NO-SABOTAGE UNDER CONDITIONAL MEAN RISK SHARING OF DEPENDENT-BY-MIXTURE INSURANCE LOSSES

Michel Denuit, Patricia Ortega-Jimenez, Christian Y. Robert

LIDAM Discussion Paper ISBA
2024 / 19

UCLouvain

LIDAM
ISBA | Louvain Institute of Data Analysis and Modeling in economics and statistics

ISBA
Voie du Roman Pays 20 - L1.04.01
B-1348 Louvain-la-Neuve
Email : lidam-library@uclouvain.be
https://uclouvain.be/en/research-institutes/lidam/isba/publication.html

NO-SABOTAGE UNDER CONDITIONAL MEAN RISK SHARING OF DEPENDENT-BY-MIXTURE INSURANCE LOSSES

MICHEL DENUIT
Institute of Statistics, Biostatistics and Actuarial Science - ISBA
Louvain Institute of Data Analysis and Modeling - LIDAM
UCLouvain
Louvain-la-Neuve, Belgium
michel.denuit@uclouvain.be

PATRICIA ORTEGA-JIMENEZ
Institute of Statistics, Biostatistics and Actuarial Science - ISBA
Louvain Institute of Data Analysis and Modeling - LIDAM
UCLouvain
Louvain-la-Neuve, Belgium
patricia.ortega@uclouvain.be

CHRISTIAN Y. ROBERT
Laboratory of Actuarial and Financial Science - LSAF
Université Lyon 1
Lyon, France
christian.robert@univ-lyon1.fr

July 10, 2024

Abstract

Conditional mean risk sharing defines an allocation rule to distribute total losses among participants in an insurance pool. Under this risk-sharing scheme, the no-sabotage condition holds when conditional expectations of individual losses given their sum are comonotonic. This property has been widely studied considering independent risks, often assuming that they possess log-concave densities. This paper considers the no-sabotage condition for dependent-by-mixture risks which do not necessarily obey log-concave distributions. Sufficient conditions derived from three different approaches are proposed in order to fulfill the no-sabotage requirement. Several examples are given to illustrate the applicability of the results.

Keywords: Conditional mean risk sharing; no-sabotage condition; conditional independence; Efron's monotonicity

1 Introduction

The conditional mean risk-sharing rule has been proposed by Denuit and Dhaene (2012) building upon the work by Landsberger and Meilijson (1994). It is a risk allocation mechanism assigning to each participant within an insurance pool the conditional expectation of the loss brought to the pool, given the total losses experienced by the entire pool. It has been axiomatized by Jiao et al. (2022). The reader is referred e.g. to Blier-Wong et al. (2022), Levantesi and Piscopo (2022), Clemente et al. (2023, 2024), Denuit and Robert (2023) and Ghossoub et al. (2024) for contributions to the study of the conditional mean risk-sharing rule, as well as to Feng (2023) for a general presentation of the topic.

Formally, let us consider a pool formed by n economic agents with respective (non-negative) insurance losses denoted as X_i, $i = 1, \ldots, n$, with $0 < \mathrm{E}[X_i] < \infty$ for every i. According to the conditional mean risk-sharing rule, the amount allocated to participant i is

$$m_i(s) = \mathrm{E}[X_i \mid S = s], \text{ where } S = \sum_{i=1}^{n} X_i.$$

The conditional expectations $m_i(\cdot)$ are thus used by participants to allocate the total loss S among them since $\sum_{i=1}^{n} m_i(S) = S$ a.s.

In the context of risk sharing, a desirable property is the so-called "no-sabotage condition", also referred to as the "comonotonicity property", as in Denuit et al. (2022). Under the conditional mean risk-sharing rule, there is no sabotage opportunity when all conditional expectations $m_i(\cdot)$ are non-decreasing. To understand why this property plays an important role, assume that the pool comprises a participant i such that $m_i(\cdot)$ decreases over a range of values of S. When S falls in that range, this participant may then be tempted to exaggerate the loss X_i reported to the pool since this would decrease his or her contribution to the total loss S. This behavior is clearly undesirable. Non-decreasingness of every $m_i(\cdot)$ thus appears to be a reasonable and useful requirement in the context of risk sharing, since it aligns interests for all participants. From a methodological point of view, the absence of sabotage is also intimately related to Pareto optimality as seen e.g. in Denuit and Dhaene (2012) and Dhaene et al. (2023).

Stochastic monotonicity of the random variable X_i in the sum S (also known as "Efron's monotonicity" after Efron (1965)) refers to the condition that the function $s \mapsto \mathrm{E}[\varphi(X_i) \mid S = s]$ is non-decreasing for any non-decreasing function φ, and is therefore stronger than the no-sabotage condition. Stochastic monotonicity is relevant in various research areas, such as in statistical estimation and testing, when observations are limited to the sum of a sample, necessitating inferences about the distribution of the X_i. Stochastic monotonicity of the random variable X_i in the sum S may also be written as the condition that $X_i|S = s_1 \preceq_{\mathrm{ST}} X_i|S = s_2$ whenever $s_1 \leq s_2$, where $\preceq_{\mathrm{ST}}$ refers to the usual (first) stochastic order. Let us recall that $X \preceq_{\mathrm{ST}} Y$ if $\mathrm{E}[g(X)] \leq \mathrm{E}[g(Y)]$ for any non-decreasing function $g(\cdot)$. Lehmann (1966) also studied the monotonicity condition when the stochastic order is replaced by the stronger likelihood ratio order. Given the variables X_1 and X_2 having densities f_1 and f_2, X_1 is said to be smaller than X_2 in the likelihood ratio order (denoted by $X_1 \preceq_{\mathrm{LR}} X_2$) if the ratio $f_1(x)/f_2(x)$ is non-increasing in x over the union of the supports of X_1 and X_2 (for details see, e.g., Shaked and Shanthikumar 2007). It is well-known that, if $X_1 \preceq_{\mathrm{LR}} X_2$, then

$X_1 \preceq_{\text{ST}} X_2$, but the reciprocal is not necessarily true.

In this paper, we focus on an important class of models for insurance losses: the common mixture models consisting in conditionally independent risks. Intuitively speaking, the dependence between risks results from the existence of an external mechanism impacting on all losses $X_1, X_2, \ldots, X_n$. The external mechanism is generally viewed as a common environment described by some positive random variable Λ. Formally, individual losses $X_1, X_2, \ldots, X_n$ are independent given the environmental parameter Λ, but become dependent when Λ is unknown. That is, random vector $\boldsymbol{X} = (X_1, X_2, \ldots, X_n)$ follows a common mixture model if its joint distribution function can be written as

$$F_{\boldsymbol{X}}(x_1, \ldots, x_n) = \int_0^\infty \left(\prod_{i=1}^n F_i(x_i|\lambda) \right) \mathrm{d}F_\Lambda(\lambda), \quad \boldsymbol{x} \in \mathbb{R}^n, \tag{1.1}$$

where $F_i(\cdot \mid \lambda)$ is the conditional distribution function of X_i given $\Lambda = \lambda$, that is $F_i(x|\lambda) = \mathrm{P}[X_i \leq x|\Lambda = \lambda]$ for $x \geq 0$. A vector following a common mixture model is said to be dependent-by-mixture. Typically, Λ represents the magnitude of a natural catastrophe or any hidden feature influencing all risks, simultaneously. The interested reader is referred e.g. to Wang (1998) or to Chapter 7 in Denuit et al. (2005) for a detailed discussion on the common mixture construction.

Two important case studies of the no-sabotage condition (or the stochastic monotonicity) for dependent-by-mixture risks have attracted attention in the literature: i) $X_1, X_2, \ldots, X_n$ are independent, i.e. Λ is a (deterministic) constant, ii) the conditional distribution functions of X_i given $\Lambda = \lambda$ are the same. Since Efron (1965), it is well known that, when $X_1, X_2, \ldots, X_n$ are independent, the stochastic monotonicity of each component of $\boldsymbol{X}$ holds when these components have log-concave densities. Lehmann (1966) also proved that, if $\boldsymbol{X}$ is a set of independent random variables having log-concave densities, except for X_i, then $X_i|S = s_1 \preceq_{\text{LR}} X_i|S = s_2$ whenever $s_1 \leq s_2$. However log-concavity is an assumption that might be too strong in some applications to insurance, because log-concave densities are associated with light-tailed asymptotic behavior, as discussed in Asmussen and Lehtomaa (2017). Denuit et al. (2024) identified situations where a non-monotonic behavior appears for heavy-tailed (Pareto-like) risks. The authors proved that the greater the heterogeneity between the heaviness of the distribution tails, the more frequent sabotage situations are. For the second case study, i.e. when $F_1(\cdot|\lambda) = \ldots = F_n(\cdot|\lambda)$ for any λ, $\boldsymbol{X}$ is then an exchangeable random vector and it is easily derived that $m_i(s) = s/n$ and that the no-sabotage condition holds. The aim of this study is to propose other configurations with sufficient conditions where the no-sabotage condition still holds.

Efron's monotonicity for general dependent variables has mainly been investigated by considering two absolutely continuous dependent risks, providing conditions on their joint probability density function. Given two random variables X_1 and X_2 with joint probability density function $f(\cdot, \cdot)$ and sum $S = X_1 + X_2$, Saumard and Wellner (2018) studied the stochastic monotonicity of X_i in S in terms of the usual stochastic order by providing conditions on the second-order partial derivatives of $-\ln f(\cdot, \cdot)$. In a similar setting, Pellerey and Navarro (2022) provided sufficient conditions to ensure the monotonicity in terms of the likelihood ratio order. Ortega-Jimenez et al. (2022) extended some of these results and provided conditions involving the marginals and the density of the survival copula, when one of the components is Exponentially distributed.

Intuitively, it is expected that sabotage opportunities should disappear under positive dependence (i.e. when all risks X_i are likely to be large or to be small together). Whatever the marginals, there is indeed an extreme case of positive dependence where the no-sabotage condition always holds true: comonotonicity. Recall that the random vector $(X_1, \dots X_n)$ is said to be comonotonic when there exists a random variable Z and non-decreasing functions g_i such that $X_i = g_i(Z)$ for $i = 1, \dots n$. Comonotonic risks X_i can always be represented as non-decreasing functions of their sums so that every $m_i(\cdot)$ is non-decreasing for any choice of marginal distributions. The intuition on the effect of positive dependence is supported by the conditions provided in Pellerey and Navarro (2022) which are related to total positivity (TP2) of the joint density and since all the copula-related examples provided by Saumard and Wellner (2018) consider parameter values implying positive dependence. Dependence-by-mixture may induce positive dependence under additional assumptions on the conditional distribution functions. E.g. if they are likelihood ratio increasing in λ, then $\boldsymbol{X}$ is TP2 in the case $n = 2$. See e.g. Section 7.2.4.3 in Denuit et al. (2005).

The remainder of the paper is organized as follows. In Section 2, we propose several approaches to derive sufficient no-sabotage conditions within the class of common mixture models. The first method uses a formula from Denuit and Robert (2022) that characterizes the no-sabotage condition through a comparison of two sums under the likelihood ratio order, where the dominating sum is derived directly from the formula. The second method expands on results by Pellerey and Navarro (2022) regarding positively dependent risks, applying their insights to common mixture models. The third method introduces conditions that depend on the monotonic behavior of the environmental parameter Λ in S and of the conditional expectations both in the sum and the environmental parameter, using both the stochastic and likelihood ratio orders. These methods have proven valuable for analyzing multiplicative background risk models in Section 3, particularly in the study of the multivariate Type II Mardia Pareto distribution, which meets the no-sabotage conditions and allows for closed-form solutions for the contributions. Finally, Section 4 discusses the results as well as some open problems related to risk reduction.

