u) = f^2(u_0) + (u - u_0)^{\mathrm{T}} \cdot Df^2(u_0) + \frac{1}{2}(u - u_0)^{\mathrm{T}} \cdot D^2 f^2(u_0) \cdot (u - u_0)$$

[5]Throughout the article, we call the argument of the function f u instead of x, which is merely for better presentation in line with the literature, specifically with Tasche (2008) and Buch et al. (2011).

can be rewritten as

$$P_{f^2}(u) = u^{\mathrm{T}} \cdot \frac{1}{2} D^2 f^2(u_0) \cdot u$$

and is positive homogeneous of degree two.

Let $f(u_0) > 0$. Thanks to Theorem 1, the original function $f(u)$ can be approximated at the point u_0 by the function

$$g(u) = \sqrt{P_{f^2}(u)} = \sqrt{u^{\mathrm{T}} \cdot \frac{1}{2} \mathrm{D}^2 f^2(u_0) \cdot u}$$

The approximation yields $g(u_0) = f(u_0)$, and all first and second-order derivatives of g and f coincide at u_0. Moreover, the function $g(u)$—like $f(u)$—is positive homogeneous of degree one. The latter property of $g(u)$ is in general not fulfilled by the straightforward second-order Taylor polynomial for $f(u)$.

The remainder $R_{f^2}(u) = f^2(u) - P_{f^2}(u)$ allows an assessment of the approximation error of $g(u)$ in terms of

$$f(u) - g(u) = \sqrt{R_{f^2}(u) + g^2(u)} - g(u) \tag{2}$$

provided $|R_{f^2}(u)| \leq g^2(u)$ which can be guaranteed in a neighborhood of u_0.[6] Proposition 1 shows that—by homogeneity—the analysis of $R_{f^2}(u)$ may be restricted to a unit sphere, i.e. an $(n-1)$-dimensional submanifold of U.

Proposition 1. *Under the definitions and assumptions of Theorem 1, the remainder $R_{f^2}(u) = f^2(u) - P_{f^2}(u)$ is positive homogeneous of degree two. Assume $n \geq 2$ and let*

$$\tilde{U} = \{u \in U \textit{ such that } \|u\| = 1\}$$

[6]For error estimation, it is common to consider the term $|f(u) - g(u)|$. Eq. (2) immediately implies that the absolute values of both sides of the equation coincide.

be the intersection of U and a unit sphere with respect to some norm $\|.\|$. Then for any $u \in U \setminus \{0\}$, we have

$$R_{f^2}(u) = \|u\|^2 \cdot R_{f^2}\left(\frac{u}{\|u\|}\right)$$

with $\frac{u}{\|u\|} \in \tilde{U}$.

As an alternative to analyzing $R_{f^2}(u)$ on a unit sphere, one could identify the set $U^* = \{u \in U \text{ such that } f(u) = f(u_0)\}$, which is an $(n-1)$-dimensional submanifold under suitable conditions on f. Each $u^* \in U^*$ allows an assessment of $f(\lambda u^*) = \lambda f(u^*) = \lambda f(u_0)$ for all $\lambda > 0$ and the remainder at λu^* is hence known. In total, the challenging task is either to identify the remainder (or thresholds of it) on a unit sphere or to identify the submanifold U^*.

2.2 Translation to a risk measurement context

Suppose the loss of a portfolio over a specified period of time is given by

$$u^{\mathrm{T}} X = \sum_{i=1}^{n} u_i \cdot X_i, \tag{3}$$

where $n \in \mathbb{N}$ denotes the number of relevant risks, $X = (X_1, ..., X_n)^{\mathrm{T}}$ is a random vector with $\mathbb{E}[X_i] < \infty$ for all $i \in \{1, ..., n\}$ and the vector $u = (u_1, ..., u_n)^{\mathrm{T}} \in \mathbb{R}^n$ reflects the exposures to each risk. Going forward, we assume that the multivariate distribution of X is fixed and that the variable u fully specifies the portfolio.[7] Moreover, in line with Tasche (2008), we assume that the X_i are scaled such that the coordinates $u = \mathbb{1}_n = (1, \ldots, 1)^{\mathrm{T}}$

[7] The assumption of a linear relationship between the portfolio return and the exposure vector u is popular in the related literature, cf. for example Gourieroux et al. (2000); Zanjani (2002), Tasche (2008), Buch et al. (2011), Mittnik (2014). An approach to generalize the relationship is presented by Boonen et al. (2017).

reflect the current portfolio. Following the notation in Tasche (2008), the function $f_{\varrho,X}$ measures the "true" aggregate risk of portfolio u,

$$\begin{aligned} f_{\varrho,X} &: U \to \mathbb{R}, \\ u = (u_1, \ldots, u_n)^{\mathrm{T}} &\mapsto f_{\varrho,X}(u) = \varrho\left(u^{\mathrm{T}} X\right), \end{aligned} \tag{4}$$

with ϱ being a risk measure which is positive homogeneous of degree one and $\mathbb{1}_n \in U \subseteq \mathbb{R}^n$.[8] Let $e_k \in \mathbb{R}^n$ denote a vector which takes the value one at the k-th position and zero elsewhere. Assuming that $e_k \in U$, let $x \in \mathbb{R}^n$ be the univariate risk measurements in accordance with ϱ, namely

$$x_k = f_{\varrho,X}(e_k) = \varrho(X_k), \quad k = 1, ..., n \tag{5}$$

The function g measures the aggregate risk of portfolio u based on the risk aggregation approach (1), depending on the risk measurement x and the matrix $R = (\varrho_{ij})_{i,j=1}^n \in \mathbb{R}^{n \times n}$:

$$g : u = (u_1, \ldots, u_n)^{\mathrm{T}} \mapsto g(u) = \sqrt{(u \circ x)^{\mathrm{T}} R\,(u \circ x)}, \tag{6}$$

[8] Following McNeil et al. (2015, p. 275 ff.), consider a probability space $(\Omega, \mathcal{F}, P)$, let $\mathcal{L}^0(\Omega, \mathcal{F}, P)$ be the set of all random variables on $(\Omega, \mathcal{F}, P)$ that are almost surely finite, and consider a linear space of random variables $\mathcal{M} \subset \mathcal{L}^0(\Omega, \mathcal{F}, P)$. A risk measure is a mapping $\varrho : \mathcal{M} \to \mathbb{R}$. Axioms on ϱ are defined as follows. Positive homogeneity: for $\lambda \geq 0, \varrho(\lambda L) = \lambda \varrho(L)$. Monotonicity: for $L_1 \leq L_2, \varrho(L_1) \leq \varrho(L_2)$. Translation invariance: for $m \in \mathbb{R}, \varrho(L + m) = \varrho(L) + m$. Subadditivity: for $L_1, L_2 \in \mathcal{M}, \varrho(L_1 + L_2) \leq \varrho(L_1) + \varrho(L_2)$. Convexity: for $0 \leq \gamma \leq 1, L_1, L_2 \in \mathcal{M}, \varrho(\gamma L_1 + (1-\gamma) L_2) \leq \gamma \varrho(L_1) + (1-\gamma)\varrho(L_2)$. Law invariance: if $L_1, L_2 \in \mathcal{M}$ have the same distribution functions, $\varrho(L_1) = \varrho(L_2)$. A risk measure satisfying the positive homogeneity, monotonicity, translation invariance and subadditivity axioms is called coherent.

where $\circ$ denotes the Hadamard product, i.e. $u \circ x = (u_1 x_1, \ldots, u_n x_n)^{\mathrm{T}} \in \mathbb{R}^n$. For the case of $f_{\varrho,X}(u)$ being twice continuously differentiable,[9] Proposition 2 states that $g(u)$ locally approximates $f_{\varrho,X}(u)$ if R is chosen based on second-order sensitivities of $f^2_{\varrho,X}(u)$.

Proposition 2. *Let ϱ be a risk measure which is positive homogeneous of degree one and let $X = (X_1, ..., X_n)^{\mathrm{T}}$ be a random vector. Let $U \subseteq \mathbb{R}^n$ be a convex cone with $\mathbb{1}_n, e_1, ..., e_n \in U$. Assume that the function $f_{\varrho,X}(u)$, as defined in formula (4), is twice continuously differentiable in a neighborhood of $\mathbb{1}_n$ and $f_{\varrho,X}(\mathbb{1}_n) > 0$. Let $x \in \mathbb{R}^n$ be defined as in (5) and assume that $x_k > 0$ for all $k = 1, \ldots, n$. Then, the matrix $R = (\varrho_{k\ell})_{k,\ell=1}^n$ defined by*

$$\varrho_{k\ell} = \frac{1}{2x_k x_\ell} \frac{\partial^2}{\partial u_k \partial u_\ell} f^2_{\varrho,X}(\mathbb{1}_n) \tag{7}$$

$$= \frac{1}{x_k x_\ell} \left(\frac{\partial f_{\varrho,X}(\mathbb{1}_n)}{\partial u_k} \frac{\partial f_{\varrho,X}(\mathbb{1}_n)}{\partial u_\ell} + f_{\varrho,X}(\mathbb{1}_n) \frac{\partial^2}{\partial u_k \partial u_\ell} f_{\varrho,X}(\mathbb{1}_n) \right) \tag{8}$$

is symmetric. In combination with this matrix R, let the function g be given as in (6). Then, $g(u)$ is defined in a neighborhood of $\mathbb{1}_n$. $g^2(u) = (g(u))^2$ is the second-order Taylor polynomial for $f^2_{\varrho,X}(u)$. Moreover, we have

$$g(\mathbb{1}_n) = f_{\varrho,X}(\mathbb{1}_n), \tag{9}$$

$$\frac{\partial}{\partial u_\ell} g(\mathbb{1}_n) = \frac{\partial}{\partial u_\ell} f_{\varrho,X}(\mathbb{1}_n), \quad 1 \le \ell \le n, \tag{10}$$

$$\frac{\partial^2}{\partial u_k \partial u_\ell} g(\mathbb{1}_n) = \frac{\partial^2}{\partial u_k \partial u_\ell} f_{\varrho,X}(\mathbb{1}_n), \quad 1 \le k, \ell \le n \tag{11}$$

We call the matrix R, whose entries are defined in (7), the "sensitivity-implied tail-correlation matrix". In connection with this matrix R, $g(u)$ approximates $f_{\varrho,X}(u)$ in

[9] Differentiability of $f_{\varrho,X}$ is commonly assumed in the context of Euler capital allocation, cf., for example, Tasche (2008). Moreover, using second-order derivatives can be essential in the context of portfolio optimization, cf. Buch et al. (2011).

the sense that it correctly determines the aggregate risk of the current portfolio, the sensitivities of the aggregate risk with respect to the exposures of all risks (starting at the current portfolio)[10] as well as the corresponding second-order sensitivities with respect to all combinations of risks.

The terms in line (8) have been studied in the literature. First-order derivatives of Value-at-Risk and Expected Shortfall (ES) can be viewed as expectations conditioned on the rare event that the aggregate portfolio loss coincides with the Value-at-Risk (or exceeds it in case of Expected Shortfall).[11] Analogously, the Hesse matrix of the Value-at-Risk can be presented in terms of the covariance matrix of risk drivers conditioned on the rare event (cf. Gourieroux et al. 2000, p. 229). Therefore, the sensitivity-implied correlation parameter $\rho_{k\ell}$ is driven by the interaction of risks conditioned on (tail) events that are relevant for the aggregate risk measurement. In general, $\rho_{k\ell}$ is not determined solely by the bivariate distribution of the random variables X_k and X_ℓ, but rather depends on the joint distribution of the whole random vector X.

In connection with the sensitivity-implied tail-correlation matrix, $g^2(u)$ is the second-order Taylor polynomial of $f^2_{\varrho,X}(u)$. Therefore, there are well-elaborated methods for estimating the Taylor remainder which can help to assess the error between $g(u)$ and $f_{\varrho,X}(u)$. Making use of Proposition 1, Appendix I demonstrates a possible procedure for $n = 2$ risks. For large n, the error estimation comes up against the challenge that a threshold for a large number of third-order derivatives of $f^2_{\varrho,X}(u)$ is needed.

[10] The first-order sensitivities are also known as the Euler capital allocation principle, cf., for example, Tasche (2008).

[11] This applies under quite general conditions concerning the multivariate risk distribution; cf. Targino et al. (2015, p. 209) for an overview.

