

# Risk sharing, measuring variability, and distortion riskmetrics

Jean-Gabriel Lauzier\*      Liyuan Lin<sup>†</sup>      Ruodu Wang<sup>‡</sup>

September 10, 2025

## Abstract

We address the problem of sharing risk among agents with preferences modelled by a general class of comonotonic additive and law-invariant functionals that need not be either monotone or convex. Such functionals are called distortion riskmetrics, which include many statistical measures of risk and variability used in portfolio optimization and insurance. The set of Pareto-optimal allocations is characterized under various settings of general or comonotonic risk sharing problems. We solve explicitly Pareto-optimal allocations among agents using the Gini deviation, the mean-median deviation, or the inter-quantile difference as the relevant variability measures. The latter is of particular interest, as optimal allocations are not comonotonic in the presence of inter-quantile difference agents; instead, the optimal allocation features a mixture of pairwise counter-monotonic structures, showing some patterns of extremal negative dependence.

**Keywords:** Signed Choquet integrals, risk sharing, inter-quantile difference, variability measures, pairwise counter-monotonicity

---

## 1 Introduction

Anne, Bob and Carole are sharing a random financial loss. After negotiating their respective expected returns, each of them prefers to minimize a statistical measure of variability of their allocated risk. While agreeing on the distribution of the total loss, and that the variance is a poor metric of riskiness, each of them has their own favourite tool for measuring risks. Anne, as an economics student, likes the Gini deviation (GD) because of its intuitive appearance as an economic index. Bob, as a computer science student, prefers the mean-median deviation (MMD)

---

\*Department of Economics, Memorial University of Newfoundland, Canada. ✉ [jlauzier@mun.ca](mailto:jlauzier@mun.ca)

<sup>†</sup>Department of Econometrics and Business Statistics, Monash University, Australia. ✉ [liyuan.lin@monash.edu](mailto:liyuan.lin@monash.edu)

<sup>‡</sup>Department of Statistics and Actuarial Science, University of Waterloo, Canada. ✉ [wang@uwaterloo.ca](mailto:wang@uwaterloo.ca)

because it minimizes the mean absolute error. Finally, Carole, as a statistics student, finds that an inter-quantile difference (IQD) is the most representative of her preference, as she does not worry about extreme events for this particular risk.<sup>1</sup> How should Anne, Bob and Carole optimally share risks among themselves?

The reader familiar with risk sharing problems may immediately realize two notable features of such a problem. First, the preferences are not monotone, different from standard decision models in the literature. Second, and most crucially, Carole’s preference is neither convex nor consistent with second-order stochastic dominance. This alludes to the possibility of non-comonotonic Pareto-optimal allocations, in contrast to the comonotonic ones, which are well studied in the literature (e.g., Landsberger and Meilijson, 1994; Jouini et al., 2008; Carlier et al., 2012; Rüschendorf, 2013).

The GD, the MMD and the IQD are measures of distributional variability. Variability is used to characterize the concept of risk as in the classic work of Markowitz (1952) and Rothschild and Stiglitz (1970). For this reason, we also call them *riskmetrics*, which also include *risk measures* in the literature, often associated with monotonicity (e.g., Föllmer and Schied, 2016). As the most popular measure of variability, the variance is known to be a coarse metric and it does not distinguish positive and negative deviations from the mean;<sup>2</sup> Embrechts et al. (2002) discussed various flaws of using variance and correlation in financial risk management. Anne’s decision criterion has been proposed in Shalit and Yitzhaki (1984), which considers an optimal portfolio problem *à la* Markowitz (1952), but with the variance replaced by the GD.<sup>3</sup> Formally, the authors analyze the investor’s problem  $\min_X \text{GD}(X)$  subject to  $\mathbb{E}[X] \geq R$ , for a given rate  $R \geq 0$  of return proportional to the investor’s risk aversion. As with mean-variance preferences (e.g., Markowitz, 2014; Maccheroni et al., 2013), the decision criterion can thus be viewed as the problem of maximizing  $\mathbb{E}[X] - \eta \text{GD}(X)$ , for  $\eta \geq 0$  being the Lagrange multiplier of the problem. While the decision criterion  $\mathbb{E}[X] - \eta \text{GD}(X)$  seems natural, it is not monotone unless  $\eta$  is less than or equal to one, in which case the investor’s preference belongs to those of Yaari (1987). The other measures MMD and IQD also have sound foundations and long history in statistics and its applications (Yule, 1911, Chapter 6). Slightly different from MMD, Konno and Yamazak (1991) studied portfolio optimization using the mean-absolute deviation from the mean. Risk sharing problems with convex risk measures are well studied (e.g., Barrieu and El Karoui, 2005, Jouini et al., 2008, Filipović and Svindland, 2008 and Ravanelli and Svindland (2014)), but the classes of riskmetrics mentioned above do not belong to convex risk

---

<sup>1</sup>For an integrable random variable  $X$ , its GD is defined as  $\mathbb{E}[|X - X'|]$  where  $X'$  is an iid copy of  $X$ , its MMD is defined as  $\mathbb{E}[|X - m|]$  where  $m$  is a median of  $X$ , and its IQD at level  $\alpha \in (0, 1/2)$  is the difference between the  $\alpha$ -quantile and the  $(1 - \alpha)$ -quantile of  $X$ . For precise definitions, see Section 2.

<sup>2</sup>The latter criticism also applies to GD, MMD and IQD, but most of our results do not assume symmetry and can accommodate non-symmetric riskmetrics.

<sup>3</sup>The authors use the term *Gini’s mean difference*.

measures in general. When the riskmetric is a deviation measure, the risk sharing problem may become non-monotone, a feature that also appears in the model of [Markowitz \(1952\)](#).

In this paper, we address the problem of sharing risk among agents that uses *distortion riskmetrics* as their preferences. Distortion riskmetrics are evaluation functionals that are characterized by comonotonic additivity and law invariance ([Wang et al., 2020a](#)). This rich family includes many measures of risk and variability, and in particular, the mean, the GD, the MMD, the IQD, and their linear combinations. Distortion riskmetrics are closely related to Choquet integrals and rank-dependent utilities widely used in decision theory (e.g., [Yaari, 1987](#); [Schmeidler, 1989](#); [Carlier and Dana, 2003](#)); for a comprehensive treatment of distortions in decisions and economics, see [Wakker \(2010\)](#). The combination of the mean and GD or that of the mean and MMD, as well as other distortion riskmetrics, are used as premium principles in the insurance literature; see [Denneberg \(1990\)](#). Several variability measures within the class of distortion riskmetrics are studied by [Grechuk et al. \(2009\)](#), [Furman et al. \(2017\)](#) and [Bellini et al. \(2022\)](#).

While we analyze the general problem of sharing risk amongst distortion riskmetrics agents, non-monotone and non-cash-additive evaluation functionals receive greater attention. This is for a few reasons. First, the special case of sharing risk with cash-additive and law-invariant functionals is well studied, and more so when the functionals are monotone, but the general case is less understood. Second, the formalism we introduce allows us to generalize the example above and consider individuals that analyze their risks with different variability measures. This is critical because we aim to understand how the act of measuring risk differently gives rise to incentives to trade it. Also, it allows for non-monetary measurement of risk. Third, technically, relaxing monotonicity and convexity allows us to deal with maximization and minimization problems of risk in a unified framework.

The following simple example, by considering the GD and MMD agent only, illustrates the structure of a Pareto-optimal allocation as either an insurance or a financial contract.

**Example 1.** Consider the problem of sharing a random loss  $X$  between Anne ( $A$ ) and Bob ( $B$ ) only. Recall that Bob evaluates its allocation  $X_B$  using the mean-median deviation  $\text{MMD}(X_B)$ . Similarly, Anne's allocation is  $X_A$  which she evaluates with the Gini deviation  $\text{GD}(X_A)$ . We will show (in [Section 6](#)) that any Pareto-optimal allocation takes the form

$$X_A = X \wedge \ell - X \wedge d, \quad X_B = X - X_A,$$

where  $\ell \geq d$  will be specified later. We can interpret this as a situation where  $X$  is Bob's potential loss and Anne provides some degree of insurance for Bob. The contract (transfer function) is thus

simply the random variable  $X_A$ . Notice that (i) when  $\ell \geq X \geq d = 0$  there is full insurance, (ii) when  $\ell = d$  there is no insurance and (iii) for other choices of  $\ell > d$ , the contract is a simple deductible  $d$  with an upper limit  $\ell$ . Further, we show that each Pareto-optimal allocation minimizes  $\lambda\text{MMD}(X_B) + (1 - \lambda)\text{GD}(X_A)$  for some  $\lambda \in [0, 1]$ . If we interpret  $P = -X > 0$  as the future (after one period) price of an asset that Bob owns one share, then case (iii) can be achieved by Bob writing to Anne a bull call spread option with long strike price  $-l$  and short strike price  $-d$ .

The previous example is interesting because it confirms the intuition that the act of measuring risk differently leads to incentives to trade it. Yet, the “shape” of the solution above is not surprising, as both the Gini deviation and mean-median deviation are convex order consistent functionals, and so exhibit risk aversion of [Rothschild and Stiglitz \(1970\)](#). Just as in the increasing distortion case, risk-minimizing (utility-maximizing) Pareto-optimal allocations are comonotonic when the distortion riskmetrics’ distortion function is concave (convex), because concavity of the distortion function is equivalent to convex order consistency.

The situation for IQD agents like Carole is more sophisticated. The distortion function of IQD is discontinuous, non-concave, non-monotone, and takes value zero on both tails of the distribution. The preference induced by IQD does not correspond to a decision criterion typically considered in the literature, whereas the preferences induced by quantiles, called quantile maximization, have been axiomatized by [Rostek \(2010\)](#). IQD is a standard measure of dispersion used in statistics such as in box plots, and its most popular special case in statistics is the inter-quartile difference, which measures the difference between the 25% and 75% quantiles of data.

The general theory of risk sharing between agents using distortion riskmetrics is laid out in [Section 3](#). A convenient feature of distortion riskmetrics is that they are convex order consistent if and only if the distortion function is concave ([Wang et al., 2020a](#), Theorem 3). This enables the characterization of Pareto-optimal allocations for such agents using the comonotonic improvement, a notion introduced in [Landsberger and Meilijson \(1994\)](#) to characterize the optimal sharing of risk among risk-averse expected utility maximizers; see also [Ludkovski and Rüschemdorf \(2008\)](#) and [Rüschemdorf \(2013\)](#). Non-concave distortion functions lead to substantial challenges and to non-comonotonic optimal allocations, with limited recent results obtained by [Embrechts et al. \(2018\)](#) and [Weber \(2018\)](#) for some increasing distortion riskmetrics.

We study optimality within the subset of comonotonic allocations, which we refer to as the comonotonic risk sharing problem, for general distortion riskmetrics which are not necessarily convex in [Section 4](#). We show that the risk possibility set of distortion riskmetrics is always a convex set when restricted to the subset of comonotonic allocations. By the Hahn-Banach Theorem, we

can always find comonotonic Pareto-optimal allocations by optimizing a linear combination of the agents' welfare. This simple but valuable result “essentially comes for free” by the comonotonic additivity and positive homogeneity of distortion riskmetrics. In particular, it does not require the convexity of the evaluation functionals. Moreover, this comonotonic setting allows us to easily incorporate heterogeneous beliefs as in the setting of [Embrechts et al. \(2020\)](#), which we study in [Section 7](#) for the interested reader.

With IQD agents, the set of optimal allocations can dramatically differ when defined on the whole set of allocations or the subset of comonotonic ones, as shown by results in [Sections 3.2](#) and [4.2](#). We show the surprising result that Pareto-optimal allocations are precisely those which solve a sum optimality problem, which is not true for other variability measures such as GD or MMD. Closed-form Pareto-optimal allocations are obtained, which can be decomposed as the sum of two pairwise counter-monotonic allocations. This observation complements the optimal allocations for quantile agents obtained by [Embrechts et al. \(2018\)](#) which are pairwise counter-monotonic.

Combining results obtained in [Sections 3](#) and [4](#), the general problem of sharing risks between IQD agents (like Carole) and agents with concave and symmetric distortion functions (like Anne and Bob) mentioned in the beginning of the paper is solved in [Section 5](#) and further illustrated in [Section 6](#). We obtain a sum-optimal allocation which features a combination of comonotonicity and pairwise counter-monotonicity. These two structures are, respectively, regarded as extremal positive and negative dependence concepts; see [Puccetti and Wang \(2015\)](#) for an overview of these dependence concepts and [Lauzier et al. \(2023\)](#) for a stochastic representation of pairwise counter-monotonicity. More specifically, there exists an event on which the obtained Pareto-optimal allocation is comonotonic, and two events on which the sum-optimal allocation is pairwise counter-monotonic. To the best of our knowledge, this is the first article to obtain such a type of sum-optimal or Pareto-optimal allocation. Moreover, none of our results relies on continuity of the distortion functions. [Section 7](#) extends our results on comonotonic risk sharing to the problem where agents have heterogeneous beliefs. We conclude the paper in [Section 8](#) with a few remarks, and all proofs are put in the appendices.

## 2 Preliminaries

### 2.1 Distortion riskmetrics

For a measurable space  $(\Omega, \mathcal{F})$  and a finite set function  $\nu : \mathcal{F} \rightarrow \mathbb{R}$  with  $\nu(\emptyset) = 0$ , the signed *Choquet integral* of a random variable  $X : \Omega \rightarrow \mathbb{R}$  is the integral

$$\int X \, d\nu = \int_0^\infty \nu(X > x) \, dx + \int_{-\infty}^0 (\nu(X > x) - \nu(\Omega)) \, dx, \quad (1)$$

provided these integrals are finite. Let  $n$  be a positive integer and let  $[n] = \{1, \dots, n\}$ . The random variables  $X_1, \dots, X_n$  are *comonotonic* if there exists a collection of increasing functions  $f_i : \mathbb{R} \rightarrow \mathbb{R}$ ,  $i \in [n]$ , and a random variable  $Z$  such that  $X_i = f_i(Z)$  a.s. for all  $i \in [n]$ . Two random variables  $X_1, X_2$  are *counter-monotonic* if  $X_1, -X_2$  are comonotonic. The random variables  $X_1, \dots, X_n$  are *pairwise counter-monotonic* if  $X_i, X_j$  are counter-monotonic for each pair of distinct  $i, j$  ([Puccetti and Wang, 2015](#); [Lauzier et al., 2023](#)). Terms like “increasing” or “decreasing” are in the non-strict sense.

Assume that  $(\Omega, \mathcal{F}, \mathbb{P})$  is an atomless probability space where almost surely equal random variables are treated as identical. When  $X(\omega)$  appears, it does not matter which version we choose. Let  $\mathcal{X}$  be a set of random variables on this space. For simplicity, we assume throughout that  $\mathcal{X} = L^\infty$ , the set of essentially bounded random variables, and we will inform the reader when a result can be extended to larger spaces. A *distortion riskmetric*  $\rho_h$  is the mapping from  $\mathcal{X}$  to  $\mathbb{R}$ ,

$$\rho_h(X) = \int X \, d(h \circ \mathbb{P}) = \int_0^\infty h(\mathbb{P}(X > x)) \, dx + \int_{-\infty}^0 (h(\mathbb{P}(X > x)) - h(1)) \, dx, \quad (2)$$

where  $h$  is in the set  $\mathcal{H}^{\text{BV}}$  of all possibly non-monotone *distortion functions*, i.e.,

$$\mathcal{H}^{\text{BV}} = \{h : [0, 1] \rightarrow \mathbb{R} \mid h \text{ is of bounded variation and } h(0) = 0\}.$$

Distortion riskmetrics are law-invariant versions of a general Choquet integral defined with regards to (possibly non-monotone) set functions; see Theorem 4.5 of [Marinacci and Montrucchio \(2004\)](#).<sup>4</sup>

We now recall some properties of distortion riskmetrics that we use throughout. Any distortion riskmetric  $\rho_h$  always satisfies the following four properties as a function  $\rho : \mathcal{X} \rightarrow \mathbb{R}$ .

1. *Law invariance:*  $\rho(X) = \rho(Y)$  for  $X \stackrel{\text{d}}{=} Y$ .

---

<sup>4</sup>To avoid any confusion, we refrain from using the term *capacity*, as those are typically defined as positive monotone set functions that are not necessarily additive. In fact, the set functions  $h \circ \mathbb{P}$  we consider in the text need not be either positive or monotone for  $h \in \mathcal{H}^{\text{BV}}$ .

2. *Positive homogeneity*:  $\rho(\lambda X) = \lambda\rho(X)$  for all  $\lambda > 0$  and  $X \in \mathcal{X}$  with  $\lambda X \in \mathcal{X}$ .
3. *Comonotonic additivity*:  $\rho(X + Y) = \rho(X) + \rho(Y)$  whenever  $X$  and  $Y$  are comonotonic and  $X + Y \in \mathcal{X}$ .
4. *Translation invariance*:  $\rho(X + c) = \rho(X) + c\rho(1)$  for all  $c \in \mathbb{R}$  and  $X \in \mathcal{X}$  with  $X + c \in \mathcal{X}$ .

As a special case of translation invariance with  $\rho(1) = 1$ ,  $\rho$  is *cash additive* if  $\rho(X + c) = \rho(X) + c$  for  $c \in \mathbb{R}$  and  $X \in \mathcal{X}$ . For a distortion riskmetric  $\rho_h$ , cash additivity means  $h(1) = 1$ . We also say *location invariance* for  $h(1) = 0$  and *reverse cash additivity* for  $h(1) = -1$ . We note that although we use the general term “cash additivity” as in the literature of risk measures, the values of random variables may be interpreted as non-monetary, such as carbon dioxide emissions, as long as they can be transferred between agents in an additive fashion.

A distortion riskmetric  $\rho_h$  may also satisfy the following properties depending on  $h$ . A random variable  $X$  is said to be dominated by a random variable  $Y$  in *convex order*, denoted by  $X \leq_{\text{cx}} Y$ , if  $\mathbb{E}[\phi(X)] \leq \mathbb{E}[\phi(Y)]$  for every convex function  $\phi : \mathbb{R} \rightarrow \mathbb{R}$ , provided that both expectations exist (allowed to be infinite).

5. *Increasing monotonicity*:  $\rho(X) \leq \rho(Y)$  whenever  $X \leq Y$ .
6. *Convex order consistency*:  $\rho(X) \leq \rho(Y)$  whenever  $X \leq_{\text{cx}} Y$ .
7. *Subadditivity*:  $\rho(X + Y) \leq \rho(X) + \rho(Y)$  for every  $X, Y \in \mathcal{X}$ .

We also say that  $\rho$  is *monotone* if either  $\rho$  or  $-\rho$  is increasing. Increasing and cash-additive functionals are *monetary risk measure* (Föllmer and Schied, 2016) or *niveloids* (Cerrei-Vioglio et al., 2014); see also Artzner et al. (1999) for coherent risk measures and Cerrei-Vioglio et al. (2011) for quasiconvex risk measures. For a distortion riskmetric  $\rho_h$ , increasing monotonicity means that  $h$  is increasing, and either subadditivity or convex order consistency is equivalent to the concavity of  $h$  by Theorem 3 of Wang et al. (2020a).

Distortion riskmetrics are precisely all law-invariant and comonotonic-additive mappings satisfying some forms of continuity; see Wang et al. (2020b) on  $L^\infty$  and Wang et al. (2020a) on general spaces. The subset of increasing normalized distortion functions is denoted by  $\mathcal{H}^{\text{DT}}$ , that is,

$$\mathcal{H}^{\text{DT}} = \{h : [0, 1] \rightarrow \mathbb{R} \mid h \text{ is increasing, } h(0) = 0 \text{ and } h(1) = 1\}.$$

If  $h \in \mathcal{H}^{\text{DT}}$ , then  $\rho_h$  is often called a *dual utility* of Yaari (1987). Recall that a Yaari agent is strongly risk averse when the distortion function  $h$  is concave (Yaari, 1987). Hence, we slightly abuse

nomenclature and simply say that a distortion riskmetric agent is risk averse when its distortion function is concave, regardless of whether it is increasing or not. This is consistent with the concept of increasing in risk introduced by [Rothschild and Stiglitz \(1970\)](#). With risk aversion,  $\rho_h$  is a *spectral risk measure* ([Acerbi, 2002](#)) in risk management, an important class of *coherent* risk measures ([Kusuoka, 2001](#)). In insurance, it is also known as *Wang's premium principle* ([Wang, 1996](#)).

Any distortion riskmetric admits a quantile representation (Lemma 1 of [Wang et al. \(2020a\)](#); see the monotone case in Theorems 4 and 6 of [Dhaene et al. \(2012\)](#)). For a concise presentation of results, we define quantiles by counting losses *from large to small*.<sup>5</sup> Formally, for  $t \in [0, 1]$ , we define the left quantile of a random variable  $X \in \mathcal{X}$  as  $Q_t^-(X) = \inf\{x \in \mathbb{R} : \mathbb{P}(X \leq x) \geq 1 - t\}$ , and the right quantile as  $Q_t^+(X) = \inf\{x \in \mathbb{R} : \mathbb{P}(X \leq x) > 1 - t\}$ , where  $\inf \emptyset = \infty$ . Further let  $\text{ess-sup} = Q_0^-$  and  $\text{ess-inf} = Q_1^+$ . The following integrals are in the sense of Lebesgue–Stieltjes.