2 No-sabotage conditions for dependent-by-mixture risks

2.1 A condition based on the size-biased representation formula

Consider risks $X_1, \dots, X_n$ with joint distribution function $F_{\boldsymbol{X}}$. For each $i \in \{1, 2, \dots, n\}$, define the random vector $\boldsymbol{Y}^{[i]} = (Y_1^{[i]}, \dots, Y_n^{[i]})$ with joint distribution function $F_{\boldsymbol{Y}^{[i]}}$ given by

$$\begin{aligned} F_{\boldsymbol{Y}^{[i]}}(y_1, \dots, y_n) &= \int_0^{y_1} \cdots \int_0^{y_n} \frac{x_i \mathrm{d}F_{\boldsymbol{X}}(x_1, \dots, x_n)}{\mathrm{E}[X_i]} \\ &= \frac{\mathrm{E}[X_i | X_1 \le y_1, \dots, X_n \le y_n]}{\mathrm{E}[X_i]} F_{\boldsymbol{X}}(y_1, \dots, y_n). \end{aligned} \tag{2.1}$$

Note that (2.1) corresponds to a particular case of general multivariate size-biased transform of the vector $\boldsymbol{X}$ (see Section 2.3 in Denuit and Robert (2022) for further detail). The random vector $\boldsymbol{Y}^{[i]}$ with distribution function (2.1) plays a central role in the characterization of the risk-sharing function, $m_i(\cdot)$, as explained in Denuit and Robert (2022).

Proposition 2.1. *(Denuit and Robert (2022)) Consider risks $X_1, \ldots, X_n$ with joint distribution function $F_{\boldsymbol{X}}$ and expected values $\mathrm{E}[X_i]$ such that $0 < \mathrm{E}[X_i] < \infty$. Let $\boldsymbol{Y}^{[i]} = (Y_1^{[i]}, \ldots, Y_n^{[i]})$ be a random vector with joint distribution function $F_{\boldsymbol{Y}^{[i]}}$ given by (2.1). The following results hold:*

(i) if $X_1, \ldots, X_n$ are continuous random variables with joint probability density function $f_{\boldsymbol{X}}$ then for any $s \geq 0$,

$$m_i(s) = \mathrm{E}\left[X_i\right] \frac{f_{Y_1^{[i]}+\ldots+Y_n^{[i]}}(s)}{f_{X_1+\ldots+X_n}(s)}.$$

(ii) if $X_1, \ldots, X_n$ are valued in $\{0, 1, 2, \ldots\}$ then for any $s \in \{0, 1, 2, \ldots\}$,

$$m_i(s) = \mathrm{E}\left[X_i\right] \frac{\mathrm{P}[Y_1^{[i]} + \ldots + Y_n^{[i]} = s]}{\mathrm{P}[X_1 + \ldots + X_n = s]}.$$

(iii) if $X_1, \ldots, X_n$ are zero-augmented random variables with positive probability masses at the origin and probability density functions over $(0, \infty)$ then $\mathrm{E}[X_i|S = 0] = 0$ and for any $s > 0$, the representation in (i) holds true.

When the risks $X_1, X_2, \ldots, X_n$ are conditionally independent, given Λ, with joint distribution function (1.1), it turns out that each $\boldsymbol{Y}^{[i]}$ also obeys a common mixture model with a change in the ith conditional distribution and in the mixing distribution. This is formally stated next.

Proposition 2.2. *(Denuit and Robert (2022)) Consider the risks $X_1, X_2, \ldots, X_n$ with joint distribution function (1.1). Then, each $\boldsymbol{Y}^{[i]}$ also possesses a joint distribution function of the form (1.1), that is,*

$$F_{\boldsymbol{Y}^{[i]}}(t_1, \ldots, t_n) = \int_0^\infty \left(\prod_{j \neq i} F_j(t_j|\lambda) \right) \widetilde{F}_i(t_i|\lambda) \mathrm{d}F_{\Lambda_i^*}(\lambda), \;\; \boldsymbol{t} \in \mathbb{R}_+^n,$$

where

$$\widetilde{F}_i(x|\lambda) = \frac{\mathrm{E}\big[X_i \mathrm{I}[X_i \leq x] \big| \Lambda = \lambda\big]}{\mathrm{E}[X_i|\Lambda = \lambda]}, \;\; x \geq 0,$$

and,

$$\mathrm{d}F_{\Lambda_i^\star}(\lambda) = \frac{\mathrm{E}[X_i|\Lambda = \lambda]}{\mathrm{E}[X_i]} \mathrm{d}F_\Lambda(\lambda).$$

We are now ready to derive the no-sabotage condition based on the representation formula given in Proposition 2.1.

Proposition 2.3. *Consider the risks $X_1, X_2, \ldots, X_n$ with joint distribution function (1.1) as defined in Proposition 2.1. Define the random vector $\boldsymbol{Y}^{[i]}$ with joint distribution function given in Proposition 2.2. Then, $m_i(\cdot)$ is non-decreasing if, and only if, $S \preceq_{LR} \sum_{j=1}^n Y_j^{[i]}$.*

Proof. Let us give the proof in the absolutely continuous case. The other cases easy follow. According to Proposition 2.1,

$$m_i(s) = \mathrm{E}\left[X_i\right] \frac{f_{Y_1^{[i]}+\ldots+Y_n^{[i]}}(s)}{f_{X_1+\ldots+X_n}(s)}, \ \ s \geq 0. \tag{2.2}$$

The no-sabotage condition is then equivalent to $X_1 + \ldots + X_n \preceq_{\mathrm{LR}} Y_1^{[i]} + \ldots + Y_n^{[i]}$, as announced. $\square$

Notice that Proposition 2.3 characterizes the no-sabotage condition for common mixture models by a comparison, in the likelihood ratio order, of two sums of conditionally independent random variables.

2.2 A condition based on total positivity

2.2.1 Conditions for a general random vector $\boldsymbol{X}$

Pellerey and Navarro (2022) considered a bivariate vector $\boldsymbol{X}$ with a joint density $f(\cdot,\cdot)$ and provided a necessary and sufficient condition for the stochastic monotonicity based on the likelihood ratio order. This condition is related to the notion of total positivity. As explained in the Introduction section, broadly speaking, positive dependence refers to the intuitive concept of random variables being likely to be large or to be small together. This concept can be formalized in many different ways, including the notion of total positivity. See e.g. Denuit et al. (2005) for a general overview of the topic in actuarial science. Let us recall that a positive bivariate function $p(\cdot,\cdot)$ is said to be totally positive of order 2, abbreviated as TP2, when

$$p\left(x_1, y_1\right) p\left(x_2, y_2\right) \geq p\left(x_1, y_2\right) p\left(x_2, y_1\right) \text{ for all } x_1 \leq x_2 \text{ and } y_1 \leq y_2.$$

If $p(\cdot,\cdot)$ possesses a second-order partial derivative then it is TP2 if, and only if, $\frac{\partial^2}{\partial x \partial y} \ln p(x,y) \geq 0$ for all x and y. Analogously, a function $g : \mathbb{R}^n \to \mathbb{R}^+$ is said to be multivariate totally positive of order 2 (MTP2) if $g(\boldsymbol{x})g(\boldsymbol{y}) \leq g(\boldsymbol{x} \vee \boldsymbol{y})g(\boldsymbol{x} \wedge \boldsymbol{y})$ for any $\boldsymbol{x}, \boldsymbol{y}$ in $\mathbb{R}^n$, where $\wedge$ and $\vee$ refer to the componentwise minimum and maximum, respectively.

Proposition 2.4. *(Pellerey and Navarro (2022)) Let (X_1, X_2) be an absolutely continuous random vector with joint probability density function $f(\cdot,\cdot)$. Then, the following conditions are equivalent:*

(a) The function $f(x, s-x)$ is TP2 in (x, s);

(b) $X_1 \mid S = s_1 \preceq_{LR} X_1 \mid S = s_2$ for all $s_1 \leq s_2$.

Based on this proposition, the following result provides equivalent conditions for X_i to be stochastically increasing in S with respect to the likelihood ratio order.

Property 2.5. *Let (X_1, X_2) be an absolutely continuous random vector with joint probability density function $f(\cdot,\cdot)$ possessing second-order partial derivatives with respect to its arguments. Then, the following conditions are equivalent:*

(a)

$$\frac{\partial^2}{\partial x \partial y} \ln f(x,y) \geq \frac{\partial^2}{\partial y^2} \ln f(x,y) \text{ for all } x \text{ and } y, \tag{2.3}$$

(b)

$$\frac{\partial^2}{\partial x \partial y} \ln f(x,y) - \frac{\partial^2}{\partial y^2} \ln f_{X_1|X_2=y}(x) \geq \frac{\mathrm{d}^2}{\mathrm{d}y^2} \ln f_{X_2}(y) \text{ for all } x \text{ and } y,$$

(c)

$$\frac{\partial}{\partial y}\left(\left(\frac{\partial}{\partial x} - \frac{\partial}{\partial y}\right) \ln f_{X_1|X_2=y}(x)\right) \geq \frac{\mathrm{d}^2}{\mathrm{d}y^2} \ln f_{X_2}(y) \text{ for all } x \text{ and } y,$$

(d) $X_1 \mid S = s_1 \preceq_{LR} X_1 \mid S = s_2$ *for all* $s_1 \leq s_2$.

Proof. According to Proposition 2.4, X_1 increases in S with respect to the likelihood ratio order if, and only if, $f(x, s-x)$ is TP2 in (x, s). For simplicity of notation, let us denote as $\partial_i^k f$ the kth derivative of f with respect to the ith argument. If $\frac{\partial^2}{\partial x \partial y} \ln f(x,y) \geq \frac{\partial^2}{\partial y^2} \ln f(x,y)$ then

$$\frac{f(x,y)\ \partial_1\partial_2 f(x,y) - \partial_1 f(x,y)\ \partial_2 f(x,y)}{f(x,y)^2} - \frac{\partial_2^2 f(x,y)\ f(x,y) - (\partial_2 f(x,y))^2}{f(x,y)^2} \geq 0. \tag{2.4}$$

Now, $f(x, s-x)$ is TP2 in (s, x) if, and only if, $\frac{\partial^2}{\partial x \partial s} \ln f(x, s-x) \geq 0$. Since

$$\begin{aligned}
\frac{\partial^2}{\partial x \partial s} \ln f(x, s-x) &= \frac{\big(\partial_1\partial_2 f(x,s-x) - \partial_2^2 f(x,s-x)\big) f(x,s-x)}{f(x,s-x)^2} \\
&\quad - \frac{\big(\partial_1 f(x,s-x) - \partial_2 f(x,s-x)\big)\partial_2 f(x,s-x)}{f(x,s-x)^2} \\
&= \frac{f(x,s-x)\ \partial_1\partial_2 f(x,s-x) - \partial_1 f(x,s-x)\partial_2 f(x,s-x)}{f(x,s-x)^2} \\
&\quad - \frac{\partial_2^2 f(x,s-x)\partial_2 f(x,s-x) - (\partial_2 f(x,s-x))^2}{f(x,s-x)^2},
\end{aligned} \tag{2.5}$$

the conclusion for the equivalence between (a) and (d) follows from Eq. (2.4) by considering $y = s - x$. Equivalences between (a), (b) and (c) are easily derived. □

Property 2.5 can be seen as an extension of the result obtained by Lehmann (1966) to the case where X_1 and X_2 are dependent. Indeed, if X_1 and X_2 are independent with respective marginal probability density functions f_{X_1} and f_{X_2} then conditions (a), (b) and (c) reduce to

$$\frac{\mathrm{d}^2}{\mathrm{d}y^2} \ln f_{X_2}(y) \leq 0 \Leftrightarrow f_{X_2} \text{ log-concave.}$$

Intuitively, Property 2.5 suggests that as far as the TP2 behavior of the joint density prevails over $\frac{\partial^2}{\partial y^2} \ln f(x,y)$, Efron's monotonicity holds in the likelihood ratio order. Notice also that the condition in Property 2.5 and the intuition are analogous to Corollary 6 of Saumard and Wellner (2018) dealing with a weaker form of Efron's monotonicity, where the usual stochastic order replaces the likelihood ratio order.