2.3 Examples

A. Risk measure is standard deviation

Assume that risk is measured by the standard deviation:

$$f_{\varrho,X}(u) = \mathrm{sd}\left(\sum_{i=1}^{n} u_i \cdot X_i\right)$$

Let Σ and $R_P = \left(\varrho_{ij}^{(P)}\right)_{i,j=1}^{n}$ denote the variance matrix and Pearson correlation matrix of the random vector $X = (X_1, ..., X_n)^{\mathrm{T}}$, and let $x = \sigma = (\mathrm{sd}(X_1), ..., \mathrm{sd}(X_n))^{\mathrm{T}}$ denote the vector of univariate standard deviations. Assume that all these moments exist. We have

$$\begin{aligned} f_{\varrho,X}(u) &= \sqrt{u^{\mathrm{T}} \cdot \Sigma \cdot u} = \sqrt{(u \circ \sigma)^{\mathrm{T}} \cdot R_P \cdot (u \circ \sigma)} = \sqrt{(u \circ x)^{\mathrm{T}} \cdot R_P \cdot (u \circ x)} \\ &= \sqrt{\sum_{i=1}^{n}\sum_{j=1}^{n} \varrho_{ij}^{(P)} u_i x_i u_j x_j} \end{aligned} \tag{12}$$

and hence

$$f_{\varrho,X}^2(u) = \sum_{i=1}^{n}\sum_{j=1}^{n} \varrho_{ij}^{(P)} u_i x_i u_j x_j \tag{13}$$

Differentiating the left-hand and right-hand sides of Eq. (13) with respect to u_k, $k \in \{1, ..., n\}$, implies

$$\frac{\partial}{\partial u_k} f_{\varrho,X}^2(u) = 2x_k \sum_{j=1}^{n} \varrho_{kj}^{(P)} u_j x_j \tag{14}$$

Differentiating both sides of Eq. (14) again with respect to u_ℓ, $\ell \in \{1, ..., n\}$ implies

$$\begin{aligned} &\frac{\partial^2}{\partial u_k \partial u_\ell} f_{\varrho,X}^2(u) = 2x_k x_\ell \varrho_{k\ell}^{(P)} \\ \Rightarrow\ &\varrho_{k\ell}^{(P)} = \frac{1}{2x_k x_\ell} \frac{\partial^2}{\partial u_k \partial u_\ell} f_{\varrho,X}^2(u) \end{aligned} \tag{15}$$

Hence, in the outlined situation, all entries of the sensitivity-implied tail-correlation matrix coincide with those of the Pearson correlation matrix, irrespective of the multivariate risk distribution.

B. Multivariate elliptical distribution

Assume that the random vector $X = (X_1, ..., X_n)^{\mathrm{T}}$ follows an elliptical distribution with a finite mean vector and dispersion matrix $\Sigma = (\sigma_{ij})_{i,j=1}^{n}$. Moreover, assume that

$$f_{\varrho,X}(u) = \varrho\left(u^{\mathrm{T}}X\right) = \tilde{\varrho}\left(u^{\mathrm{T}}X\right) - \mathbb{E}\left[u^{\mathrm{T}}X\right] \tag{16}$$

with $\tilde{\varrho}$ being a positive homogeneous, translation-invariant and law-invariant risk measure. McNeil et al. (2015, pp. 295) show that $f_{\varrho,X}(u)$ can be presented analogous to line (12). If the covariance matrix of X exists, the sensitivity-implied tail-correlation matrix coincides with the Pearson correlation matrix. In general, the entries of the sensitivity-implied tail-correlation matrix are $\rho_{k\ell} = \sigma_{k\ell}/\sqrt{\sigma_{kk}\sigma_{\ell\ell}}$.[12]

C. Independent gamma distributions and multivariate mixed gamma distribution

We now outline situations in which the sensitivity-implied tail-correlation matrix can be calculated without Monte Carlo simulations, but instead in a way which is numerically less elaborate and not subject to sampling error.

Assume that the entries of the random vector $X = (X_1, ..., X_n)^{\mathrm{T}}$ are independent and gamma distributed with shape parameter $\gamma_i > 0$ and rate parameter $\vartheta_i > 0$. If all rate parameters are equal, $\vartheta_1 = ... = \vartheta_n =: \vartheta$, the aggregate loss $\sum_{i=1}^{n} X_i$ is gamma distributed with shape parameter $\gamma_1 + ... + \gamma_n$ and rate parameter ϑ. For the case that

[12] McNeil et al. (2015, p. 200-205) discuss estimating these quantities.

the rate parameters are not all the same, Moschopoulos (1985, p. 543) provides an analytical representation of the distribution function of X, which we denote by

$$F_X(x) = F_{\Gamma^+}(x; \gamma_1, ..., \gamma_n, \vartheta_1, ..., \vartheta_n) \tag{17}$$

For scalars $u_i > 0$, the product $u_i \cdot X_i$ is gamma distributed with shape parameter γ_i and rate parameter ϑ_i/u_i. Hence, the distribution function of $u^{\mathrm{T}}X = \sum_{i=1}^n u_i X_i$, with $u_i > 0$ for all i, is given by

$$F_{u^{\mathrm{T}}X}(x) = F_{\Gamma^+}(x; \gamma_1, ..., \gamma_n, \vartheta_1/u_1, ..., \vartheta_n/u_n) \tag{18}$$

If the risk measure ϱ is law-invariant, line (18) offers a starting point for calculating the sensitivity-implied tail-correlation matrix. Specifically, the Value-at-Risk of $u^{\mathrm{T}}X$,

$$\mathrm{VaR}_{1-\alpha}\left(u^{\mathrm{T}}X\right) = F^{-1}_{u^{\mathrm{T}}X}(1-\alpha), \tag{19}$$

can be calculated by inverting $F_{u^{\mathrm{T}}X}(x)$ from line (18) numerically using the Newton method. First and second-order derivatives of $\mathrm{VaR}_{1-\alpha}\left(\sum_{i=1}^n u_i X_i\right)$ with respect to scalars u_i can also be calculated numerically.

Furman et al. (2020) introduce the class of multivariate mixed gamma distributions. This class of distributions is dense in the class of all continuous distributions with non-negative support.[13] Hence, mixed gamma distributions are flexible in terms of the shape of the univariate distributions and the stochastic dependencies between them. Appendix D demonstrates that the distribution function of $u^{\mathrm{T}}X$ has an analytical representation if the

[13] More precisely, for any random vector in the class of continuous distributions with non-negative support, a sequence of mixed gamma distributed random vectors can be constructed which converges in distribution to the given random vector.

random vector X is mixed gamma distributed. Hence, the calculation of a law-invariant risk measure and its sensitivities is again possible without Monte-Carlo simulations.

2.4 Properties of the sensitivity-implied tail-correlation matrix

Apart from the situations in the examples A and B from section 2.3, the sensitivity-implied tail-correlation matrix does not necessarily satisfy the properties of the Pearson correlation matrix.

Firstly, the sensitivity-implied matrix is not always positive semi-definite (psd). If it is not psd, there are exposure vectors u such that $(u \circ x)^{\mathrm{T}} R (u \circ x)$ is negative, and $g(u)$ is hence not defined in real numbers. Apart from this issue, the missing psd'ness may result in optimization problems involving $g(u)$ in the target function not being convex. Section 4.3 shows a situation in which diversification based on the true multivariate risk distribution does not necessarily increase value. In this situation, the sensitivity-implied matrix is not psd and the optimization problem is not convex—neither on the basis of the function $f_{\varrho,X}(u)$ nor on the basis of $g(u)$. Proposition 3 shows that the psd'ness of the sensitivity-implied matrix can be guaranteed by the risk measure satisfying the convexity axiom as defined in footnote 8.

Proposition 3. *Under the assumptions and definitions in Proposition 2, the matrix R with entries defined in (7) is positive semi-definite if the risk measure ϱ satisfies the convexity axiom.*

Given that we generally assume the risk measure to be positive homogeneous, the convexity axiom is always fulfilled if the risk measure is coherent.[14] Hence, in connection

[14] Cf. McNeil et al. (2015, p. 276).

with Expected Shortfall,[15] the sensitivity-implied matrix is always psd. Consequently, the counter-example in section 4.3 is based on Value-at-Risk.

Secondly, in contrast to the Pearson correlation matrix, the sensitivity-implied tail-correlation matrix does not generally have ones on its diagonal. Appendix F provides an example based on a discrete distribution where the diagonal elements of the sensitivity-implied matrix can become arbitrarily large. Even negative diagonal entries are possible, as the example in section 4.3—building on the mixed gamma distribution—demonstrates. Proposition 4 states a sufficient condition for the diagonal elements being one.

Proposition 4. *Under the assumptions and definitions in Proposition 2, assuming that $f_{\varrho,X}(u)$ is twice continuously differentiable on U, the entries of the sensitivity-implied tail-correlation matrix satisfy*

$$\rho_{kk} = \frac{\frac{\partial^2}{(\partial u_k)^2} f^2_{\varrho,X}(\mathbb{1}_n)}{\frac{\partial^2}{(\partial u_k)^2} f^2_{\varrho,X}(e_k)} \tag{20}$$

Hence, we have $\rho_{kk} = 1$ for all $k \in \{1, ..., n\}$ if $f^2_{\varrho,X}(u)$ is quadratic in u on some open and convex subspace $\tilde{U} \subseteq \mathbb{R}^n$ containing $\mathbb{1}_n$ and e_k for all $k = 1, ..., n$.

The next Proposition shows that the distance between the diagonal elements and one informs about the relative error between $g(u)$ and $f_{\varrho,X}(u)$ for stand-alone risks.

Proposition 5. *Based on the definitions in lines (4) and (6), assuming that $f_{\varrho,X}(e_k) > 0$, we have*

$$\frac{f^2_{\varrho,X}(e_k) - g^2(e_k)}{f^2_{\varrho,X}(e_k)} = 1 - \rho_{kk} \tag{21}$$

[15] Expected Shortfall is known as a coherent risk measure; see, for example, McNeil et al. (2015, p. 283).

Assuming $\rho_{kk} \geq 0$, this translates into

$$\frac{f_{\varrho,X}(e_k) - g(e_k)}{f_{\varrho,X}(e_k)} = 1 - \sqrt{\rho_{kk}} \tag{22}$$

Proposition 5 is not restricted to the sensitivity-implied tail-correlation matrix, as it holds for any matrix R underlying the function $g(u)$. Hence, the traditional notion of the diagonal elements of a tail-correlation matrix being one refers to $g(u)$ accurately reflecting stand-alone risks. Based on the sensitivity-implied matrix, $g(u)$ can reflect stand-alone risks with an error that can even become arbitrarily large (cf. Appendix F).

2.5 Estimation

Estimating the sensitivity-implied tail-correlation matrix in situations other than the examples in section 2.3 can be based on kernel estimation. The literature proposes consistent estimators for the first-order derivatives of Value-at-Risk (Tasche 2009, p. 584) and Expected Shortfall (Scaillet 2004, p. 118 f.). Similarly, Gourieroux et al. (2000) derive consistent estimates for a portfolio's Value-at-Risk and its second-order derivatives, which the authors apply for daily stock return data. At least for Value-at-Risk, therefore, all items in line (8) can be consistently estimated. Slutsky's theorem (cf. Casella & Berger 2002, p. 239 f.) implies that the composition of consistent estimators in terms of line (8) provides a consistent estimator for $\varrho_{k\ell}$. Appendix J applies the previously mentioned kernel estimators to the example specified in section 4.2.