**Lemma 1.** *For  $h \in \mathcal{H}^{\text{BV}}$  and  $X \in \mathcal{X}$  such that  $\rho_h(X)$  is well-defined (it may take values  $\pm\infty$ ),*

(i) *if  $h$  is right-continuous, then  $\int X \, d(h \circ \mathbb{P}) = \int_0^1 Q_t^+(X) \, dh(t)$ ;*

(ii) *if  $h$  is left-continuous, then  $\int X \, d(h \circ \mathbb{P}) = \int_0^1 Q_t^-(X) \, dh(t)$ ;*

(iii) *if  $t \mapsto Q_t^-(X)$  is continuous on  $(0, 1)$ , then  $\int X \, d(h \circ \mathbb{P}) = \int_0^1 Q_t^-(X) \, dh(t) = \int_0^1 Q_t^+(X) \, dh(t)$ .*

There are two main advantages of working with non-monotone distortion functions. First, as monotonicity is not assumed, results on maxima and minima are symmetric; we only need to analyze one of them. Second, distortion riskmetrics include many more functionals in risk management, such as variability measures, which never have a monotone distortion function. Some properties of non-monotone risk functionals are studied by [Wang and Wei \(2020\)](#). We will make extensive use of three variability measures which appeared in the introduction. They are well defined on spaces larger than  $L^\infty$ , although we state our main results on  $\mathcal{X} = L^\infty$ .

The first measure of variability we use extensively is the Gini deviation (GD)

$$\text{GD}(X) = \frac{1}{2} \mathbb{E}[|X^* - X^{**}|] = \int X \, d(h_{\text{GD}} \circ \mathbb{P})$$

for  $X \in L^1$ , where  $X^*$ ,  $X^{**}$  are any independent copies of  $X$ , and  $h_{\text{GD}}(t) = t - t^2$  for  $t \in [0, 1]$ . The specific choice of  $X^*$ ,  $X^{**}$  is not relevant. Its distortion function is depicted in [Figure 1 \(a\)](#). As our

---

<sup>5</sup>It will be clear from [Theorem 2](#) that this nontraditional choice of notation significantly simplifies the presentation of several results; this is also the case in [Embrechts et al. \(2018\)](#).

second measure of variability, the mean-median deviation (MMD) is defined by

$$\text{MMD}(X) = \min_{x \in \mathbb{R}} \mathbb{E}[|X - x|] = \mathbb{E}[|X - Q_{1/2}^-(X)|] = \int X \, d(h_{\text{MMD}} \circ \mathbb{P})$$

for  $X \in L^1$  and  $h_{\text{MMD}}(t) = t \wedge (1 - t)$ ,  $t \in [0, 1]$ ; see Figure 1 (b). The mean-median deviation is sometimes called the mean (or average) absolute deviation from the median and is well known for its statistical robustness. Both the mean-median deviation and the Gini deviation have concave distortions and thus are convex order consistent. Lastly, the inter-quantile difference (IQD) is defined by

$$\text{IQD}_\alpha(X) = Q_\alpha^-(X) - Q_{1-\alpha}^+(X) = \int X \, d(h_{\text{IQD}} \circ \mathbb{P})$$

for  $X \in L^0$  and  $h_{\text{IQD}}(t) = \mathbb{1}_{\{\alpha < t < 1-\alpha\}}$ ,  $t \in [0, 1]$  and  $\alpha \in [0, 1/2)$ . See Figure 1 (c) for its distortion function. We further set  $\text{IQD}_\alpha = 0$  for  $\alpha \in [1/2, \infty)$ , but this is only for the purpose of unifying the presentation of some results. Our formulation of IQD is slightly different from the definition used by Bellini et al. (2022) where  $\text{IQD}_\alpha$  is defined as  $Q_\alpha^+ - Q_{1-\alpha}^-$ , but this difference is minor. The two definitions coincide when the quantile function is continuous. For  $X \in \mathcal{X}$  and  $\alpha \in [0, 1/2)$ , a convenient formula (see Theorem 1 of Bellini et al. (2022)) is

$$\text{IQD}_\alpha(X) = Q_\alpha^-(X) + Q_\alpha^-(-X), \tag{3}$$

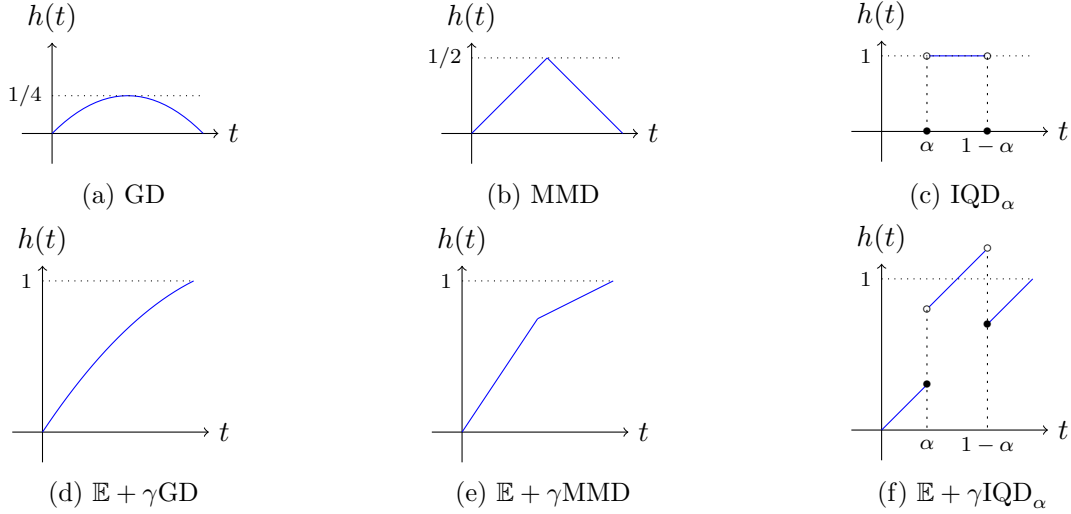
and this is due to  $Q_{1-\alpha}^+(X) = -Q_\alpha^-(-X)$ .

Consider now a preference functional  $\mathcal{I}$  of the form

$$\mathcal{I}(X) = \theta \mathbb{E}(X) + \gamma D(X)$$

for  $\theta \geq 0$ ,  $\gamma \in \mathbb{R}$  and  $D(X)$  a variability measure. The version of  $\mathcal{I}$  with  $\theta = 1$  and  $\gamma < 0$  is widely used in modern portfolio theory (as an objective to maximize). There, the random variable  $X$  denotes the gains, the parameter  $\gamma$  indicates the degree of risk aversion and  $D(X)$  is a variability measure chosen to replace the variance. This yields the ‘‘Mean- $D$ ’’ preferences nomenclature common in the literature. The version of  $\mathcal{I}$  with  $X$  being a loss,  $\theta \geq 1$  and  $\gamma \geq 0$  is common in the insurance/reinsurance literature, where it is called a distortion-deviation premium principle. For instance, Denneberg (1990) suggests the premium principle  $\theta = 1$  and  $D(X) = \text{MMD}(X)$ . The functional  $\mathcal{I}$  is a distortion riskmetric as long as  $D$  is one, and so we adopt the convention of denoting such functional by  $\rho_h$  and interpreting  $X$  as losses. Panels (d)-(f) of Figure 1 illustrate the distortion functions of  $\mathbb{E} + \gamma D$ .

Figure 1: Distortion functions for GD, MMD, IQD and  $\mathbb{E} + \gamma D$ , where  $\gamma = 1/2$



## 2.2 Risk sharing problems

There are  $n$  agents sharing a total loss  $X \in \mathcal{X}$ . Recall that, for all results, we consider  $\mathcal{X} = L^\infty$ . Agents can have different preferences, which may be due to their own statistical modelling; see e.g., [Amarante \(2022\)](#). Suppose that agent  $i \in [n]$  has a preference modelled by a distortion riskmetric  $\rho_{h_i}$  with smaller values preferred. Given  $X \in \mathcal{X}$ , we define the set of *allocations* of  $X$  as

$$\mathbb{A}_n(X) = \left\{ (X_1, \dots, X_n) \in \mathcal{X}^n : \sum_{i=1}^n X_i = X \right\}. \quad (4)$$

The *inf-convolution*  $\square_{i=1}^n \rho_{h_i}$  of  $n$  distortion riskmetrics  $\rho_{h_1}, \dots, \rho_{h_n}$  is defined as

$$\square_{i=1}^n \rho_{h_i}(X) := \inf \left\{ \sum_{i=1}^n \rho_{h_i}(X_i) : (X_1, \dots, X_n) \in \mathbb{A}_n(X) \right\}, \quad X \in \mathcal{X}.$$

That is, the inf-convolution of  $n$  distortion riskmetrics is the infimum over aggregate welfare for all possible allocations. For a general treatment of inf-convolution in risk sharing problems, see [Rüschendorf \(2013\)](#).

Let  $\rho_{h_1}, \dots, \rho_{h_n}$  be distortion riskmetrics and  $X \in \mathcal{X}$ . The allocation  $(X_1, \dots, X_n)$  is *sum optimal* in  $\mathbb{A}_n(X)$  if  $\square_{i=1}^n \rho_{h_i}(X) = \sum_{i=1}^n \rho_{h_i}(X_i)$ . An allocation  $(X_1, \dots, X_n) \in \mathbb{A}_n(X)$  is *Pareto optimal* in  $\mathbb{A}_n(X)$  if for any  $(Y_1, \dots, Y_n) \in \mathbb{A}_n(X)$  satisfying  $\rho_{h_i}(Y_i) \leq \rho_{h_i}(X_i)$  for all  $i \in [n]$ , we have  $\rho_{h_i}(Y_i) = \rho_{h_i}(X_i)$  for all  $i \in [n]$ . Note that sum-optimal allocations are always Pareto optimal.

Part of our analysis is conducted for the constrained problem where the allocations are confined

to the set of comonotonic allocations, that is,

$$\mathbb{A}_n^+(X) = \{(X_1, \dots, X_n) \in \mathbb{A}_n(X) : X_1, \dots, X_n \text{ are comonotonic}\},$$

We first make a useful observation about  $\mathbb{A}_n^+(X)$  below, which is a simple generalization of [Denneberg \(1994, Proposition 4.5\)](#) in the case of  $n = 2$ .

**Proposition 1.** *The random variables  $X_1, \dots, X_n$  are comonotonic if and only if there exist increasing functions  $f_i : \mathbb{R} \rightarrow \mathbb{R}$  such that  $X_i = f_i(\sum_{k=1}^n X_k)$  a.s.,  $i \in [n]$  and  $\sum_{i=1}^n f_i(x) = x$  for  $x \in \mathbb{R}$ .*

Proposition 1 implies that if  $(X_1, \dots, X_n) \in \mathbb{A}_n^+(X)$ , then we can set  $X = Z$  in the definition of comonotonicity while guaranteeing that  $\sum_{i=1}^n X_i = X$  a.s.

The *comonotonic inf-convolution*  $\boxplus_{i=1}^n \rho_{h_i}$  of distortion riskmetrics  $\rho_{h_1}, \dots, \rho_{h_n}$  is defined as

$$\boxplus_{i=1}^n \rho_{h_i}(X) := \inf \left\{ \sum_{i=1}^n \rho_{h_i}(X_i) : (X_1, \dots, X_n) \in \mathbb{A}_n^+(X) \right\}.$$

Let  $\rho_{h_1}, \dots, \rho_{h_n}$  be distortion riskmetrics and  $X \in \mathcal{X}$ . An allocation  $(X_1, \dots, X_n)$  is *sum optimal* in  $\mathbb{A}_n^+(X)$  when  $\boxplus_{i=1}^n \rho_{h_i}(X) = \sum_{i=1}^n \rho_{h_i}(X_i)$ . An allocation  $(X_1, \dots, X_n) \in \mathbb{A}_n^+(X)$  is *Pareto optimal* in  $\mathbb{A}_n^+(X)$  if for any  $(Y_1, \dots, Y_n) \in \mathbb{A}_n^+(X)$  satisfying  $\rho_{h_i}(Y_i) \leq \rho_{h_i}(X_i)$  for all  $i \in [n]$ , we have  $\rho_{h_i}(Y_i) = \rho_{h_i}(X_i)$  for all  $i \in [n]$ .

The set of comonotonic allocations  $\mathbb{A}_n^+(X)$  is a strict subset of the set of all possible allocations  $\mathbb{A}_n(X)$ . Hence, the sequel refers to the problem of sharing risk in  $\mathbb{A}_n(X)$  and  $\mathbb{A}_n^+(X)$  as *unconstrained* and *comonotonic* risk sharing, respectively.

### 3 Unconstrained risk sharing

This section tackles the unconstrained risk sharing problem. It is divided into two subsections. The first aims at providing general results and the second subsection characterizes the unconstrained risk sharing problem with IQD agents. There, we show that sum-optimal allocations involve pairwise counter-monotonicity, an extreme form of negative dependence between the agents' risk.

#### 3.1 Pareto optimality, sum optimality, and comonotonic improvement

In all results, we will always assume that agents have preferences modelled by  $\rho_{h_1}, \dots, \rho_{h_n}$  where  $h_1, \dots, h_n \in \mathcal{H}^{\text{BV}}$ , with one exception which will be specified clearly. The value of  $h(1)$  is

important for a distortion riskmetric  $\rho_h$  because, by translation invariance, it pins down the value attributed to a sure gain or loss.

**Proposition 2.** *Let  $X \in \mathcal{X}$ . Then*

- (i) *If a Pareto-optimal allocation in either  $\mathbb{A}_n^+(X)$  or  $\mathbb{A}_n(X)$  exists then  $h_i(1)$ ,  $i \in [n]$ , are all 0, all positive, or all negative;*
- (ii) *If  $\boxplus_{i=1}^n \rho_{h_i}(X) > -\infty$ , then  $h_1(1) = \dots = h_n(1)$ .*

The proof of Proposition 2 highlights the role of translation invariance. For (i), we thus assume by contradiction that  $(X_1, \dots, X_n)$  is Pareto optimal but that  $h_i(1)$ ,  $i \in [n]$ , are not all zero or all of the same sign. We can organize a (cash) transfer  $(c_1, \dots, c_n)$  between agents such that  $\sum_{i=1}^n c_i = 0$  and the allocation  $(X_1 + c_1, \dots, X_n + c_n)$  strictly improves upon  $(X_1, \dots, X_n)$ ; this should not be possible. This condition implies that, in order for the risk-sharing problem to be meaningful, all agents must agree on whether they like or dislike an increase of their allocation. In the former case,  $X_1, \dots, X_n$  may represent a good like monetary gains, and in the latter case, they may represent bad outcomes, like carbon dioxide emissions. For (ii), when the value of  $h(1)$  differs between agents, a similar type of transfer strictly reduces the sum of welfare  $\sum_{i=1}^n \rho_{h_i}$ , and so the value attained by the inf-convolution  $\boxplus_{i=1}^n \rho_{h_i}$  is arbitrarily small.

For  $h \in \mathcal{H}^{\text{BV}}$ , we write  $\tilde{h} = h/|h(1)|$  if  $h(1) \neq 0$  and  $\tilde{h} = h$  if  $h(1) = 0$ . If  $h(1) \neq 0$ , then  $\tilde{h}(1) = \pm 1$ . Note that replacing  $h_i$  with its normalized version  $\tilde{h}_i$  does not change the preference of agent  $i$ . Hence, we sometimes consider in our proofs distortion riskmetrics that are either all cash additive or all reverse cash additive. While this normalization does change the value attained by the inf-convolution, it is without loss of generality for characterizing Pareto optimality.

We now state our first theorem, a generalization of Proposition 1 of [Embrechts et al. \(2018\)](#) stated for monetary risk measures.

**Theorem 1.** *Suppose that  $h_i(1) \neq 0$  for some  $i \in [n]$ . An allocation  $(X_1, \dots, X_n) \in \mathbb{A}_n(X)$  is Pareto optimal in  $\mathbb{A}_n(X)$  if and only if  $\sum_{i=1}^n \rho_{\tilde{h}_i}(X_i) = \square_{i=1}^n \rho_{\tilde{h}_i}(X)$ .*

Theorem 1 states that Pareto optimality and sum optimality can be translated into each other via normalization whenever the distortion riskmetrics are not location invariant. The picture for location-invariant distortion riskmetrics is, however, drastically different, as we only have one direction. The next statement considers this setting. Its proof is straightforward and thus omitted.

**Proposition 3.** *Suppose that  $h_i(1) = 0$  for all  $i \in [n]$ . For an allocation  $(X_1, \dots, X_n) \in \mathbb{A}_n(X)$ , it holds that (i)  $\Rightarrow$  (ii):*

(i)  $\sum_{i=1}^n \lambda_i \rho_{h_i}(X_i) = \square_{i=1}^n (\lambda_i \rho_{h_i})(X)$  for some  $(\lambda_1, \dots, \lambda_n) \in (0, \infty)^n$ ;

(ii)  $(X_1, \dots, X_n)$  is Pareto optimal in  $\mathbb{A}_n(X)$ .

The weights  $(\lambda_1, \dots, \lambda_n)$  in (i) are often called *Negishi weights* (Negishi, 1960). One might be interested in the converse statement of Proposition 3, asking whether the Pareto optimality of  $(X_1, \dots, X_n)$  implies the existence of a set of  $(\lambda_1, \dots, \lambda_n) \in [0, \infty)^n \setminus \{\mathbf{0}\}$  such that  $\sum_{i=1}^n \lambda_i \rho_{h_i}(X_i) = \square_{i=1}^n (\lambda_i \rho_{h_i})(X)$ . We see in this paper that this claim holds in three cases: when agents have  $h_i(1) > 0$  or  $h_i(1) < 0$  (Theorem 1); when all agents are IQD (Theorem 2); when they have concave distortion functions (a combination of Propositions 4 and 7). However, we do not know whether this holds true for general distortion functions with  $h_1(1) = \dots = h_n(1) = 0$ ; see also the discussion after Proposition 7.

In view of Proposition 3, we say that an allocation  $(X_1, \dots, X_n)$  of  $X$  is  $\boldsymbol{\lambda}$ -optimal if

$$\square_{i=1}^n \rho_{\lambda_i h_i}(X) = \sum_{i=1}^n \rho_{\lambda_i h_i}(X_i). \quad (5)$$

where  $\boldsymbol{\lambda} = (\lambda_1, \dots, \lambda_n)$ . Clearly,  $\boldsymbol{\lambda}$ -optimality is equivalent to sum optimality when the preferences are specified as  $(\lambda_1 \rho_{h_1}, \dots, \lambda_n \rho_{h_n})$ , and conversely, sum optimality is  $(1, \dots, 1)$ -optimality. Therefore, we encounter no additional technical complications when solving either of them.

The following result follows from the well-known result of comonotonic improvement of Landsberger and Meilijson (1994)<sup>6</sup> and the fact that distortion riskmetrics are convex order consistent when the distortion functions  $h_i$  are concave (Theorem 3 of Wang et al. (2020a)). A *comonotonic improvement* of  $(X_1, \dots, X_n) \in \mathbb{A}_n(X)$  is a random vector  $(Y_1, \dots, Y_n) \in \mathbb{A}_n^+(X)$  such that  $Y_i \leq_{\text{cx}} X_i$  for all  $i \in [n]$ . Such a comonotonic improvement always exists for any  $(X_1, \dots, X_n)$ .

**Proposition 4.** *Suppose that  $h_1, \dots, h_n$  are concave. It holds that  $\square_{i=1}^n \rho_{h_i} = \boxplus_{i=1}^n \rho_{h_i}$ . Moreover, for any  $X \in \mathcal{X}$ , if there exists a Pareto-optimal allocation in  $\mathbb{A}_n(X)$ , then there exists a comonotonic Pareto-optimal allocation in  $\mathbb{A}_n(X)$ .*

Next, we prove that if  $h_1, \dots, h_n$  are strictly concave, then the set of optimal allocations in  $\mathbb{A}_n(X)$  is exactly that of those in  $\mathbb{A}_n^+(X)$ . This is because comonotonic improvements lead to a strict increase in welfare when the probability distortions  $h_i$  are strictly concave. We state this result formally in Corollary 1 as a consequence of the following ancillary lemma:

**Lemma 2.** *For two random variables  $X, Y \in \mathcal{X}$ , the following are equivalent:*

<sup>6</sup>See Rüschemdorf (2013) for this result on general spaces.

- (i)  $X \stackrel{d}{=} Y$ ;
- (ii)  $\rho_h(X) = \rho_h(Y)$  for all concave  $h \in \mathcal{H}^{\text{BV}}$ ;
- (iii)  $\rho_h(X) \leq \rho_h(Y)$  for all concave  $h \in \mathcal{H}^{\text{BV}}$ , in which the equality holds for a strictly concave  $h$ ;
- (iv)  $X \leq_{\text{cx}} Y$  and  $\rho_h(X) = \rho_h(Y)$  for a strictly concave  $h \in \mathcal{H}^{\text{BV}}$ .

**Corollary 1.** *If  $X \leq_{\text{cx}} Y$  and  $X \not\stackrel{d}{=} Y$ , then  $\rho_h(X) < \rho_h(Y)$  for any strictly concave  $h$ .*

*Remark 1.* The equivalence in Lemma 2 holds true for any random variables  $X, Y$  with finite mean, by requiring that  $\rho_h(X)$  and  $\rho_h(Y)$  are finite for the strictly concave function  $h$  in (iii) and (iv). This follows by noting that we did not use the boundedness of  $X$  and  $Y$  in the proof.