Remark 2.6. For some multivariate distributions, the condition needed for the stochastic monotonicity in the likelihood ratio order may be very close to the no-sabotage condition. To see this, let us consider the characterization of the conditional expectation considered in Property 1 of Gourieroux et al. (2000) and the case where $\boldsymbol{X}$ is multivariate Normal with mean vector $\boldsymbol{\mu}$ and variance-covariance matrix $\boldsymbol{\Omega}$. Given a n-dimensional continuous random vector $\boldsymbol{X}$, it is proved in Gourieroux et al. (2000) that, for any $\mathbf{a} \in \mathbb{R}^n$ and any $p \in (0,1)$,

$$\frac{\partial \mathrm{VaR}\left[\mathbf{a}^\top \boldsymbol{X}; p\right]}{\partial \mathbf{a}} = \mathrm{E}\left[\boldsymbol{X} \mid \mathbf{a}^\top \boldsymbol{X} = \mathrm{VaR}\left[\mathbf{a}^\top \boldsymbol{X}; p\right]\right], \tag{2.6}$$

where VaR stands for the Value-at-risk which is defined, for a given a risk X and a probability level $p \in (0,1)$, as $\mathrm{VaR}[X;p] = F_X^{-1}(p)$ (see, for instance, Definition 2.3.1 in Denuit et al. (2005)). If $\boldsymbol{X}$ is multivariate Normal, then it is easily derived that

$$\left.\frac{\partial}{\partial \mathbf{a}} \mathrm{VaR}\left[\mathbf{a}^\top \boldsymbol{X}; p\right]\right|_{\mathbf{a}=\mathbf{1}} = \boldsymbol{\mu} + \frac{\boldsymbol{\Omega}\mathbf{1}}{\left(\mathbf{1}^\top \boldsymbol{\Omega}\mathbf{1}\right)^{1/2}} z_p,$$

where z_p is the quantile of level p of the standard Normal distribution and $\mathbf{1}$ the unit vector. It is then direct to see that the no-sabotage condition holds as far as $\boldsymbol{\Omega}\mathbf{1} \geq 0$. That is, $m_i(\cdot)$ is increasing if $\sum_{j=1}^n \mathrm{Cov}\left[X_i, X_j\right] \geq 0$. This condition is similar to the condition given in Section 6.4.1 of Ortega-Jimenez (2022) and is known to be sufficient to ensure stochastic monotonicity in the likelihood ratio order.

2.2.2 Conditions for common mixture models

Positive dependence properties for common mixture models have been discussed e.g. in Section 7.2.4.3 in Denuit et al. (2005). For instance, Property 7.2.18 in Denuit et al. (2005) states that, if $\boldsymbol{X}$ is an n-dimensional random vector with distribution function given by (1.1) and every $X_i \mid \Lambda = \lambda$ increases in λ with respect to the likelihood ratio order, then the joint probability density function of $\boldsymbol{X}$ is MTP_2. Note that this assumption is true if $f_i(x \mid \lambda)$ is TP2 in x and λ, as seen in Definition 3.3.52 of Denuit et al. (2005). Under a bivariate framework, this condition can be motivated by the following representation of the second-order mixed partial derivative.

Property 2.7. *Let X_1, X_2 be two absolutely continuous risks which are independent given the environmental parameter Λ. Let us consider that the second-order derivatives of the joint distribution function $f(\cdot,\cdot)$ exists and that $f_i(\cdot \mid \lambda)$, which denotes the conditional density function of X_i given $\Lambda = \lambda$, are assumed to be differentiable. Then:*

$$\frac{\partial^2}{\partial x \partial y} \ln f(x,y) = \mathrm{Cov}\left[(\ln f_1)'\left(x \mid \Lambda_{x,y}\right), (\ln f_2)'\left(y \mid \Lambda_{x,y}\right)\right], \text{ for all } x \text{ and } y, \tag{2.7}$$

where, for each x, y the random variable $\Lambda_{x,y}$ obeys the distribution function $F_{\Lambda_{x,y}}(\cdot)$ defined as $\mathrm{d}F_{\Lambda_{x,y}}(\lambda) = \frac{f_1(x|\lambda) f_2(y|\lambda)}{f(x,y)} \mathrm{d}F_\Lambda(\lambda)$ for all $\lambda \geq 0$.

Proof. The joint density can be written as $f(x,y) = \mathrm{E}[f_1(x \mid \Lambda) f_2(y \mid \Lambda)]$. Then:

$$\frac{\partial^2}{\partial x \partial y} \ln f(x,y) = \frac{\left(\mathrm{E}\left[f_1'(x \mid \Lambda) f_2'(y \mid \Lambda)\right] f(x,y) - \mathrm{E}\left[f_1'(x \mid \Lambda) f_2(y \mid \Lambda)\right] \mathrm{E}\left[f_1(x \mid \Lambda) f_2'(y \mid \Lambda)\right]\right)}{f^2(x,y)}$$

Taking into account that $(\ln f_i)'(x \mid \lambda) = \frac{f_i'(x|\lambda)}{f_i(x|\lambda)}$, we can write:

$$\begin{aligned}\frac{\partial^2}{\partial x \partial y} \ln f(x,y) &= \mathrm{E}\left[(\ln f_1)'(x \mid \Lambda)(\ln f_2)'(y \mid \Lambda)\frac{f_1(x \mid \Lambda) f_2(y \mid \Lambda)}{f(x,y)}\right] - \\ &\quad \mathrm{E}\left[(\ln f_1)'(x \mid \Lambda)\frac{f_1(x \mid \Lambda) f_2(y \mid \Lambda)}{f(x,y)}\right] \mathrm{E}\left[(\ln f_2)'(y \mid \Lambda)\frac{f_1(x \mid \Lambda) f_2(y \mid \Lambda)}{f(x,y)}\right].\end{aligned}$$

Since $\mathrm{d}F_{\Lambda_{x,y}}(\lambda) = \frac{f_1(x|\lambda) f_2(y|\lambda)}{f(x,y)} \mathrm{d}F_\Lambda(\lambda)$,

$$\begin{aligned}\mathrm{E}\left[(\ln f_1)'(x \mid \Lambda)(\ln f_2)'(y \mid \Lambda)\frac{f_1(x \mid \Lambda) f_2(y \mid \Lambda)}{f(x,y)}\right] &= \mathrm{E}\left[(\ln f_1)'(x \mid \Lambda_{x,y})(\ln f_2)'(y \mid \Lambda_{x,y})\right], \\ \mathrm{E}\left[(\ln f_i)'(x \mid \Lambda)\frac{f_i(x \mid \Lambda) f_j(y \mid \Lambda)}{f(x,y)}\right] &= \mathrm{E}\left[(\ln f_i)'(x \mid \Lambda_{x,y})\right], \text{ for } i \neq j.\end{aligned}$$

Therefore, expression (2.7) follows. □

From expression (2.7), it is straightforward that, as stated in Denuit et al. (2005), if for $i = 1, 2$, $f_i(x \mid \lambda)$ is TP2 in x and λ, then $\lambda \mapsto (\ln f_i)'(x \mid \lambda)$ are monotonic functions and therefore $\frac{\partial^2}{\partial x \partial y} \ln f(x,y) \geq 0$.

Example 2.8. *Let the vector $\boldsymbol{X} = (X_1, X_2)$ follow a bivariate frailty model. That is, given two baseline survival functions $\bar{G}_1, \bar{G}_2$ and a random variable Λ taking values in $\Omega \subseteq \mathbb{R}^+$ with distribution function F_Λ, then $\boldsymbol{X}$ has joint survival function defined as*

$$\bar{F}(x_1, x_2) = E\left[\prod_{i=1}^{2} \bar{G}_i^\Lambda(x_i)\right] = \int_\Omega \bar{G}_1^\lambda(x_1)\, \bar{G}_2^\lambda(x_2)\, \mathrm{d}F_\Lambda(\lambda), \quad \text{for all } x_1, x_2 \in \mathbb{R}^+.$$

It is easy to check that, if the densities exist, $f_i(x \mid \lambda) = g_i(x)\bar{G}_i^{\lambda-1}(x)$ with $g_i(x) = -\bar{G}_i'(x)$. Since $\frac{\partial^2}{\partial x \partial y} \ln f_i(x \mid \lambda) = \frac{g_i(x)}{\bar{G}_i(x)} \geq 0$, then the joint density function of bivariate frailty models verify the TP2 condition.

The following result provides an equivalent condition for X_i to be stochastically increasing in S with respect to the likelihood ratio order for common mixture models.

Proposition 2.9. *Let X_1, X_2 be two absolutely continuous risks which are independent given the environmental parameter Λ. Let us consider that the second-order derivatives of the joint probability density function $f(\cdot,\cdot)$ exists and that $f_i(\cdot \mid \lambda)$, which denotes the conditional density function of X_i given $\Lambda = \lambda$, are assumed to be twice differentiable. Then, the following conditions are equivalent:*

(a) $\mathrm{Cov}[(\ln f_1)'(x|\Lambda_{x,y}) - (\ln f_2)'(y|\Lambda_{x,y}), (\ln f_2)'(y|\Lambda_{x,y})] \geq \mathrm{E}[(\ln f_2)''(y|\Lambda_{x,y})]$ *for all x and y.*

(b) $X_1 \mid S = s_1 \preceq_{LR} X_1 \mid S = s_2$ *for all $s_1 \leq s_2$.*

Proof. From Property 2.7, we know that

$$\frac{\partial^2}{\partial x \partial y} \ln f(x,y) = \mathrm{Cov}\left[(\ln f_1)'\left(x \mid \Lambda_{x,y}\right), (\ln f_2)'\left(y \mid \Lambda_{x,y}\right)\right].$$

Proceeding in the same way, we get

$$\frac{\partial}{\partial y} \ln f\left(x,y\right) = \frac{\mathrm{E}[f_1\left(x|\Lambda\right) f'\left(y|\Lambda\right)]}{f\left(x,y\right)} = \frac{\mathrm{E}[(\ln f_2)'\left(y|\Lambda\right) f_1\left(x|\Lambda\right) f_2\left(y|\Lambda\right)]}{f\left(x,y\right)}$$

and

$$\begin{aligned}
\frac{\partial^2}{\partial y^2} \ln f\left(x,y\right) &= \frac{\mathrm{E}[(\ln f_2)''\left(y|\Lambda\right) f_1\left(x|\Lambda\right) f_2\left(y|\Lambda\right)]}{f\left(x,y\right)} + \frac{\mathrm{E}\left[\left[(\ln f_2)'\left(y|\Lambda\right)\right]^2 f_1\left(x|\Lambda\right) f_2\left(y|\Lambda\right)\right]}{f\left(x,y\right)} \\
&\quad - \left[\frac{\mathrm{E}[(\ln f_2)'\left(y|\Lambda\right) f_1\left(x|\Lambda\right) f_2\left(y|\Lambda\right)]}{f\left(x,y\right)}\right]^2 \\
&= \mathrm{E}[(\ln f_2)''\left(y|\Lambda_{x,y}\right)] + \mathrm{Var}[(\ln f_2)'\left(y|\Lambda_{x,y}\right)].
\end{aligned}$$

It follows that the condition

$$\frac{\partial^2}{\partial x \partial y} \ln f\left(x,y\right) \geq \frac{\partial^2}{\partial y^2} \ln f\left(x,y\right)$$

is equivalent to the condition

$$\mathrm{Cov}[(\ln f_1)'\left(x|\Lambda_{x,y}\right) - (\ln f_2)'\left(y|\Lambda_{x,y}\right), (\ln f_2)'\left(y|\Lambda_{x,y}\right)] \geq \mathrm{E}[(\ln f_2)''\left(y|\Lambda_{x,y}\right)].$$

□

2.3 Conditions based on conditional expectations given the environmental parameter

In this section, we consider a slightly more general setting where individual losses X_i are impacted by a non-negative random variable Λ, but for which no assumption is made about the dependence structure within the random vector $(\Lambda, \boldsymbol{X})$. The next result derives sufficient conditions to meet the no-sabotage requirement in this general setting.