The literature proposes a wide range of variance reduction techniques in the context of estimating expectations conditioned on rare events (i.e. first-order derivatives) from Monte-Carlo simulations. Several papers estimate marginal risk contributions for credit

portfolios which are subject to systematic risk factors by employing Importance Sampling (Glasserman 2005), kernel estimation in combination with Importance Sampling (Tasche 2009), or the Fast Fourier Transform technique for risk aggregation (cf. Siller 2013). Targino et al. (2015) employ Sequential Monte Carlo simulation to estimate the Euler capital allocation of a portfolio with stochastic dependencies being modeled by a copula. To efficiently estimate the sensitivity-implied tail-correlation matrix, the above-mentioned kernel estimators use the conditioned covariance, i.e. the expectation of products of random variables. To the best of our knowledge, an algorithm designed to address this problem is not immediately available and is left for future research.[16]

3 Traditional tail-correlation matrices

Traditionally, a tail-correlation matrix is assumed to be a symmetric matrix with ones on its diagonal. For $n = 2$ risks and risk measurements $x_1 > 0, x_2 > 0$, there is only one free correlation parameter $\varrho_{1,2}$, which is to be set such that approach (1) correctly determines the aggregate risk of the current portfolio (cf. Campbell et al. 2002, p. 89):

$$\begin{aligned} & g\left(\mathbb{1}_2\right) = f_{\varrho,X}\left(\mathbb{1}_2\right) \\ \Rightarrow \quad & x_1^2 + 2\varrho_{1,2}x_1x_2 + x_2^2 = \left(f_{\varrho,X}\left(\mathbb{1}_2\right)\right)^2 \\ \Rightarrow \quad & \varrho_{1,2} = \frac{\left(f_{\varrho,X}\left(\mathbb{1}_2\right)\right)^2 - x_1^2 - x_2^2}{2x_1x_2} \end{aligned} \tag{23}$$

For $n \geq 3$ risks, Mittnik (2014, p. 70 f.) proposes defining the matrix R based on a set of ℓ benchmarking portfolio weight vectors $w_1, ..., w_\ell \in \mathbb{R}^n$, assuming that R is symmetric

[16] To enhance the efficiency of kernel estimation, Epperlein & Smillie (2006, p. 71) and Tasche (2009, p. 584) use an adjustment to ensure that the Euler allocation adds up to $f_{\varrho,X}\left(\mathbb{1}_n\right)$. In our context, an additional restriction about the row sums of the Hesse matrix of $f^2_{\varrho,X}(u_0)$ can be used, cf. Eq. (A.3).

with ones on the diagonal. The $\frac{n\cdot(n-1)}{2}$ free correlation parameters in R can be determined in light of the errors

$$f_{\varrho,X}^2\left(w_k\right) - g^2\left(w_k\right) \tag{24}$$

for the benchmark portfolios w_k, $k = 1, ..., \ell$. Mittnik (2014, p. 70 f.) shows that the term in (24) is linear in the correlation parameters ϱ_{ij}. For $\ell = n(n-1)/2$, the correlation parameters can be determined such that the risk assessments of $g(u)$ and $f_{\varrho,X}(u)$ coincide for all benchmark portfolios and the identification is called "exact". For $\ell > n(n-1)/2$, there are more benchmark portfolios than correlation parameters in $g(u)$. In general, it is then not possible to choose correlation parameters bringing the errors in (24) to zero for all $k = 1, ..., \ell$, but the parameters are described by an overdetermined equation system.

Mittnik (2014, p. 71) proposes identifying the correlation parameters by a least-squares estimator minimizing the mean squared error (MSE)

$$\frac{1}{\ell}\sum_{k=1}^{\ell}\left(f_{\varrho,X}^2\left(w_k\right) - g^2\left(w_k\right)\right)^2. \tag{25}$$

Going forward, we call the calibration with $\ell > n(n-1)/2$ calibration portfolios "least-squares". The matrix R calibrated as explained so far in this section is called the "*benchmark-implied tail-correlation matrix*".[17]

[17]For two reasons, we depart from the usual term "*VaR-implied tail-correlation matrix*". Firstly, the definitions in lines (23) and (25) are compatible with risk measures other than Value-at-Risk (e.g. with Expected Shortfall). Secondly, contrasting the benchmark-implied tail-correlation matrix with the sensitivity-implied tail-correlation matrix makes it clearer how the two concepts differ: the first matrix is induced by the risk measurements of a finite set of benchmarking portfolio weight vectors, whereas the second is induced by sensitivities of the risk measurement at a single calibration portfolio.

Devineau & Loisel (2009, section 5) define R as the "minimal standard R" which solves the optimization problem

$$\begin{aligned} \|R\| &\to \min \\ \text{subject to } f(\mathbb{1}_n) &= g(\mathbb{1}_n), \end{aligned} \tag{26}$$

with the norm $\|\cdot\|$ being defined as $\|D\| = \sqrt{\text{trace}(D \cdot D^{\mathrm{T}})}$. Devineau & Loisel (2009) employ the approach only for $n = 2$ risks. For this case, the authors state that the problem in (26) is solved by (23).[18] For $n \geq 3$ risks, the matrix R calibrated according to (26) does not reflect which of the risks are more or less interdependent, since the calibration is only based on the diversified risk measurement and the stand-alone risk measurements. Our numerical examples in section 4 will illustrate this issue. More generally, our examples will demonstrate that the function $g(u)$ in connection with a traditional tail-correlation matrix can misstate the slope and curvature of $f_{\varrho,X}(u)$, and that these misstatements can induce severe distortions in portfolio optimization.

4 Numerical examples

4.1 Set-up

We consider an insurance company with n lines of business (lobs). The scalars $u_1, ..., u_n$ represent the volume of lob i in terms of the number of insurance contracts. We assume that the u_i are scaled, for example, in 100,000 contracts such that we may disregard the integer restriction. Moreover, we assume that the diversification within each lob does not

[18] It thereby becomes clear that Devineau & Loisel (2009) restrict R to have ones on the diagonal.

vary in u_i such that the claims costs of lob i are modeled by $u_i \cdot X_i$. In line with Solvency II regulations, risk is measured by the 99.5% Value-at-Risk of unexpected losses.

The connections between the volume u_i and the premium p_i of each lob $i \in \{1, ..., n\}$ are determined by an isoelastic demand function,[19]

$$u_i(p_i) = n_i \cdot p_i^{\epsilon_i}, \tag{27}$$

where $n_i > 0$ calibrates demand to market size and $\epsilon_i < -1$ is the price elasticity of demand which is constant in p_i. We consider a representative insurer whose objective is to maximize its economic value added (EVA).[20] In our model, the insurer's EVA is the expected profit minus the cost of capital, which is modeled by a hurdle rate r_h times the 99.5% Value-at-Risk of the aggregate risk. In our baseline calibration, we set $\epsilon_i = -9$ for all lobs i,[21] and $r_h = 5\%$.[22]

On the one hand, we consider the EVA in connection with the risk measurement based on the true multivariate risk distribution:

$$\begin{aligned} \mathrm{EVA}_{\mathrm{true}}(u) &= \sum_{i=1}^{n} u_i \cdot (p_i(u_i) - \mathbb{E}[X_i]) - r_h \cdot f_{\varrho,X}(u) \\ &= \sum_{i=1}^{n} u_i \cdot (p_i(u_i) - \mathbb{E}[X_i]) - r_h \cdot \mathrm{VaR}_{99.5\%} \left(\sum_{i=1}^{n} u_i \cdot (X_i - \mathbb{E}[X_i]) \right) \end{aligned} \tag{28}$$

[19] To simplify the notation, p_i is also scaled. If u_i are specified per 100,000 contracts, p_i is the premium income per 100,000 contracts.

[20] The objective is analogous to the analysis of Chen et al. (2019). It can be justified by assuming that the insurer jointly decides on its level of equity capital and on the volumes $u_1, ..., u_n$; the regulatory capital requirement is based on VaR and binding. The objective can easily be modified to a situation in which the insurer sticks to a fixed capital requirement ratio (in terms of equity capital over capital requirement). In the context of Solvency II regulations, this ratio is relevant when insurers transmit information about their solvency level, cf. Gatzert & Heidinger (2020). In related analyses, the Economic Value Added has been employed by Stoughton & Zechner (2007) and Diers (2011).

[21] According to the empirical results of Yow & Sherris (2008, p. 318), this may reflect the price elasticity of compulsory third party or motor insurance.

[22] Zanjani (2002, p. 297) estimates that the discounted cost of holding capital is 5% in commercial automobile insurance.

with $p_i(u_i)$ denoting the inverse of the demand function in Eq. (27). We call the portfolio u, which maximizes $\mathrm{EVA}_{\mathrm{true}}(u)$, the "true optimal portfolio". On the other hand, we identify which portfolio u an insurer chooses if the risk measurement is conducted in connection with a tail-correlation matrix R, i.e. the portfolio maximizing

$$\begin{aligned}\mathrm{EVA}_R(u) &= \sum_{i=1}^{n} u_i \cdot (p_i(u_i) - \mathbb{E}[X_i]) - r_h \cdot g(u) \\ &= \sum_{i=1}^{n} u_i \cdot (p_i(u_i) - \mathbb{E}[X_i]) - r_h \cdot \sqrt{(u \circ x)^{\mathrm{T}} R (u \circ x)} \end{aligned} \tag{29}$$

The chosen set-up allows us to distinguish the distortions caused by the function $g(u)$ in terms of EVA and in terms of the insurer's safety level. The insurer's safety level is measured by the true VaR confidence level which corresponds to $g(u)$, i.e. the solution $\tilde{\alpha}$ of

$$\mathrm{VaR}_{1-\tilde{\alpha}}\left(\sum_{i=1}^{n} u_i \cdot (X_i - \mathbb{E}[X_i])\right) = g(u) \tag{30}$$

4.2 Relevance of first-order sensitivities

This section demonstrates that an inappropriate calibration of the matrix R can lead to biased first-order sensitivities of the aggregate risk measurement and can induce a suboptimal portfolio.

We model the basic losses of $n = 5$ lobs using stochastically independent and gamma distributed random variables $\tilde{X}_1, ..., \tilde{X}_5$.[23] Specifically, we assume that $\tilde{X}_1$, $\tilde{X}_2 \sim \Gamma(\frac{1}{3}, \frac{2}{3})$, $\tilde{X}_3, \tilde{X}_4 \sim \Gamma(2, 2)$ and $\tilde{X}_5 \sim \Gamma(1, 2)$, where $\Gamma(\gamma, \vartheta)$ denotes the gamma distribution with shape parameter γ and rate parameter ϑ. In addition, lobs 1, 2 and 5 are exposed to

[23] The calculations of $f_{\varrho,X}(u)$ and its sensitivities are conducted according to section 2.3, example C, and are hence unaffected by sampling error. For calculating $F_{u^{\mathrm{T}}X}(x)$ in (18), we make use of the R package `coga`, which implements the methodology of Hu et al. (2020).

a common risk factor $Y \sim \Gamma(1,1)$, which is independent from the $\tilde{X}_i$. The total claims costs of the three lobs are $X_i = \tilde{X}_i + 0.5Y$ for $i \in \{1,2,5\}$ and $X_j = \tilde{X}_j$ for $j \in \{3,4\}$.[24]

The vector of stand-alone capital requirements is

$$x = (4.679, 4.679, 2.715, 2.715, 2.715)^{\mathrm{T}} \tag{31}$$

and the Pearson correlation matrix is

$$R_P = \begin{pmatrix} 1 & 0.25 & 0 & 0 & 0.354 \\ 0.25 & 1 & 0 & 0 & 0.354 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0.354 & 0.354 & 0 & 0 & 1 \end{pmatrix} \tag{32}$$

The insurer's current portfolio is $u = \mathbb{1}_5 = (1,1,1,1,1)^{\mathrm{T}}$ and the corresponding aggregate risk measurement—based on the true multivariate risk distribution—is

$$f_{\varrho,X}(\mathbb{1}_5) = \mathrm{VaR}_{0.995}(X_1 + ... + X_5) - \mathbb{E}[X_1 + ... + X_5] = 8.115 \tag{33}$$

[24]The expected claims costs of all three lobs are 1, e.g. we calculate $\mathbb{E}[X_1] = \mathbb{E}[\tilde{X}_1] + 0.5\mathbb{E}[Y] = \left(\frac{1}{3}\right) / \left(\frac{2}{3}\right) + 0.5 \cdot 1/1 = 1$. Moreover, the variances are $\mathrm{var}[X_1] = \mathrm{var}[X_2] = \mathrm{var}[\tilde{X}_1] + 0.5^2\mathrm{var}[Y] = \left(\frac{1}{3}\right) / \left(\frac{2}{3}\right)^2 + 0.5^2 \cdot 1/1^2 = 1$ and $\mathrm{var}[X_3] = \mathrm{var}[X_4] = \mathrm{var}[X_5] = 0.5$. The risks of lobs 1 and 2 have a relatively heavy tail, the risks of lobs 3, 4 and 5 have a relatively light tail. The ratios of the 99.5% VaR and the 90% VaR are 3.89 for lobs 1 and 2 and 2.87 for lobs 3, 4 and 5. For comparison, Bernard et al. (2018, p. 847) assume the distribution $200 \cdot \mathrm{LogNormal}(0,1)$ for aggregate non-life insurance risks. This implies a corresponding VaR ratio of 5.88. For the aggregate market risk, Bernard et al. (2018, p. 847) assume a normal distribution, for which the corresponding VaR ratio is 2.01. This value is achieved by the gamma distribution when setting the shape parameter to infinity.

The true sensitivities of the aggregate risk measurement—i.e. the Euler allocation—are obtained as

$$Df_{\varrho,X}(u)\Big|_{u=\mathbb{1}_5} = (2.830, 2.830, 0.416, 0.416, 1.623)^{\mathrm{T}} \tag{34}$$

According to line (34), the first two lobs have the strongest impact on the insurer's aggregate risk. The fifth lob follows next—due to its positive correlation with the risks of the first lobs. The third and fourth lobs are less influential due to their independence from the other risks.