**Proposition 5.** *Suppose that  $h_1, \dots, h_n$  are strictly concave and  $X \in \mathcal{X}$ .*

- (i) *Every Pareto-optimal allocation in  $\mathbb{A}_n(X)$  is comonotonic.*
- (ii) *If for each  $i \in [n]$ ,  $h_i = a_i h_1$  for some  $a_i > 0$  then an allocation is Pareto optimal in  $\mathbb{A}_n(X)$  if and only if it is comonotonic.*

Both Proposition 4 and 5 can easily be extended to  $L^p$  for  $p \geq 1$  instead of  $\mathcal{X} = L^\infty$  provided that every  $\rho_{h_i}$  is finite when defined on  $L^p$ .<sup>7</sup>

### 3.2 IQD agents and negatively dependent optimal allocations

We characterize the sum-optimal allocations on general spaces when agents evaluate their risk with the IQD measure of variability. We start with the problem of sharing risk among IQD agents only. In this setting, agent  $i \in [n]$  has  $\text{IQD}_{\alpha_i}$  as their preference where  $\alpha_i \in [0, 1/2)$ .

For a random variable  $X$  on the probability space  $(\Omega, \mathcal{F}, \mathbb{P})$ , we define tail events as in Wang and Zitikis (2021). For  $\beta \in [0, 1]$ , we say that an event  $A \in \mathcal{F}$  is a *right (resp. left)  $\beta$ -tail event* of  $X$  if  $\mathbb{P}(A) = \beta$  and  $X(\omega) \geq X(\omega')$  (resp.  $X(\omega) \leq X(\omega')$ ) holds for a.s. all  $\omega \in A$  and  $\omega' \in A^c$ , where  $A^c$  stands for the complement of  $A$ . One could also follow the definition of low tail-event in Jouini et al. (2008) to define the  $\beta$ -tail event; that is,  $A$  a right (resp. left)  $\beta$ -tail event if  $\mathbb{P}(A) = \beta$  and  $\text{ess-inf}_A(X) \geq \text{ess-sup}_{A^c}(X)$  (resp.  $\text{ess-sup}_A(X) \leq \text{ess-inf}_{A^c}(X)$ ), where  $\text{ess-inf}_A(X) = \sup\{m \in \mathbb{R} : \mathbb{P}(X \geq m | A) = 1\}$  and  $\text{ess-sup}_A(X) = \inf\{m \in \mathbb{R} : \mathbb{P}(X \leq m | B) = 1\}$ . The two definitions are equivalent.

**Theorem 2.** *For  $X \in \mathcal{X}$  and the IQD risk sharing problem in  $\mathbb{A}_n(X)$  with  $\alpha_1, \dots, \alpha_n \in [0, 1/2)$ , let  $\alpha = \sum_{i=1}^n \alpha_i$ .*

<sup>7</sup>Conditions for the finiteness of  $\rho_h$  on  $L^p$  are provided in Proposition 1 of Wang et al. (2020a).

(i) An allocation of  $X$  is Pareto optimal if and only if it is sum optimal.

(ii) For  $\lambda_1, \dots, \lambda_n \geq 0$  and  $\lambda = \bigwedge_{i=1}^n \lambda_i$ ,

$$\square_{i=1}^n (\lambda_i \text{IQD}_{\alpha_i}) = \left( \bigwedge_{i=1}^n \lambda_i \right) \text{IQD}_{\sum_{i=1}^n \alpha_i} = \lambda \text{IQD}_{\alpha}. \quad (6)$$

In particular,  $\square_{i=1}^n \text{IQD}_{\alpha_i} = \text{IQD}_{\alpha}$ .

(iii) A class of Pareto-optimal allocations of  $X \in \mathcal{X}$  for IQD agents is given by

$$X_i = (X - c) \mathbb{1}_{A_i \cup B_i} + a_i (X - c) (1 - \mathbb{1}_{A \cup B}) + c_i, \quad i \in [n], \quad (7)$$

where, by setting  $\beta = \alpha \wedge (1/2)$ ,

(a)  $\{A_i\}_{i=1}^n$  and  $\{B_i\}_{i=1}^n$  are partitions of a right  $\beta$ -tail event  $A$  and a left  $\beta$ -tail event  $B$  of  $X$  with  $A, B$  disjoint, respectively, satisfying  $\mathbb{P}(A_i) = \mathbb{P}(B_i) = \alpha_i \beta / \alpha$ ,  $i \in [n]$ ;

(b)  $a_i \geq 0$  for  $i \in [n]$  and  $\sum_{i=1}^n a_i = 1$ ;

(c)  $c \in [Q_{1/2}^-(X), Q_{1/2}^+(X)]$  and  $\sum_{i=1}^n c_i = c$ .

*Remark 2.* The allocation (7) satisfies  $\sum_{i=1}^n \lambda_i \text{IQD}_{\alpha_i}(X_i) = \square_{i=1}^n (\lambda_i \text{IQD}_{\alpha_i})(X)$  by setting  $a_i = 0$  for  $i \in [n]$  such that  $\lambda_i > \lambda$ .

The surprising ingredient of Theorem 2, part (i) is that, for IQD agents, sum optimality is indeed equivalent to Pareto optimality, which is the case for cash-additive distortion riskmetrics (Theorem 1). However, for general agents with  $h_1(1) = \dots = h_n(1) = 0$ , Pareto optimality is not necessarily equivalent to sum optimality, because different choices of  $(\lambda_1, \dots, \lambda_n)$  in Proposition 3 lead to different Pareto-optimal allocations, which are not necessarily sum optimal (see Proposition 3 as well as Section 6). As a consequence of this result, Pareto-optimal allocations for IQD agents are precisely those for agents using the mean-risk preferences with risk measured by IQD,

$$\rho_{h_i}(X_i) = \mathbb{E}[X_i] + \text{IQD}_{\alpha_i}(X_i), \quad i \in [n],$$

because both solve the same sum optimality problem by noting that  $\sum_{i=1}^n \mathbb{E}[X_i] = \mathbb{E}[X]$  for any allocation  $(X_1, \dots, X_n)$  of  $X$ .

In part (ii) of Theorem 2, we see that the inf-convolution of several IQD is an IQD. Related to this observation, Embrechts et al. (2018) showed that the inf-convolution of several quantiles is again a quantile.

Figure 2: A Pareto-optimal allocation in (7), where the shaded area represents the allocation to agent 1 minus  $c_1$ , that is,  $X_1 - c_1 = (X - c)\mathbb{1}_{A_i \cup B_i} + a_i(X - c)\mathbb{1}_{(A \cup B)^c}$

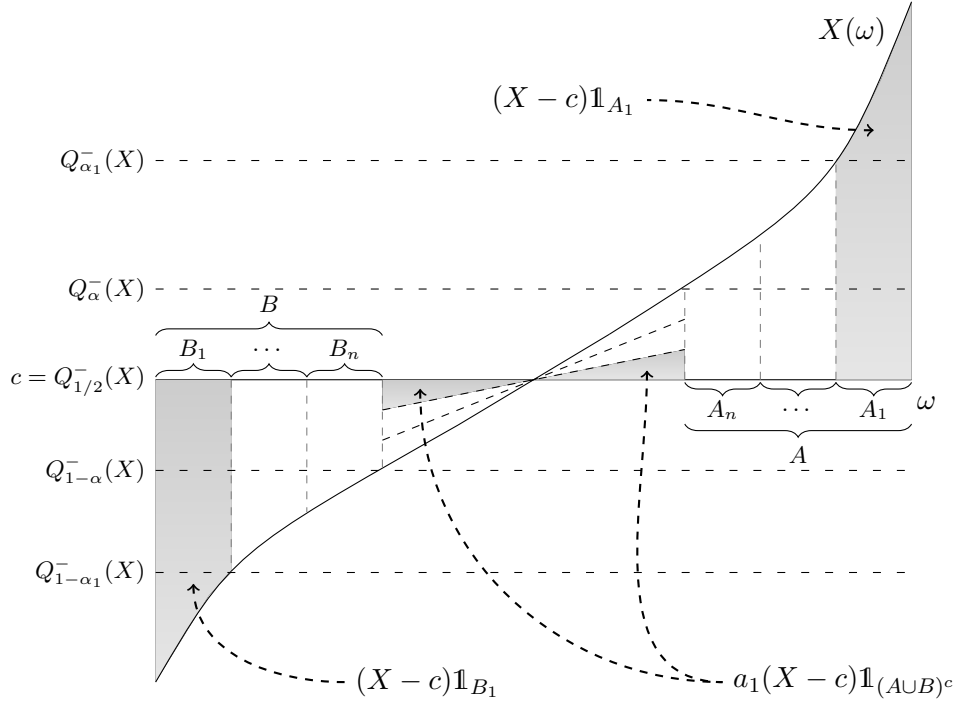


Figure 2 illustrates an example of the Pareto-optimal allocation (7) in Theorem 2, part (iii). The dependence structure of this allocation warrants some further discussion. Without loss of generality, assume  $c_1 = \dots = c_n = 0$  (this implies that a median of  $X$  is  $c = 0$ ), and assume that  $X$  is continuously distributed. Note that (a.s.)  $X > 0$  if event  $A$  occurs and  $X < 0$  if event  $B$  occurs.

First, suppose  $\alpha \geq 1/2$  so that  $\mathbb{P}(A \cup B) = 1$ . In this case, we have  $X_i = X\mathbb{1}_{A_i \cup B_i}$  for  $i \in [n]$ . The random vector  $(X\mathbb{1}_{A_i}, X\mathbb{1}_{A_j})$  for  $i \neq j$  is counter-monotonic because  $A_i \cap A_j = \emptyset$  and  $X > 0$  on  $A$ . This implies  $(X\mathbb{1}_{A_1}, \dots, X\mathbb{1}_{A_n})$  is *pairwise counter-monotonic*. From the above analysis, we can see that conditional on  $A$ ,  $(X_1, \dots, X_n)$  is pairwise counter-monotonic, and so is it conditional on  $B$ ; that is  $(X_1, \dots, X_n)$  is a mixture of two pairwise counter-monotonic vectors. Moreover,  $(X_1, \dots, X_n)$  is also the sum of the two pairwise counter-monotonic vectors  $(X\mathbb{1}_{A_1}, \dots, X\mathbb{1}_{A_n})$  and  $(X\mathbb{1}_{B_1}, \dots, X\mathbb{1}_{B_n})$ . We can check that  $(X_i(\omega) - X_j(\omega))(X_i(\omega') - X_j(\omega')) < 0$  a.s. for  $\omega \in A_i$  and  $\omega' \in A_j$ , and  $(X_i(\omega) - X_j(\omega))(X_i(\omega') - X_j(\omega')) > 0$  a.s. for  $\omega \in A_i$  and  $\omega' \in B_j$ . Therefore, the allocation  $(X_1, \dots, X_n)$  is not comonotonic, yet it is not pairwise counter-monotonic either. This is illustrated by the “vertical slicing” in Figure 2, where on  $A$  and  $B$  pairwise counter-monotonicity holds.

As discussed above, we can describe  $(X_1, \dots, X_n)$  as either the sum or the mixture of two

pairwise counter-monotonic vectors. Pairwise counter-monotonicity is a form of extreme negative dependence that extends the concept of counter-monotonicity to the case of  $n \geq 3$  agents; see [Puccetti and Wang \(2015\)](#) and [Lauzier et al. \(2023\)](#) for more details. This observation is in contrast to the optimal allocations for quantile agents in Theorem 1 of [Embrechts et al. \(2018\)](#), which are indeed pairwise counter-monotonic.

If  $0 < \alpha < 1/2$ , then the term  $a_i X \mathbb{1}_{(A \cup B)^c}$  appears in the allocation of every agent. Note that conditional on  $(A \cup B)^c$ ,  $(X_1, \dots, X_n)$  becomes comonotonic. This is illustrated by “proportional slicing” in the middle part of Figure 2. This local comonotonicity will become crucial in Section 5, where we study the risk sharing problem among several IQD agents and risk-averse agents. The lack of global comonotonicity but having some local comonotonicity is similar to the shape of optimal allocations obtained by [Liebrich \(2024\)](#), but in the latter setting this phenomenon is generated by the heterogeneity of the reference measures. A similar pattern appears in [Weber \(2018\)](#).

As hinted by Propositions 4 and 5, solving Pareto-optimal allocations for risk-averse agents requires us to study comonotonic risk sharing, which is the topic of the next section.

## 4 Comonotonic risk sharing

We now turn to the important case of comonotonic risk sharing. As before, we first provide theoretical results and then proceed to analyze further the special case of sharing risks with IQD agents.

### 4.1 Pareto optimality, sum optimality, and explicit allocations

The next result is similar to Theorem 1, but for comonotonic risk sharing. We omit its proof because it does not provide new insights.

**Proposition 6.** *Suppose that  $h_i(1) \neq 0$  for some  $i \in [n]$ . Then,  $(X_1, \dots, X_n) \in \mathbb{A}_n^+(X)$  is Pareto optimal in  $\mathbb{A}_n^+(X)$  if and only if  $\sum_{i=1}^n \rho_{\tilde{h}_i}(X_i) = \boxplus_{i=1}^n \rho_{\tilde{h}_i}(X)$ .*

We now show that  $\lambda$ -optimality in  $\mathbb{A}_n^+(X)$  pins down Pareto optimality. This result is stated in a stronger form than Proposition 3 for the corresponding notions of optimality in  $\mathbb{A}_n(X)$ .

**Proposition 7.** *Suppose that  $h_i(1) = 0$  for all  $i \in [n]$ . For an allocation  $(X_1, \dots, X_n) \in \mathbb{A}_n^+(X)$ , it holds that (i) $\Rightarrow$ (ii) $\Rightarrow$ (iii):*

$$(i) \sum_{i=1}^n \lambda_i \rho_{h_i}(X_i) = \boxplus_{i=1}^n (\lambda_i \rho_{h_i})(X) \text{ for some } (\lambda_1, \dots, \lambda_n) \in (0, \infty)^n;$$

$$(ii) (X_1, \dots, X_n) \text{ is Pareto optimal in } \mathbb{A}_n^+(X);$$

(iii)  $\sum_{i=1}^n \lambda_i \rho_{h_i}(X_i) = \boxplus_{i=1}^n (\lambda_i \rho_{h_i})(X)$  for some  $(\lambda_1, \dots, \lambda_n) \in [0, \infty)^n \setminus \{\mathbf{0}\}$ .

Comonotonicity plays an important role in the proof of Proposition 7. The risk possibility set of the set of comonotonic allocations is defined as  $S = \{(\rho_{h_1}(X_1), \dots, \rho_{h_n}(X_n)) : (X_1, \dots, X_n) \in \mathbb{A}_n^+(X)\}$ . The comonotonic additivity of distortion riskmetrics guarantees that the risk possibility set  $S$  is a convex set when restricted to  $\mathbb{A}_n^+(X)$ . This needs not be true on  $\mathbb{A}_n(X)$ . In this case, we cannot use the Hahn-Banach Theorem to obtain the existence of the Negishi weights  $(\lambda_1, \dots, \lambda_n)$ , which is the reason why we did not state a “converse statement” in Proposition 3. Propositions 4 and 7 together yield that if all agents have concave distortion functions, then any Pareto-optimal allocation in  $\mathbb{A}_n(X)$ , which yields the same welfare for all agents as a Pareto-optimal allocation in  $\mathbb{A}_n^+(X)$ , must satisfy (iii). If their distortion functions are strictly concave, then by Proposition 5, every Pareto-optimal allocation can be found through an inf-convolution.

We now aim to characterize further the set of Pareto-optimal allocations in  $\mathbb{A}_n^+(X)$ . The following result generalizes Proposition 5 of Embrechts et al. (2018) for dual utilities.

**Theorem 3.** *Suppose that  $h_1(1) = \dots = h_n(1)$ . Then*

$$\boxplus_{i=1}^n \rho_{h_i} = \rho_{h_\wedge},$$

where  $h_\wedge(t) = \min\{h_1(t), \dots, h_n(t)\}$ , and  $\rho_{h_\wedge}$  is finite on  $\mathcal{X}$ . Moreover, a sum-optimal allocation  $(X_1, \dots, X_n)$  in  $\mathbb{A}_n^+(X)$  is given by  $X_i = f_i(X)$ ,  $i = 1, \dots, n$ , where

$$f_i(x) = \int_0^x g_i(t) dt, \quad \text{and} \quad g_i(x) = \frac{1}{|M_x|} \mathbf{1}_{\{i \in M_x\}}, \quad x \in \mathbb{R}, \quad (8)$$

and where  $M_x = \{j \in [n] : h_j(\mathbb{P}(X > x)) = h_\wedge(\mathbb{P}(X > x))\}$ . The sum-optimal allocation is unique up to constant shifts almost surely if and only if  $|M_x| = 1$  for  $\mu_X$ -almost every  $x$ , where  $\mu_X$  is the distribution measure of  $X$ .

A key step in the proof of Theorem 3 is the following lemma, which gives a convenient alternative formula for  $\rho_h(f(X))$ . The lemma generalizes Lemma 2.1 of Cheung and Lo (2017) for dual utilities in the context of optimal reinsurance design.

**Lemma 3.** *For any  $h \in \mathcal{H}^{\text{BV}}$ , random variable  $X$  bounded from below, and increasing Lipschitz function  $f$  with right-derivative  $g$ , we have*

$$\rho_h(f(X)) = \int_0^\infty g(x) h(\mathbb{P}(X > x)) dx + \int_{-\infty}^0 g(x) (h(\mathbb{P}(X > x)) - h(1)) dx. \quad (9)$$

The results in Theorem 3 can be extended to domains like  $\{X \in L^p : X_- \in L^\infty\}$  for  $p \geq 0$  as long as  $\rho_{h_1}, \dots, \rho_{h_n}$  are finite on this domain. This is because Lemma 3 only requires boundedness from below. The next example illustrates the uniqueness statement in Theorem 3, which gives not only unique sum-optimal allocations in  $\mathbb{A}_n^+(X)$ , but also unique Pareto-optimal ones, up to constant shifts.

**Example 2.** Suppose that  $\rho_{h_1} = \beta_1\mathbb{E} + \gamma_1\text{GD}$ ,  $\rho_{h_2} = \beta_2\mathbb{E} + \gamma_2\text{MMD}$  and  $\rho_{h_3} = \beta_3\mathbb{E} + \gamma_3\text{IQD}_\alpha$  for some  $\beta_i, \gamma_i > 0$ ,  $i = 1, 2, 3$ , and  $\alpha \in [0, 1/2)$ . For any continuously distributed  $X \in \mathcal{X}$ , the Pareto-optimal allocation in  $\mathbb{A}_3^+(X)$  is unique up to constant shifts. To see this, by Proposition 6, any Pareto-optimal allocation  $(X_1, X_2, X_3)$  in  $\mathbb{A}_3^+(X)$  satisfies  $\sum_{i=1}^3 \rho_{\tilde{h}_i}(X_i) = \boxplus_{i=1}^3 \rho_{\tilde{h}_i}(X)$ . Noting that for each  $1 \leq i < j \leq 3$ ,  $\tilde{h}_i(t) = \tilde{h}_j(t)$  for at most two points  $t \in (0, 1)$ , by Theorem 3, the allocation  $(X_1, X_2, X_3)$  is unique up to constant shifts.

By replacing  $h_i$  with  $\lambda_i h_i$  for some  $\lambda_i \geq 0$ , we obtain the following corollary, which helps to identify  $\lambda$ -optimal allocations in conjunction with Theorem 3.

**Corollary 2.** *Let  $\lambda \in \mathbb{R}_+^n \setminus \{\mathbf{0}\}$  be a vector and  $\boxplus_{i=1}^n \rho_{\lambda_i h_i}$  be finite. Then  $\boxplus_{i=1}^n \rho_{\lambda_i h_i} = \rho_{h_\lambda}$ , where  $h_\lambda(t) = \min\{\lambda_1 h_1(t), \dots, \lambda_n h_n(t)\}$  for  $t \in [0, 1]$ .*

By Proposition 2, the inf-convolution  $\boxplus_{i=1}^n \rho_{\lambda_i h_i}$  being finite implies that  $\lambda_i h_i(1)$  are equal for all  $i \in [n]$ . Corollary 2 is thus only useful for the case of location-invariant distortion riskmetrics ( $h_i(1) = 0$ ,  $i \in [n]$ ), as otherwise we simply normalize  $\lambda_i = 1$ ,  $i \in [n]$ . Theorem 3's characterization of  $\lambda$ -optimality in  $\mathbb{A}_n^+(X)$  extends to location-invariant distortion riskmetrics by setting  $M_x = \{i \in [n] : \lambda_i h_i(\mathbb{P}(X > x)) = h_\lambda(\mathbb{P}(X > x))\}$  in (8).

For cash-additive and reverse cash-additive distortion riskmetrics, Proposition 6 and Theorem 3 together yield a full characterization of Pareto-optimal allocations in  $\mathbb{A}_n^+$ . It remains to characterize those for location-invariant distortion riskmetrics. The next proposition gives an answer for a large class of such distortion riskmetrics.

**Proposition 8.** *Suppose  $h_i(1) = 0$  and  $h_i(t) > 0$  for all  $i \in [n]$  and all  $t \in (0, 1)$ . Then the allocation  $(X_1, \dots, X_n) \in \mathbb{A}_n^+(X)$  is Pareto optimal if and only if there exists  $K \subseteq [n]$  and a vector  $\lambda \in (0, \infty)^K$  such that  $(X_i)_{i \in K}$  solves  $\boxplus_{i \in K} \rho_{\lambda_i h_i}(X)$ , and  $X_i$ ,  $i \notin K$  are constants.*

The assumption that  $h_i(t) > 0$  for all  $i \in [n]$  and all  $t \in (0, 1)$  is critical for the characterization of Proposition 8. This condition has a natural interpretation, as it is equivalent to  $\rho_{h_i}(X) > 0$  for all non-degenerate  $X$  and it is satisfied by many variability measures; it is part of the definition of deviation measures of Rockafellar et al. (2006). But this assumption rules out IQD, which we study in the next section.