Proposition 2.10. *Consider the random vector $(\Lambda, \boldsymbol{X})$ with non-negative components and define $S = \sum_{i=1}^n X_i$. If*

(a) $(s, \lambda) \mapsto \mathrm{E}[X_i \mid S = s, \Lambda = \lambda]$ *is non-decreasing,*

(b) $\Lambda \mid S = s_1 \preceq_{ST} \Lambda \mid S = s_2$ *for all* $s_1 \leq s_2$,

then $m_i(\cdot)$ *is non-decreasing.*

Proof. We can rewrite

$$m_i(s) = \mathrm{E}\big[\mathrm{E}[X_i \mid S = s, \Lambda]\big|S = s\big] = \int_0^\infty \mathrm{E}[X_i \mid S = s, \Lambda = \lambda]\mathrm{d}F_{\Lambda|S=s}(\lambda).$$

For $s_2 \geq s_1$, we have

$$\begin{aligned} m_i(s_2) &= \int_0^\infty \mathrm{E}[X_i \mid S = s_2, \Lambda = \lambda]\mathrm{d}F_{\Lambda|S=s_2}(\lambda) \\ &\geq \int_0^\infty \mathrm{E}[X_i \mid S = s_1, \Lambda = \lambda]\mathrm{d}F_{\Lambda|S=s_2}(\lambda). \end{aligned}$$

Now, the function $g_s(\cdot)$ defined as $g_s(\lambda) = \mathrm{E}[X_i|S = s, \Lambda = \lambda]$ is non-decreasing in λ. Therefore, since $\Lambda \mid S = s_1 \preceq_{\mathrm{ST}} \Lambda \mid S = s_2$,

$$\begin{aligned} m_i(s_2) &\geq \int_0^\infty g_{s_1}(\lambda)\mathrm{d}F_{\Lambda|S=s_2}(\lambda) \\ &\geq \int_0^\infty g_{s_1}(\lambda)\mathrm{d}F_{\Lambda|S=s_1}(\lambda) \\ &= m_i(s_1), \end{aligned}$$

which ends the proof. □

Proposition 2.10 can be interpreted as follows. Condition (a) ensures that X_i gets larger on average when Λ increases. It can be seen as a conditional version of the no-sabotage condition. Condition (b) means that Λ is stochastically increasing in the sum. Both conditions result in a positive relationship between Λ and $X_1, X_2, \ldots, X_n$ ensuring that the conditional version of the no-sabotage condition stated under condition (a) also holds unconditionally, when Λ is not known anymore.

Under a common mixture model, $X_1, X_2, \ldots, X_n$ are independent given Λ so that condition (a) in Proposition 2.10 corresponds to no-sabotage in the independent case. Several theoretical results can be applied to check this. For instance, no-sabotage holds when losses are independent and identically distributed, possess log-concave densities or when the size of the pool is large, as seen in Denuit and Robert (2021). Condition (b) in Proposition 2.10 then ensures that the unconditional version of no-sabotage still holds provided Λ stochastically increases in S.

Remark 2.11. The result stated under Proposition 2.10 is still valid when the random variable Λ is replaced with a random vector $\boldsymbol{\Lambda}$. Precisely, if $(s, \boldsymbol{\lambda}) \mapsto \mathrm{E}[X_i \mid S = s, \boldsymbol{\Lambda} = \boldsymbol{\lambda}]$ is non-decreasing and $\boldsymbol{\Lambda}$ increases in S with respect to the multivariate version of $\preceq_{\mathrm{ST}}$ then $m_i(\cdot)$ is non-decreasing.

The result stated under Proposition 2.10 holds true when $\preceq_{\mathrm{ST}}$ is replaced with any stronger stochastic order relation in condition (b). This is the case with $\preceq_{\mathrm{LR}}$, for instance, with the added benefit that the increasingness condition can be reversed, as shown next.

Proposition 2.12. *Consider a random vector* $(\Lambda, \boldsymbol{X})$ *with non-negative components and define* $S = \sum_{i=1}^n X_i$. *Assume that* $X_1, X_2, \ldots, X_n$ *are either absolutely continuous, valued in* $\{0, 1, 2, \ldots\}$ *or zero-augmented absolutely continuous random variables. If* *(a)* $(s, \lambda) \mapsto \mathrm{E}[X_i \mid S = s, \Lambda = \lambda]$ *is non-decreasing in both* s *and* λ *when the other variable is fixed.*

(b) $S \mid \Lambda = \lambda_1 \preceq_{LR} S \mid \Lambda = \lambda_2$ *for all* $\lambda_1 \leq \lambda_2$*,*

then $m_i(\cdot)$ *is non-decreasing.*

Proof. First, suppose that all random variables involved are absolutely continuous. Let $(s, \lambda) \mapsto f(s, \lambda)$ denote the joint probability density function of (S, Λ), $s \mapsto f(s|\lambda)$ denote the conditional probability density function of S given $\Lambda = \lambda$, and $\lambda \mapsto f(\lambda|s)$ denote the conditional probability density function of Λ given $S = s$. Condition (b) means that

$$\begin{aligned} &\frac{f(s_1|\lambda_2)}{f(s_1|\lambda_1)} \leq \frac{f(s_2|\lambda_2)}{f(s_2|\lambda_1)} \text{ for } \lambda_1 \leq \lambda_2 \text{ and } s_1 \leq s_2 \\ \Leftrightarrow\ &\frac{f(s_1, \lambda_2)}{f(s_1, \lambda_1)} \leq \frac{f(s_2, \lambda_2)}{f(s_2, \lambda_1)} \text{ for } \lambda_1 \leq \lambda_2 \text{ and } s_1 \leq s_2 \\ \Leftrightarrow\ &\frac{f(\lambda_2|s_1)}{f(\lambda_1|s_1)} \leq \frac{f(\lambda_2|s_2)}{f(\lambda_1|s_2)} \text{ for } \lambda_1 \leq \lambda_2 \text{ and } s_1 \leq s_2 \end{aligned}$$

which is equivalent to $\Lambda \mid S = s_1 \preceq_{\mathrm{LR}} \Lambda \mid S = s_2$ for all $s_1 \leq s_2$ and is stronger than condition (b) in Proposition 2.10. This ends the proof of the absolutely continuous case. The proof is similar when $X_1, \ldots, X_n$ are valued in $\{0, 1, 2, \ldots\}$. In this case, it suffices to replace all densities with corresponding probability mass functions. If $X_1, \ldots, X_n$ are zero-augmented random variables with positive probability masses at the origin and probability density functions over $(0, \infty)$, the proof follows as for the continuous case by Proposition 2.2 (iii) in Denuit (2019) since we know that $m_i(0) = 0$. □

Notice that condition (b) in Proposition 2.12 is equivalent to require that the joint probability density function of (S, Λ) to be TP2.

Remark 2.13 (Conditional mean independence)**.** Let us now consider the particular case where $(s, \lambda) \mapsto \mathrm{E}[X_i \mid S = s, \Lambda = \lambda]$ appearing in condition (a) of Propositions 2.10-2.12 is constant with respect to λ. This is related to the concept of conditional mean independence. Given a triplet (U_1, U_2, W) of random variables, recall that conditional mean independence of U_1 given W conditioning on U_2 holds if

$$\mathrm{E}[U_1|U_2, W] = \mathrm{E}[U_1|U_2]. \tag{2.8}$$

Here, and throughout this paper, all equalities between random variables are assumed to hold almost surely (that is, with probability 1). Concept (2.8) is used in econometrics to decide whether additional features need to be included in a regression model. If $(s, \lambda) \mapsto \mathrm{E}[X_i \mid S = s, \Lambda = \lambda]$ is constant with respect to λ then X_i is conditionally mean independent of Λ conditioning on S. This is useful in particular when results derived by Furman et al. (2018) to characterize the linearity of $m_i(s)$ apply conditionally to Λ. See, for instance, the compound mixed Poisson model considered in Section 4.2.5 of Denuit and Robert (2022).

3 Application to Type II Mardia Pareto distributions

3.1 No-sabotage condition in multiplicative background risk model

In the multiplicative background risk model, also known as the scale mixture model, risks $X_1, \ldots, X_n$ can be written as $X_i = \Lambda W_i$ for $i = 1, \ldots, n$, where $W_1, \ldots, W_n, \Lambda$ are independent non-negative random variables. The common factor Λ impacting on every X_i plays the role of the environmental parameter. It can be interpreted as the background or systemic risk while $W_1, \ldots, W_n$ are the stand-alone or idiosyncratic risks, specific and independent among participants.

The multiplicative background risk model has been discussed extensively in the actuarial literature. In particular, multiplicative background risk models with W_i obeying the Exponential distribution have attracted a lot of attention. See, for instance, Furman et al. (2018) or Sarabia et al. (2018). A relevant example is the multivariate Type II Mardia Pareto distribution, discussed in Section 3.2, where Λ^{-1} follows a Gamma distribution. See also Furman et al. (2021). For further discussion on this construction, we refer the reader to Asimit et al. (2016) and references therein.

Let us first consider the third approach presented in Section 2 and the sufficient conditions provided in Proposition 2.10. Note that Condition (b) in Proposition 2.10 is obviously satisfied for the multiplicative background risk model. Let us now denote $m_{W_i}(s) = \mathrm{E}[W_i \mid \sum_{j=1}^n W_j = s]$ and assume that $m_{W_i}(\cdot)$ admits a derivative $m'_{W_i}(\cdot)$. Condition (a) in Proposition 2.10 requires

$$m_i(s \mid \lambda) = \mathrm{E}[X_i \mid S = s, \Lambda = \lambda] = \lambda m_{W_i}\left(\frac{s}{\lambda}\right)$$

to be non-decreasing in (s, λ) where $S = \Lambda \sum_{i=1}^n W_i$. For any fixed λ, $m_i(s \mid \lambda)$ is non decreasing in s if $m'_{W_i}(s) \geq 0$. Condition (a) in Proposition 2.10 then holds provided furthermore that

$$\frac{\partial}{\partial \lambda} m_i(s \mid \lambda) = m_{W_i}(s) - \frac{s}{\lambda} m'_{W_i}\left(\frac{s}{\lambda}\right) \geq 0 \Leftrightarrow m'_{W_i}(t) \leq \frac{m_{W_i}(t)}{t} \text{ for all } t.$$

The conditions imposed on the derivative of the conditional expectation $m_{W_i}(\cdot)$ seem to be farly restrictive.