We calibrate the sensitivity-implied tail-correlation matrix as defined in Proposition 2 with the calibration portfolio being the insurer's current portfolio $u = \mathbb{1}_5$ by numerical differentiation as discussed in section 2.3, example C. In Appendix J, moreover, we estimate the matrix from Monte-Carlo simulations. To calibrate the benchmark-implied tail-correlation matrix in line with Mittnik (2014), we need to choose a set of $\ell \geq \frac{5\cdot(5-1)}{2} = 10$ calibration portfolios. We consider three of those sets. Firstly, we conduct a *pairwise* calibration based on all equally weighted two-risk portfolios. Hence, we set $w_1 = (1,1,0,0,0)^{\mathrm{T}}$, $w_2 = (1,0,1,0,0)^{\mathrm{T}}$ ... $w_{10} = (0,0,0,1,1)^{\mathrm{T}}$. Secondly, we consider an *exact* calibration: We take $w_1, ..., w_9$ as before and set $w_{10} = (1,1,1,1,1)^{\mathrm{T}}$. Thirdly, we use the *least-squares* estimator minimizing (25) in connection with all 26 portfolios consisting of two, three, four or five assets. Finally, we calculate the "minimal standard" tail-correlation matrix as proposed by Devineau & Loisel (2009) for $u = \mathbb{1}_5$. Table 1 presents all five calculated tail-correlation matrices. The upper part of Figure 4 in Appendix K visualizes the relative error between $g(u)$—in connection with three different tail-correlation matrices—and $f_{\varrho,X}(u)$ depending on the exposure vector u. In connection with the sensitivity-implied tail-correlation matrix, $g(u)$ is relatively accurate as long as u is close at $\mathbb{1}_5$ and underestimates the true risk by about -6% for

Table 1: Calculated tail-correlation matrices.

Sensitivity-implied

	Lob 1	Lob 2	Lob 3	Lob 4	Lob 5
Lob 1	0.999	-0.035	-0.049	-0.049	0.243
Lob 2	-0.035	0.999	-0.049	-0.049	0.243
Lob 3	-0.049	-0.049	0.649	0.013	-0.036
Lob 4	-0.049	-0.049	0.013	0.649	-0.036
Lob 5	0.243	0.243	-0.036	-0.036	1.021

Benchmark-implied

	Pairwise (Mittnik 2014)					Exact (Mittnik 2014)				
	Lob 1	Lob 2	Lob 3	Lob 4	Lob 5	Lob 1	Lob 2	Lob 3	Lob 4	Lob 5
Lob 1	1.000	-0.071	-0.174	-0.174	0.110	1.000	-0.071	-0.174	-0.174	0.110
Lob 2	-0.071	1.000	-0.174	-0.174	0.110	-0.071	1.000	-0.174	-0.174	0.110
Lob 3	-0.174	-0.174	1.000	-0.174	-0.174	-0.174	-0.174	1.000	-0.174	-0.174
Lob 4	-0.174	-0.174	-0.174	1.000	-0.174	-0.174	-0.174	-0.174	1.000	1.376
Lob 5	0.110	0.110	-0.174	-0.174	1.000	0.110	0.110	-0.174	1.376	1.000

	Least-squares (Mittnik 2014)					Minimum standard (Devineau & Loisel 2009)				
	Lob 1	Lob 2	Lob 3	Lob 4	Lob 5	Lob 1	Lob 2	Lob 3	Lob 4	Lob 5
Lob 1	1.000	-0.029	-0.131	-0.131	0.163	1.000	-0.0004	-0.0002	-0.0002	-0.0002
Lob 2	-0.029	1.000	-0.131	-0.131	0.163	-0.0004	1.0000	-0.0002	-0.0002	-0.0002
Lob 3	-0.131	-0.131	1.000	-0.120	-0.123	-0.0002	-0.0002	1.0000	-0.0001	-0.0001
Lob 4	-0.131	-0.131	-0.120	1.000	-0.123	-0.0002	-0.0002	-0.0001	1.0000	-0.0001
Lob 5	0.163	0.163	-0.123	-0.123	1.000	-0.0002	-0.0002	-0.0001	-0.0001	1.0000

$u = (0.25, 0.25, 1.75, 1.75, 1)^{\mathrm{T}}$, which is in the top left corner of the considered plane of exposures. In connection with a pairwise calibrated tail-correlation matrix, $g(u)$ is specifically biased at $u = (0.57, 0.57, 1.11, 1.11, 1)^{\mathrm{T}}$ with a relative error of -22%; with the least-squares estimator, $g(u)$ has an error of -10.5% at $u = (0.44, 0.44, 0.67, 0.67, 1)^{\mathrm{T}}$.

Table 2 reports the aggregate risk of the current portfolio, $u = \mathbb{1}_5$, and the Euler allocations in connection with all considered tail-correlation matrices. The results for the traditional tail-correlation matrices clearly depend on the type of calibration. Moreover, based on a benchmark-implied matrix, the aggregate risk of the current portfolio can be underestimated by 9% (least-squares) or even 19% (pairwise calibration).[25] In connection with all considered calibrations, the traditional tail-correlation matrices lead to substantially biased Euler allocations. For instance, the risk of lob 4 can be underestimated by

[25] As explained in section 3, the benchmark-implied tail-correlation matrix leads to a correct risk assessment of the current portfolio only if the number of benchmark portfolios coincides with the number of correlation parameters and the current portfolio is one of the benchmark portfolios.

Table 2: Chosen portfolios of the model insurer. The risk measurement is conducted based on the true multivariate risk measurement, the sensitivity-implied tail-correlation matrix or a traditional tail-correlation matrix. The bracket terms show the relative error compared to the true risk measurement.

Type of calculation	True distribution	Sensitivity-implied	Benchmark-implied Pairwise	Exact	Least-squares	Min. Std.
Aggregate risk measurement	8.115	8.115 (±0%)	6.557 (−19%)	8.115 (±0%)	7.360 (−9%)	8.115 (±0%)
Euler allocation						
Lob 1	2.830	2.830 (±0%)	2.641 (−7%)	2.134 (−25%)	2.716 (−4%)	2.696 (−5%)
Lob 2	2.830	2.830 (±0%)	2.641 (−7%)	2.134 (−25%)	2.716 (−4%)	2.696 (−5%)
Lob 3	0.416	0.416 (±0%)	0.059 (−86%)	0.048 (−88%)	0.306 (−27%)	0.908 (+118%)
Lob 4	0.416	0.416 (±0%)	0.059 (−86%)	1.456 (+250%)	0.306 (−27%)	0.908 (+118%)
Lob 5	1.623	1.623 (±0%)	1.157 (−29%)	2.343 (+44%)	1.318 (−19%)	0.908 (−44%)
Chosen portfolio						
Lob 1	1.000	1.000	1.102	1.156	1.052	1.041
Lob 2	1.000	1.000	1.102	1.156	1.052	1.041
Lob 3	1.000	1.000	1.157	1.133	1.054	0.861
Lob 4	1.000	1.000	1.157	0.784	1.054	0.861
Lob 5	1.000	1.000	1.190	0.876	1.115	1.247
True EVA	0.676	0.676 (±0.0%)	0.670 (−0.9%)	0.666 (−1.5%)	0.674 (−0.2%)	0.668 (−1.1%)
True VaR conf. level	0.50%	0.50%	1.30%	0.65%	0.79%	0.60%

86% (pairwise calibration) or overestimated by 250% (exact calibration). In contrast, the use of the sensitivity-implied tail-correlation matrix leads to an accurate measurement of the aggregate risk and Euler allocations.

In terms of the EVA analysis, we set the demand function parameters to $n_1 = n_2 = 9.497$, $n_3 = n_4 = 3.474$ and $n_5 = 5.826$. These parameter values imply that the true EVA-optimal strategy is $u = \mathbb{1}_5$. The same strategy maximizes the EVA in line (29) if R is the sensitivity-implied tail-correlation matrix. However, the distorted risk measurement based on the traditional tail-correlation matrices lead the insurer away from the truly optimal strategy. For example, in the case of a pairwise calibration, the insurer chooses $u_{\text{new}} = (1.102, 1.102, 1.157, 1.157, 1.190)^{\mathrm{T}}$. Based on the true multivariate risk distribu-

Table 3: Chosen portfolios of 500 randomly parameterized insurers. The risk measurement is conducted based on the true multivariate risk measurement, the sensitivity-implied tail-correlation matrix or a traditional tail-correlation matrix. The table reports the means across the 500 insurers as well as the 5% and 95% percentiles.

Type of calculation	True distribution	Sensitivity-implied	Benchmark-implied Pairwise	Exact	Least-squares	Min. Std.
RMSE between chosen portfolio and "true" optimal portfolio						
Mean	0.000	0.004	0.140	0.154	0.067	0.137
$p_{5\%}$	0.000	0.001	0.105	0.102	0.052	0.103
$p_{95\%}$	0.000	0.010	0.178	0.205	0.084	0.173
Relative loss in EVA of chosen portfolio versus "true" optimal portfolio						
Mean	0.000%	0.002%	0.867%	1.445%	0.216%	1.022%
$p_{5\%}$	0.000%	0.000%	0.340%	0.366%	0.090%	0.385%
$p_{95\%}$	0.000%	0.008%	1.559%	3.107%	0.381%	1.784%
True VaR confidence level						
Mean	0.500%	0.500%	1.253%	0.653%	0.772%	0.587%
$p_{5\%}$	0.500%	0.498%	1.051%	0.572%	0.712%	0.503%
$p_{95\%}$	0.500%	0.503%	1.453%	0.728%	0.827%	0.686%

tion, the aggregate risk of the chosen portfolio is $f_{\varrho,X}(u_{\text{new}}) = 9.135$. The benchmark-implied correlation matrix in connection with a pairwise calibration, however, understates the risk of this portfolio by $g(u_{\text{new}})/f_{\varrho,X}(u_{\text{new}}) - 1 = 7.339/9.135 - 1 = -19.7\%$. The true VaR-confidence level of the chosen strategy is clearly too high and amounts to 1.3%. In addition, the true EVA of the chosen portfolio, $\text{EVA}_{\text{true}}(u_{\text{new}})$, is 0.9% lower than the maximal EVA, i.e. $\text{EVA}_{\text{true}}(\mathbb{1}_5)$.

In the analyses so far, the calibrations of the tail-correlation matrices were centered at the true optimal portfolio, $u = \mathbb{1}_5$. Next, we study how insurance companies with different properties—and different true optimal portfolios—choose their portfolios if their risk measurement is based on the tail-correlation matrices calibrated at $\mathbb{1}_5$. We modify eleven parameters—namely the values of the demand function parameters, $n_1, ..., n_5$, $\epsilon_1, ..., \epsilon_5$, as well as the hurdle rate r_h—by multiplying them with scalars. To this end, we randomly choose eleven scalars as independent realizations of uniform random variables on the interval 0.6 to 1.4. This process is executed 500 times to generate 500 heterogeneous insurance companies. The true optimal portfolios of the 500 insurers differ from $u = \mathbb{1}_5$

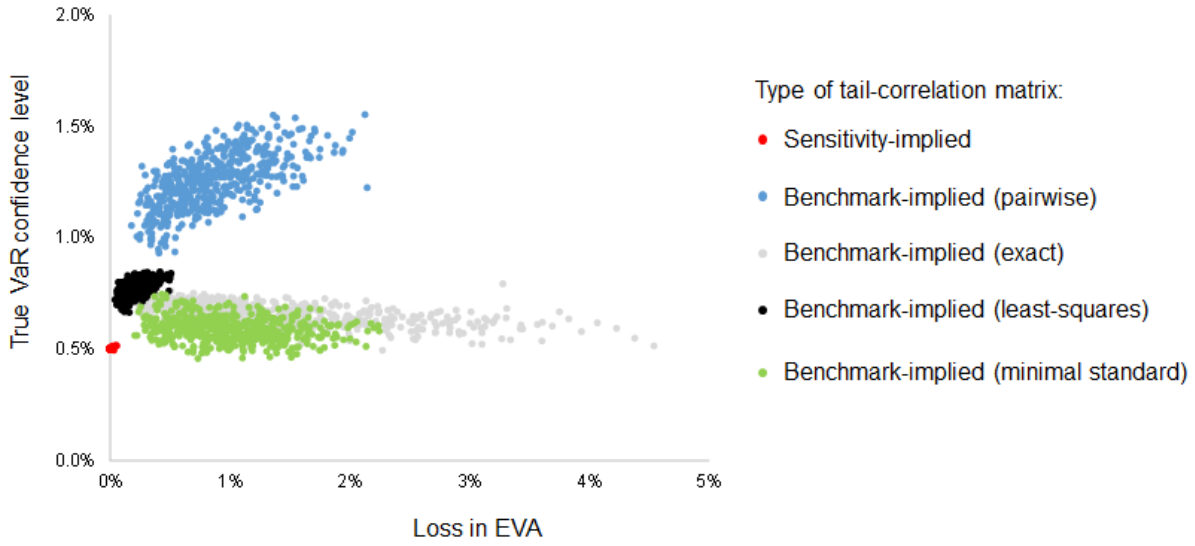

Figure 1: Loss in EVA and true VaR confidence level for 500 randomly parameterized insurers. The risk measurement is conducted based on the sensitivity-implied tail-correlation matrix or a traditional tail-correlation matrix. The results show that portfolio optimization in connection with the sensitivity-implied tail-correlation matrix hardly induces distortions in terms of the VaR confidence level or the achieved EVA.