## 4.2 IQD agents and positively dependent optimal allocations

We start with the comonotonic risk sharing problem among IQD agents. The following proposition gives the corresponding statements, parallel to Theorem 2, on Pareto optimality and inf-convolution in this setting. The sum-optimal allocations are given by Theorem 3.

**Proposition 9.** *Consider  $X \in \mathcal{X}$  and the IQD risk sharing problem in  $\mathbb{A}_n^+(X)$  with  $\alpha_1, \dots, \alpha_n \in [0, 1/2)$ .*

(i) *An allocation of  $X$  is Pareto optimal if and only if it is sum optimal.*

(ii) *For  $\lambda_1, \dots, \lambda_n \geq 0$ ,*

$$\boxplus_{i=1}^n (\lambda_i \text{IQD}_{\alpha_i}) = \left( \bigwedge_{i=1}^n \lambda_i \right) \text{IQD}_{\bigvee_{i=1}^n \alpha_i}.$$

*In particular,  $\boxplus_{i=1}^n \text{IQD}_{\alpha_i} = \text{IQD}_{\bigvee_{i=1}^n \alpha_i}$ .*

Comparing Theorem 2 with Proposition 9, we note that for  $\alpha_1, \dots, \alpha_n \in (0, 1/2)$ , we have  $\sum_{i=1}^n \alpha_i > \bigvee_{i=1}^n \alpha_i$ , which implies that

$$\boxplus_{i=1}^n (\lambda_i \text{IQD}_{\alpha_i})(X) - \boxminus_{i=1}^n (\lambda_i \text{IQD}_{\alpha_i})(X) > 0 \quad (10)$$

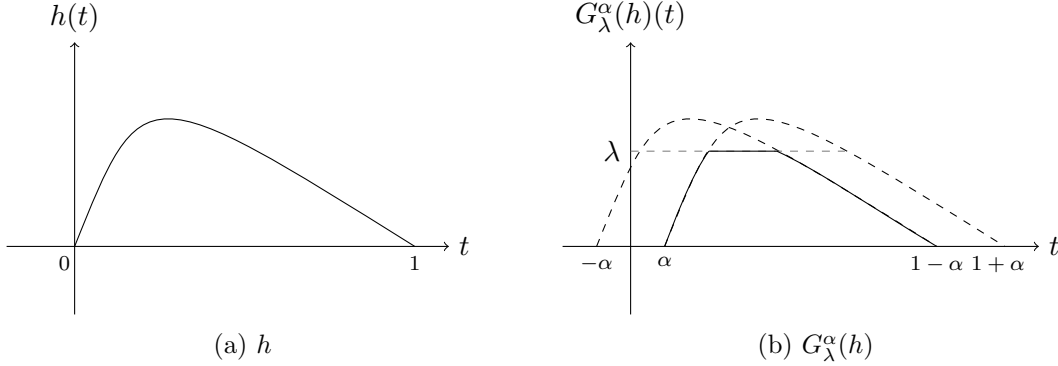
for any continuously distributed  $X$ . This further implies that the Pareto-optimal allocations in  $\mathbb{A}_n(X)$  are disjoint from those in  $\mathbb{A}_n^+(X)$ . The difference in (10) can be interpreted as the welfare gain of allowing agents to share risks in non-comonotonic arrangements. For IQD agents, comonotonic allocations are not Pareto optimal in general, and therefore, some form of “gambling behaviour” in the allocation is beneficial to all agents, although the agents are neither risk averse nor risk seeking in the sense of Rothschild and Stiglitz (1970). This is similar to the case of quantile agents, discussed by Embrechts et al. (2018).

## 5 Several IQD and risk-averse agents

Combining results established in Sections 3 and 4, we are now able to tackle the unconstrained risk sharing problem for IQD and risk-averse agents. We consider agents from the following two sets: the IQD agents, modelled by distortion functions in

$$\mathcal{H}^{\text{IQD}} = \{t \mapsto \mathbb{1}_{\{\alpha < t < 1-\alpha\}} : \alpha \in [0, 1/2)\}$$

Figure 3: An illustration of the transform  $G_\lambda^\alpha$



and the risk-averse agents, modelled by distortion functions in

$$\mathcal{H}^C = \{h \in \mathcal{H}^{\text{BV}} \mid h(1) = 0, h \text{ is concave}\}.$$

That is,  $\mathcal{H}^{\text{IQD}}$  is the set of all distortion functions for IQD variability measures and  $\mathcal{H}^C$  is the set of location-invariant concave distortion functions  $h \in \mathcal{H}^{\text{BV}}$ . Notice that each  $h \in \mathcal{H}^C$  is increasing in  $[0, s]$  and decreasing in  $[s, 1]$  for some  $s \in (0, 1)$ . Define the mapping  $G_\lambda^\alpha : \mathcal{H}^C \rightarrow \mathcal{H}^{\text{BV}}$  for  $\alpha \in [0, 1/2)$  and  $\lambda \geq 0$  as

$$G_\lambda^\alpha(h)(t) = (h(t - \alpha) \wedge h(t + \alpha) \wedge \lambda) \mathbb{1}_{\{\alpha < t < 1 - \alpha\}} \quad \text{for } t \in [0, 1].$$

The mapping  $G_\lambda^\alpha$  transforms a concave distortion function to another distortion function with value 0 on  $[0, \alpha] \cup [1 - \alpha, 1]$ . See Figure 3 for an illustration of this transform. For  $\alpha \geq 1/2$ , we define  $G_\lambda^\alpha(h) = 0$ .

We will see in the next proposition that the function  $G_\lambda^\alpha$  plays an important role because of the inf-convolution of  $\lambda \text{IQD}_\alpha$  and  $\rho_h$  for  $h \in \mathcal{H}^C$  satisfies

$$(\lambda \text{IQD}_\alpha) \square \rho_h = \rho_{G_\lambda^\alpha(h)}.$$

This formula is a special case of (11) in Theorem 4 below.

**Theorem 4.** *Let  $N \subseteq [n]$  and  $I = [n] \setminus N$ . Suppose that  $h_i \in \mathcal{H}^C$  for  $i \in N$  and  $h_i \in \mathcal{H}^{\text{IQD}}$  for  $i \in I$  with IQD parameter  $\alpha_i$ . Denote by  $\alpha = \sum_{i \in I} \alpha_i$ .*

(i) For  $\lambda_1, \dots, \lambda_n \geq 0$ , denoting by  $\lambda = \bigwedge_{i \in I} \lambda_i$  and  $h = \bigwedge_{i \in N} (\lambda_i h_i)$ , we have

$$\bigsqcup_{i=1}^n (\lambda_i \rho_{h_i}) = \rho_{G_\lambda^\alpha(h)}. \quad (11)$$

(ii) A Pareto-optimal allocation is given by

$$X_i = (X - c) \mathbf{1}_{A_i \cup B_i} + Y_i + c_i, \quad (12)$$

where, by denoting by  $\beta = \alpha \wedge (1/2)$ ,

(a)  $\{A_i\}_{i=1}^n$  and  $\{B_i\}_{i=1}^n$  are partitions of a right  $\beta$ -tail event  $A$  and a left  $\beta$ -tail event  $B$  of  $X$  with  $A, B$  disjoint, respectively, satisfying  $\mathbb{P}(A_i) = \mathbb{P}(B_i) = \alpha_i \beta / \alpha$  for  $i \in I$  and  $A_i = B_i = \emptyset$  for  $i \in N$ ;

(b)  $(Y_1, \dots, Y_n)$  is a Pareto-optimal allocation of  $(X - c) \mathbf{1}_{(A \cup B)^c}$  for preferences with distortion functions  $\hat{h}_1, \dots, \hat{h}_n$  where  $\hat{h}_i = h_i$  if  $i \in N$  and  $\hat{h}_i(t) = \mathbf{1}_{\{t \in (0, 1)\}}$  for  $i \in I$ .

(c)  $c \in [Q_{1/2}^-(X), Q_{1/2}^+(X)]$  and  $\sum_{i=1}^n c_i = c$ .

Similarly to the allocation in Theorem 2, the allocation is counter-monotonic when conditioning on tail events. On the tails, the risk is allocated to only one IQD agent at a time. Once again, this is because IQD agents do not care about tail risks. What is left of the risk is then distributed optimally among agents using the techniques introduced in Section 4. The resulting allocation is quite unusual, but it can be implemented in financial markets through derivatives such as call options and digital options. Section 6.2 further analyzes the allocations found in Theorem 4 and gives explicit examples of the underlying financial contracts.

*Remark 3.* Let  $N \subseteq [n]$  and  $I = [n] \setminus N$ . Suppose that  $h_i \in \mathcal{H}^C$  for  $i \in N$  and  $h_i \in \mathcal{H}^{\text{IQD}}$  for  $i \in I$  with IQD parameter  $\alpha_i$ . For any  $\lambda_1, \dots, \lambda_n \geq 0$  it is  $\bigsqcup_{i=1}^n (\lambda_i \rho_{h_i}) = \rho_{h_\lambda}$ , where  $h_\lambda = \bigwedge_{i \in [n]} \lambda_i h_i$ . The distortion function  $h_\lambda$  takes value 0 on  $[0, \bigvee_{i \in I} \alpha_i] \cup [\bigvee_{i \in I} \alpha_i, 1]$ ; on the other hand, the distortion function  $G_\lambda^\alpha(h)$  from Theorem 4 takes value 0 on  $[0, \sum_{i \in I} \alpha_i] \cup [\sum_{i \in I} \alpha_i, 1]$ .

## 6 GD, MMD and IQD agents

We now provide examples of the results obtained in Section 3 and 4. Some calculation details are put in Appendix E. The following two subsections come back to the risk sharing problem with several IQD agents and explains further the allocations found in Section 3.2. The last two

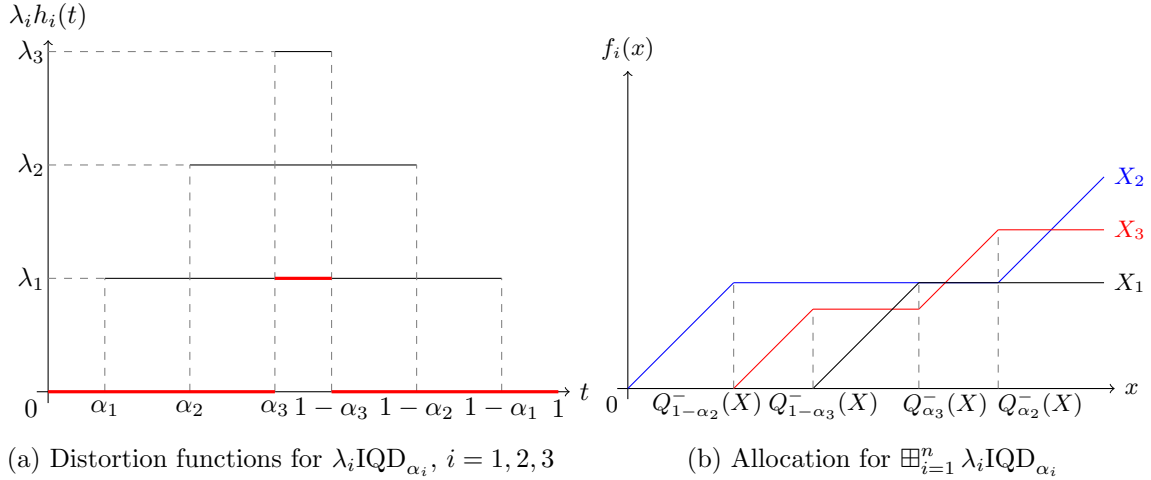
subsections analyze the risk sharing problem when agents consider the Gini and mean-median deviations as the relevant statistical measures of risk.

## 6.1 Several IQD agents

The difference between the two sum-optimal allocations found in Theorem 2 and Proposition 9 is important.

In contrast, Figure 4 illustrates some comonotonic allocations that are  $\lambda$ -optimal (and also Pareto optimal and sum optimal; see Proposition 9) when restricted to the subset  $\mathbb{A}_n^+(X)$ . The solution for  $\boxplus_{i=1}^n (\lambda_i \text{IQD}_{\alpha_i})$  is not unique as  $|M_x|$  can be larger than 1. The figure depicts a particular case when simultaneously  $\alpha_1 < \alpha_2 < \alpha_3$  and  $\lambda_1 < \lambda_2 < \lambda_3$ . The left panel shows the distortion function of each agent multiplied by the corresponding  $\lambda$ , and the lower envelope  $h_\lambda(t)$ . Figure 4(b) presents a sum-optimal allocation where all three agents take non-zero risks. Comonotonic sum-optimal allocations are not unique, because the allocation where agent 3 takes all risks in the  $\alpha_3$ -tails and agent 1 takes the rest is also sum optimal. As discussed before, comonotonic sum-optimal allocations are generally not sum optimal in  $\mathbb{A}_n(X)$ .

Figure 4: Distortion functions and the sum-optimal allocation for  $\boxplus_{i=1}^n \lambda_i \text{IQD}_{\alpha_i}$



## 6.2 The GD, MMD, and IQD problem

We now turn to the allocations characterized by Theorem 4. Consider the problem of sharing risk between Anne, Bob and Carole, i.e., the case when there is only one GD agent, one MMD and IQD agent. Let  $\alpha < 1/2$  and  $\lambda_1, \lambda_2, \lambda_3 > 0$  and consider the inf-convolution

$$\inf_{(X_1, X_2, X_3) \in \mathbb{A}(X)} \{ \lambda_1 \text{GD}(X_1) + \lambda_2 \text{MMD}(X_2) + \lambda_3 \text{IQD}_\alpha(X_3) \}.$$

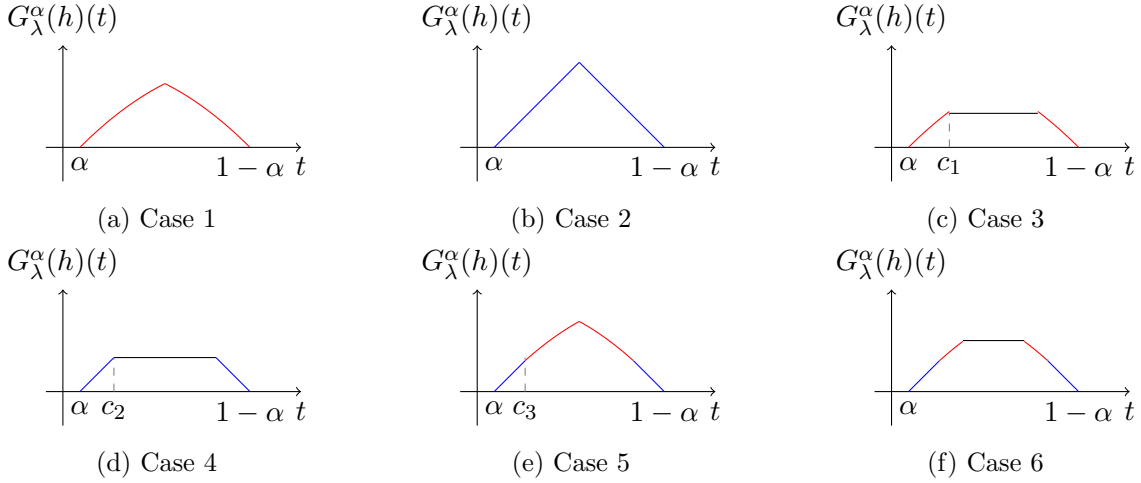
Without loss of generality we assume  $Q_{1/2}^-(X) = 0$  for the convenience of presentation, so that  $c$  in Theorem 4 is taken as 0.

Let  $A$  be a right  $\alpha$ -tail event and  $B \subseteq A^c$  be a left  $\alpha$ -tail event of  $X$ , where  $A$  and  $B$  are disjoint sets. All the  $\alpha$ -tail risks must go to the IQD agent. That is, every sum-optimal allocation requires that the IQD agent takes the whole risk on  $A \cup B$ .

It remains to share risk “in the middle”, that is, on the event  $(A \cup B)^c$ . We denote by  $Y = X\mathbb{1}_{(A \cup B)^c}$ , which has an optimal allocation  $(Y_1, \dots, Y_n)$  in Theorem 4 which is comonotonic on  $(A \cup B)^c$ . This is done in the same fashion as we do later for comonotonic risk sharing, with the caveat that the IQD agent might take on some risk depending on the weights  $\lambda_1, \lambda_2$  and  $\lambda_3$ . Define  $c_1 = 1/2 - \sqrt{1/4 - \lambda_3/\lambda_1} + \alpha$ ,  $c_2 = \lambda_3/\lambda_2 + \alpha$  and  $c_3 = 1 - \lambda_2/\lambda_1 + \alpha$ . If  $c_1 \in (\alpha, 1/2)$ , then  $\lambda_1 h_{\text{GD}}(t)$  and  $\lambda_3 h_{\text{IQD}}(t)$  cross twice on  $(0, 1)$ , once at  $c_1 - \alpha$  and then once again at  $1 - c_1 + \alpha$ . If  $c_2 \in (\alpha, 1/2)$ , then  $\lambda_2 h_{\text{MMD}}(t)$  and  $\lambda_3 h_{\text{IQD}}(t)$  cross twice on  $(0, 1)$ , once at  $c_2 - \alpha$  and then once again at  $1 - c_2 + \alpha$ . Similarly, if  $c_3 \in (\alpha, 1/2)$  then  $\lambda_1 h_{\text{GD}}(t)$  and  $\lambda_2 h_{\text{MMD}}(t)$  cross at  $c_3 - \alpha$  and  $1 - c_3 + \alpha$ . Note that  $c_2 > \alpha$  and  $\alpha \leq c_1 \leq 1/2 + \alpha$  whenever  $1/4 \geq \lambda_3/\lambda_1$ .

We have six cases to handle; the details can be found in Appendix E. Figure 5 plots the function  $G_\lambda^\alpha(h)$  for  $h = \min\{\lambda_1 h_{\text{GD}}, \lambda_2 h_{\text{MMD}}\}$ . The red, blue, and black parts represent the risk that goes to the GD agent, the MMD agent, and the IQD agent, respectively.

Figure 5: The function  $G_\lambda^\alpha(h)$



We present the Pareto-optimal allocations  $(X_1, X_2, X_3)$  in the six cases below. These allocations are generally not comonotonic, but they are comonotonic on the event  $(A \cup B)^c$ . Recall that  $Y$  stands for  $X\mathbb{1}_{(A \cup B)^c}$ .

**Case 1**,  $c_1 \geq 1/2$  and  $c_3 \leq \alpha$ :  $X_1 = Y$ ,  $X_2 = 0$  and  $X_3 = X\mathbb{1}_{A \cup B}$ .

**Case 2**,  $c_2 \geq 1/2$  and  $c_3 \geq 1/2$ :  $X_1 = 0$ ,  $X_2 = Y$  and  $X_3 = X\mathbb{1}_{A \cup B}$ .

**Case 3**,  $c_3 \leq \alpha < c_1 < 1/2$ :  $X_1 = X - X_3$ ,  $X_2 = 0$  and  $X_3 = X\mathbf{1}_{A \cup B} + Y \wedge Q_{c_1}^-(X) - Y \wedge Q_{1-c_1}^-(X)$ .

**Case 4**, either  $\alpha < c_1 < c_3 < 1/2$  or  $\alpha < c_2 < 1/2 < c_3$ :  $X_1 = 0$ ,  $X_2 = X - X_3$  and

$$X_3 = X\mathbf{1}_{A \cup B} + Y \wedge Q_{c_2}^-(X) - Y \wedge Q_{1-c_2}^-(X).$$

**Case 5**,  $\alpha < c_3 < 1/2 < c_1$ :  $X_2 = X - X_1 - X_3$ ,  $X_3 = X\mathbf{1}_{A \cup B}$  and  $X_1 = Y \wedge Q_{c_3}^-(X) - Y \wedge Q_{1-c_3}^-(X)$ .

**Case 6**,  $\alpha < c_3 \leq c_1 < 1/2$ :  $X_1 = Y \wedge Q_{c_3}^-(X) - Y \wedge Q_{c_1}^-(X) + Y \wedge Q_{1-c_1}^-(X) - Y \wedge Q_{1-c_3}^-(X)$ ,

$$X_2 = Y - Y \wedge Q_{c_3}^-(X) + Y \wedge Q_{1-c_3}^-(X) \quad \text{and} \quad X_3 = X\mathbf{1}_{A \cup B} + Y \wedge Q_{c_1}^-(X) - Y \wedge Q_{1-c_1}^-(X).$$

The allocation in Case 6 shows a particularly rich feature, and we depict it in Figure 6(a).

*Remark 4* (Interpreting the Pareto-optimal allocation by derivative contracts). We consider Case 6, the most sophisticated case. Let  $P = -X > 0$  be the future (after one period) price of an asset to be shared by the three agents. Due to symmetry of the preferences,  $(-X_1, -X_2, -X_3)$  in Case 6 is a Pareto-optimal allocation for  $P$ . The corresponding sharing strategy can be obtained by trading bull call spreads (differences of call options) and digital options. We will describe the position for Anne first. Anne purchases two bull call spreads (a and b) and a digital put option (c) from Carole, and sells a digital call option (d) to Carole.

a: A bull call spread with long call strike price  $Q_{c_1}^-(P)$  and short call strike  $Q_{c_3}^-(P)$ .

b: A bull call spread with long call strike price  $Q_{1-c_3}^-(P)$  and short call strike  $Q_{1-c_1}^-(P)$ .

c:  $Q_{1-c_1}^-(P) - Q_{1-c_3}^-(P)$  units of a digital put with strike price  $Q_{1-\alpha}^-(P)$ .

d:  $Q_{c_3}^-(P) - Q_{c_1}^-(P)$  units of a digital put with strike price  $Q_{\alpha}^-(P)$ .