Let us now rather consider the first approach presented in Section 2. In order to deduce a sufficient condition, we must first present an intermediate result. Let us establish the following technical results that will be useful in the remainder of this section.

Lemma 3.1. *A probability density function is such that $f(x/y)$ is TP2 in x and y if, and only if, $g(x) = f(\mathrm{e}^x)$ is log-concave.*

Proof. Clearly, $f(x/y) = f(\exp(\ln x - \ln y))$. Hence, as the logarithm is a strictly increasing function, considering $g(x) = f(\mathrm{e}^x)$, it means that $f(x/y)$ is TP2 in (x, y) if and only if $g(z - w)$ is TP2 in z, w. That is, if g is log-concave. □

We recall that, for a random variable Z, its size-biased version $\widetilde{Z}$ has distribution function $\mathrm{P}[\widetilde{Z} \leq x] = \frac{\mathrm{E}\left[Z\,\mathrm{I}[Z \leq x]\right]}{\mathrm{E}[Z]}$, for $x \geq 0$. If both Z and $\widetilde{Z}$ have respective probability density functions $f_Z(\cdot)$ and $f_{\widetilde{Z}}(\cdot)$ and finite expectations, then $f_{\widetilde{Z}}(x) = \frac{x}{\mathrm{E}[Z]} f_Z(x)$. Therefore, taking into account that $\frac{d^2}{dx^2}\log(\mathrm{e}^x f_Z(\mathrm{e}^x)) = \frac{d^2}{dx^2}\log(f_Z(\mathrm{e}^x))$, by Lemma 3.1, the following corollary follows.

Corollary 3.2. *Let Z be a random variable with probability density function $f_Z(\cdot)$ and let $f_{\widetilde{Z}}(\cdot)$ be the probability density function of the size-biased transform $\widetilde{Z}$. Then, $f_Z(x/y)$ is TP2 in (x, y) if and only if $f_{\widetilde{Z}}(x/y)$ is TP2 in (x, y).*

Note that Lemma 3.1 is useful to determine whether the likelihood ration order is preserved under a product. In particular, if $g(x) = f(\mathrm{e}^x)$ is log-concave and f is derivable, then $yf'(y)/f(y)$ is decreasing in $(0, \infty)$. Therefore, if a random variable Λ verifies the assumption of Lemma 3.1 and W_1, W_2 are absolutely continuous positive independent random variables such that $W_1 \preceq_{\mathrm{LR}} W_2$, by Proposition 1 in Arevalillo and Navarro (2024), $W_1\Lambda \preceq_{\mathrm{LR}} W_2\Lambda$.

Now, we can obtain alternative sufficient conditions for no-sabotage based on Proposition 2.3.

Property 3.3. *Let us consider a random vector $\boldsymbol{X} = (X_1, \ldots, X_n)$ such that $X_i = \Lambda W_i$, $i = 1, \ldots, n$, for independent non-negative random variables $W_1, \ldots, W_n, \Lambda$. If $\sum_{j=1}^n W_j$ and $\sum_{j=1, j\neq i}^n W_j$ have log-concave densities and if the probability density functions of Λ and $\sum_{j=1}^n W_j$ verify the assumption of Lemma 3.1, then $m_i(\cdot)$ is non-decreasing.*

Proof. Note that, considering Λ^* as defined in Proposition 2.2, that is, distributed according to $\mathrm{d}F_{\Lambda^*}(\lambda) = \frac{\lambda}{\mathrm{E}[\Lambda]}\mathrm{d}F_\Lambda(\lambda)$, then $\Lambda^* \overset{\mathrm{d}}{=} \widetilde{\Lambda}$. Therefore, under the scale mixture model, the characterization of Proposition 2.3 can be rewritten as $S^{[i]} \succeq_{\mathrm{LR}} S$, where

$$S = \Lambda\left(\sum_{j=1}^n W_j\right) \text{ and } S^{[i]} = \widetilde{\Lambda}\left(\widetilde{W_i} + \sum_{j\neq i} W_j\right)$$

with all the random variables being independent. Now, $S^{[i]} \succeq_{\mathrm{LR}} S$ is verified if the following inequalities hold:

(I1) $\widetilde{\Lambda}\left(\widetilde{W_i} + \sum_{j\neq i} W_j\right) \succeq_{\mathrm{LR}} \widetilde{\Lambda}\left(\sum_{j=1}^n W_j\right)$,

(I2) $\widetilde{\Lambda}\left(\sum_{j=1}^n W_j\right) \succeq_{\mathrm{LR}} \Lambda\left(\sum_{j=1}^n W_j\right)$.

Since $\sum_{j\neq i} W_j$ has log-concave density and $\widetilde{W_i} \succeq_{\mathrm{LR}} W_i$, we deduce from Theorem 1.C.9. in Shaked and Shanthikumar (2007) that $\widetilde{W_i} + \sum_{j\neq i} W_j \succeq_{\mathrm{LR}} \sum_{j=1}^n W_j$. Since the probability density function of Λ verifies the assumption of Lemma 3.1, so does $\widetilde{\Lambda}$ (see Corollary 3.2) and (I1) is true by Proposition 1 in Arevalillo and Navarro (2024). Since the probability density function of $\sum_{j=1}^n W_j$ also verifies the assumption of Lemma 3.1, (I2) holds by Proposition 1 in Arevalillo and Navarro (2024). $\square$

3.2 No-sabotage under Type II Mardia Pareto distributions

The class of Type II Mardia Pareto distributions is a typical example of heavy-tailed multivariate distributions derived from a common mixture model. Let $\alpha > 0$ and $\boldsymbol{\sigma} = (\sigma_1, \ldots, \sigma_n)$ with $\sigma_i > 0$ for all $i = 1, \ldots, n$. A random vector $\boldsymbol{X}$ is said to have Type II Mardia Pareto distribution, which is henceforth denoted as $\boldsymbol{X} \sim MP^{(n)}(\boldsymbol{\sigma}, \alpha)$, if its joint survival function is given by

$$\overline{F}_{\boldsymbol{X}}(\boldsymbol{x}) = \left(1 + \sum_{i=1}^{n} \frac{x_i}{\sigma_i}\right)^{-\alpha}, \quad x_i \geq 0, \quad i = 1, 2, \ldots, n, \tag{3.1}$$

see e.g. Arnold (1983) for a detailed presentation. It is well-known that the following multiplicative background risk model generates Type II Mardia Pareto distributions: $X_i = \Lambda W_i$ for $i = 1, \ldots, n$ where the random variables $W_1, \ldots, W_n, \Lambda$ are independent and such that $W_i \sim \text{Exponential}\,(\sigma_i)$, $i = 1, \ldots, n$, and $\Lambda^{-1} \sim \text{Gamma}(\alpha, 1)$. If $\boldsymbol{X} \sim MP^{(n)}(\boldsymbol{\sigma}, \alpha)$, then the marginals follow a Type II Pareto univariate distibution. Hence $X_i \sim P(II)(\sigma_i, \alpha, 0)$ for every $i = 1, \ldots, n$, i.e.

$$\overline{F}_i\,(x) = \left(1 + \frac{x}{\sigma_i}\right)^{-\alpha}, \quad x \geq 0, \quad i = 1, 2, \ldots, n,$$

To study whether the probability density functions of the random variables verify the assumption of Lemma 3.1, let us first give the following technical lemma which will be needed after.

Lemma 3.4. *Let* $Y \sim \text{Exponential}(\lambda)$ *and let* X *be an independent positive random variable with density* f_X. *Let us consider* $g_X(x) = f_X(\mathrm{e}^x)$. *And let* f_S *be the density function of* $S = X + Y$ *with* $g_S(x) = f_S(\mathrm{e}^x)$. *If* $\frac{d^2}{dy^2} \ln g_X(y) \leq -\lambda \mathrm{e}^y$, *then* $\frac{d^2}{dy^2} \ln g_S(y) \leq -\lambda \mathrm{e}^y$.

Proof. Starting from

$$f_S(s) = \int_0^s f_X(x) \lambda \mathrm{e}^{-\lambda(s-x)} \mathrm{d}x = \lambda \mathrm{e}^{-\lambda s} \int_0^s f_X(x) \mathrm{e}^{\lambda x} \mathrm{d}x$$

we see that

$$g_s(s) = f_S(\mathrm{e}^s) = \lambda \mathrm{e}^{-\lambda \mathrm{e}^s} \int_{-\infty}^{s} f_X(\mathrm{e}^y) \mathrm{e}^{\lambda \mathrm{e}^y} \mathrm{e}^y \mathrm{d}y.$$

So, we have

$$\frac{d^2}{dy^2} \ln g_s(s) = -\lambda \mathrm{e}^s + \frac{d^2}{dy^2} \ln \left(\int_{-\infty}^{s} f_X(\mathrm{e}^y) \mathrm{e}^{\lambda \mathrm{e}^y} \mathrm{e}^y \mathrm{d}y \right).$$

Now, let us note that, by the assumptions, $\frac{d^2}{dy^2} \ln f_X(\mathrm{e}^y) + \lambda \mathrm{e}^y \leq 0$ and therefore, $h_1(y) = f_X(\mathrm{e}^y) \mathrm{e}^{\lambda \mathrm{e}^y} \mathrm{e}^y$ is log concave. By Lemma 3 in Bagnoli and Bergstrom (2006), it follows that $h_2(s) = \int_{-\infty}^{s} f_X(\mathrm{e}^y) \mathrm{e}^{\lambda \mathrm{e}^y} \mathrm{e}^y \mathrm{d}y$ is also log-concave and we can finally conclude that $\frac{d^2}{dy^2} \ln g_s(s) \leq -\lambda \mathrm{e}^s$. $\square$

As a consequence of this result, the probability density function of the sum of n Exponentially distributed random variables verifies the assumption of Lemma 3.1. The proof is given in Appendix A. A direct consequence of this statement is that, by Property 3.3, the no-sabotage condition holds when losses obey the multiplicative background risk model with idiosyncratic Exponentially distributed risk factors provided the background risk possesses a probability density function satisfying the assumption of Lemma 3.1. This is the case in particular for the Type II Mardia Pareto distribution, as proved in Appendix A. Hence, it is direct to provide the following property.

Property 3.5. *Let $\boldsymbol{X} \sim MP^{(n)}(\boldsymbol{\sigma}, \alpha)$ with $\alpha > 0$ and $\boldsymbol{\sigma} = (\sigma_1, \ldots, \sigma_n)$ with $\sigma_i > 0$ for all $i = 1, \ldots, n$. Then $m_i(\cdot)$ is non-decreasing for every $i = 1, 2, \ldots, n$.*

Therefore, if a pool is formed by risk-holders with risks following the Type II Mardia Pareto distribution, the conditional mean risk sharing rule satisfies the no-sabotage requirement.