with a root mean squared error (RMSE) of 0.260 on average across the 500 insurers. The risk of the true optimal portfolios is not measured accurately by any of the tail-correlation matrices. Nevertheless, the sensitivity-implied matrix guides insurers to a portfolio which achieves almost the same EVA as the true optimal portfolio (cf. Table 3: for 95% of insurers, the relative loss in EVA does not exceed 0.01%). In contrast, measuring risk based on a traditional tail-correlation matrix goes along with a considerable loss in EVA. For example, the "exact" calibration of the benchmark-implied matrix leads to a relative loss in EVA of 1.445% on average and of 3.107% or higher for those 5% of insurers with the highest loss. Moreover, when using the sensitivity-implied matrix, the true VaR confidence level of insurers' aggregate risk is close to the supposed value of 0.5%. With respect to traditional tail-correlation matrices, each of the four calibration methods guides more than 95% of insurers to strategies with a true VaR confidence level above 0.5%. Figure 1 visualizes the loss in EVA and the true VaR confidence level of the

portfolios that the 500 insurers choose based on the tail-correlation matrices. The figure depicts that the sensitivity-implied matrix leads to substantially smaller distortions than any of the traditional tail-correlation matrices.

4.3 Relevance of second-order sensitivities

This section studies the implications of biased second-order sensitivities of the aggregate risk measurement based on a traditional tail-correlation matrix with ones on the diagonal. The three lobs' claims costs, X_1, X_2 and X_3, now follow a mixed gamma distribution with the parameters defined in Table 4.[26] With a large weight in terms of p_κ, the distribution

Table 4: Parameters of the mixed gamma distribution.

i	ϑ_i	γ_{k_1}	γ_{k_2}	γ_{k_3}
1	0.5	0.5	9.5	0.5
2	0.5	0.5	0.5	9.5
3	0.5	0.5	4.5	4.5
p_κ		0.99	0.005	0.005

consists of $n = 3$ independent and identically distributed risks. However, conditioning on a high aggregate loss, the risks X_1 and X_2 are negatively correlated. In this set-up, the marginal capital requirement of X_1 decreases when increasing the exposure to X_1. Hence, the Hesse matrix of $f_{\varrho,X}$ with respect to u has negative entries on the diagonal:

$$D^2 f_{\varrho,X}(u) = \begin{pmatrix} -3.850 & 3.632 & 0.218 \\ 3.632 & -3.850 & 0.218 \\ 0.218 & 0.218 & -0.437 \end{pmatrix}$$

[26] Appendix D provides more details about this distribution.

The aggregate Value-at-Risk can be reduced by shifting the exposures from $u = (1, 1, 1)^{\mathrm{T}}$ to $u = (1 + h, 1 - h, 1)^{\mathrm{T}}$ for a small value of h.[27] We embed this distribution into the EVA-optimization problem as studied in section 4.2. By setting $n_1 = n_2 = 128.082$ and $n_3 = 90.209$, all first-order derivatives of the function $\mathrm{EVA}_{\mathrm{true}}(u)$ are zero at $u = \mathbb{1}_3$. The Hesse matrix of $\mathrm{EVA}_{\mathrm{true}}(u)$ is indefinite at $u = \mathbb{1}_3$, reflecting the fact that it is a saddle point, as illustrated on the left side of Figure 2. To keep the example graphically fully tractable, we assume from now on that $u_3 = 1$ is fixed and only u_1 and u_2 are decision variables. The function $\mathrm{EVA}_{\mathrm{true}}(u_1, u_2 | u_3 = 1)$ then has a global maximum at $(u_1, u_2) = (1.8365, 0.5998)$ and, due to symmetry, another global maximum at $(0.5998, 1.8365)$; cf. points B and B' in Figure 2.

We calibrate two tail-correlation matrices, R_1 and R_2, both with the calibration portfolio $u = \mathbb{1}_3$. R_1 is the sensitivity-implied tail-correlation matrix. R_2 is calibrated such that the function $g(u)$ reflects the true first-order sensitivities:

$$Dg\left(\mathbb{1}_3\right) = Df\left(\mathbb{1}_3\right) \tag{35}$$

In accordance with the traditional view of tail-correlation matrices, we restrict R_2 to be a symmetric matrix with ones on its diagonal. The three correlation parameters in R_2 are thereby uniquely defined by (35).[28]

The right side of Figure 2 depicts the EVA function in connection with the matrix R_2. By construction, all first-order derivatives of this EVA function are zero at $u = \mathbb{1}_3$. However, the Hesse matrix of this EVA function is negative definite, and this function hence has

[27] This can be seen by approximating $f_{\varrho,X}(u)$ by a Taylor polynomial of degree 2 and noting that $\partial/\partial u_1 f(u) = \partial/\partial u_2 f(u)$ at $u = (1, 1, 1)^{\mathrm{T}}$.

[28] Specifically, the three correlation parameters are the solution of a linear equation system. As shown by Eq. (L.2) in Appendix L, $Dg\left(\mathbb{1}_3\right)$ is linear in the correlation parameters when fixing $\sqrt{x^{\mathrm{T}} R_2\, x}$ at $f_{\varrho,X}\left(\mathbb{1}_3\right)$.

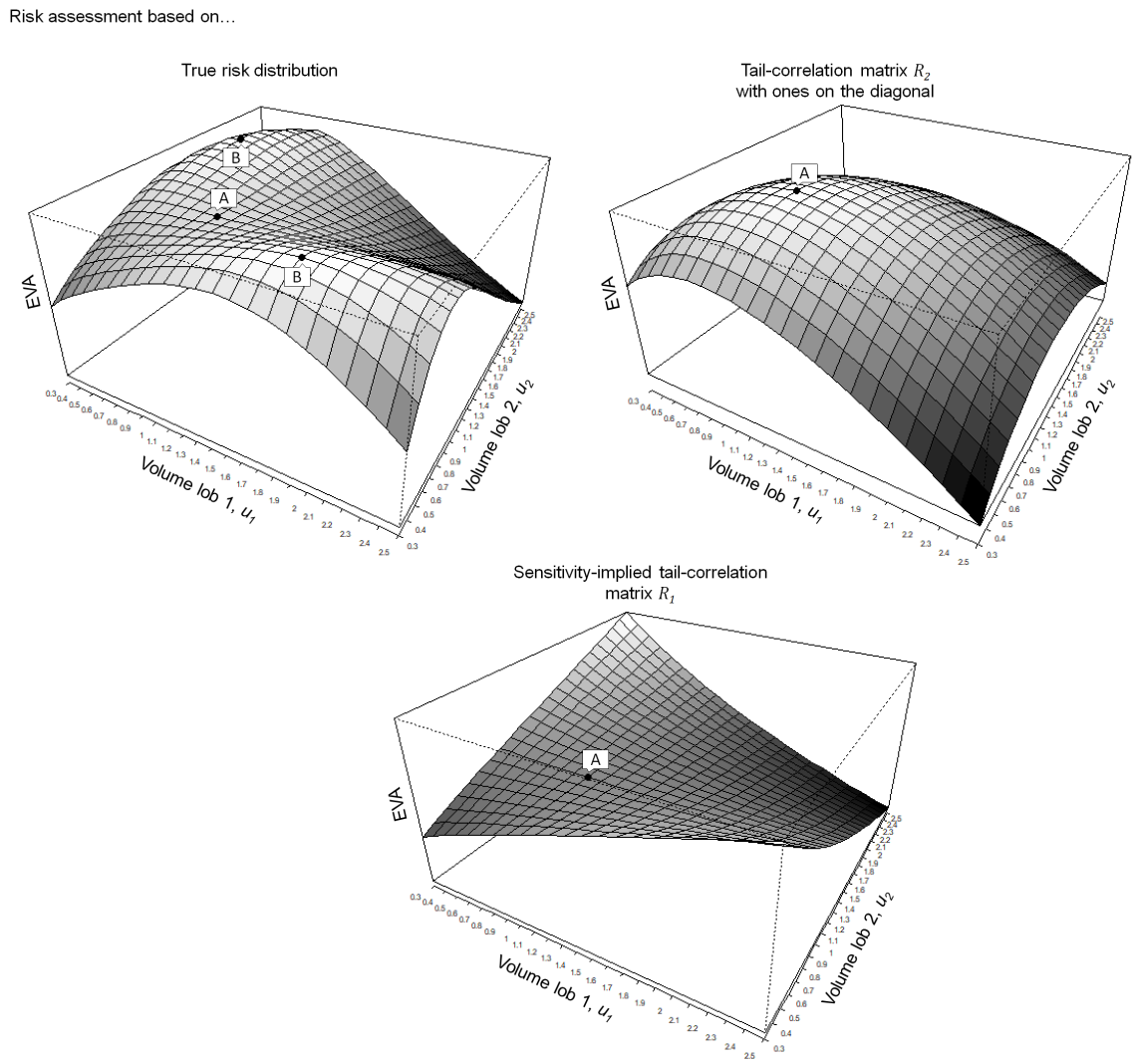

Figure 2: EVA based on volumes u_1 and u_2 and for fixed $u_3 = 1$. Point A reflects $u_1 = u_2 = 1$; points B and B' are optimal based on the true risk distribution.

a global maximum at $u = (1, 1, 1)^{\mathrm{T}}$—in contrast to the true EVA function. The lower part of Figure 2 shows the EVA function in connection with the sensitivity-implied tail-correlation matrix R_1. This EVA function correctly approximates the true EVA function at the calibration portfolio $u = \mathbb{1}_3$, and the company hence does not mistake the saddle point for an optimum.

Finally, as previously for the example in section 4.2, the lower part of Figure 4 in Appendix K visualizes the relative error between $g(u)$ and $f_{\varrho,X}(u)$ for the sensitivity-implied and two benchmark-implied tail-correlation matrices. The sensitivity-implied

tail-correlation matrix again induces the largest errors in the corners of the considered plane of exposures (i.e. -12% for $u = (0.25, 1.75, 1)^{\mathrm{T}}$). In connection with a pairwise or least-squares calibration, $g(u)$ has the largest displayed errors at $u = (0.76, 0.76, 1)^{\mathrm{T}}$, amounting to -12% (pairwise) and -7.5% (least-squares).

5 Conclusion

This paper demonstrates that the traditional notion of (tail-) correlation matrices having ones on their diagonal can make it impossible to fit them in accordance with the true risk distribution. Those misstatements can distort portfolio management decisions in terms of risk and return. We show that the square-root formula for risk aggregation is structurally related to a second-order Taylor polynomial of a positive homogeneous risk measure. Based on this result, we propose so-called "sensitivity-implied" tail-correlation matrices, which approximate the risk measurement according to the true distribution up to second-order derivatives with respect to exposures. We see several areas of application.

In the context of regulation, the proposed method may help to circumvent moral hazard effects arising from misstated marginal capital requirements, as empirically detected by Chen et al. (2019). Our example in section 4.2 indicates that the matrix R would not have to be calibrated for every insurer individually. Instead, the sensitivity-implied tail-correlation leads to a relatively stable risk measurement and steering signals even if the insurers' optimal portfolios differ from the calibration portfolio of R.

In the context of a firm's internal economic capital assessment, the use of the correlation-based risk aggregation is sometimes called the "hybrid approach" (Hull 2018, p. 594). In comparison with a risk aggregation based on a Monte-Carlo simulation, the correlation-

based approach facilitates the risk measurement process, since changes in the univariate risk assessments do not require new simulations of the entire firm. Once the matrix R has been calibrated, various methods (including scenario analyses, expert surveys, etc.) can be used for the measurement of the univariate risks.

Finally, the proposed tail-correlation matrix can be helpful for portfolio optimization in general when risk is to be measured using a positive homogeneous risk measure. For instance, optimization problems with a VaR or ES constraint thereby become structurally identical to mean-variance portfolio optimization. However, given that the sensitivity-implied matrix is locally calibrated, the solution of the simplified problem may only be useful if the optimal portfolio is not too far away from the calibration portfolio.