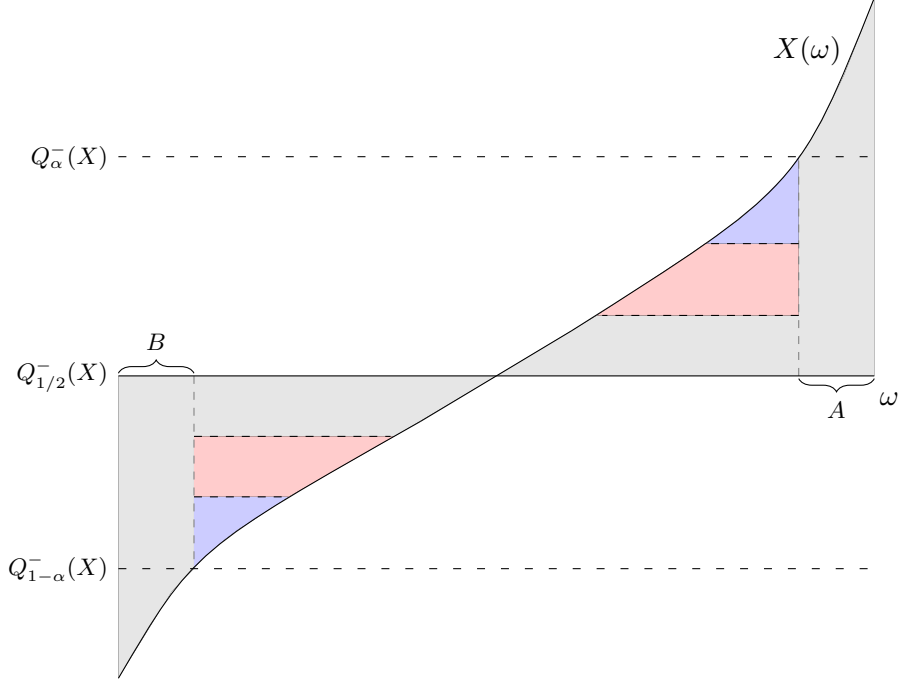
We can check that the above derivative portfolio precisely gives the payoff  $-X_1 + y$  in Case 6 with  $P = -X$ , where  $y$  is a constant (see Figure 6(b)). We can construct Bob's strategy using bull call spreads and digital options similarly. Carole writes the above derivative contracts to Anne and Bob and keeps the asset to herself. The prices of these contracts can be negotiated among the agents, and they do not affect the optimality of the allocation because of the translation invariance of variability measures.

### 6.3 Insurance between two GD and MMD agents

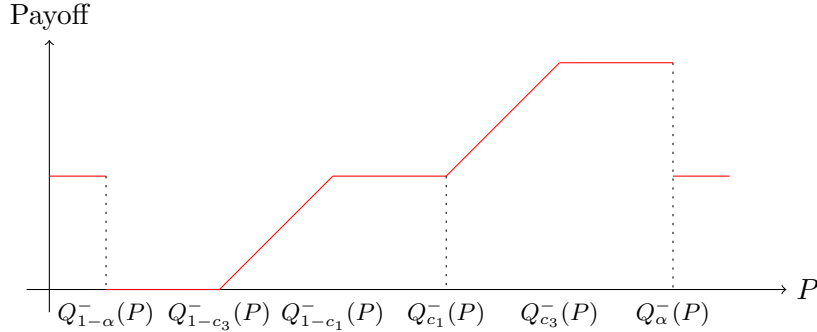
We next solve the insurance example (Example 1) presented in the introduction. Consider two individuals, Anne and Bob, who evaluate their risk with GD and MMD, respectively. That is,

Figure 6: Interpreting the Pareto-optimal allocation in Case 6

(a) A Pareto-optimal allocation for Anne, Bob and Carole in Case 6, where the red, blue and gray areas represent the allocations to Anne (GD), Bob (MMD) and Carole (IQD) respectively, up to constant shifts



(b) Payoff of Anne from the derivative contracts

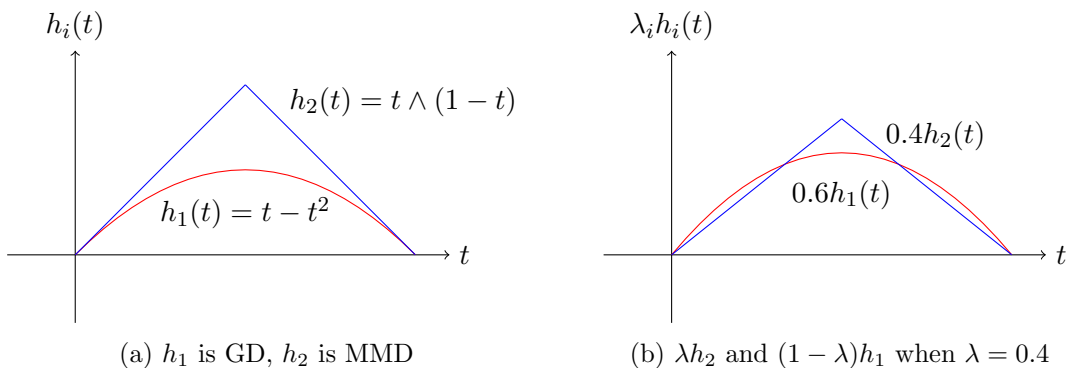


set  $h_1 = h_{\text{GD}}$  and  $h_2 = h_{\text{MMD}}$ . (Or, they could use  $\mathbb{E} + \lambda_1 \text{GD}$  and  $\mathbb{E} + \lambda_2 \text{MMD}$ , which would not change our analysis.) This setting is simpler than the three-agent problem in Section 6.2, and it offers a clearer visualization of the Pareto-optimal allocation.

Both  $h_1$  and  $h_2$  are strictly concave, and, by Proposition 5, any Pareto-optimal allocation in  $\mathbb{A}_n(X)$  is comonotonic. By Proposition 7, each Pareto-optimal allocation can be found by solving the inf-convolution  $\boxplus_{i=1}^2 (\lambda_i \rho_{h_i})$  for some Negishi weights  $(\lambda_1, \lambda_2) \in [0, \infty)^2 \setminus \{\mathbf{0}\}$ . Consider the normalized ones  $\lambda_1 = \lambda \in [0, 1]$  and  $\lambda_2 = 1 - \lambda$ . Figure 7 depicts the functions  $h_i(t)$  and  $\lambda_i h_i(t)$ .

By positive homogeneity it is  $\lambda \rho_{h_1}(X_1) = \rho_{\lambda h_1}(X_1)$  for Anne, and similarly for Bob. By Corollary 2, we have  $\boxplus_{i=1}^2 \rho_{\lambda_i h_i} = \rho_{h_\lambda}$ , where  $h_\lambda(t) = \min\{\lambda h_1(t), (1 - \lambda) h_2(t)\}$ . That is, the sum-

Figure 7: Distortion functions of GD and MMD agents



optimal allocation gives all the marginal  $t$ -quantile risk to the individual with the lowest  $\lambda_i h_i(t)$ .

The condition of Theorem 3 is satisfied, and so the  $(\lambda_1, \lambda_2)$ -optimal allocation is unique up to constant shifts. Any Pareto-optimal allocation takes the form

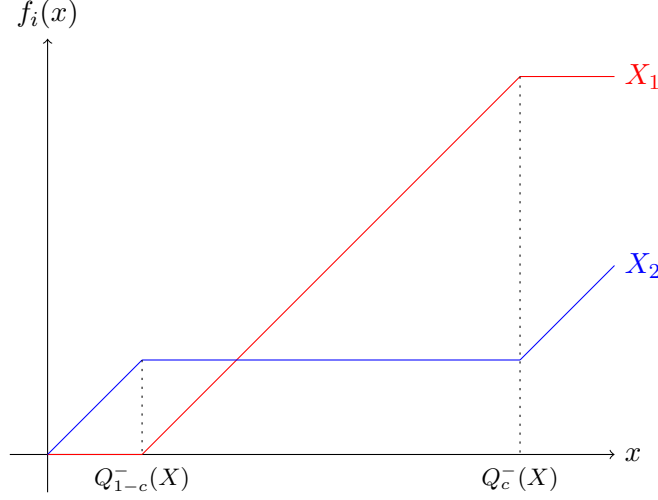
$$X_1 = X \wedge Q_c^-(X) - X \wedge Q_{1-c}^-(X) + k \quad \text{and} \quad X_2 = X - X_1,$$

where  $c \in [0, 1/2]$  and  $k \in \mathbb{R}$  is a constant. We can interpret this as a situation where the GD individual insures the potential losses  $X$  of the MMD one. The ceded risk is the random variable  $X_1$ , while its price is  $k$ , the latter which needs to be negotiated between the two agents. Next, we argue that the mapping  $\lambda \mapsto c$  is surjective.

(i) If  $\lambda < 1/2$  we have  $\lambda h_1 < (1 - \lambda)h_2$  everywhere and so  $c = 0$ . That is, the ceded risk is  $X$  and the GD agent provides full insurance. (ii) Similarly, if  $\lambda > 2/3$  we have  $\lambda h_1 > (1 - \lambda)h_2$  everywhere and so  $c = 1/2$ . It is the MMD agent that retains all the risk and no insurance is provided. Finally, (iii) if  $1/2 < \lambda < 2/3$  then  $\lambda h_1 > (1 - \lambda)h_2$  on both  $(0, (2\lambda - 1)/\lambda)$  and  $((1 - \lambda)/\lambda, 1)$  and  $\lambda h_1 < (1 - \lambda)h_2$  on  $((2\lambda - 1)/\lambda, (1 - \lambda)/\lambda)$ . Hence,  $c = (2\lambda - 1)/\lambda$  and the ceded risk has a simple deductible  $Q_{1-c}^-(X)$  and an upper limit  $Q_c^-(X)$ . This type of allocation is depicted in Figure 8.

The constant  $k$  can take any value because by location invariance, for any  $k \in \mathbb{R}$ , we have  $\rho_{h_i}(X_i + k) = \rho_{h_i}(X_i) + h_i(1)k = \rho_{h_i}(X_i)$  and the price of the insurance does not affect Pareto optimality. This observation remains true if agents use  $\mathbb{E} + \lambda_i \rho_{h_i}$  instead of  $\rho_{h_i}$ .

Figure 8: A Pareto-optimal allocation for the MMD and GD pair



#### 6.4 Risk sharing with several mixed GD-MMD agents

We conclude with the problem of sharing risk among many agents  $i \in [n]$  evaluating their risks with the variability measure

$$\rho_{h_i}(X_i) = \int X_i d((a_i h_{\text{GD}} + (1 - a_i) h_{\text{MMD}}) \circ \mathbb{P}) = a_i \text{GD}(X_i) + (1 - a_i) \text{MMD}(X_i),$$

$a_i \in [0, 1]$ . It is easily verified that for every  $i \in [n]$  the distortion function  $h_i = a_i h_{\text{GD}} + (1 - a_i) h_{\text{MMD}}$  is strictly concave and satisfies  $h_i(1) = 0$ . We can therefore invoke Theorem 3, Corollary 2 and Proposition 8 to characterize the set of Pareto-optimal allocations. Consider the usual normalization of the Negishi weights  $\sum_{i=1}^n \lambda_i = 1$  with  $\lambda_i > 0$  and notice that

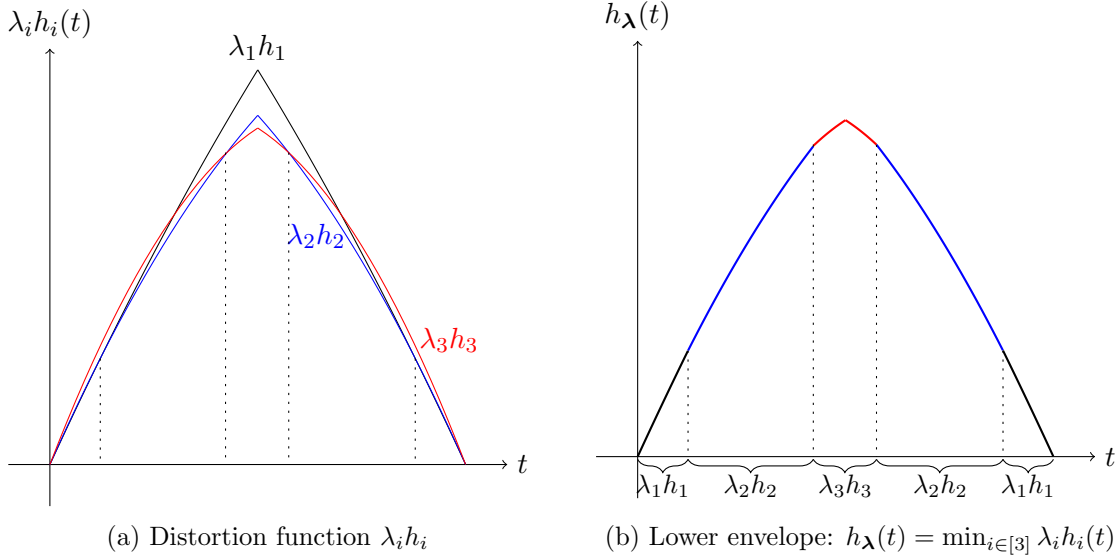
$$\bigoplus_{i=1}^n \rho_{\lambda_i h_i} = \rho_{h_\lambda},$$

where  $h_\lambda(t) = \min\{\lambda_1 h_1(t), \dots, \lambda_n h_n(t)\}$ .

Deriving every agent's allocation (contract) in a closed-form solution is a bit more cumbersome. Yet, Theorem 3 and Corollary 2 still fully pin down the shape of the optimal allocation, and we can visualize it easily. Consider the case when  $0 < \lambda_1 a_1 < \lambda_2 a_2 < \dots < \lambda_n a_n$  and set  $M_x = \{i \in [n] : \lambda_i h_i(\mathbb{P}(X > x)) = h_\lambda(\mathbb{P}(X > x))\}$  as before. We have that  $|M_x| = 1$   $\mu_X$ -almost surely, so the sum-optimal allocation is unique up to constant shifts for any  $\lambda$ . Figure 9 shows an example with three agents.

As we obtained in the previous application,  $h_\lambda$  induces a partition of the state space on which

Figure 9: Distortion functions for mixed GD-MMD agents, where  $a_1 = 1/4$ ,  $a_2 = 1/2$ ,  $a_3 = 3/4$  and  $\lambda = (0.31, 0.32, 0.37)$



only one agent takes the full marginal risk. That is, the Pareto-optimal allocation's shape is similar to the payoff obtained with a collection of straight deductibles insurance contracts with upper limits. For instance, the part of the risk that goes to agent 2 is

$$X_2 = X \wedge b - X \wedge a + X \wedge d - X \wedge c$$

for  $0 < a < b < c < d < \infty$  implicitly defined through the lower envelope  $h_\lambda(t)$ .

## 7 Heterogeneous beliefs in comonotonic risk sharing

We considered throughout an atomless probability space  $(\Omega, \mathcal{F}, \mathbb{P})$ . This assumption entails that every individual  $i \in [n]$  agrees on the fundamentals of the risk to be shared. We now show that all our results on comonotonic risk sharing can be extended to incorporate heterogeneous beliefs with almost no extra effort; this is not true for the unconstrained setting of risk sharing in Section 3. Our characterization of comonotonic risk sharing extends the main results of Liu (2020), which focus on dual utilities. See also Embrechts et al. (2020), Boonen and Ghossoub (2020) and Liebrich (2024) for risk sharing with risk measures and heterogeneous beliefs.

Let  $(\Omega, \mathcal{F})$  be a measurable space that allows for atomless probability measures and denote by  $\mathbb{P}_i$  the atomless probability measure that agent  $i \in [n]$  considers. That is, every individual  $i \in [n]$  believes the probability space  $(\Omega, \mathcal{F}, \mathbb{P}_i)$  is the true one. Let  $\mathcal{P}$  be the set of atomless probability

measures on the measurable space  $(\Omega, \mathcal{F})$  and let  $\ll$  denote absolute continuity. As before, every individual evaluates their risk with the distortion riskmetric

$$\rho_{h_i}^{\mathbb{P}_i}(X) = \int X \, d(h_i \circ \mathbb{P}_i).$$

For a probability measure  $\mathbb{P}$ , we define the corresponding left quantile as  $Q_t^{\mathbb{P}}(X) = \inf\{x \in \mathbb{R} : \mathbb{P}(X \leq x) \geq 1 - t\}$ .

The next lemma is instrumental in proving our last result:

**Lemma 4.** *Let  $\mathbb{P}_0, \mathbb{P} \in \mathcal{P}$  be such that  $\mathbb{P}_0 \ll \mathbb{P}$  and  $h \in \mathcal{H}^{\text{BV}}$ , and suppose that  $X \in \mathcal{X}$  is continuously distributed under  $\mathbb{P}$ . The function  $g(t) = h(\mathbb{P}_0(X > Q_t^{\mathbb{P}}(X)))$ ,  $t \in [0, 1]$ , satisfies  $\rho_h^{\mathbb{P}_0}(f(X)) = \rho_g^{\mathbb{P}}(f(X))$  for any increasing functions  $f : \mathbb{R} \rightarrow \mathbb{R}$ .*

Lemma 4 states that if a belief  $\mathbb{P}_0$  is absolutely continuous with respect to a probability measure  $\mathbb{P}$  and if a random variable  $X$  is continuously distributed under  $\mathbb{P}$ , then we can always find a distortion function  $g$  such that the two distortion riskmetrics  $\rho_h^{\mathbb{P}_0}$  and  $\rho_g^{\mathbb{P}}$  are exactly the same for every random variable  $Y = f(X)$  comonotonic with  $X$ .

Our last proposition states that when every belief is sufficiently “well-behaved”, then the comonotonic risk sharing problem with heterogeneous beliefs is equivalent to a comonotonic risk sharing problem with homogeneous belief  $\mathbb{P}$ .

**Proposition 10.** *Let  $\mathbb{P}_1, \dots, \mathbb{P}_n \in \mathcal{P}$ ,  $h_1, \dots, h_n \in \mathcal{H}^{\text{BV}}$  be given and let  $X \in \mathcal{X}$  admit a density under all  $\mathbb{P}_1, \dots, \mathbb{P}_n$ . There exist a probability measure  $\mathbb{P} \in \mathcal{P}$  and a collection of distortion functions  $g_1, \dots, g_n \in \mathcal{H}^{\text{BV}}$  such that the allocation  $(X_1, \dots, X_n) \in \mathbb{A}_n^+(X)$  is Pareto optimal for  $(\rho_{h_1}^{\mathbb{P}_1}, \dots, \rho_{h_n}^{\mathbb{P}_n})$  if and only if it is Pareto optimal for  $(\rho_{g_1}^{\mathbb{P}}, \dots, \rho_{g_n}^{\mathbb{P}})$ .*

The proof follows immediately by noticing that we can find a probability measure  $\mathbb{P}$  such that  $X$  admits a density under  $\mathbb{P}$  and for which  $\mathbb{P}_i \ll \mathbb{P}$ ,  $i \in [n]$ , and then invoking Lemma 4. The key insight is that restricting our allocation to be (globally) comonotonic eliminates, by construction, any potential “side-bets” originating from the heterogeneity of the beliefs of the agents, as seen in Liebrich (2024). Characterizing the set of Pareto-optimal allocations then simply boils down to making suitable changes of measure and/or distortion functions and solving the simplified problem with homogeneous beliefs.

## 8 Conclusion

We summarize the paper with a few remarks on the results that we obtained. While we focused on the case  $\mathcal{X} = L^\infty$ , all the results of this article generalize to larger spaces provided all the decision criteria  $\rho_{h_i}$  and the inf-convolution  $\square_{i=1}^n \rho_{h_i}(X)$  are finite on the larger space. We emphasized when key results can be readily generalized, but this finiteness property must often be verified case-by-case. For example, the results on risk sharing with IQD agents can be extended to  $L^0$  because the IQD is bounded from below. This property does not generalize to other functionals.

The unconstrained risk sharing problem for non-concave distortion functions typically leads to non-comonotonic sum-optimal allocations without explicit forms, and they can be difficult to analyze. Although we obtained several results on necessary or sufficient conditions for Pareto and sum optimality (Theorem 1 and Propositions 2-5), a full characterization of the Pareto-optimal or sum-optimal allocations for arbitrary distortion riskmetrics is beyond the current techniques.

The case of IQD agents is, nevertheless, special, although they do not have concave distortion functions. For this setting, we can fully characterize all Pareto-optimal allocations via sum-optimal ones, and the inf-convolution for such distortion riskmetrics admit concise formulas (Theorem 2 and Proposition 9):

$$\square_{i=1}^n (\lambda_i \text{IQD}_{\alpha_i}) = \left( \bigwedge_{i=1}^n \lambda_i \right) \text{IQD}_{\sum_{i=1}^n \alpha_i} \quad \text{and} \quad \boxplus_{i=1}^n (\lambda_i \text{IQD}_{\alpha_i}) = \left( \bigwedge_{i=1}^n \lambda_i \right) \text{IQD}_{\bigvee_{i=1}^n \alpha_i},$$

and their particular instances

$$\square_{i=1}^n \text{IQD}_{\alpha_i} = \text{IQD}_{\sum_{i=1}^n \alpha_i} \quad \text{and} \quad \boxplus_{i=1}^n \text{IQD}_{\alpha_i} = \text{IQD}_{\bigvee_{i=1}^n \alpha_i}.$$

These formulas may be compared with the quantile inf-convolutions formulas obtained by [Embrechts et al. \(2018\)](#) and [Liu et al. \(2022\)](#)

$$\square_{i=1}^n Q_{\alpha_i}^- = Q_{\sum_{i=1}^n \alpha_i}^-, \quad \square_{i=1}^n Q_{\alpha_i}^+ = Q_{\sum_{i=1}^n \alpha_i}^+, \quad \boxplus_{i=1}^n Q_{\alpha_i}^- = Q_{\bigvee_{i=1}^n \alpha_i}^-, \quad \text{and} \quad \boxplus_{i=1}^n Q_{\alpha_i}^+ = Q_{\bigvee_{i=1}^n \alpha_i}^+.$$

These results show that the representative agent (using the inf-convolution as its reference) for risk sharing among several IQD agents is again an IQD agent, and similarly, the representative agent among several quantile agents is again a quantile agent.