This is an important result, as it may be false in the case where the components of $\boldsymbol{X}$ have the same distributions, but are independent. Indeed, let us consider the case $n = 2$, $\boldsymbol{X} \sim MP^{(2)}((\sigma_1, \sigma_2), \alpha)$ with $\sigma_1 = 1$, $\sigma_2 = 10$ and $\alpha = 3$, and $\boldsymbol{Y}$ with $Y_i \stackrel{\mathrm{d}}{=} X_i$ for $i = 1, 2$ and Y_1, Y_2 mutually independent. That is, $Y_i \sim P(II)(\sigma_i, \alpha, 0)$ for $i = 1, 2$ with the same values for the parameters σ_1, σ_2 and α. Figure 1 provides contributions for the first risk holder and may be used to compare these contributions if the components are independent ($m_{Y_1}(s)$) or not ($m_{X_1}(s)$). It is noteworthy that $m_{X_1}(\cdot)$ has been plotted for the Type II Mardia Pareto distribution using the analytical expression given in Proposition 3.6, by (3.3). Figure 1 shows that there are sabotage opportunities under the case where the risks are independent, but not under the case where they are dependent-by-mixture. Note, however that Property 3.6. in Denuit et al. (2024) ensures that $m_{X_1}(s) \to \infty$ as $s \to \infty$, and therefore $m_{X_1}(\cdot)$ decreases over a range of central values despite being asymptotically increasing in s.

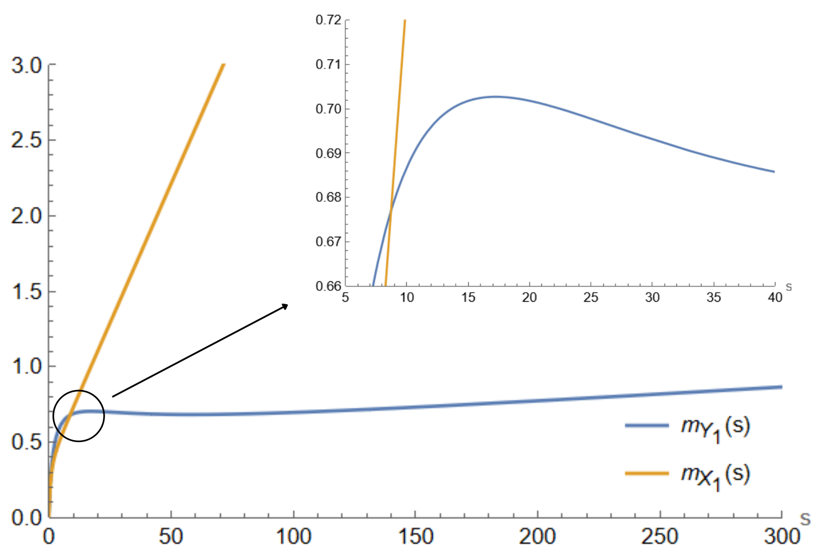

Figure 1: Conditional expectations $m_{X_1}(\cdot)$ and $m_{Y_1}(\cdot)$, with $\boldsymbol{X} \sim MP^{(2)}((\sigma_1, \sigma_2), \alpha)$ (orange), independent $X_i \sim P(II)(\theta_i, \alpha, 0)$ for $i = 1, 2$ (blue) considering $\alpha = 3$, $\sigma_1 = 1$, $\sigma_2 = 10$.

To conclude this section, we provide an analytical expression for CMRS rules using divided differences and several expressions from Chiragiev and Landsman (2007). The proof of the result is included in Appendix B.

Proposition 3.6. *Let* $\mathbf{X} = (X_1, \ldots, X_n)^T \sim MP^{(n)}(\sigma, \alpha)$ *with* $\alpha > 0$*,* $\sigma = (\sigma_1, \ldots, \sigma_n)$*,* $\sigma_i > 0$ *for all* $i = 1, \ldots, n$ *and* $\sigma_i \neq \sigma_j$ *for all* $i \neq j$*. Then:*

$$m_k(s) = \frac{\frac{\alpha\, \sigma_k^{n-2}}{\prod_{\substack{j \neq k \\ 1 \leq j \leq n}} (\sigma_j - \sigma_k)} s \left(1 + \frac{s}{\sigma_k}\right)^{-(\alpha+1)} + \sum_{\substack{i \neq k \\ 1 \leq i \leq n}} \frac{\sigma_k \sigma_i^{n-1} \left(\left(1+\frac{s}{\sigma_i}\right)^{-\alpha} - \left(1+\frac{s}{\sigma_k}\right)^{-\alpha}\right)}{(\sigma_i - \sigma_k) \prod_{\substack{j \neq i \\ 1 \leq j \leq n}} (\sigma_j - \sigma_i)}}{(-1)^{n+1} \sum_{i=1}^{n} \frac{\alpha \sigma_i^{n-2} \left(1+\frac{s}{\sigma_i}\right)^{-(\alpha+1)}}{\prod_{j \neq i} (\sigma_i - \sigma_j)}} \tag{3.2}$$

In particular, considering risk pooling with two participants, if $(X_1, X_2) \sim MP^{(2)}\big((\sigma_1, \sigma_2), \alpha\big)$,

$$m_1(s) = \frac{\alpha\, s \left(1 + \frac{s}{\sigma_1}\right)^{-(\alpha+1)} + \frac{\sigma_1 \sigma_2}{(\sigma_1 - \sigma_2)} \left(\left(1 + \frac{s}{\sigma_2}\right)^{-\alpha} - \left(1 + \frac{s}{\sigma_1}\right)^{-\alpha}\right)}{\alpha \left(\left(1 + \frac{s}{\sigma_1}\right)^{-(\alpha+1)} - \left(1 + \frac{s}{\sigma_2}\right)^{-(\alpha+1)}\right)}. \tag{3.3}$$

4 Conclusions

This paper identified sufficient conditions for the common mixture models which ensure that the no-sabotage requirement is satisfied under the conditional mean risk-sharing rule. Several approaches are proposed to derive such conditions. A first approach is motivated by the representation formula of the conditional mean risk-sharing rule for common mixture models derived in Denuit and Robert (2022). The no-sabotage condition is characterized by the comparison with respect to the likelihood ratio order $\preceq_{\text{LR}}$ of two sums, where the dominating one comes naturally from the representation formula. A second approach starts from a result by Pellerey and Navarro (2022) for positive dependent risks, and extends their analysis to the case of mixing models. Note that this approach deals with bivariate random vectors but, applying the results to the vector $(X_i, S - X_i)$, it can be extrapolated to a higher dimensional framework. Thirdly, conditions are provided based on the monotonicity of the environmental parameter Λ in S with respect to the usual stochastic order $\preceq_{\text{ST}}$, or of S in Λ with respect to the likelihood ratio order $\preceq_{\text{LR}}$, together with the monotonicity of the conditional expectations both in the sum and Λ. These approaches appear to be useful for multiplicative background risk model and the particular case of the multivariate Type II Mardia Pareto distribution has been studied in detail. This model verifies the no-sabotage condition and closed-form expressions are also provided for the contributions.

A promising field of applications of the analytical results derived in this paper is cyber risk insurance. Faure and Nieuwesteeg (2018) and Eling et al. (2021) carefully discussed cyber risk management, including retention, pooling and transfer. In particular, cyber risk pooling can be implemented through the establishment of a mutual insurer, like the new cyber mutual insurer MIRIS (acronym for Mutual Insurance and Reinsurance for Information Systems) based in Brussels. Heavy-tailedness in cyber risk severities is well documented in the literature. See for instance Maillart and Sornette (2010) or Eling and Wirfs (2019). In addition, as discussed in Eling (2020), risk management in the field of cyber risk is highly influenced by systemic risk and the inclusion of models which consider heavy tails and tail dependence among cyber risk is necessary. Modeling this systemic risk through a common mixture model appears to be a natural idea. For instance, Peters et al. (2023) study cyber event severity distributions, aggregate cyber-related losses by business sector and explore the copula dependence structure for pairs of sectors. These authors found that the survival Joe copula provided the best fit in most scenarios. Joe's copula is Archimedean and, therefore, can be written as a common mixture model. This suggests the suitability of the results here derived for cyber insurance applications.

To end with, let us discuss an open question related to risk reduction by conditional mean risk sharing. Recall that the random variable Y is said to be smaller than the random variable Z in the convex order, denoted as $Y \preceq_{\mathrm{CX}} Z$, when $\mathrm{E}[g(Y)] \leq \mathrm{E}[g(Z)]$ holds for every convex function g such that the expectations exist. Jensen inequality then shows that $m_i(S) \preceq_{\mathrm{CX}} X_i$ which guarantees that joining the pool is the preferred option for all risk-averse agents when total losses are allocated according to the conditional mean risk-sharing rule, since this means paying the preferred $m_i(S)$ rather than X_i. This is always true, whatever the dependence structure.

A more delicate problem is whether enlarging the pool is regarded as beneficial by all risk-averse agents. Denoting as $S_n = \sum_{i=1}^{n} X_i$ and $m_{i,n}(S_n) = \mathrm{E}[X_i|S_n]$ for $i = 1, 2, \ldots, n$ where the subscript "n" has been added because we vary the size of the pool, this amounts to determine whether $m_{i,n+1}(S_{n+1}) \preceq_{\mathrm{CX}} m_{i,n}(S_n)$ holds true. It is known from Denuit and Robert (2023) that this is the case when losses X_i are independent. For comonotonic losses, X_i and $m_{i,n}(S_n)$ are identically distributed so that there is indifference between joining the pool or not or with respect to enlarging the pool. A natural question is whether risk reduction still holds for conditionally independent losses X_i under appropriate conditions. We leave this problem to a subsequent work.

Acknowledgements

Michel Denuit and Patricia Ortega-Jimenez gratefully acknowledge funding from the FWO and F.R.S.-FNRS under the Excellence of Science (EOS) programme, project ASTeRISK (40007517).

References

- Arevalillo, J. M., Navarro, J. (2024). Assessment of extreme records in environmental data through the study of stochastic orders for scale mixtures of skew normal vectors. Environmental and Ecological Statistics 31, 151-179.

- Arnold, B.C. (1983). Pareto Distributions. International Co-operative Publishing House.

- Asimit, A. V., Vernic, R., Zitikis, R. (2016). Background risk models and stepwise portfolio construction. Methodology and Computing in Applied Probability 18, 805-827.

- Asmussen, S., Lehtomaa, J. (2017). Distinguishing log-concavity from heavy tails. Risks 5, 10.

- Bagnoli, M., Bergstrom, T. (2006). Log-concave probability and its applications. In "Rationality and Equilibrium: A Symposium in Honor of Marcel K. Richter" (pp. 217-241). Springer.

- Blier-Wong, C., Cossette, H., Marceau, E. (2022). Generating function method for the efficient computation of expected allocations. arXiv preprint arXiv:2207.02654.

- Chiragiev, A., Landsman, Z. (2007). Multivariate Pareto portfolios: TCE-based capital allocation and divided differences. Scandinavian Actuarial Journal 4, 261-280.

- Clemente, G. P., Levantesi, S., Piscopo, G. (2023). Optimal cashback in a cooperative framework for peer-to-peer insurance coverages. Annals of Operations Research, in press.

- Clemente, G. P., Levantesi, S., Piscopo, G. (2024). Risk sharing rule and safety loading in a peer to peer cooperative insurance model. Decisions in Economics and Finance, in press.

- Denuit, M. (2019). Size-biased transform and conditional mean risk sharing, with application to P2P insurance and tontines. ASTIN Bulletin 49, 591-617.

- Denuit, M., Dhaene, J. (2012). Convex order and comonotonic conditional mean risk sharing. Insurance: Mathematics and Economics, 51, 265-270.

- Denuit, M., Dhaene, J., Goovaerts, M., Kaas, R. (2005). Actuarial Theory for Dependent Risks: Measures, Orders and Models. John Wiley & Sons.