Appendix

A Proof of Theorem 1

Given that the function $f(u)$ is positive homogeneous of degree one, $f^2(u)$ is positive homogeneous of degree two: for all $\lambda > 0$ we have

$$f^2(\lambda u) = \lambda^2 f^2(u) \tag{A.1}$$

and Euler's theorem for homogeneous functions implies

$$\frac{1}{2} u^{\mathrm{T}} \cdot \mathrm{D} f^2(u) = f^2(u) \tag{A.2}$$

Differentiating both sides of Eq. (A.1) with respect to u implies that $\mathrm{D} f^2(u)$ is positive homogeneous of degree one: for all $\lambda > 0$ we have

$$\begin{aligned} & \mathrm{D} f^2(\lambda u) \cdot \lambda = \lambda^2 \mathrm{D} f^2(u) \\ \Rightarrow \quad & \mathrm{D} f^2(\lambda u) = \lambda \mathrm{D} f^2(u) \end{aligned}$$

Euler's theorem for homogeneous functions thus implies that

$$u^{\mathrm{T}} \cdot \mathrm{D}^2 f^2(u) = \mathrm{D} f^2(u) \tag{A.3}$$

The second-order Taylor polynomial of $f^2(u)$ can be presented as

$$
\begin{aligned}
P_{f^2}(u) &= f^2(u_0) + (u-u_0)^{\mathrm{T}} \cdot \mathrm{D}f^2(u_0) + \frac{1}{2}(u-u_0)^{\mathrm{T}} \cdot \mathrm{D}^2 f^2(u_0) \cdot (u-u_0) \\
&\overset{(A.2)}{=} \frac{1}{2}u_0^{\mathrm{T}} \cdot \mathrm{D}f^2(u_0) + (u-u_0)^{\mathrm{T}} \cdot \mathrm{D}f^2(u_0) + \frac{1}{2}(u-u_0)^{\mathrm{T}} \cdot \mathrm{D}^2 f^2(u_0) \cdot (u-u_0) \\
&\overset{(A.3)}{=} \frac{1}{2}u_0^{\mathrm{T}} \cdot \mathrm{D}^2 f^2(u_0) \cdot u_0 + (u-u_0)^{\mathrm{T}} \cdot \mathrm{D}^2 f^2(u_0) \cdot u_0 \\
&\quad + \frac{1}{2}(u-u_0)^{\mathrm{T}} \cdot \mathrm{D}^2 f^2(u_0) \cdot (u-u_0) \\
&= \frac{1}{2}u_0^{\mathrm{T}} \cdot \mathrm{D}^2 f^2(u_0) \cdot u_0 + (u-u_0)^{\mathrm{T}} \cdot \mathrm{D}^2 f^2(u_0) \cdot u_0 \\
&\quad + \frac{1}{2}u^{\mathrm{T}} \cdot \mathrm{D}^2 f^2(u_0) \cdot u - u^{\mathrm{T}} \cdot \mathrm{D}^2 f^2(u_0) \cdot u_0 + \frac{1}{2}u_0^{\mathrm{T}} \cdot \mathrm{D}^2 f^2(u_0) \cdot u_0 \\
&= \frac{1}{2}u^{\mathrm{T}} \cdot \mathrm{D}^2 f^2(u_0) \cdot u
\end{aligned}
$$

Finally, for all $\lambda > 0$ we have $P_{f^2}(\lambda u) = \frac{1}{2}(\lambda u)^{\mathrm{T}} \cdot \mathrm{D}^2 f^2(u_0) \cdot (\lambda u) = \lambda^2 P_{f^2}(u)$.

B Proof of Proposition 1

For all $\lambda > 0$ we have $R_{f^2}(\lambda x) = f^2(\lambda x) - P_{f^2}(\lambda x) = \lambda^2(f^2(x) - P_{f^2}(x)) = \lambda^2 R_{f^2}(x)$. For any $u \in U \setminus \{0\}$ this implies

$$
R_{f^2}(u) = R_{f^2}\left(\|u\| \cdot \frac{u}{\|u\|}\right) = \|u\|^2 \cdot R_{f^2}\left(\frac{u}{\|u\|}\right)
$$

C Proof of Proposition 2

Function $f_{\varrho,X}(u)$ with $u_0 = \mathbb{1}_n$ fulfills the assumptions of Theorem 1. Hence, $P_{f^2_{\varrho,X}}(u) = u^{\mathrm{T}} \cdot \frac{1}{2}\mathrm{D}^2 f^2_{\varrho,X}(\mathbb{1}_n) \cdot u$ is the second-order Taylor polynomial of $f^2_{\varrho,X}(u)$. We have

$$\begin{aligned} g^2(u) &\overset{(6)}{=} (u \circ x)^{\mathrm{T}} R\,(u \circ x) = \sum_{i,j=1}^{n} \rho_{i,j} u_i x_i u_j x_j \overset{(7)}{=} \sum_{i,j=1}^{n} \frac{1}{2x_i x_j} \frac{\partial^2}{\partial u_i \partial u_j} f^2_{\varrho,X}(\mathbb{1}_n) u_i x_i u_j x_j \\ &= \sum_{i,j=1}^{n} \frac{1}{2} \frac{\partial^2}{\partial u_i \partial u_j} f^2_{\varrho,X}(\mathbb{1}_n) u_i u_j = P_{f^2_{\varrho,X}}(u) \end{aligned}$$

Hence, the assertions in lines (9), (10) and (11) follow.

D Distribution of portfolio loss for mixed gamma distributed risks

According to Furman et al. (2020, p. 8 f.), the n-dimensional mixed gamma distribution is defined as follows: let $\boldsymbol{\kappa} = (\kappa_1, ..., \kappa_n)$ be a vector of discrete random variables which can assume non-negative integer values, and let $p_{\boldsymbol{\kappa}}(\boldsymbol{k}) = \mathbb{P}(\kappa_1 = k_1, ..., \kappa_n = k_n)$ denote the probability mass function of $\boldsymbol{\kappa}$ with $\boldsymbol{k} = (k_1, ..., k_n) \in \mathbb{N}_0^n$. Let $f_\Gamma(x; \gamma, \vartheta)$ denote the density function of the univariate gamma distribution with shape parameter γ and rate parameter ϑ. The random vector $\boldsymbol{\Gamma}^{(\boldsymbol{\kappa})} = (\Gamma_1^{(\kappa_1)}, ..., \Gamma_n^{(\kappa_n)})$ is distributed n-variate mixed gamma if its density function is given by

$$f_{\boldsymbol{\Gamma}^{(\boldsymbol{\kappa})}}(x_1, ..., x_n) = \sum_{\boldsymbol{k} \in \mathbb{N}_0^n} p_{\boldsymbol{\kappa}}(\boldsymbol{k}) \prod_{i=1}^{n} f_\Gamma(x_i; \gamma_{k_i}, \vartheta_i) \tag{D.1}$$

where the shape parameters are determined by $\gamma_{k_i} = \gamma_i + k_i$ with $\gamma_i > 0$. Recall that $F_{\Gamma^+}(x; \gamma_1, ..., \gamma_n, \vartheta_1, ..., \vartheta_n)$ in (17) is the distribution function of the sum of independent

gamma distributed random variables with different shape and rate parameters. Assuming that there is only a finite number of vectors $\boldsymbol{k}$ with positive probability $p_{\boldsymbol{\kappa}}(\boldsymbol{k})$, the distribution function of $X = \Gamma_1^{(\kappa_1)} + ... + \Gamma_n^{(\kappa_1)}$ is given by

$$\begin{aligned} F_X(x) &= \int \cdots \int_{\{y \in \mathbb{R}_+^n \text{ such that } \sum_{i=1}^n y_i \leq x\}} \sum_{\boldsymbol{k} \in \mathbb{N}_0^n} p_{\boldsymbol{\kappa}}(\boldsymbol{k}) \prod_{i=1}^n f_\Gamma(y_i; \gamma_{k_i}, \vartheta_i) \mathrm{d}y_1 .... \mathrm{d}y_n \\ &= \sum_{\boldsymbol{k} \in \mathbb{N}_0^n} p_{\boldsymbol{\kappa}}(\boldsymbol{k}) \int \cdots \int_{\{y \in \mathbb{R}_+^n \text{ such that } \sum_{i=1}^n y_i \leq x\}} \prod_{i=1}^n f_\Gamma(y_i; \gamma_{k_i}, \vartheta_i) \mathrm{d}y_1 .... \mathrm{d}y_n \\ &= \sum_{\boldsymbol{k} \in \mathbb{N}_0^n} p_{\boldsymbol{\kappa}}(\boldsymbol{k}) F_{\Gamma^+}(x; \gamma_{k_1}, ..., \gamma_{k_n}, \vartheta_1, ..., \vartheta_n), \end{aligned}$$

and the distribution function of the linear combination $u_1 \Gamma_1^{(\kappa_1)} + ... + u_n \Gamma_n^{(\kappa_1)}$ is given by

$$\sum_{\boldsymbol{k} \in \mathbb{N}_0^n} p_{\boldsymbol{\kappa}}(\boldsymbol{k}) F_{\Gamma^+}(x; \gamma_{k_1}, ..., \gamma_{k_n}, \vartheta_1/u_1, ..., \vartheta_n/u_n)$$

E Proof of Proposition 3

We show that $f_{\varrho,X}$ is convex on U. Let $u_1, u_2 \in U$ and $\gamma \in [0,1]$. Then, by positive homogeneity and convexity of ϱ,

$$\begin{aligned} f_{\varrho,X}(\gamma u_1 + (1-\gamma) u_2) &= \varrho\left(\gamma u_1^{\mathrm{T}} X + (1-\gamma) u_2^{\mathrm{T}} X\right) \\ &\leq \gamma \varrho\left(u_1^{\mathrm{T}} X\right) + (1-\gamma) \varrho\left(u_2^{\mathrm{T}} X\right) = \gamma f_{\varrho,X}(u_1) + (1-\gamma) f_{\varrho,X}(u_2) \end{aligned}$$

Due to $f_{\varrho,X}$ being continuous with $f_{\varrho,X}(\mathbb{1}_n) > 0$, there is some $\widetilde{U}$ open with $\mathbb{1}_n \in \widetilde{U} \subseteq U$ and $f(u) > 0$ for all $u \in \widetilde{U}$. Given that $h : [0,\infty) \to [0,\infty), x \mapsto x^2$ is non-decreasing and convex, the composition $h(f_{\varrho,X}(u)) = f_{\varrho,X}^2(u)$ is convex on $\widetilde{U}$. Hence, the Hesse matrix of

$f^2_{\varrho,X}(u)$ at $u = \mathbb{1}_n$, $D^2 f^2_{\varrho,X}(\mathbb{1}_n)$, is positive semidefinite. Moreover, the sensitivity-implied tail-correlation matrix R is positive semidefinite: for any $v \in \mathbb{R}^n$ we have

$$v^{\mathrm{T}} R v \overset{(7)}{=} \sum_{i,j=1}^{n} \frac{1}{2x_i x_j} \frac{\partial^2}{\partial u_i \partial u_j} f^2_{\varrho,X}(\mathbb{1}_n) v_i v_j = \frac{1}{2}(v \circ \bar{x})^{\mathrm{T}} D^2 f^2_{\varrho,X}(\mathbb{1}_n)(v \circ \bar{x}) \geq 0,$$

with the entries of the vector $\bar{x}$ being defined as $\bar{x}_i = x_i^{-1}$ for all $i = 1, ..., n$.

F Example: No upper threshold for diagonal elements of sensitivity-implied matrix

Consider two independent random variables defining losses X_1, X_2 with $0 < c \leq 1$,

$$P(X_1 = c) = 96\%, \quad P(X_1 = 1) = 4\%, \quad P(X_2 = 2c) = 96\%, \quad P(X_2 = 2) = 4\%,$$

and the risk measure $\varrho = \mathrm{VaR}_{95\%}$. Then

$$x_1 = \varrho(X_1) = c, \quad x_2 = \varrho(X_2) = 2c,$$

$$f_{\varrho,X}(u) = \min\{cu_1 + 2u_2, u_1 + 2cu_2\} = \begin{cases} u_1 + 2cu_2, & u_1 \leq 2u_2 \\ cu_1 + 2u_2, & u_1 > 2u_2 \end{cases}$$

Hence, for $u_1 \leq 2u_2$, we have

$$D f^2_{\varrho,X}(u) = 2(u_1 + 2cu_2) \begin{pmatrix} 1 \\ 2c \end{pmatrix}, \quad D^2 f^2_{\varrho,X}(u) = \begin{pmatrix} 2 & 4c \\ 4c & 8c^2 \end{pmatrix},$$

such that the sensitivity-implied tail-correlation matrix at $(u_1, u_2)^{\mathrm{T}} = (1,1)^{\mathrm{T}}$ according to (7) is

$$R = \begin{pmatrix} \frac{1}{c^2} & \frac{1}{c} \\ \frac{1}{c} & 1 \end{pmatrix}$$

As c can be small, the first diagonal element of R can be arbitrarily large. Note that $g(u) = u_1 + 2cu_2$ coincides with $f_{\varrho,X}(u)$ for $u_1 \leq 2u_2$. For $c = 1$, $f^2_{\varrho,X}(u)$ is quadratic, and the diagonal elements of R are hence 1 (cf. Proposition 4).