When the distortion functions are concave, or, when we constrain ourselves to the set of comonotonic allocations, the risk sharing problem becomes much more tractable, and we obtain

explicit allocations which are Pareto optimal or sum optimal (Theorem 3). This builds on the comonotonic improvement *à la* Landsberger and Meilijson (1994), when the distortion riskmetrics are convex order consistent. A high-level summary is that all results that were established for increasing distortion riskmetrics, in particular, Yaari (1987)’s dual utilities, can be extended in parallel to non-increasing ones without extra efforts (these results are summarized in Propositions 6-8). This opens up various application areas where risks are traditionally studied with only increasing distortion riskmetrics.

Combining the results for IQD agents and for risk-averse agents, we are able to solve risk sharing problems among these agents, whose Pareto-optimal allocations are found explicitly (Theorem 4). Various examples of risk sharing among these agents are presented in Section 6.

It remains unclear to us whether our analysis can be generalized to distortion riskmetrics other than IQD, which are not convex (i.e., with non-concave distortion functions), and how large the class of such tractable risk functionals is. As far as we are aware, the unconstrained risk sharing problems for non-convex risk measures and variability measures have very limited explicit results (e.g., Embrechts et al. (2018), Weber (2018) and Liu et al. (2022)), and further investigation is needed for a better understanding of the challenges and their solutions.

## Acknowledgments

We thank the Editor, an Associate Editor, and two anonymous referees for helpful comments on the paper. Ruodu Wang is supported by the Natural Sciences and Engineering Research Council of Canada (CRC-2022-00141, RGPIN-2024-03728).

## References

- Acerbi, C. (2002). Spectral measures of risk: A coherent representation of subjective risk aversion. *Journal of Banking and Finance*, **26**(7), 1505–1518.
- Amarante, M. (2022). A unified framework for Bayesian and non-Bayesian decision making and inference. *Mathematics of Operations Research*, **47**(4), 2721–2742.
- Artzner, P., Delbaen, F., Eber, J.-M. and Heath, D. (1999). Coherent measures of risk. *Mathematical Finance*, **9**(3), 203–228.
- Barrieu, P. and El Karoui, N. (2005). Inf-convolution of risk measures and optimal risk transfer. *Finance and Stochastics*, **9**, 269–298.
- Bellini, F., Fadina, T., Wang, R. and Wei, Y. (2022). Parametric measures of variability induced by risk measures. *Insurance: Mathematics and Economics*, **106**, 270–284.

- Boonen, T. J. and Ghossoub, M. (2020). Bilateral risk sharing with heterogeneous beliefs and exposure constraints. *ASTIN Bulletin*, **50**(1), 293–323.
- Carlier, G. and Dana, R.-A. (2003). Core of convex distortions of a probability. *Journal of Economic Theory*, **113**, 199–222.
- Carlier, G., Dana, R.-A. and Galichon, A. (2012). Pareto efficiency for the concave order and multivariate comonotonicity. *Journal of Economic Theory*, **147**, 207–229.
- Carreia-Vioglio, S., Maccheroni, F., Marinacci, M. and Montrucchio, L. (2011). Risk measures: Rationality and diversification. *Mathematical Finance*, **21**(4), 743–774.
- Carreia-Vioglio, S., Maccheroni, F., Marinacci, M. and Rustichini, A. (2014). Niveloids and their extensions: Risk measures on small domains. *Journal of Mathematical Analysis and Applications*, **413**(1), 343–360.
- Cheung, K. C. and Lo, A. (2017). Characterizations of optimal reinsurance treaties: a cost-benefit approach. *Scandinavian Actuarial Journal*, **2017**(1), 1–28.
- Denneberg, D. (1990). Premium calculation: Why standard deviation should be replaced by absolute deviation. *ASTIN Bulletin*, **20**(2), 181–190.
- Denneberg, D. (1994). *Non-additive Measures and Integral*. Kluwer, Dordrecht.
- Dhaene, J. and Kukush, A., Linders, D. and Tang, Q. (2012). Remarks on quantiles and distortion risk measures. *European Actuarial Journal*, **2**(2), 319–328.
- Embrechts, P., Liu, H. and Wang, R. (2018). Quantile-based risk sharing. *Operations Research*, **66**(4), 936–949.
- Embrechts, P., Liu, H., Mao, T. and Wang, R. (2020). Quantile-based risk sharing with heterogeneous beliefs. *Mathematical Programming Series B*, **181**(2), 319–347.
- Embrechts, P., McNeil, A. and Straumann, D. (2002). Correlation and dependence in risk management: properties and pitfalls. In *Risk Management: Value at Risk and Beyond* (Ed. Dempster, M. A. H.), 176–223, Cambridge University Press.
- Filipovic, D and Svindland, G. (2008). Optimal capital and risk allocations for law- and cash-invariant convex functions. *Finance and Stochastics*, **12**(3), 423–439.
- Föllmer, H. and Schied, A. (2016). *Stochastic Finance. An Introduction in Discrete Time*. Fourth Edition. Walter de Gruyter, Berlin.
- Furman, E., Wang, R. and Zitikis, R. (2017). Gini-type measures of risk and variability: Gini shortfall, capital allocation and heavy-tailed risks. *Journal of Banking and Finance*, **83**, 70–84.
- Grechuk, B., Molyboha, A. and Zabaranin, M. (2009). Maximum entropy principle with general deviation measures. *Mathematics of Operations Research*, **34**(2), 445–467.
- Jouini, E., Schachermayer, W. and Touzi, N. (2008). Optimal risk sharing for law invariant monetary utility functions. *Mathematical Finance*, **18**(2), 269–292.
- Konno, H. and Yamazaki, H. (1991). Mean-absolute deviation portfolio optimization model and its applications to Tokyo stock market. *Management Science*, **37**(5), 519–531.
- Landsberger, M. and Meilijson, I. (1994). Co-monotone allocations, Bickel-Lehmann dispersion and the

- Arrow-Pratt measure of risk aversion. *Annals of Operations Research*, **52**(2), 97–106.
- Lauzier, J.G., Lin, L. and Wang, R. (2023). Pairwise counter-monotonicity. *Insurance: Mathematics and Economics*, **111**, 279–287.
- Liu, H. (2020). Weighted comonotonic risk sharing under heterogeneous beliefs. *ASTIN Bulletin*, **50**(2), 647–673.
- Liu, F., Mao, T., Wang, R. and Wei, L. (2022). Inf-convolution, optimal allocations, and model uncertainty for tail risk measures. *Mathematics of Operations Research*, **47**(3), 2494–2519.
- Liebrich, FB. (2024). Risk sharing under heterogeneous beliefs without convexity. *Finance and Stochastics* **28**, 999–1033.
- Ludkovski, M. and Rüschendorf, L. (2008). On comonotonicity of Pareto optimal risk sharing. *Statistics and Probability Letters*, **78**(10), 1181–1188.
- Liu, P., Wang, R. and Wei, L. (2020). Is the inf-convolution of law-invariant preferences law-invariant? *Insurance: Mathematics and Economics*, **91**, 144–154.
- Markowitz, H. (1952). Portfolio selection. *The Journal of Finance*, **7**(1), 77–91.
- Markowitz, H. (2014). Mean–variance approximations to expected utility. *European Journal of Operational Research*, **234**(2), 346–355.
- Maccheroni, F., Marinacci, M. and Ruffino, D. (2013). Alpha as ambiguity: Robust mean-variance portfolio analysis. *Econometrica*, **81**(3), 1075–1113.
- Marinacci, M. and Montrucchio L. (2004). Introduction to the mathematics of ambiguity. In *Uncertainty in Economic Theory*, I. Gilboa, Ed. Routledge, New York, NY, USA. 46–107.
- Negishi, T. (1960). Welfare economics and existence of an equilibrium for a competitive economy. *Metroeconomica*, **12**(2-3), 92–97.
- Kusuoka, S. (2001). On law invariant coherent risk measures. *Advances in Mathematical Economics*, **3**, 83–95.
- Puccetti, G. and Wang R. (2015). Extremal dependence concepts. *Statistical Science*, **30**(4), 485–517.
- Ravanelli, C. and Svindland, G. (2014). Comonotone Pareto optimal allocations for law invariant robust utilities on  $L^1$ . *Finance and Stochastics*, **18**(1), 249–269.
- Rockafellar, R. T., Uryasev, S. and Zabaranin, M. (2006). Generalized deviation in risk analysis. *Finance and Stochastics*, **10**, 51–74.
- Rostek, M. (2010). Quantile maximization in decision theory. *Review of Economic Studies*, **77**, 339–371.
- Rothschild, M. and Stiglitz, J. E. (1970). Increasing risk: I. A definition. *Journal of Economic Theory*, **2**(3), 225–243.
- Rüschendorf, L. (2013). *Mathematical Risk Analysis. Dependence, Risk Bounds, Optimal Allocations and Portfolios*. Springer, Heidelberg.
- Schmeidler, D. (1989). Subjective probability and expected utility without additivity. *Econometrica*, **57**(3), 571–587.
- Shaked, M. and Shanthikumar, J. G. (2007). *Stochastic Orders*. Springer Series in Statistics.
- Shalit, H. and Yitzhaki, S. (1984). Mean-Gini, portfolio theory, and the pricing of risky assets. *The Journal*

of *Finance*, **39**(5), 1449–1468.

Wakker, P. P. (2010). *Prospect Theory: For Risk and Ambiguity*. Cambridge University Press.

Wang, Q., Wang, R. and Wei, Y. (2020a). Distortion riskmetrics on general spaces. *ASTIN Bulletin*, **50**(4), 827–851.

Wang, S. (1996). Premium calculation by transforming the layer premium density. *ASTIN Bulletin*, **26**, 71–92.

Wang, R., Wei, Y. and Willmot, G. E. (2020b). Characterization, robustness and aggregation of signed Choquet integrals. *Mathematics of Operations Research*, **45**(3), 993–1015.

Wang, R. and Zitikis, R. (2021). An axiomatic foundation for the Expected Shortfall. *Management Science*, **67**, 1413–1429.

Wang, R. and Wei, Y. (2020). Risk functionals with convex level sets. *Mathematical Finance*, **30**(4), 1337–1367.

Weber, S. (2018). Solvency II, or how to sweep the downside risk under the carpet. *Insurance: Mathematics and Economics*, **82**, 191–200.

Yaari, M. E. (1987). The dual theory of choice under risk. *Econometrica*, **55**(1), 95–115.

Yule, G. U. (1911). *An Introduction to the Theory of Statistics*. Charles Griffin and Company.

## A Proofs of results in Section 3

*Proof of Proposition 2.* (i) Let  $(X_1, \dots, X_n)$  be a Pareto-optimal allocation in  $\mathbb{A}_n(X)$ . We will show, without loss of generality, that any of the three following hypotheses leads to a contradiction of the Pareto optimality of  $(X_1, \dots, X_n)$ : (1) if simultaneously  $h_1(1) = 0$  and  $h_2(1) > 0$ ; (2) if simultaneously  $h_1(1) < 0$  and  $h_2(1) > 0$  and (3) if simultaneously  $h_1(1) = 0$  and  $h_2(1) < 0$ .

Consider the allocation  $(X_1 + c, X_2 - c, X_3, \dots, X_n)$ . Clearly, the allocation belongs to  $\mathbb{A}_n(X)$ . Recall that by translation invariance it is  $\rho_{h_1}(X_1 + c) = \rho_{h_1}(X_1) + ch_1(1)$  and  $\rho_{h_2}(X_2 - c) = \rho_{h_2}(X_2) - ch_2(1)$ .

Suppose (1) first so that  $h_1(1) = 0$  and  $h_2(1) > 0$ . Setting  $c > 0$  we have that  $\rho_{h_1}(X_1 + c) = \rho_{h_1}(X_1)$  and  $\rho_{h_2}(X_2 - c) < \rho_{h_2}(X_2)$  contradicting the Pareto optimality of  $(X_1, \dots, X_n)$ . For (2), we have  $\rho_{h_1}(X_1 + c) < \rho_{h_1}(X_1)$  and  $\rho_{h_2}(X_2 - c) < \rho_{h_2}(X_2)$  as  $h_1(1) < 0$  and  $h_2(1) > 0$ . For (3), the case when  $h_1(1) = 0$  and  $h_2(1) < 0$ , we can choose  $c < 0$ , which leads to a similar contradiction of the Pareto optimality of  $(X_1, \dots, X_n)$ .

The case when  $(X_1, \dots, X_n)$  is Pareto optimal in  $\mathbb{A}_n^+(X)$  is identical, and we conclude that  $h_i(1)$  are either all zero, all positive, or all negative.

(ii) We show that if there exist  $i, j \in [n]$  such that  $h_i(1) \neq h_j(1)$ , then  $\boxplus_{i=1}^n \rho_{h_i}(X) = -\infty$  for any  $X \in \mathcal{X}$ . Without loss of generality, let  $h_1(1) < h_2(1)$  and consider a  $c > 0$ . Given  $X \in \mathcal{X}$ , for

any allocation  $(X_1, \dots, X_n) \in \mathbb{A}_n^+(X)$  we have that

$$\rho_{h_1}(X_1 + c) + \rho_{h_2}(X_2 - c) = \rho_{h_1}(X_1) + \rho_{h_2}(X_2) + c(h_1(1) - h_2(1)).$$

Consider now the allocation  $(X_1 + c, X_2 - c, X_3, \dots, X_n)$ . Taking the limit  $c \rightarrow \infty$  we have  $\sum_{i=1}^n \rho_{h_i}(X_i) = -\infty$  and so  $\boxplus_{i=1}^n \rho_{h_i}(X) = -\infty$ .  $\square$

*Proof of Theorem 1.* For the “if” part, it is clear that  $(X_1, \dots, X_n)$  is Pareto optimal for the normalized preferences  $\rho_{\tilde{h}_1}, \dots, \rho_{\tilde{h}_n}$ . Hence,  $(X_1, \dots, X_n)$  is also Pareto optimal for agents using  $\rho_{h_1}, \dots, \rho_{h_n}$  as their preferences, as the normalization does not change the preferences.

Next, we show the “only if” part. Let  $(X_1, \dots, X_n) \in \mathbb{A}_n(X)$  be a Pareto-optimal allocation in  $\mathbb{A}_n(X)$ . By Proposition 2, we have  $h_i(1), i \in [n]$ , are either all positive or all negative; that is  $\tilde{h}_i(1), i \in [n]$ , are all 1 or  $-1$ . We first consider the case where  $\tilde{h}_i(1) = 1$  for  $i \in [n]$ . Assume by contradiction that  $\sum_{i=1}^n \rho_{\tilde{h}_i}(X_i) > \boxplus_{i=1}^n \rho_{\tilde{h}_i}(X)$ . There exists an allocation  $(Y_1, \dots, Y_n) \in \mathbb{A}_n(X)$  such that  $\sum_{i=1}^n \rho_{\tilde{h}_i}(Y_i) < \sum_{i=1}^n \rho_{\tilde{h}_i}(X_i)$ . Set  $c_i = \rho_{\tilde{h}_i}(X_i) - \rho_{\tilde{h}_i}(Y_i), i = 1, \dots, n$  and notice that  $c = \sum_{i=1}^n c_i > 0$ . Hence,

$$(Y_1 + c_1 - c/n, \dots, Y_n + c_n - c/n) \in \mathbb{A}_n(X)$$

and by translation invariance for every  $i \in [n]$  it is

$$\rho_{\tilde{h}_i}(Y_i + c_i - c/n) = \rho_{\tilde{h}_i}(Y_i + c_i) - c/n < \rho_{\tilde{h}_i}(Y_i + c_i) = \rho_{\tilde{h}_i}(X_i),$$

contradicting the Pareto optimality of  $(X_1, \dots, X_n)$ . The case  $\tilde{h}_i(1) = -1, i \in [n]$ , is analogous.  $\square$

*Proof of Lemma 2.* The implications (i) $\Rightarrow$ (ii) $\Rightarrow$ (iii) $\Rightarrow$ (iv) are all straightforward, where (iii) $\Rightarrow$ (iv) follows from the fact that  $X \leq_{\text{cx}} Y$  is equivalent to  $\rho_h(X) \leq \rho_h(Y)$  holding for all concave  $h \in \mathcal{H}^{\text{BV}}$  by Theorem 2 of Wang et al. (2020b).

We next show (iv) $\Rightarrow$ (i). Define

$$L_X(t) = \int_0^t Q_s^-(X) ds \quad \text{for } X \in \mathcal{X} \text{ and } t \in [0, 1].$$

Let  $h \in \mathcal{H}^{\text{BV}}$  be strictly concave,  $X \leq_{\text{cx}} Y$ , and  $\rho_h(X) = \rho_h(Y)$ . The convex order relation implies  $L_X(t) \leq L_Y(t)$  for all  $t \in [0, 1]$  with equality at  $t = 0$  and  $t = 1$  (Theorem 3.A.5 of Shaked and Shanthikumar (2007)). Note that  $h$  is continuous and almost everywhere differentiable on  $(0, 1)$  with possible discontinuity at 0 and 1. Let  $h'$  be the left derivative of  $h$  on  $(0, 1)$ , and let

$a = \lim_{x \downarrow 0} h(x) - h(0)$  and  $b = \lim_{x \uparrow 1} h(x) - h(1)$ ; both  $a$  and  $b$  are nonnegative due to concavity of  $h$ .

By decomposing  $h$  into two possible jumps at 0, 1 and a continuous part, we can write, using Lemma 1,

$$\rho_h(X) = \int_{(0,1)} Q_t^-(X) dh(t) - bQ_1^+(X) + aQ_0^-(X) = \int_0^1 h'(t)Q_t^-(X) dt - bQ_1^+(X) + aQ_0^-(X).$$

Since  $h$  is concave, the convex order condition implies  $\int_{(0,1)} Q_t^-(X) dh(t) \leq \int_{(0,1)} Q_t^-(Y) dh(t)$  (Theorem 2 of Wang et al. (2020b)). Moreover,  $Q_1^+(X) \geq Q_1^+(Y)$  and  $Q_0^-(X) \leq Q_0^-(Y)$ . Therefore, in order for  $\rho_h(X) = \rho_h(Y)$ , all three inequalities above are equalities. In particular we have  $\int_0^1 h'(t)Q_t^-(X) dt = \int_0^1 h'(t)Q_t^-(Y) dt$ . It follows that

$$\int_0^1 h'(t)Q_t^-(X) dt = \int_0^1 h'(t) dL_X(t) = \int_0^1 h'(t) dL_Y(t). \quad (13)$$

Let us show  $(L_X(t) - L_Y(t))h'(t) \rightarrow 0$  as  $t \downarrow 0$  or  $t \uparrow 1$ . Since  $X$  and  $Y$  are bounded,  $L_X(t) - L_Y(t)$  is Lipschitz as  $t \downarrow 0$  and  $t \uparrow 1$ . Since  $L_X(0) = L_Y(0)$  and  $L_X(1) = L_Y(1)$ , the Lipschitz property gives  $|L_X(t) - L_Y(t)| \leq c((1-t) \wedge t)$  for some  $c > 0$ . Note that  $h'$  is integrable and strictly decreasing. Since  $h'$  is decreasing, as  $t \downarrow 0$ ,  $h'(t)$  either has a finite limit or has the limit  $\infty$ . In the first case, clearly  $th'(t) \rightarrow 0$ . In the second case, for  $t > 0$  small enough, we have  $h'(t) > 0$  and  $\int_0^t h'(t) dt \geq th'(t)$ . Integrability of  $h'$  implies  $\int_0^t h'(t) dt \rightarrow 0$ , which implies  $th'(t) \rightarrow 0$ . Similarly,  $(1-t)h'(t) \rightarrow 0$  as  $t \uparrow 1$ . Hence, we have the desired limits.