- Denuit, M., Dhaene, J., Robert, C. (2022). Risk-sharing rules and their properties, with applications to peer-to-peer insurance. Journal of Risk and Insurance 89, 615-667.

- Denuit, M., Ortega-Jiménez, P., Robert, C.Y. (2024). Conditional expectation given the sum of independent random variables with regularly varying densities. UCL-LIDAM Discussion Paper ISBA 2024/06 .

- Denuit, M., Robert, C.Y. (2021). Efron's asymptotic monotonicity property in the Gaussian stable domain of attraction. Journal of Multivariate Analysis 186, 104803.

- Denuit, M., Robert, C.Y. (2022). Conditional tail expectation decomposition and conditional mean risk sharing for dependent and conditionally independent losses. Methodology and Computing in Applied Probability 24, 1953–1985.

- Denuit, M., Robert, C. Y. (2023). From risk reduction to risk elimination by conditional mean risk sharing of independent losses. Insurance: Mathematics and Economics 108, 46-59.

- Dhaene, J., Robert, C.Y., Cheung, K.C., Denuit, M. (2023). An axiomatic theory for comonotonicity-based risk sharing. LIDAM Discussion Papers ISBA 2023028, Université catholique de Louvain, Institute of Statistics, Biostatistics and Actuarial Sciences (ISBA).

- Eling, M. (2020). Cyber risk research in business and actuarial science. European Actuarial Journal 10, 303-333.

- Eling, M., McShane, M., Nguyen, T. (2021). Cyber risk management: History and future research directions. Risk Management and Insurance Review 24, 93-125.

- Eling, M., Wirfs, J. (2019). What are the actual costs of cyber risk events? European Journal of Operational Research 272, 1109-1119.

- Faure, M.G., Nieuwesteeg, B. (2018). The law and economics of cyber risk pooling. New York University Journal of Law & Business 14, 923-963.

- Feng, R. (2023). Decentralized Insurance: Technical Foundation of Business Models. Springer.

- Furman, E., Kuznetsov, A., Zitikis, R. (2018). Weighted risk capital allocations in the presence of systematic risk. Insurance: Mathematics and Economics 79, 75-81.

- Furman, E., Kye, Y., Su, J. (2021). Multiplicative background risk models: Setting a course for the idiosyncratic risk factors distributed phase-type. Insurance: Mathematics and Economics 96, 153-167.

- Ghossoub, M., Principi, G., Wang, R. (2024). Allocation mechanisms in decentralized exchange markets with frictions. arXiv preprint arXiv:2404.10900.

- Gourieroux, C., Laurent, J. P., Scaillet, O. (2000). Sensitivity analysis of values at risk. Journal of Empirical Finance, 7, 225-245.

- Jiao, Z., Kou, S., Liu, Y., Wang, R. (2022). An axiomatic theory for anonymized risk sharing. arXiv preprint arXiv:2208.07533.

- Landsberger, M., Meilijson, I. (1994). Co-monotone allocations, Bickel-Lehmann dispersion and the Arrow-Pratt measure of risk aversion. Annals of Operations Research 52, 97-106.

- Lehmann, E. L. (1966). Some concepts of dependence. The Annals of Mathematical Statistics 37, 1137-1153.

- Levantesi, S., Piscopo, G. (2022). Mutual peer-to-peer insurance: The allocation of risk. Journal of Co-operative Organization and Management 10, 100154.

- Maillart, T., Sornette, D. (2010). Heavy-tailed distribution of cyber-risks. The European Physical Journal B 75, 357-364.

- Nelsen, R.B., 2007. An Introduction to Copulas. Springer.

- Oudghiri, Y. (2021). Generalizations of Efron's theorem. Statistics & Probability Letters 177, 109158.

- Ortega-Jiménez, P. (2022). Stochastic orders applied to insurance and economics. Doctoral dissertation, Universidad de Cadiz.

- Ortega-Jiménez, P., Pellerey, F., Sordo, M. A., Sarez-Llorens, A. (2022). Note on Efron's monotonicity property under given copula structures. In International Conference on Soft Methods in Probability and Statistics (pp. 303-310). Cham: Springer International Publishing.

- Pellerey, F., Navarro, J. (2022). Stochastic monotonicity of dependent variables given their sum. Test 31, 543-561.

- Peters, G. W., Malavasi, M., Sofronov, G., Shevchenko, P. V., Track, S., Jang, J. (2023). Cyber loss model risk translates to premium mispricing and risk sensitivity. The Geneva Papers on Risk and Insurance-Issues and Practice 48, 372-433.

- Sarabia, J. M., Gomez-Deniz, E., Prieto, F., Jorda, V. (2018). Aggregation of dependent risks in mixtures of exponential distributions and extensions. ASTIN Bulletin 48, 1079-1107.

- Shaked, M., Shanthikumar, J. G.. (2007). Stochastic Orders. Springer.

- Wang, S. (1998). Aggregation of correlated risk portfolios: Models and algorithms. Proceedings of the Casualty Actuarial Society 85, 848-939.

Appendix

A Assumption in Lemma 3.1 for some distribution families

Property A.1. *Let $W_1, \dots, W_n$ be independent random variables such that $W_i \sim \text{Exponential}(1/\lambda_i)$. Then, the probability density function of $\sum_{i=1}^n W_i$ verifies the assumption of Lemma 3.1.*

Proof. Let us consider the variables sorted in terms of the parameters verifying $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n$. Let us denote $S_i = \sum_{j=1}^i W_j$, by f_j the density of W_j and g_j the function $g_j(x) = f_j(\mathrm{e}^x)$. Now, proceed by induction:

- For $i = 2$. We have that $\frac{d^2}{dy^2} \ln f_1(\mathrm{e}^y) = -\lambda_1 \mathrm{e}^y$, then, as $\lambda_1 \geq \lambda_2$, by Lemma 3.4, $\frac{d^2}{dy^2} \ln g_2(y) \leq -\lambda_2 \mathrm{e}^y$.
- Assume that the result holds for i and let us establish its validity for $i+1$. Let us assume that $\frac{d^2}{dy^2} \ln g_i(y) \leq -\lambda_i \mathrm{e}^y$. Then, by Lemma 3.4, we have that $\frac{d^2}{dy^2} \ln g_{i+1}(y) \leq -\lambda_i \mathrm{e}^y$. Since $\lambda_i \geq \lambda_{i+1}$, we have that $\frac{d^2}{dy^2} \ln g_{i+1}(y) \leq -\lambda_{i+1} \mathrm{e}^y$.

Hence, $\frac{d^2}{dy^2} \ln g_n(y) \leq -\lambda_n \mathrm{e}^y < 0$ holds true. This means that g_n is log-concave. The result then follows from the equivalence stated under Lemma 3.1. □

Property A.2. *Let $\gamma \sim \text{Gamma}(\alpha, 1)$. Then the probability density function of $X = \frac{1}{\gamma}$ satisfies the assumption of Lemma 3.1.*

Proof. If $\gamma \sim \text{Gamma}(\alpha, 1)$, then $f_\gamma(x) = \frac{x^{\alpha-1}\mathrm{e}^{-x}}{\Gamma(\alpha-1)}$. Note that $\mathrm{P}\left[\frac{1}{X} < x\right] = \mathrm{P}\left[\frac{1}{x} < X\right]$ and, therefore, if $X \sim f_X$, then $f_{1/X}(x) = \frac{1}{x^2} f_X(1/x)$. Hence:

$$\partial_t \partial_w \log f_{1/\gamma}\left(\frac{w}{t}\right) = \partial_t \partial_w \log \left(\left(\frac{t}{w}\right)^2 f_\gamma\left(\frac{t}{w}\right)\right) = \partial_t \partial_w \log \left(\frac{t}{w}\right)^2 + \partial_w \partial_t \log f_\gamma \left(\frac{t}{w}\right)$$

Now, let us note that $\partial_t \partial_w \log \left(\frac{t}{w}\right)^2 = 0$ and

$$\begin{aligned}
\partial_w \partial_t \log f_\gamma\left(\frac{t}{w}\right) &= \partial_w \left(\frac{1}{w} \frac{\frac{(\alpha-1)\left(\frac{t}{w}\right)^{\alpha-2}\mathrm{e}^{-\frac{t}{w}} - \left(\frac{t}{w}\right)^{\alpha-1}\mathrm{e}^{-\frac{t}{w}}}{\Gamma(\alpha-1)}}{\frac{\left(\frac{t}{w}\right)^{\alpha-1} e^{-\frac{t}{w}}}{\Gamma(\alpha-1)}} \right) \\
&= \partial_w \left(\frac{1}{w} \left((\alpha - 1)\left(\frac{w}{t}\right) - 1 \right) \right) \\
&= \partial_w \left((\alpha - 1)\left(\frac{1}{t}\right) - \frac{1}{w} \right) \\
&= \frac{1}{w^2} \geq 0.
\end{aligned}$$

This ends the proof. □

B Analytical expression of contributions for Type II Mardia Pareto distribution using divided differences

In this section, we provide the proof of Proposition 3.6, which provides a closed-form expression of contributions for Type II Mardia Pareto distributions. In this proof we use the notations of Chiragiev and Landsman (2007). We can write $m_k(s) = \frac{1}{f_S(s)} \int_0^s x f_{X_k,S}(x,s)\mathrm{d}x$. From expression (4.6) in Chiragiev and Landsman (2007) we have

$$f_S(s) = \sum_{i=1}^{n} \frac{\alpha \sigma_i^{n-2} \left(1 + (s/\sigma_i)\right)^{-(\alpha+1)}}{\prod_{j\neq i} (\sigma_i - \sigma_j)}.$$

Let us use the computations appearing in the proof of Theorem 1 of Chiragiev and Landsman (2007) for the mixture representation of the density functions. Recall that the density g of a random variable following the Gamma(α,β) distribution is given by $g(x) = \frac{x^{\alpha-1}\mathrm{e}^{-\beta x}\beta^{\alpha}}{\Gamma(\alpha)}$. In the following expressions, we denote $\lambda_i = \lambda/\sigma_i$ and $\pi(\lambda)$ the density of $\Lambda \sim$ Gamma$(\alpha, 1)$ for $\alpha > 2$:

$$\begin{aligned}
\int_0^s x f_{X_k,S}(x,s)\mathrm{d}x &= \int_0^\infty \int_0^s x f_{(X_k,S)|\lambda}(x,s)\mathrm{d}x\pi(\lambda)\mathrm{d}\lambda \\
&= \int_0^\infty \int_0^s x f_{X_k|\lambda}(x) f_{S-X_k|\lambda}(s-x)\mathrm{d}x\pi(\lambda)\mathrm{d}\lambda \\
&= \int_0^\infty (-1)^{n-1} \left(\prod_{j=1}^n \lambda_j\right) h_k(s,\lambda_1;\lambda_2;...;\lambda_{k-1};\lambda_{k+1};...;\lambda_n)\pi(\lambda)\mathrm{d}\lambda,
\end{aligned} \tag{B.1}$$

where $h_k(s,\lambda_1;\lambda_2;...;\lambda_{k-1};\lambda_{k+1};...;\lambda_n)$ is the nth-order divided difference in λ of the function h_k given by

$$h_k(s,\lambda_i) = \frac{s\mathrm{e}^{-\lambda_k s}}{\lambda_k - \lambda_i} + \frac{\mathrm{e}^{-\lambda_k s}}{(\lambda_k - \lambda_i)^2} - \frac{\mathrm{e}^{-\lambda_i s}}{(\lambda_k - \lambda_i)^2}.$$