G Proof of Proposition 4

Given that $f^2_{\varrho,X}(u)$ is positive homogeneous of degree two, Theorem 1 implies for all $u \in U$

$$u^{\mathrm{T}} \cdot \frac{1}{2} D^2 f^2_{\varrho,X}(u) \cdot u = f^2_{\varrho,X}(u) \tag{G.1}$$

We have

$$\rho_{kk} \overset{(7)}{=} \frac{1}{2} \frac{\frac{\partial^2}{(\partial u_k)^2} f^2_{\varrho,X}(\mathbb{1}_n)}{f^2_{\varrho,X}(e_k)} \overset{(G.1)}{=} \frac{1}{2} \frac{\frac{\partial^2}{(\partial u_k)^2} f^2_{\varrho,X}(\mathbb{1}_n)}{e_k^{\mathrm{T}} \cdot \frac{1}{2} D^2 f^2_{\varrho,X}(e_k) \cdot e_k} = \frac{\frac{\partial^2}{(\partial u_k)^2} f^2_{\varrho,X}(\mathbb{1}_n)}{\frac{\partial^2}{(\partial u_k)^2} f^2_{\varrho,X}(e_k)}$$

If $f^2_{\varrho,X}(u)$ is quadratic on $\tilde{U}$ containing $\mathbb{1}_n$ and e_k for all $k = 1, ..., n$, then $\frac{\partial^2}{(\partial u_k)^2} f^2_{\varrho,X}(u)$ is constant in u on $\tilde{U}$ and hence $\rho_{kk} = 1$ for all $k = 1, ..., n$.

H Proof of Proposition 5

According to the definitions in lines (5) and (6) and with the assumption $f_{\varrho,X}(e_k) > 0$, we derive

$$\frac{g^2(e_k)}{f^2_{\varrho,X}(e_k)} = \frac{x_k^2 \; \varrho_{kk}}{x_k^2} = \varrho_{kk}$$

Assuming $\rho_{kk} \geq 0$ we can take the square-root on the left-hand side and right-hand side of the equation. The assertions of Proposition 5 immediately follow.

I Example: Assessing the error between $g(u)$ and $f_{\varrho,X}(u)$ with $n = 2$ risks

Suppose we want to estimate $f_{\varrho,X}(u) - g(u)$ for $n = 2$ risks and some u which is entrywise non-negative. In light of Proposition 1 in connection with the norm $\|u\| = 0.5 \cdot \sum_{i=1}^{n} |u_i|$, we consider $u = \mathbb{1}_2 + \alpha \cdot (1, -1)^{\mathrm{T}}$ with $\alpha \in [-1, 1]$. We define

$$\tilde{f}^2 : [-1, 1] \to \mathbb{R}, \alpha \mapsto f^2_{\varrho,X}\left(\mathbb{1}_2 + \alpha \cdot (1, -1)^{\mathrm{T}}\right)$$

$$\tilde{g}^2 : [-1, 1] \to \mathbb{R}, \alpha \mapsto g^2\left(\mathbb{1}_2 + \alpha \cdot (1, -1)^{\mathrm{T}}\right)$$

$$R_{\tilde{f}^2} : [-1, 1] \to \mathbb{R}, \alpha \mapsto \tilde{f}^2(\alpha) - \tilde{g}^2(\alpha)$$

Due to the mean value theorem, the remainder can be presented based on the third order derivative of $\tilde{f}^2(\alpha)$, i.e.

$$R_{\tilde{f}^2}(\alpha) = (\tilde{f}^2)^{(3)}(\bar{\alpha}) \frac{\alpha^3}{3!} \tag{I.1}$$

for some $\bar{\alpha}$ between 0 and α.[29] Therefore, a lower and upper threshold for $R_{\tilde{f}^2}(\alpha)$ are given by

$$\min\left\{(\tilde{f}^2)^{(3)}(\bar{\alpha}) \text{ with } \bar{\alpha} \text{ between 0 and } \alpha\right\} \cdot \frac{\alpha^3}{3!} \tag{I.2}$$

and

$$\max\left\{(\tilde{f}^2)^{(3)}(\bar{\alpha}) \text{ with } \bar{\alpha} \text{ between 0 and } \alpha\right\} \cdot \frac{\alpha^3}{3!} \tag{I.3}$$

[29] Cf. Folland (2001, p. 88).

Moreover, at $\alpha = 1$ and $\alpha = -1$, the mean value theorem implies that there are some $0 < \bar{\alpha}_1 < 1$ and $-1 < \bar{\alpha}_2 < 0$ such that

$$\begin{aligned}
(\tilde{f}^2)^{(3)}(\bar{\alpha}_1)\frac{1^3}{3!} &= R_{\tilde{f}^2}(1) = (f^2_{\varrho,X}(2\cdot e_1) - g^2(2\cdot e_1)) \overset{\text{Eq. (21)}}{=} 4\cdot(1-\rho_{11})\cdot x_1^2 \\
(\tilde{f}^2)^{(3)}(\bar{\alpha}_2)\frac{(-1)^3}{3!} &= R_{\tilde{f}^2}(-1) = (f^2_{\varrho,X}(2\cdot e_2) - g^2(2\cdot e_2)) \overset{\text{Eq. (21)}}{=} 4\cdot(1-\rho_{22})\cdot x_2^2
\end{aligned}$$

If α is close to 1 or -1, we approximate $\bar{\alpha}$ from Eq. (I.1) by $\bar{\alpha}_1$, or $\bar{\alpha}_2$ respectively, and thus have

$$R_{\tilde{f}^2}(\alpha) \approx \begin{cases} (\tilde{f}^2)^{(3)}(\bar{\alpha}_1)\frac{\alpha^3}{3!} = 4\cdot(1-\rho_{11})\cdot x_1^2\cdot\alpha^3 & \text{if } \alpha > 0 \\ (\tilde{f}^2)^{(3)}(\bar{\alpha}_2)\frac{\alpha^3}{3!} = -4\cdot(1-\rho_{22})\cdot x_2^2\cdot\alpha^3 & \text{if } \alpha < 0 \end{cases} \tag{I.4}$$

As an example, consider $n = 2$ independent Gamma distributed risks with shape and rate parameters $\gamma_1 = \vartheta_1 = 1$ for risk 1 and $\gamma_2 = \vartheta_2 = 2$ for risk 2. We consider $f_{\varrho,X}(u) = \mathrm{VaR}_{99.5\%}\left(u^{\mathrm{T}}X\right) - \mathbb{E}\left(u^{\mathrm{T}}X\right)$. We have evaluated the remainder $R_{\tilde{f}^2}$ for $\alpha \in [-1, 1]$. Hence, we have evaluated the error between $g^2(u)$ and $f^2_{\varrho,X}(u)$ (as well as the corresponding thresholds) for all weighted sums of the portfolios $u_1 = (2, 0)^{\mathrm{T}}$ and $u_2 = (0, 2)^{\mathrm{T}}$. Afterwards, we have translated the error and thresholds into the error between $g(u)$ and $f_{\varrho,X}(u)$ according to Eq. (2). Figure 3 depicts the latter error (red curve) and estimates of it (gray and black curves) as a percentage of $f_{\varrho,X}(u)$. The black curve in Figure 3 shows the error estimate according to line (I.4). While this estimate does not require the calculation of any third-order derivatives, it is not necessarily accurate or conservative if α is in the middle between -1 and 0 or between 0 and 1. The gray

curves in Figure 3 reflect the thresholds according to (I.2) and (I.3). We have calculated third-order derivatives numerically by

$$(\tilde{f}^2)^{(3)}(\alpha) \approx \frac{(\tilde{f}^2)'(\alpha - \Delta\alpha) - 2 \cdot (\tilde{f}^2)'(\alpha) + (\tilde{f}^2)'(\alpha + \Delta\alpha)}{(\Delta\alpha)^2}$$

with $\Delta\alpha = 0.1$. As shown in Figure 3, the thresholds from (I.2) and (I.3) are conservative, but not necessarily close at $R_{\tilde{f}^2}(\alpha)$ if α is away from 0.

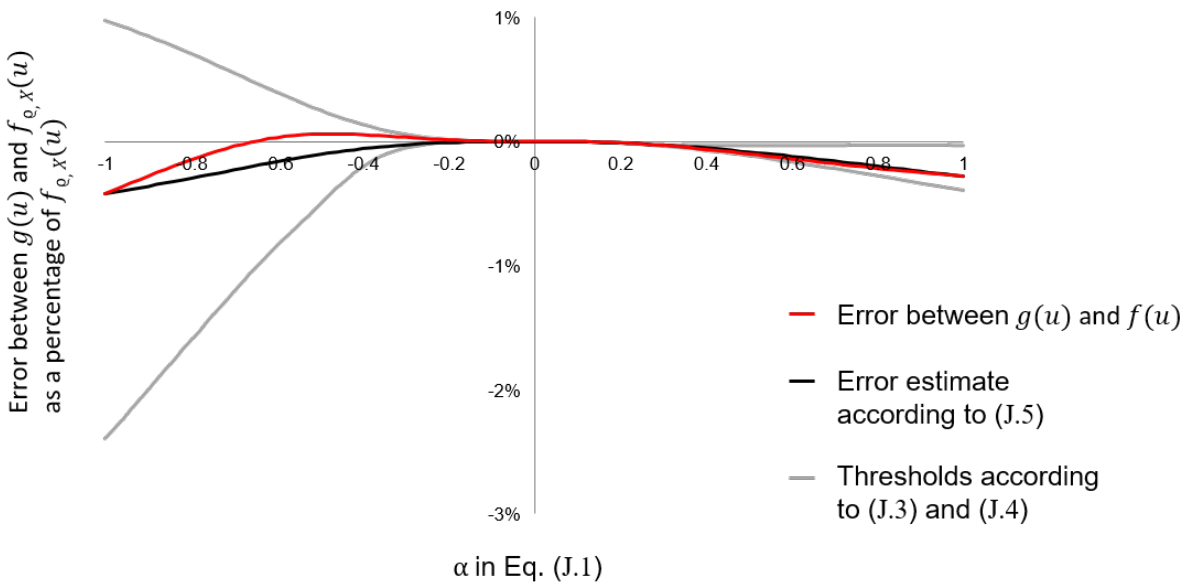

Figure 3: Error between $g(u)$ and $f_{\varrho,X}(u)$ and corresponding estimates as a percentage of $f_{\varrho,X}(u)$.

J Example: Estimating the sensitivity-implied tail-correlation matrix from Monte-Carlo simulations

We consider the example from section 4.2. We perform a simulation-based estimation of the sensitivity-implied tail-correlation matrix similarly to the example in Tasche (2009, pp. 586 ff.). We draw $n_{\text{sim}} = 50,000$ simulations of the random vector $(X_1, ..., X_5)^{\text{T}}$ and repeat the simulation 50 times. The risk measurements $f_{\varrho,X}(\mathbb{1}_5)$ and $f_{\varrho,X}(e_k)$ are

Table 5: Estimation results from Monte-Carlo simulations for some ingredients of the sensitivity-implied tail-correlation matrix, cf. line (8); $n_{\text{sim}} = 50,000$.