Using integration by parts, (13) and the above limits yield  $\int_0^1 (L_X(t) - L_Y(t)) d(-h')(t) = 0$ . Since  $L_X(t) \leq L_Y(t)$  for all  $t \in (0, 1)$  and  $-h'$  is strictly increasing on  $(0, 1)$ , we conclude that  $L_X(t) = L_Y(t)$  for almost every  $t \in (0, 1)$ . Since  $L_X$  and  $L_Y$  are continuous, this further implies that  $L_X = L_Y$ . This is sufficient for  $X \stackrel{d}{=} Y$ .  $\square$

*Proof of Proposition 5.* (i) follows from Corollary 1 observing that comonotonic improvements strictly improve welfare. For (ii), the ‘‘only if’’ part is directly shown by (i). We only show the ‘‘if’’ part. As the normalization of  $h_i$ ,  $i \in [n]$ , will not change the preferences, we only consider the case when  $a_i = a_j = a$  for all  $i, j \in [n]$ . Let  $(X_1, \dots, X_n) \in \mathbb{A}_n^+(X)$ . By comonotonic additivity and positive homogeneity it is  $\sum_{i=1}^n \rho_{a_i h_1}(X_i) = a \rho_{h_1}(X)$ . Let  $(Y_1, \dots, Y_n) \in \mathbb{A}_n(X)$ . By subadditivity we have  $\sum_{i=1}^n \rho_{a_i h_1}(Y_i) \geq a \rho_{h_1}(\sum_{i=1}^n Y_i) = a \rho_{h_1}(X)$ . Hence, a comonotonic allocation  $(X_1, \dots, X_n)$  always solves  $\square_{i=1}^n \rho_{a_i h_1}(X)$ , and thus it is Pareto optimal.  $\square$

*Proof of Theorem 2.* We first prove part (ii) and then use it to prove part (i). Let us first verify

$\square_{i=1}^n(\lambda_i \text{IQD}_{\alpha_i}) \geq \lambda \text{IQD}_{\alpha}$ . Using (3) and the fact that an IQD is nonnegative if  $\alpha < 1/2$ , then for  $X \in \mathcal{X}$ ,

$$\begin{aligned} \square_{i=1}^n(\lambda_i \text{IQD}_{\alpha_i}) &\geq \lambda \square_{i=1}^n \text{IQD}_{\alpha_i}(X) \\ &= \lambda \inf \left\{ \sum_{i=1}^n Q_{\alpha_i}^-(X_i) + \sum_{i=1}^n Q_{\alpha_i}^-(-X_i) : (X_1, \dots, X_n) \in \mathbb{A}_n(X) \right\} \\ &\geq \lambda \square_{i=1}^n Q_{\alpha_i}^-(X) + \lambda \square_{i=1}^n Q_{\alpha_i}^-(-X) \\ &= \lambda Q_{\sum_{i=1}^n \alpha_i}^-(X) + \lambda Q_{\sum_{i=1}^n \alpha_i}^-(-X) = \lambda \text{IQD}_{\alpha}(X), \end{aligned}$$

where the second-last equality is due to Corollary 2 of Embrechts et al. (2018). If  $\alpha \geq 1/2$ , then  $\square_{i=1}^n(\lambda_i \text{IQD}_{\alpha_i}) \geq 0 = \lambda \text{IQD}_{\alpha}$  holds automatically.

Next, we verify  $\square_{i=1}^n(\lambda_i \text{IQD}_{\alpha_i}) \leq \lambda \text{IQD}_{\alpha}$  by showing that the construction of the allocation  $(X_1, \dots, X_n)$  of  $X \in \mathcal{X}$  in (7) satisfies  $\sum_{i=1}^n \lambda_i \text{IQD}_{\alpha_i}(X_i) = \lambda \text{IQD}_{\alpha}(X)$ . This will prove part (ii) as well as Remark 2. First, it is straightforward to verify  $(X_1, \dots, X_n) \in \mathbb{A}_n(X)$ . Since IQD is location invariant, we can, without loss of generality, assume  $c = c_1 = \dots = c_n = 0$ ; i.e., 0 is a median of  $X$ . Note that this leads to the simplified form

$$X_i = X \mathbb{1}_{A_i \cup B_i} + a_i X (1 - \mathbb{1}_{A \cup B}), \quad i \in [n].$$

If  $\alpha \geq 1/2$ , then  $\beta = 1/2$  and  $\mathbb{P}(A^c \cup B^c) = 0$ . Thus,  $\mathbb{P}(X_i > 0) = \mathbb{P}(A_i) \leq \alpha_i$  and symmetrically,  $\mathbb{P}(X_i < 0) \leq \mathbb{P}(B_i) \leq \alpha_i$ . Hence,  $\text{IQD}_{\alpha_i}(X_i) = 0$ .

Next, assume  $\alpha < 1/2$ . We have  $\mathbb{P}(\{X > Q_{\alpha}^-(X)\} \cap A^c) = 0$  by Lemma A.3 of Wang and Zitikis (2021). For  $i \in [n]$ , if  $a_i > 0$ , we can compute

$$\mathbb{P}(X_i > a_i Q_{\alpha}^-(X)) \leq \mathbb{P}(A_i) + \mathbb{P}(\{X_i > a_i Q_{\alpha}^-(X)\} \setminus A_i) = \alpha_i + \mathbb{P}(\{X > Q_{\alpha}^-(X)\} \cap A^c) = \alpha_i.$$

If  $a_i = 0$ ,  $X_i = X \mathbb{1}_{A_i \cup B_i}$ . Hence,  $\mathbb{P}(X_i > a_i Q_{\alpha}^-(X)) = \mathbb{P}(X_i > 0) \leq \mathbb{P}(A_i) = \alpha_i$ . In conclusion,  $\mathbb{P}(X_i > a_i Q_{\alpha}^-(X)) \leq \alpha_i$  for  $i \in [n]$ , which implies  $Q_{\alpha_i}^-(X_i) \leq a_i Q_{\alpha}^-(X)$ . Using a symmetric argument, we get  $Q_{1-\alpha_i}^+(X_i) \geq a_i Q_{1-\alpha}^+(X)$ . It follows that

$$\text{IQD}_{\alpha_i}(X_i) \leq a_i Q_{\alpha}^-(X) - a_i Q_{1-\alpha}^+(X) = a_i \text{IQD}_{\alpha}(X).$$

Therefore,  $\sum_{i=1}^n \lambda_i \text{IQD}_{\alpha_i}(X_i) \leq \sum_{i=1}^n \lambda_i a_i \text{IQD}_{\alpha}(X)$ . Taking  $a_i = 0$  for all  $i \in [n]$  with  $\lambda_i > \lambda$  gives the desired inequality  $\sum_{i=1}^n \lambda_i \text{IQD}_{\alpha_i}(X_i) \leq \lambda \text{IQD}_{\alpha}(X)$ .

Putting the above arguments together, we prove (ii), that is,  $\square_{i=1}^n \lambda_i \text{IQD}_{\alpha_i}(X) = \lambda \text{IQD}_{\alpha}(X)$ .

In particular,

$$\text{IQD}_{\alpha_i}(X_i) = a_i \text{IQD}_{\alpha}(X) \quad \text{and} \quad \sum_{i=1}^n \text{IQD}_{\alpha_i}(X_i) = \text{IQD}_{\alpha}(X) = \square_{i=1}^n \text{IQD}_{\alpha_i}(X), \quad (14)$$

and thus  $(X_1, \dots, X_n)$  is sum optimal.

Next, we show part (i). The ‘‘if’’ statement follows from Proposition 3, and we will show the ‘‘only if’’ statement. Take any Pareto-optimal allocation  $(Y_1, \dots, Y_n)$  of  $X$ . Write  $x = \text{IQD}_{\alpha}(X)$ ,  $y_i = \text{IQD}_{\alpha_i}(Y_i)$  for  $i \in [n]$ , and  $y = \sum_{i=1}^n y_i$ . It suffices to show  $y = x$ . By (14), we have that  $x \leq y$  always holds. If  $y = 0$ , there is nothing to show; next we assume  $y > 0$ . Let  $(X_1, \dots, X_n)$  be an allocation in (7) with  $a_i = y_i/y$  for  $i \in [n]$ , which sums up to 1. By (14), we have  $\text{IQD}_{\alpha_i}(X_i) = a_i \text{IQD}_{\alpha}(X) = a_i x$ . If  $x < y$ , then  $\text{IQD}_{\alpha_i}(X_i) = xy_i/y \leq y_i = \text{IQD}_{\alpha_i}(Y_i)$  for  $i \in [n]$ , and strict inequality holds as soon as  $y_i > 0$ . In this case,  $(X_1, \dots, X_n)$  Pareto dominates  $(Y_1, \dots, Y_n)$ , conflicting Pareto optimality of  $(Y_1, \dots, Y_n)$ . Hence, we obtain  $x = y$ .

Finally, part (iii) on Pareto optimality of  $(X_1, \dots, X_n)$  follows by combining (i) and (14).  $\square$

## B Proofs of results in Section 4

*Proof of Proposition 7.* (i)  $\Rightarrow$  (ii) is analogous to Theorem 1.

(ii)  $\Rightarrow$  (iii) Let  $S = \{(\rho_{h_1}(X_1), \dots, \rho_{h_n}(X_n)) : (X_1, \dots, X_n) \in \mathbb{A}_n^+(X)\}$  be the risk possibility set of the set of comonotonic allocations. By comonotonic additivity and positive homogeneity,  $S$  is a convex set. Let  $(X_1, \dots, X_n)$  be a Pareto-optimal allocation and  $\mathbf{x} = (\rho_{h_1}(X_1), \dots, \rho_{h_n}(X_n))$ . Notice now that  $\mathbf{x}$  always belongs to the boundary of  $S$ . Let  $V = \{(v_1, \dots, v_n) : v_i \leq \rho_{h_i}(X_i) \text{ for } i \in [n]\}$  where  $(X_1, \dots, X_n)$  is Pareto optimal. It is clear that  $V \cap S = \{\mathbf{x}\}$ .

Therefore, by the Separating Hyperplane Theorem, there exists  $(\lambda_1, \dots, \lambda_n) \in \mathbb{R}^n \setminus \mathbf{0}$  such that  $\sum_{i=1}^n \lambda_i \rho_{h_i}(X_i) = \inf_{\mathbf{u} \in S} \sum_{i=1}^n \lambda_i u_i = \boxplus_{i=1}^n \lambda_i \rho_{h_i}(X_i)$ .

We are left to show that  $\lambda_i \geq 0$  for every  $i \in [n]$ . Let  $\mathbf{v} = \mathbf{x} - (1, 0, \dots, 0)$ . We have  $\mathbf{v} \in V$ . Hence, we have  $\lambda_1 \geq 0$  as  $\sum_{i=1}^n \lambda_i v_i \leq \sum_{i=1}^n \lambda_i \rho_{h_i}(X_i)$ . Similarly, we obtain  $\lambda_i \geq 0$  for all  $i \in [n]$ .  $\square$

*Proof of Theorem 3.* Let  $h_{\wedge}(t) = \min\{h_1(t), \dots, h_n(t)\}$ . We first show that  $\boxplus_{i=1}^n \rho_{h_i} = \rho_{h_{\wedge}}$ . Comonotonic additivity of  $\rho_{h_{\wedge}}$  implies that  $\boxplus_{i=1}^n \rho_{h_i} \geq \rho_{h_{\wedge}}$ . Conversely, notice that for every  $i \in [n]$  the function  $f_i$  in (8) is Lipschitz continuous and non-decreasing because  $g_i$  is nonnegative and bounded.

Using Lemma 3, we get

$$\rho_{h_i}(f_i(X)) = \int_0^\infty g_i(s)h_i(\mathbb{P}(X > s)) ds + \int_{-\infty}^0 g_i(s)(h_i(\mathbb{P}(X > s)) - h_i(1)) ds. \quad (15)$$

It follows that

$$\begin{aligned} \sum_{i=1}^n \rho_{h_i}(f_i(X)) &= \sum_{i=1}^n \int_0^\infty g_i(s)h_i(\mathbb{P}(X > s)) ds + \int_{-\infty}^0 g_i(s)(h_i(\mathbb{P}(X > s)) - h_i(1)) ds \\ &= \int_0^\infty \sum_{i=1}^n g_i(s)h_i(\mathbb{P}(X > s)) ds + \int_{-\infty}^0 \sum_{i=1}^n g_i(s)(h_i(\mathbb{P}(X > s)) - h_i(1)) ds \\ &= \int_0^\infty h_\wedge(\mathbb{P}(X > s)) ds + \int_{-\infty}^0 (h_\wedge(\mathbb{P}(X > s)) - h_\wedge(1)) ds \\ &= \rho_{h_\wedge}(X) \geq \boxplus_{i=1}^n \rho_{h_i}(X). \end{aligned}$$

Hence,  $\boxplus_{i=1}^n \rho_{h_i} = \rho_{h_\wedge}$ .

Next, we show that the solution is unique up to constant shifts almost surely if and only if  $|M_x| = 1$  for  $\mu_X$ -almost every  $x$ , where  $\mu_X$  is the distribution measure of  $X$ .

Since the above argument of  $\sum_{i=1}^n \rho_{h_i}(f_i(X)) = \boxplus_{i=1}^n \rho_{h_i}(X)$  only requires  $\sum_{i \in M_x} g_i(x) = 1$  for almost every  $x$ , any allocation  $(f_1(X), \dots, f_n(X))$  in (8) with  $g_i$  replaced by

$$g_i(x) = \mathbb{1}_{\{i=\min M_x\}} \text{ or } g_i(x) = \mathbb{1}_{\{i=\max M_x\}}, \quad x \in \mathbb{R},$$

also satisfies sum optimality. Therefore, if  $|M_x| = 1$  does not hold  $\mu_X$ -almost surely, there are multiple optimal allocations that are not constant shifts from each other.

Conversely, we show that if  $|M_x| = 1$  for  $\mu_X$ -almost every  $x$  then every sum-optimal allocation is almost surely equal to the one in (8).

For any increasing and Lipschitz function  $k$  with right-derivative  $w$  and two distortion functions with  $h \geq g$  and  $h(1) = g(1)$ , we have, by Lemma 3,

$$\rho_h(k(X)) - \rho_g(k(X)) = \int_{-\infty}^\infty w(s)(h(\mathbb{P}(X > s)) - g(\mathbb{P}(X > s))) ds.$$

This means  $\rho_h(k(X)) = \rho_g(k(X))$  if and only if  $w(s) = 0$  almost surely for  $s$  such that  $h(\mathbb{P}(X > s)) > g(\mathbb{P}(X > s))$ . Note that if  $(k_1(X), \dots, k_n(X)) \in \mathbb{A}_n^+(X)$  is sum optimal, then

$$\sum_{i=1}^n \rho_{h_i}(k_i(X)) = \rho_{h_\wedge}(X) = \sum_{i=1}^m \rho_{h_\wedge}(k_i(X)).$$

This implies that  $w_i(x) = 0$  as soon as  $h_i(\mathbb{P}(X > x)) > h_\wedge(\mathbb{P}(X > x))$ , where  $w_i$  is the right-derivative of  $k_i$ . Moreover,  $w_i(x) = 1$  if  $h_i(\mathbb{P}(X > x)) = h_\wedge(\mathbb{P}(X > x))$  since  $\sum_{j=1}^n w_j(x) = 1$  for almost every  $x$ . Thus,  $w_i$  is uniquely determined  $\mu_X$ -a.s., implying that  $k_i$  is unique  $\mu_X$ -a.s. up to a constant shift.  $\square$

*Proof of Lemma 3.* Without loss of generality we assume  $X \geq 0$  and  $f(X) \geq 0$ . Denote by  $\nu = h \circ \mathbb{P}$ . We have

$$\begin{aligned} \rho_h(f(X)) - \int_0^\infty g(x)h(\mathbb{P}(X > x)) dx &= \int_0^\infty \nu(f(X) > y) dy - \int_0^\infty g(x)\nu(X > x) dx \\ &= \int_0^\infty g(x)\nu(f(X) > f(x)) dx - \int_0^\infty g(x)\nu(X > x) dx \\ &= \int_0^\infty g(x)(\nu(f(X) > f(x)) - \nu(X > x)) dx. \end{aligned}$$

Note that  $\mathbb{P}(f(X) > f(x)) \leq \mathbb{P}(X > x)$  for all  $x$ . If  $\mathbb{P}(f(X) > f(x)) < \mathbb{P}(X > x)$ , then there exists  $z > x$  such that  $f(z) = f(x)$ . This implies that  $g(x) = 0$  for any point  $x$  with  $\nu(f(X) > f(x)) - \nu(X > x) \neq 0$ . Therefore,

$$\rho_h(f(X)) - \int_0^\infty g(x)h(\mathbb{P}(X > x)) dx = 0.$$

The case of general  $X$  bounded from below can be obtained by constant shifts on both  $X$  and  $f$ .  $\square$

*Proof of Proposition 8.* It is clear that since  $h_i(1) = 0$  and  $h_i(t) > 0$  for all  $i \in [n]$  and all  $t \in (0, 1)$ , we have that  $\rho_{h_i}(X) \geq 0$  for all  $i \in [n]$ , with equality only if  $X$  is a constant. We first show the “if” statement. Suppose, by contradiction, that  $(X_1, \dots, X_n) \in \mathbb{A}_n^+(X)$  is not Pareto optimal but that it solves  $\boxplus_{i \in K} \rho_{\lambda_i h_i}(X - \sum_{i \notin K} X_i)$  for  $K = \{i \in [n] : X_i \notin \mathbb{R}\}$  and  $\lambda \in (0, \infty)^K$ . Our contradiction hypothesis implies that there exists a  $(Y_1, \dots, Y_n) \in \mathbb{A}_n^+(X)$  such that simultaneously  $\rho_{h_i}(Y_i) \leq \rho_{h_i}(X_i)$  for every  $i \in [n]$  and  $\rho_{h_j}(Y_j) < \rho_{h_j}(X_j)$  for some  $j \in [n]$ . Notice that if  $i \notin K$  it is

$$0 \leq \rho_{h_i}(Y_i) \leq \rho_{h_i}(X_i) = 0$$

and so it must be the case that  $\rho_{h_i}(Y_i) < \rho_{h_i}(X_i)$  for some  $i \in K$ , a contradiction with the hypothesis that  $(X_i)_{i \in K}$  solves  $\boxplus_{i \in K} \rho_{\lambda_i h_i}(X - \sum_{i \notin K} X_i) = \boxplus_{i \in K} \rho_{\lambda_i h_i}(X)$ , where the equality follows because of location invariance of  $\boxplus_{i \in K} \rho_{\lambda_i h_i}$ .

Conversely, let  $(X_1, \dots, X_n) \in \mathbb{A}_n^+(X)$  be Pareto optimal and define  $K = \{i \in [n] : X_i \notin \mathbb{R}\}$ ; this gives that  $\sum_{i \notin K} X_i$  is a constant. Recall that  $\rho_{h_i}(X_i) = 0$  for every  $i \notin K$ , and  $\rho_{h_i}(X_i) > 0$  for every  $i \in K$ . It is clear that  $(X_i)_{i \in K}$  is a Pareto-optimal allocation of  $X - \sum_{i \notin K} X_i$  for the

collection  $(\rho_{h_i})_{i \in K}$ . By Proposition 7, there exists a  $\lambda \in [0, \infty)^K \setminus \{\mathbf{0}\}$  such that  $\sum_{i \in K} \lambda_i \rho_{h_i}(X_i) = \boxplus_{i \in K} (\lambda_i \rho_{h_i})(X - \sum_{j \notin K} X_j) = \boxplus_{i \in K} \rho_{\lambda_i h_i}(X)$ . As  $\rho_{h_i}(X_i) > 0$  for  $i \in K$ , we have  $\boxplus_{i \in K} (\lambda_i \rho_{h_i})(X) > 0$ . It must be the case that  $\lambda_i > 0$  for all  $i \in K$ , as otherwise, we have  $\boxplus_{i \in K} (\lambda_i \rho_{h_i})(X) = 0$ , a contradiction.  $\square$

*Proof of Proposition 9.* Part (ii) follows directly from Corollary 2, so it remains to show part (i). Suppose that  $(X_1, \dots, X_n) \in \mathbb{A}_n^+(X)$  is Pareto optimal. Then there exists  $(\lambda_1, \dots, \lambda_n) \in [0, \infty)^n$ , with  $\lambda = \bigwedge_{i=1}^n \lambda_i > 0$ , such that

$$\sum_{i=1}^n (\lambda_i \text{IQD}_{\alpha_i})(X_i) = \lambda \text{IQD}_{\bigvee_{i=1}^n \alpha_i}(X).$$

Using the fact that an IQD is nonnegative and part (ii), we get

$$\lambda \boxplus_{i=1}^n \text{IQD}_{\alpha_i}(X) \leq \sum_{i=1}^n (\lambda \text{IQD}_{\alpha_i})(X_i) \leq \sum_{i=1}^n (\lambda_i \text{IQD}_{\alpha_i})(X_i) = \lambda \text{IQD}_{\bigvee_{i=1}^n \alpha_i}(X) = \lambda \boxplus_{i=1}^n \text{IQD}_{\alpha_i}(X),$$

and so  $(X_1, \dots, X_n)$  is sum optimal.  $\square$

## C Proofs of results in Section 5

We first present a lemma that we will use in the proof of Theorem 4.