From Remark 1 in Chiragiev and Landsman (2007), this divided difference can be expressed as

$$h_k(s,\lambda_1;\lambda_2;...;\lambda_{k-1};\lambda_{k+1};...;\lambda_n) = \sum_{\substack{i\neq k \\ 1\leq i\leq n}} \frac{h_k(s,\lambda_i)}{\prod_{\substack{j\neq k,i \\ 1\leq j\leq n}} (\lambda_i - \lambda_j)} = \sum_{\substack{i\neq k \\ 1\leq i\leq n}} \frac{h_k(s,\lambda/\sigma_i)}{\lambda^{n-2}\prod_{\substack{j\neq k,i \\ 1\leq j\leq n}} \left(\frac{1}{\sigma_i} - \frac{1}{\sigma_j}\right)}.$$

Hence, expression (B.1) can be rewritten as

$$\sum_{\substack{i\neq k \\ 1\leq i\leq n}} \int_0^\infty \frac{(-1)^{n-1}\lambda^n \left(\prod_{j=1}^n \frac{1}{\sigma_j}\right)}{\lambda^{n-2}\prod_{\substack{j\neq k,i \\ 1\leq j\leq n}} \left(\frac{1}{\sigma_i} - \frac{1}{\sigma_j}\right)} \left(\frac{s\mathrm{e}^{-\frac{\lambda}{\sigma_k}s}}{\lambda\left(\frac{1}{\sigma_k} - \frac{1}{\sigma_i}\right)} + \frac{\mathrm{e}^{-\frac{\lambda}{\sigma_k}s}}{\lambda^2\left(\frac{1}{\sigma_k} - \frac{1}{\sigma_i}\right)^2} - \frac{\mathrm{e}^{-\frac{\lambda}{\sigma_i}s}}{\lambda^2\left(\frac{1}{\sigma_k} - \frac{1}{\sigma_i}\right)^2}\right)\pi(\lambda)\mathrm{d}\lambda$$

$$= \sum_{\substack{i\neq k \\ 1\leq i\leq n}} \frac{(-1)^n \left(\prod_{j=1}^n \frac{1}{\sigma_j}\right)}{\prod_{\substack{j\neq i \\ 1\leq j\leq n}} \left(\frac{1}{\sigma_i} - \frac{1}{\sigma_j}\right)} \int_0^\infty \left(\lambda s\mathrm{e}^{-\frac{\lambda}{\sigma_k}s} + \frac{\mathrm{e}^{-\frac{\lambda}{\sigma_k}s}}{\left(\frac{1}{\sigma_k} - \frac{1}{\sigma_i}\right)} - \frac{\mathrm{e}^{-\frac{\lambda}{\sigma_i}s}}{\left(\frac{1}{\sigma_k} - \frac{1}{\sigma_i}\right)}\right)\lambda^{\alpha-1}\frac{e^{-\lambda}}{\Gamma(\alpha)}\mathrm{d}\lambda. \tag{B.2}$$

Now, we can split $\int_0^s x f_{X_k,S}(x,s)\mathrm{d}x = (I)+(II)+(III)$, with

$$
\begin{aligned}
(I) &= \sum_{\substack{i\neq k\\ 1\leq i\leq n}} \frac{(-1)^n\left(\prod_{j=1}^n \frac{1}{\sigma_j}\right)}{\prod_{\substack{j\neq i\\ 1\leq j\leq n}}\left(\frac{1}{\sigma_i}-\frac{1}{\sigma_j}\right)} \int_0^\infty s\mathrm{e}^{-\frac{\lambda}{\sigma_k}s}\lambda^\alpha \frac{\mathrm{e}^{-\lambda}}{\Gamma(\alpha)}\mathrm{d}\lambda \\
&= \sum_{\substack{i\neq k\\ 1\leq i\leq n}} \frac{(-1)^n\left(\prod_{j=1}^n \frac{1}{\sigma_j}\right)}{\prod_{\substack{j\neq i\\ 1\leq j\leq n}}\left(\frac{1}{\sigma_i}-\frac{1}{\sigma_j}\right)} \left(1+\frac{s}{\sigma_k}\right)^{-(\alpha+1)} s\ \alpha \quad \underbrace{\int_0^\infty \frac{\lambda^\alpha e^{-(1+\frac{s}{\sigma_k})\lambda}\left(1+\frac{s}{\sigma_k}\right)^{\alpha+1}}{\Gamma(\alpha+1)}\mathrm{d}\lambda}_{=1,\text{ since is the density of Gamma}(\alpha+1,1+\frac{s}{\sigma_k}).} \\
&= (-1)^n\left(\prod_{j=1}^n \frac{1}{\sigma_j}\right)\left(1+\frac{s}{\sigma_k}\right)^{-(\alpha+1)} s\ \alpha \sum_{\substack{i\neq k\\ 1\leq i\leq n}} \frac{1}{\prod_{\substack{j\neq i\\ 1\leq j\leq n}}\left(\frac{1}{\sigma_i}-\frac{1}{\sigma_j}\right)} \\
&= (-1)^n\left(\prod_{j=1}^n \frac{1}{\sigma_j}\right)\left(1+\frac{s}{\sigma_k}\right)^{-(\alpha+1)} s\ \alpha \sum_{\substack{i\neq k\\ 1\leq i\leq n}} \frac{1}{\left(\prod_{j=1}^n \frac{1}{\sigma_j}\right)\frac{1}{\sigma_i^{n-2}}\prod_{\substack{j\neq i\\ 1\leq j\leq n}}(\sigma_j-\sigma_i)} \\
&= (-1)^n\left(1+\frac{s}{\sigma_k}\right)^{-(\alpha+1)} s\ \alpha \sum_{\substack{i\neq k\\ 1\leq i\leq n}} \frac{\sigma_i^{n-2}}{\prod_{\substack{j\neq i\\ 1\leq j\leq n}}(\sigma_j-\sigma_i)} \qquad \text{(B.3)}
\end{aligned}
$$

$$
\begin{aligned}
(II) &= \sum_{\substack{i\neq k\\ 1\leq i\leq n}} \frac{(-1)^n\left(\prod_{j\neq k} \frac{1}{\sigma_j}\right)}{\prod_{\substack{j\neq i\\ 1\leq j\leq n}}\left(\frac{1}{\sigma_i}-\frac{1}{\sigma_j}\right)} \int_0^\infty \frac{\mathrm{e}^{-\frac{\lambda}{\sigma_k}s}}{\left(\frac{1}{\sigma_k}-\frac{1}{\sigma_i}\right)}\lambda^{\alpha-1}\frac{\mathrm{e}^{-\lambda}}{\Gamma(\alpha)}\mathrm{d}\lambda \\
&= (-1)^n\left(\prod_{j=1}^n \frac{1}{\sigma_j}\right)\left(1+\frac{s}{\sigma_k}\right)^{-\alpha} \sum_{\substack{i\neq k\\ 1\leq i\leq n}} \frac{1}{\left(\frac{1}{\sigma_k}-\frac{1}{\sigma_i}\right)\prod_{\substack{j\neq i\\ 1\leq j\leq n}}\left(\frac{1}{\sigma_i}-\frac{1}{\sigma_j}\right)} \\
&= (-1)^n\left(1+\frac{s}{\sigma_k}\right)^{-\alpha}\sigma_k \sum_{\substack{i\neq k\\ 1\leq i\leq n}} \frac{\sigma_i^{n-1}}{(\sigma_i-\sigma_k)\prod_{\substack{j\neq i\\ 1\leq j\leq n}}(\sigma_j-\sigma_i)} \qquad \text{(B.4)}
\end{aligned}
$$

$$
\begin{aligned}
(III) &= -\sum_{\substack{i\neq k\\ 1\leq i\leq n}} \frac{(-1)^n\left(\prod_{j=1}^n \frac{1}{\sigma_j}\right)}{\prod_{\substack{j\neq i\\ 1\leq j\leq n}}\left(\frac{1}{\sigma_i}-\frac{1}{\sigma_j}\right)} \int_0^\infty \frac{\mathrm{e}^{-\frac{\lambda}{\sigma_i}s}}{\lambda\left(\frac{1}{\sigma_k}-\frac{1}{\sigma_i}\right)}\lambda^{\alpha-1}\frac{\mathrm{e}^{-\lambda}}{\Gamma(\alpha)}\mathrm{d}\lambda \\
&= (-1)^{n+1}\left(\prod_{j=1}^n \frac{1}{\sigma_j}\right) \sum_{\substack{i\neq k\\ 1\leq i\leq n}} \frac{\left(1+\frac{s}{\sigma_i}\right)^{-\alpha}}{\left(\frac{1}{\sigma_k}-\frac{1}{\sigma_i}\right)\prod_{\substack{j\neq i\\ 1\leq j\leq n}}\left(\frac{1}{\sigma_i}-\frac{1}{\sigma_j}\right)} \\
&= (-1)^{n+1}\sigma_k \sum_{\substack{i\neq k\\ 1\leq i\leq n}} \frac{\sigma_i^{n-1}\left(1+\frac{s}{\sigma_i}\right)^{-\alpha}}{(\sigma_i-\sigma_k)\prod_{\substack{j\neq i\\ 1\leq j\leq n}}(\sigma_j-\sigma_i)}. \qquad \text{(B.5)}
\end{aligned}
$$

By Properties (2.2) and (2.3) of the divided differences in Chiragiev and Landsman (2007),

$\sum_{i=1}^{n} \frac{1}{\prod_{j\neq i}\left(\frac{1}{\sigma_i}-\frac{1}{\sigma_j}\right)} = 0$ and therefore $\sum_{\substack{i\neq k\\1\leq i\leq n}} \frac{1}{\prod_{\substack{j\neq i\\1\leq j\leq n}}\left(\frac{1}{\sigma_i}-\frac{1}{\sigma_j}\right)} = \frac{-1}{\prod_{\substack{j\neq k\\1\leq j\leq n}}\left(\frac{1}{\sigma_k}-\frac{1}{\sigma_j}\right)}$. Hence

$$\begin{aligned}(I) &= (-1)^{n+1}\left(\prod_{j=1}^{n}\frac{1}{\sigma_j}\right)\frac{\alpha}{\prod_{\substack{j\neq k\\1\leq j\leq n}}\left(\frac{1}{\sigma_k}-\frac{1}{\sigma_j}\right)}\left(1+\frac{s}{\sigma_k}\right)^{-(\alpha+1)} s \\ &= (-1)^{n+1}\frac{\alpha\ \sigma_k^{n-2}}{\prod_{\substack{j\neq k\\1\leq j\leq n}}(\sigma_j-\sigma_k)} s\left(1+\frac{s}{\sigma_k}\right)^{-(\alpha+1)}. \end{aligned} \tag{B.6}$$

Therefore, we can write $\int_0^s x f_{X_k,S}(x,s)\mathrm{d}x$ as

$$(-1)^{n+1}\left(\frac{\alpha\ \sigma_k^{n-2}}{\prod_{\substack{j\neq k\\1\leq j\leq n}}(\sigma_j-\sigma_k)} s\left(1+\frac{s}{\sigma_k}\right)^{-(\alpha+1)} + \sigma_k\sum_{\substack{i\neq k\\1\leq i\leq n}}\frac{\sigma_i^{n-1}\left(\left(1+\frac{s}{\sigma_i}\right)^{-\alpha}-\left(1+\frac{s}{\sigma_k}\right)^{-\alpha}\right)}{(\sigma_i-\sigma_k)\prod_{\substack{j\neq i\\1\leq j\leq n}}(\sigma_j-\sigma_i)}\right)$$

and (3.2) follows. This ends the proof.