	$f(\mathbb{1}_5)$	$f_{\varrho,X}(e_1)$	$\frac{\partial}{\partial u_1} f_{\varrho,X}(\mathbb{1}_5)$	$\frac{\partial^2}{\partial u_1 \partial u_2} f_{\varrho,X}(\mathbb{1}_5)$
	True values			
	8.115	4.679	2.830	-1.080
Bandwidth factor, c	RMSE			
1	0.081	0.067	0.193	1.183
2	0.110	0.125	0.154	0.460
3	0.203	0.270	0.152	0.270
4	0.349	0.486	0.181	0.237
5	0.543	0.774	0.238	0.296

estimated based on the Gaussian kernel and recursion formula (4.2) in Gourieroux et al. (2000, p. 234). The gradient $Df_{\varrho,X}(\mathbb{1}_5)$ is estimated by the Nadaraya-Watson kernel estimator for conditional expectations, cf. Tasche (2009, p. 584, Eq. (11)). The Hesse matrix $D^2 f_{\varrho,X}(\mathbb{1}_5)$ is estimated according to Gourieroux et al. (2000, p. 235, Eq. (4.4)). In each estimation, we set the bandwidth initially in line with the classical proportionality rule, i.e. $h = \left(\frac{4}{3}\right)^{1/5} \cdot \sigma_P \cdot n_{\text{sim}}^{-1/5}$ with σ_P being the estimated standard deviation of $\sum_{i=1}^{5} X_i$. Moreover, we vary the bandwidth by multiplying it by factors $c \in \{1, 2, 3, 4, 5\}$. Table 5 reports the RMSE of the aggregate VaR, the VaR of X_1, the first entry of $Df_{\varrho,X}(\mathbb{1}_5)$ and entry $(1, 2)$ of $D^2 f_{\varrho,X}(\mathbb{1}_5)$, each depending on the bandwidth factor. To calculate the sensitivity-implied matrix, we use an increased bandwidth by factor 3 for the gradient and by factor 4 for the Hesse matrix, given that these choices seem to stabilize the estimates (cf. Table 5). On average, the entries of the estimated sensitivity-implied tail-correlation matrix have a RMSE of 0.050. This RMSE can be reduced to 0.032 if $n_{\text{sim}} = 500,000$ simulations are used.

K Graphical illustration of the approximation error

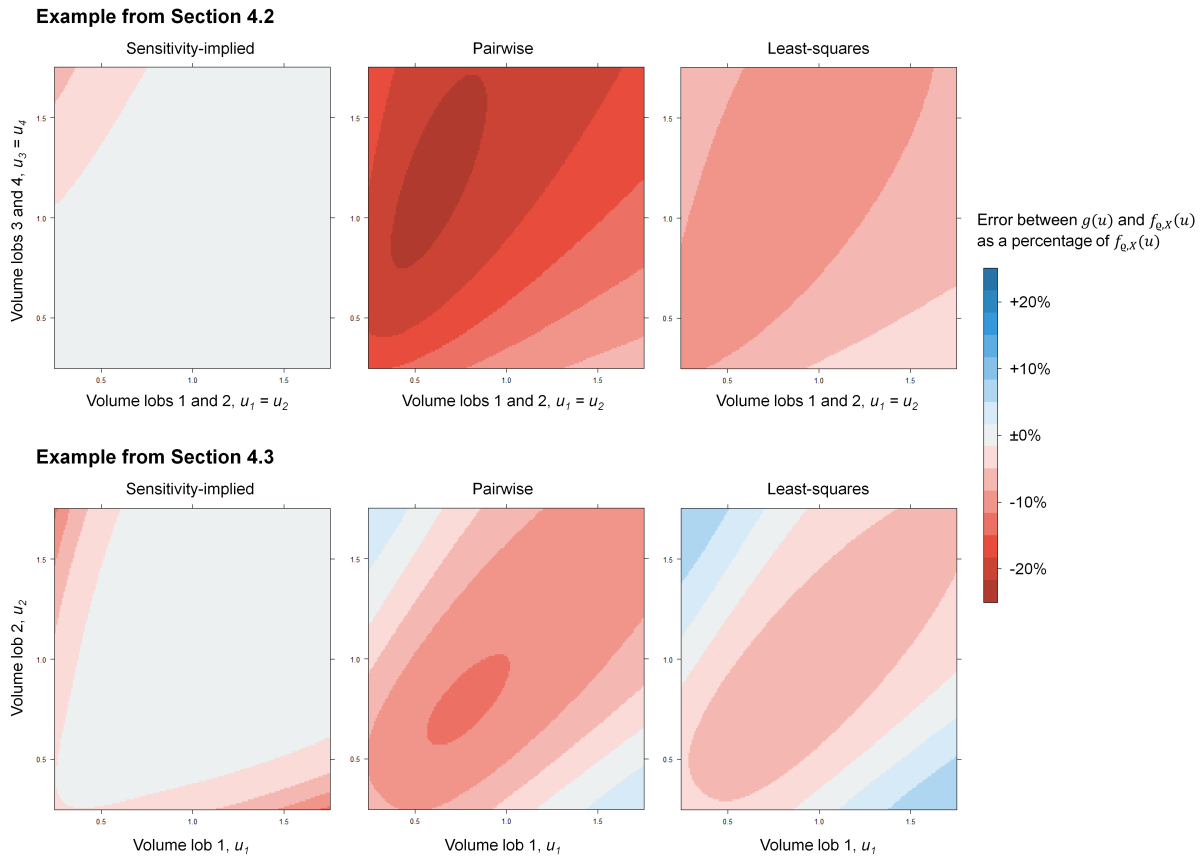

Figure 4: Contour plot of the relative error between $g(u)$ and $f_{\varrho,X}(u)$ according to the examples in sections 4.2 and 4.3. Upper part: x-axis reflects value of $u_1 = u_2$ between 0.25 and 1.75; y-axis reflects corresponding values of $u_3 = u_4$; $u_5 = 1$ is held constant. For example, the upper left corner of the figures correspond to $u = (0.25, 0.25, 1.75, 1.75, 1)^{\mathrm{T}}$. Lower part: x-axis reflects the exposures u_1; y-axis reflects the exposures u_2; $u_3 = 1$ is held constant.

L Gradient of the function $g(u)$

Let $n \in \mathbb{N}$, R be a symmetrix matrix and $x \in \mathbb{R}^n$ such that $x^{\mathrm{T}} R x > 0$. We consider the function $g(u)$ as defined in line (6). The first-order partial derivative of g with respect to an entry u_k of u is obtained as

$$\frac{\partial}{\partial u_k} g(u) = \frac{\sum_{i=1}^{n} \varrho_{ki} u_i x_i}{\sqrt{(u \circ x)^{\mathrm{T}} R (u \circ x)}} \cdot x_k \tag{L.1}$$

In matrix notation and at $u = \mathbb{1}_n$, the gradient of g is determined as

$$D\,g\,(\mathbb{1}_n) = \frac{(Rx) \circ x}{\sqrt{x^{\mathrm{T}} R\, x}} \tag{L.2}$$

References

Ang, A. & Chen, J. (2002), 'Asymmetric correlations of equity portfolios', *Journal of Financial Economics* **63**(3), 443–494.

Bernard, C., Denuit, M. & Vanduffel, S. (2018), 'Measuring portfolio risk under partial dependence information', *Journal of Risk and Insurance* **85**(3), 843–863.

Boonen, T. J., Tsanakas, A. & Wüthrich, M. V. (2017), 'Capital allocation for portfolios with non-linear risk aggregation', *Insurance: Mathematics and Economics* **72**, 95–106.

Braun, A., Schmeiser, H. & Schreiber, F. (2017), 'Portfolio optimization under Solvency II: Implicit constraints imposed by the market risk standard formula', *Journal of Risk and Insurance* **84**(1), 177–207.

Breuer, T., Jandačka, M., Rheinberger, K. & Summer, M. (2010), 'Does adding up of economic capital for market-and credit risk amount to conservative risk assessment?', *Journal of Banking & Finance* **34**(4), 703–712.

Buch, A., Dorfleitner, G. & Wimmer, M. (2011), 'Risk capital allocation for RORAC optimization', *Journal of Banking & Finance* **35**(11), 3001–3009.

Campbell, R., Koedijk, K. & Kofman, P. (2002), 'Increased correlation in bear markets', *Financial Analysts Journal* **58**(1), 87–94.

Casella, G. & Berger, R. (2002), *Statistical Inference, 2nd Edition*, Duxbury.

Chen, T., Goh, J. R., Kamiya, S. & Lou, P. (2019), 'Marginal cost of risk-based capital and risk-taking', *Journal of Banking & Finance* **103**, 130–145.

Christiansen, M. C., Denuit, M. M. & Lazarc, D. (2012), 'The Solvency II square-root formula for systematic biometric risk', *Insurance: Mathematics and Economics* **50**(2), 257–265.

Devineau, L. & Loisel, S. (2009), 'Risk aggregation in Solvency II: How to converge the approaches of the internal models and those of the standard formula?', *Bulletin Français d'Actuariat* **9**(18), 107–145.

Diers, D. (2011), 'Management strategies in multi-year enterprise risk management', *The Geneva Papers on Risk and Insurance - Issues and Practice* **36**(1), 107–125.

Eckert, J. & Gatzert, N. (2018), 'Risk-and value-based management for non-life insurers under solvency constraints', *European Journal of Operational Research* **266**(2), 761–774.

Epperlein, E. & Smillie, A. (2006), 'Cracking VAR with kernels', *RISK* **19**, 70–74.

European Insurance and Occupational Pensions Authority (EIOPA) (2014), The underlying assumptions in the standard formula for the Solvency Capital Requirement calculation, EIOPA-14-322.

Folland, G. B. (2001), *Advanced Calculus*, Pearson.

Furman, E., Kye, Y. & Su, J. (2020), 'A reconciliation of the top-down and bottom-up approaches to risk capital allocations: Proportional allocations revisited', *North American Actuarial Journal* pp. 1–22.

Gatzert, N. & Heidinger, D. (2020), 'An empirical analysis of market reactions to the first solvency and financial condition reports in the european insurance industry', *Journal of Risk and Insurance* **87**(2), 407–436.

Glasserman, P. (2005), 'Measuring marginal risk contributions in credit portfolios', *FDIC Center for Financial Research Working Paper* (2005-01).

Gourieroux, C., Laurent, J.-P. & Scaillet, O. (2000), 'Sensitivity analysis of values at risk', *Journal of empirical finance* **7**(3-4), 225–245.

Hu, C., Pozdnyakov, V. & Yan, J. (2020), 'Density and distribution evaluation for convolution of independent gamma variables', *Computational Statistics* **35**(1), 327–342.

Hull, J. C. (2018), *Risk Management and Financial Institutions. Fifth Edition*, Wiley Finance Series.

Li, J., Zhu, X., Lee, C.-F., Wu, D., Feng, J. & Shi, Y. (2015), 'On the aggregation of credit, market and operational risks', *Review of Quantitative Finance and Accounting* **44**(1), 161–189.

Longin, F. & Solnik, B. (2001), 'Extreme correlation of international equity markets', *Journal of Finance* **56**(2), 649–676.

Markowitz, H. M. (1952), 'Portfolio selection', *Journal of Finance* **7**(1), 77–91.

Mathur, S. (2015), Risk aggregation and capital management, *in* H. K. Baker & G. Filbeck, eds, 'Investment Risk Management', Oxford University Press, Oxford, chapter 14, pp. 261–279.

McNeil, A. J., Frey, R. & Embrechts, P. (2015), *Quantitative Risk Management: Concepts, Techniques and Tools - Revised Edition*, Princeton Series in Finance.

Mittnik, S. (2014), 'VaR-implied tail-correlation matrices', *Economics Letters* **122**(1), 69–73.

Moschopoulos, P. G. (1985), 'The distribution of the sum of independent gamma random variables', *Annals of the Institute of Statistical Mathematics* **37**(3), 541–544.

Pfeifer, D. & Strassburger, D. (2008), 'Solvency II: stability problems with the SCR aggregation formula', *Scandinavian Actuarial Journal* **1**, 61–77.

Scaillet, O. (2004), 'Nonparametric estimation and sensitivity analysis of expected shortfall', *Mathematical Finance* **14**(1), 115–129.

Siller, T. (2013), 'Measuring marginal risk contributions in credit portfolios', *Quantitative Finance* **13**(12), 1915–1923.

Stoughton, N. M. & Zechner, J. (2007), 'Optimal capital allocation using RAROC and EVA', *Journal of Financial Intermediation* **16**(3), 312–342.

Targino, R. S., Peters, G. W. & Shevchenko, P. V. (2015), 'Sequential monte carlo samplers for capital allocation under copula-dependent risk models', *Insurance: Mathematics and Economics* **61**, 206–226.

Tasche, D. (2008), Capital allocation to business units and sub-portfolios: The Euler principle, *in* A. Resti, ed., 'Pillar II in the New Basel Accord: The Challenge of Economic Capital', Risk Books, London, pp. 423–453.

Tasche, D. (2009), 'Capital allocation for credit portfolios with kernel estimators', *Quantitative Finance* **9**(5), 581–595.

Yow, S. & Sherris, M. (2008), 'Enterprise risk management, insurer value maximization, and market frictions', *Astin Bulletin* **38**(1), 293–339.

Zanjani, G. (2002), 'Pricing and capital allocation in catastrophe insurance', *Journal of Financial Economics* **65**(2), 283–305.