**Lemma 5.** *For  $\alpha \in [0, 1/2)$ ,  $\lambda > 0$  and  $h \in \mathcal{H}^C$  it is*

$$(\lambda \text{IQD}_{\alpha}) \square \rho_h = \rho_{G_{\lambda}^{\alpha}(h)}. \quad (16)$$

*Proof of Lemma 5.* We first verify that  $\lambda \text{IQD}_{\alpha}(X_1) + \rho_h(X_2) \geq \rho_{G_{\lambda}^{\alpha}(h)}(X)$  for any  $(X_1, X_2) \in \mathbb{A}_2(X)$ . As both  $\text{IQD}_{\alpha}$  and  $\rho_h$  are location invariant, we can, without loss of generality, assume the allocation  $(X_1, X_2)$  satisfies  $Q_{1/2}^-(X_1) = 0$ . Let  $A$  be a right  $\alpha$ -tail event of  $X_1$  and  $B \subseteq A^c$  be a left  $\alpha$ -tail event of  $X_1$ . Hence,  $\mathbb{P}(A) = \mathbb{P}(B) = \alpha$  and  $X_1(\omega_B) \leq X_1(\omega) \leq X_1(\omega_A)$  a.s. for  $\omega_A \in A$ ,  $\omega_B \in B$  and  $\omega \in (A \cup B)^c$ . Let  $X_1^* = X_1 \mathbb{1}_{(A \cup B)^c}$  and  $h^* = h \wedge \lambda$ . Recall that  $\text{IQD}_0 = Q_0^- - Q_1^+$  is the range functional. It is straightforward to verify that  $\text{IQD}_{\alpha}(X_1) = \text{IQD}_0(X_1^*)$  and that  $h^*$  is concave. Further, notice that  $\lambda \text{IQD}_0 \geq \rho_{h^*}$ ,  $\rho_h \geq \rho_{h^*}$  and  $\rho_{h^*}$  is subadditive. Therefore,

$$\lambda \text{IQD}_{\alpha}(X_1) + \rho_h(X_2) = \lambda \text{IQD}_0(X_1^*) + \rho_h(X_2) \geq \rho_{h^*}(X_1^*) + \rho_{h^*}(X_2) \geq \rho_{h^*}(X_1^* + X_2).$$

As  $Q_{1/2}^-(X_1) = 0$ , we have, in the a.s. sense,  $X_1 \geq 0$  on  $A$  and  $X_1 \leq 0$  on  $B$ ; that is,

$X_1^* + X_2 = X$  on  $(A \cup B)^c$ ,  $X_1^* + X_2 \geq X$  on  $B$ , and  $X_1^* + X_2 \leq X$  on  $A$ . For any  $x \in \mathbb{R}$ , we have

$$\mathbb{P}(X_1^* + X_2 > x) \geq \mathbb{P}(X > x, (A \cup B)^c) + \mathbb{P}(X > x, B) \geq \mathbb{P}(X > x) - \mathbb{P}(A) = \mathbb{P}(X > x) - \alpha,$$

and similarly,  $\mathbb{P}(X_1^* + X_2 \leq x) \geq \mathbb{P}(X \leq x) - \alpha$ . Therefore,

$$\mathbb{P}(X > x) - \alpha \leq \mathbb{P}(X_1^* + X_2 > x) \leq \mathbb{P}(X > x) + \alpha.$$

Let  $s \in \mathbb{R}$  be such that  $x \mapsto h^*(\mathbb{P}(X_1^* + X_2 > x))$  is increasing on  $(-\infty, s]$  and decreasing on  $[s, \infty)$ . Such  $s$  exists since  $h^*$  is first increasing and then decreasing. By treating  $h^*(t) = 0$  if  $t$  is outside  $[0, 1]$ , we have

$$\begin{aligned} \rho_{h^*}(X_1^* + X_2) &= \int_{-\infty}^s h^*(\mathbb{P}(X_1^* + X_2 > x)) dx + \int_s^{\infty} h^*(\mathbb{P}(X_1^* + X_2 > x)) dx \\ &\geq \int_{-\infty}^s h^*(\mathbb{P}(X > x) - \alpha) dx + \int_s^{\infty} h^*(\mathbb{P}(X > x) + \alpha, 1) dx \\ &\geq \int_{-\infty}^{\infty} \min \{h^*(\mathbb{P}(X > x) + \alpha), h^*(\mathbb{P}(X > x) - \alpha)\} dx = \rho_{G_\lambda^\alpha(h)}(X), \end{aligned}$$

where the last equality follows because

$$G_\lambda^\alpha(h)(t) = (h(t - \alpha) \wedge h(t + \alpha) \wedge \lambda) \mathbb{1}_{\{\alpha < t < 1 - \alpha\}} = h^*(t - \alpha) \wedge h^*(t + \alpha)$$

and  $G_\lambda^\alpha(h)(1) = 0$ . Therefore, we have  $\lambda \text{IQD}_\alpha(X_1) + \rho_h(X_2) \geq \rho_{G_\lambda^\alpha(h)}(X)$ .

Next, we give an allocation  $(X_1, X_2) \in \mathbb{A}_2(X)$  that attains the lower bound  $\rho_{G_\lambda^\alpha(h)}(X)$ . Define the function  $f(s) = h^*(\mathbb{P}(X > x) + \alpha) - h^*(\mathbb{P}(X > x) - \alpha)$  where  $h^*(t) = 0$  if  $t$  is outside  $[0, 1]$ . Since  $h^*$  is concave, the function  $s \mapsto f(s)$  is increasing on the set of  $s$  with  $\mathbb{P}(X > s) \in [\alpha, 1 - \alpha]$ . Moreover,  $f(s) \leq 0$  for  $s \leq Q_{1-\alpha}^-(X)$  and  $f(s) \geq 0$  for  $s \geq Q_\alpha^+(X)$ . Hence, there exists  $s^* \in [Q_{1-\alpha}^-(X), Q_\alpha^+(X)]$  such that  $f(s) \geq 0$  for  $s < s^*$  and  $f(s) \leq 0$  for  $s > s^*$ .

Let  $A$  be a right  $\alpha$ -tail event of  $X$  and  $B \subseteq A^c$  be a left  $\alpha$ -tail event of  $X$ . Write  $T = A \cup B$ . Let  $(Y_1, Y_2) \in \mathbb{A}_2^+(X \mathbb{1}_{T^c} + s^* \mathbb{1}_T)$  be a  $(\lambda, 1)$ -optimal allocation for  $(\text{IQD}_\alpha, \rho_h)$ . Define  $X_1 = (X - s^*) \mathbb{1}_T + Y_1$  and  $X_2 = Y_2$ ; clearly  $(X_1, X_2) \in \mathbb{A}_2(X)$ . By Theorem 3, we have

$$\lambda \text{IQD}_\alpha(X_1) + \rho_h(X_2) = \lambda \text{IQD}_0(Y_1) + \rho_h(Y_2) = \rho_{h^*}(X \mathbb{1}_{T^c} + s^* \mathbb{1}_T).$$

Note that

$$\begin{aligned}
\rho_{h^*}(X\mathbb{1}_{T^c} + s^*\mathbb{1}_T) &= \int_{Q_{1-\alpha}^-(X)}^{Q_\alpha^+(X)} h^*(\mathbb{P}(X\mathbb{1}_{T^c} + s^*\mathbb{1}_T > x)) dx \\
&= \int_{Q_{1-\alpha}^-(X)}^{s^*} h^*(\mathbb{P}(X > x, T^c) + 2\alpha) dx + \int_{s^*}^{Q_\alpha^+(X)} h^*(\mathbb{P}(X > x, T^c)) dx \\
&= \int_{Q_{1-\alpha}^-(X)}^{s^*} h^*(\mathbb{P}(X > x) + \alpha) dx + \int_{s^*}^{Q_\alpha^+(X)} h^*(\mathbb{P}(X > x) - \alpha) dx \\
&= \int_{Q_{1-\alpha}^-(X)}^{Q_\alpha^+(X)} \min\{h^*(\mathbb{P}(X > x) + \alpha), h^*(\mathbb{P}(X > x) - \alpha)\} dx = \rho_{G_\lambda^\alpha(h)}(X),
\end{aligned}$$

where the second-last equality is due to the definition of  $s^*$ . Therefore, the lower bound  $\rho_{G_\lambda^\alpha(h)}(X)$  can be attained. Thus,  $(\lambda\text{IQD}_\alpha)\square\rho_h = \rho_{G_\lambda^\alpha(h)}(X)$ .  $\square$

*Proof of Theorem 4.* As the cases  $I = [n]$  and  $N = [n]$  follow from Theorems 2 and 3 respectively, we assume that the sets  $I$  and  $N$  are non-empty.

(i) The equality  $\square_{i=1}^n(\lambda_i\rho_{h_i}) = \rho_{G_\lambda^\alpha(h)}$  follows from Lemma 5, Theorems 2 and 3, and the fact that the inf-convolution is associative (Lemma 2 of Liu et al. (2020)), which together yield

$$\square_{i=1}^n(\lambda_i\rho_{h_i}) = \left(\square_{i \in I}(\lambda_i\rho_{h_i})\right) \square \left(\square_{i \in N}(\lambda_i\rho_{h_i})\right) = (\lambda\text{IQD}_\alpha)\square\rho_h = \rho_{G_\lambda^\alpha(h)}.$$

(ii) Without loss of generality, we assume  $c = c_1 = \dots = c_n = 0$  and let  $Y = X\mathbb{1}_{(A \cup B)^c}$ . If  $\alpha \geq 1/2$ , it is straightforward to check that  $(X_1, \dots, X_n)$  is Pareto optimal as  $\rho_{h_i}(X_i) = 0$  for  $i \in [n]$ . Now, we assume  $\alpha < 1/2$ .

We first show that  $\rho_{h_i}(X_i) \leq \rho_{\hat{h}_i}(Y_i)$  for all  $i \in [n]$ . Note that  $\rho_{h_i}(X_i) = \rho_{h_i}(Y_i) = \rho_{\hat{h}_i}(Y_i)$  for all  $i \in N$ . We are left to show  $\text{IQD}_{\alpha_i}(X_i) \leq \text{IQD}_0(Y_i)$  for all  $i \in I$ . As  $X(\omega) \leq 0$  a.s. for  $\omega \in B_i$ ,

$$\mathbb{P}(X_i \leq Q_0^-(Y_i)) = \mathbb{P}(X\mathbb{1}_{A_i \cup B_i} + Y_i \leq Q_0^-(Y_i)) \geq \mathbb{P}(B_i) + \mathbb{P}((A_i \cup B_i)^c) = \alpha_i + 1 - 2\alpha_i = 1 - \alpha_i.$$

That is,  $Q_{\alpha_i}^-(X_i) \leq Q_0^-(Y_i)$ . Similarly,  $Q_{1-\alpha_i}^+(X_i) \geq Q_1^+(Y_i)$ . Hence,  $\rho_{h_i}(X) = \text{IQD}_{\alpha_i}(X_i) \leq \text{IQD}_0(Y_i) = \rho_{\hat{h}_i}(Y_i)$  for all  $i \in I$ .

Let  $(Y'_1, \dots, Y'_n)$  be a comonotonic improvement of  $(Y_1, \dots, Y_n)$ . The definition of comonotonic improvement and Pareto optimality of  $(Y_1, \dots, Y_n)$  imply that  $\rho_{\hat{h}_i}(Y_i) = \rho_{\hat{h}_i}(Y'_i)$  for all  $i \in [n]$ . First, if there exist some  $i \in N$  such that  $h_i(t) = 0$  on  $[0, 1]$ , then Pareto optimality of  $(Y'_1, \dots, Y'_n)$  implies that  $\rho_{\hat{h}_i}(Y'_i) = 0$  for each  $i \in [n]$ . This in turn implies that  $\rho_{h_i}(X_i) = 0$  for each  $i \in [n]$ , and hence  $(X_1, \dots, X_n)$  is Pareto optimal. Below, we assume for each  $i \in N$ ,  $h_i(t) > 0$  for some

$t \in (0, 1)$ , which gives that  $h_i(t) > 0$  for all  $t \in (0, 1)$  due to concavity.

As  $\hat{h}_i(1) = 0$  and  $\hat{h}_i(t) > 0$  for all  $i \in [n]$  and  $t \in (0, 1)$ , by Proposition 8, Pareto optimality of  $(Y'_1, \dots, Y'_n)$  implies that there exist  $K \subseteq [n]$  and a vector  $\lambda \in (0, \infty)^K$  such that  $(Y'_i)_{i \in K}$  solves  $\boxplus_{i \in K} \rho_{\lambda_i \hat{h}_i}(Y)$ , and  $Y'_i, i \notin K$  are constants. Denote by  $h^* = \bigwedge_{i \in N \cap K} (\lambda_i h_i)$  and  $\lambda^* = \bigwedge_{i \in I \cap K} \lambda_i > 0$ ; here, we set  $\inf \emptyset = \infty$ . Putting together several observations above, we get

$$\sum_{i \in K} \lambda_i \rho_{h_i}(X_i) \leq \sum_{i \in K} \rho_{\lambda_i \hat{h}_i}(Y_i) = \sum_{i \in K} \rho_{\lambda_i \hat{h}_i}(Y'_i) = \boxplus_{i \in K} \rho_{\lambda_i \hat{h}_i}(Y) = \rho_{h^* \wedge \lambda^*}(Y), \quad (17)$$

where the first inequality holds because  $\rho_{h_i}(X_i) \leq \rho_{\hat{h}_i}(Y_i)$  for all  $i \in [n]$ , the first equality holds because  $\rho_{\hat{h}_i}(Y_i) = \rho_{\hat{h}_i}(Y'_i)$  for all  $i \in [n]$ , the second equality is due to  $\lambda$ -optimality of  $(Y'_i)_{i \in K}$  whose component-wise sum is  $Y$  plus a constant, and the last equality is due to Theorem 3. Furthermore, for  $i \notin K$ , we have  $0 \leq \rho_{h_i}(X_i) \leq \rho_{\hat{h}_i}(c_i) = 0$ ; that is  $\rho_{h_i}(X_i) = 0$ . Note that

$$\rho_{h^* \wedge \lambda^*}(Y) = \rho_{h^* \wedge \lambda^*}(X \mathbb{1}_{(A \cup B)^c}) = \rho_{G_{\lambda^*}^\alpha(h^*)}(X). \quad (18)$$

Take  $\beta \geq \lambda^*$ . If  $i \in N \setminus K$ , then  $X_i = Y'_i$  is a constant. Write  $Z = \sum_{i \in I \cup K} X_i$ . Using (17) and (18), we get

$$\sum_{i \in K} \lambda_i \rho_{h_i}(X_i) + \sum_{i \in I \setminus K} \beta \rho_{h_i}(X_i) \leq \rho_{G_{\lambda^*}^\alpha(h^*)}(X) = \rho_{G_{\lambda^*}^\alpha(h^*)} \left( X - \sum_{i \in N \setminus K} X_i \right) = \rho_{G_{\lambda^*}^\alpha(h^*)}(Z). \quad (19)$$

Using part (i), we have

$$\left( \square_{i \in K} (\lambda_i \rho_{h_i}) \right) \square \left( \square_{i \in I \setminus K} (\beta \rho_{h_i}) \right) = \rho_{G_{\lambda^*}^\alpha(h^*)}.$$

Therefore, (19) implies that  $(X_i)_{i \in I \cup K} \in \mathbb{A}_n(Z)$  minimizes  $\sum_{i \in K} \lambda_i \rho_{h_i}(X_i) + \sum_{i \in I \setminus K} \beta \rho_{h_i}(X_i)$ . Since also  $\rho(X_i) = 0$  for  $i \notin K$ , we conclude that  $(X_1, \dots, X_n)$  is Pareto optimal.  $\square$

## D Proof of of results in Section 7

*Proof of Lemma 4.* Let  $g(t) = h(\mathbb{P}_0(X > Q_t^{\mathbb{P}}(X)))$  for  $t \in [0, 1]$ , where  $Q_t^{\mathbb{P}}(X)$  is the left quantile under the measure  $\mathbb{P}$ . We first show that  $g(\mathbb{P}(X > x)) = h(\mathbb{P}_0(X > x))$  for all  $x \in \mathbb{R}$ . It is clear that  $g(\mathbb{P}(X > x)) = h(\mathbb{P}_0(X > Q_{\mathbb{P}(X > x)}^{\mathbb{P}}(X)))$ . By the definition of  $Q_t^{\mathbb{P}}$ , we have  $Q_{\mathbb{P}(X > x)}^{\mathbb{P}}(X) \leq x$ . For  $x \in \mathbb{R}$ , if  $Q_{\mathbb{P}(X > x)}^{\mathbb{P}}(X) = x$ , then it is clear that  $g(\mathbb{P}(X > x)) = h(\mathbb{P}_0(X > x))$ . If  $Q_{\mathbb{P}(X > x)}^{\mathbb{P}}(X) < x$ ,

we have  $\mathbb{P}(Q_{\mathbb{P}(X>x)}^{\mathbb{P}}(X) < X \leq x) = 0$ . As  $\mathbb{P}_0 \ll \mathbb{P}$ , we have  $\mathbb{P}_0(Q_{\mathbb{P}(X>x)}^{\mathbb{P}}(X) < X \leq x) = 0$ . Hence,

$$h(\mathbb{P}_0(X > Q_{\mathbb{P}(X>x)}^{\mathbb{P}}(X))) = h\left(\mathbb{P}_0(x \geq X > Q_{\mathbb{P}(X>x)}^{\mathbb{P}}(X)) + \mathbb{P}_0(X > x)\right) = h(\mathbb{P}_0(X > x)).$$

Taking  $t \uparrow 1$ , we obtain  $g(1) = h(1)$ .

Next, we will show  $\rho_h^{\mathbb{P}_0}(f(X)) = \rho_g^{\mathbb{P}}(f(X))$  for any increasing function  $f : \mathbb{R} \rightarrow \mathbb{R}$ . Denote by  $f^{-1}(x) = \inf\{y : f(y) > x\}$  the pseudo-inverse function of  $f$ . As  $\mathbb{P}(X = x) = 0$  for all  $x \in \mathbb{R}$  and  $\mathbb{P}_0 \ll \mathbb{P}$ , we have  $\mathbb{P}(X = x) = \mathbb{P}_0(X = x)$  for all  $x \in \mathbb{R}$ . Hence,

$$\begin{aligned} \rho_h^{\mathbb{P}_0}(f(X)) &= \int_0^\infty h(\mathbb{P}_0(f(X) > x)) dx + \int_{-\infty}^0 (h(\mathbb{P}_0(f(X) > x)) - h(1)) dx \\ &= \int_0^\infty h(\mathbb{P}_0(X > f^{-1}(x)) + \mathbb{P}_0(X = f^{-1}(x))\mathbf{1}_{\{f(f^{-1}(x)) > x\}}) dx \\ &\quad + \int_{-\infty}^0 (h(\mathbb{P}_0(X > f^{-1}(x)) + \mathbb{P}_0(X = f^{-1}(x))\mathbf{1}_{\{f(f^{-1}(x)) > x\}}) - h(1)) dx \\ &= \int_0^\infty h(\mathbb{P}_0(X > f^{-1}(x))) dx + \int_{-\infty}^0 (h(\mathbb{P}_0(X > f^{-1}(x))) - h(1)) dx \\ &= \int_0^\infty g(\mathbb{P}(X > f^{-1}(x))) dx + \int_{-\infty}^0 (g(\mathbb{P}(X > f^{-1}(x))) - h(1)) dx = \rho_g^{\mathbb{P}}(f(X)), \end{aligned}$$

as desired.  $\square$

*Proof of Proposition 10.* Let  $\mathbb{P} = 1/n \sum_{i=1}^n \mathbb{P}_i$  and  $g_i(t) = h_i(\mathbb{P}_i(X > Q_t^{\mathbb{P}}(X)))$  for  $t \in [0, 1]$ . It is clear that  $X$  also has a density function under  $\mathbb{P}$  and  $\rho_{h_i}^{\mathbb{P}_i}(f(X)) = \rho_{g_i}^{\mathbb{P}}(f(X))$  for increasing functions  $f$  and  $i \in [n]$  by Lemma 4. Hence,  $(\rho_{h_1}^{\mathbb{P}_1}, \dots, \rho_{h_n}^{\mathbb{P}_n})$  and  $(\rho_{g_1}^{\mathbb{P}}, \dots, \rho_{g_n}^{\mathbb{P}})$  have the same class of Pareto-optimal allocations.  $\square$

## E Omitted details in Section 6

We present the functions  $G_\lambda^\alpha(h)$  for Cases 1 to 6 in Section 6.2 which yield the allocations that we present in that section.

**Case 1:** When  $c_1 \geq 1/2$  and  $c_3 \geq 1/2$  it is

$$G_\lambda^\alpha(h)(t) = \lambda_2((t - \alpha) \wedge (1 - t - \alpha)) \mathbf{1}_{\{\alpha < t < 1 - \alpha\}}.$$

**Case 2:** When  $c_2 \geq 1/2$  and  $c_3 \leq \alpha$  it is

$$G_\lambda^\alpha(h)(t) = \lambda_3([(t - \alpha)(1 + \alpha - t)] \wedge [(t + \alpha)(1 - \alpha - t)]) \mathbf{1}_{\{\alpha < t < 1 - \alpha\}}.$$

**Case 3:** When either  $\alpha < c_2 < c_3 < 1/2$  or  $\alpha < c_1 < 1/2 < c_3$  it is

$$G_\lambda^\alpha(h)(t) = (\lambda_2[(t - \alpha) \wedge (1 - t - \alpha)] \wedge \lambda_1) \mathbb{1}_{\{\alpha < t < 1 - \alpha\}}.$$

**Case 4:** When  $c_3 \leq \alpha < c_2 < 1/2$  it is

$$G_\lambda^\alpha(h)(t) = (\lambda_3[(t - \alpha)(1 + \alpha - t)] \wedge [(t + \alpha)(1 - \alpha - t)] \wedge \lambda_1) \mathbb{1}_{\{\alpha < t < 1 - \alpha\}}.$$

**Case 5:** When  $\alpha < c_3 < 1/2 < c_2$ , it is

$$G_\lambda^\alpha(h)(t) = \begin{cases} 0, & t \in [0, \alpha] \cup [1 - \alpha, 1], \\ \lambda_2(t - \alpha), & t \in (\alpha, c_3), \\ \lambda_3(t - \alpha)(1 - t + \alpha), & t \in [c_3, 1/2), \\ \lambda_3(t + \alpha)(1 - t - \alpha), & t \in [1/2, 1 - c_3), \\ \lambda_2(1 - \alpha - t), & t \in [1 - c_3, 1 - \alpha]. \end{cases}$$

**Case 6:** When  $\alpha < c_3 \leq c_2 < 1/2$  it is

$$G_\lambda^\alpha(h)(t) = \begin{cases} 0, & t \in [0, \alpha] \cup [1 - \alpha, 1], \\ \lambda_2(t - \alpha), & t \in (\alpha, c_3), \\ \lambda_3(t - \alpha)(1 - t + \alpha), & t \in [c_3, c_2), \\ \lambda_1, & t \in [c_2, 1 - c_2), \\ \lambda_3(t + \alpha)(1 - t - \alpha), & t \in [1 - c_2, 1 - c_3), \\ \lambda_2(1 - t - \alpha), & t \in [1 - c_3, 1 - \alpha]. \end{cases}$$