

**INF-CONVOLUTION OF RISK MEASURES ON REARRANGEMENT
INVARIANT SPACES**

by

Shengzhong Chen, Bachelor of Science, Ryerson University, 2017

A thesis presented to Ryerson University
in partial fulfilment of the
requirements for the degree of
Master of Science
in the program of
Applied Mathematics

Toronto, Ontario, Canada, 2019

©Shengzhong Chen, 2019

Author's Declaration

I hereby declare that I am the sole author of this thesis. I authorize Ryerson University to lend this thesis to other institutions or individuals for the purpose of scholarly research.

I further authorize Ryerson University to reproduce this thesis by photocopying or by other means, in total or in part, by the request of the other institutions or individuals for the purpose of the scholarly research.

I understand that my thesis may be made electronically available to the public.

INF-CONVOLUTION OF RISK MEASURES ON REARRANGEMENT INVARIANT SPACES

Shengzhong Chen

Master of Science, 2019

Applied Mathematics Program

Ryerson University

ABSTRACT

The problem of optimal capital and risk allocation among economic agents, has played a pre-dominant role in the respective academic and industrial research areas for decades. Typically as risk occurs in face of randomness the risks which are to be measured are identified with real-valued random variables on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Consider a model space \mathcal{X} , and n economic agents with initial endowments $X_1, \dots, X_n \in \mathcal{X}$ who assess the riskiness of their positions by means of law-invariant convex risk measures $\rho_i : \mathcal{X} \rightarrow (-\infty, \infty]$. In order to minimize total and individual risk, the agents redistribute the aggregate endowment $X = X_1 + \dots + X_n$ among themselves. An optimal capital and risk allocation Y_1, \dots, Y_n satisfies $Y_1 + \dots + Y_n = X$ and

$$\rho_1(Y_1) + \dots + \rho_n(Y_n) = \inf \left\{ \sum_{i=1}^n \rho_i(X_i) : X_i \in \mathcal{X}, i = 1, \dots, n, \text{ and } \sum_{i=1}^n X_i = X \right\}, \quad (0.1)$$

where $\square_{i=1}^n \rho_i(X) = \inf \left\{ \sum_{i=1}^n \rho_i(X_i) : X_i \in \mathcal{X}, i = 1, \dots, n, \text{ and } \sum_{i=1}^n X_i = X \right\}$ is the inf-convolution of ρ_1, \dots, ρ_n . In 2008, Filipović and Svindland proved that if \mathcal{X} is an $L^p(\mathbb{P})$ for some $1 \leq p \leq \infty$ and ρ_i satisfy a suitable continuity condition (i.e. Fatou property), then Problem (0.1) always admits a solution. To reflect the fact of randomness of risk, we should consider the model space \mathcal{X} chosen for risk evaluations to be as general as possible. The main contribution of this thesis is Theorem 4.10 has been published in [9]. It extends Filipović and Svindland's result from L^p spaces to general rearrangement invariant (r.i.) spaces.

ACKNOWLEDGEMENTS

Firstly, I would like to express my gratitude to my supervisors Dr. Foivos Xanthos and Dr. Niushan Gao. They have given me great support over the years from my BSc degree all the way to my MSc degree. Thank you for creating a wonderful learning environment for me, and thanks for all your inspiration, guidance and encouragement in developing this thesis.

Secondly, I want to thank the faculty of Mathematics at Ryerson University for giving me an opportunity to study here and for providing me with financial assistance. In particular, I want to thanks to Dr. Dzungminh Ha and Dr. Silvana Ilie for their tremendous support in the past years.

Finally, I want to thanks to all my families for their continuous and unconditional support ever since I came to Canada.

Table of Contents

Declaration	ii
Abstract	iii
Acknowledgements	iv
1 Axioms of risk measures and Fatou-type properties	1
1.1 Introduction	1
1.2 Model spaces	2
1.3 Axioms and Fatou-type properties of risk measures	5
2 Fenchel-Moreau Duality theorem	8
2.1 Convex and lower semicontinuous functions	8
2.2 Fenchel-Moreau theorem and its applications	11
3 Inf-convolution of risk measures on L^1	26
4 Inf-convolution of risk measures on rearrangement invariant spaces	39
4.1 Orlicz space	39
4.2 Dual representation of risk measures on Orlicz space	42
4.3 Inf-convolution of risk measures on r.i. spaces	43
Bibliography	47

Chapter 1

Axioms of risk measures and Fatou-type properties

1.1 Introduction

The theory of risk measures is well-established and has been extensively studied in the growing field of mathematical finance in the past decades. Measuring the risk of a financial position is a complex process which is connected with many features in the financial market. In essence, a risk measure can be viewed as a rule to assign a certain indicator of risk — typically a capital requirement — to a given financial position of a financial institution. The problem of optimal capital and risk allocation among several economic agents, or business units, has played a predominant role in the respective academic and industrial research areas for decades. These optimal risk allocation problems can be interpreted as problems that not only to minimize the total risk but also to determine the optimal allocation. In general, it is a difficult task to find the optimal allocation explicitly and the minimum of total risk may not be attainable. Föllmer and Schied ([15]), and Frittelli and Gianin ([16]), respectively, has drawn the attention to study these problems using a new kind of mathematical approach, inf-convolution of risk measures. One important result in ([14], [21], [9]) showed that the capital and risk allocation problem on L^1 space may always admits an

optimal allocation among the agents if each risk measure involved carries some desirable properties, e.g., convexity, continuity, and law-invariance. These properties are essential ingredients of some complex optimal risk allocation problems. Motivated by this result, we want to investigate the existence of such optimal allocation on a more general framework of model spaces, rearrangement invariant (r.i.) spaces.

This thesis is organized as follows. In Chapter 1, we discuss the choices of model spaces and introducing the axioms and Fatou-type properties of risk measures. In Chapter 2, we study the Fenchel-Moreau duality theorem which has played an important role in the applications of convex risk measures in the field of risk management. In Chapter 3, we introduce the inf-convolution of risk measures and deliver the proof of the most useful result of theorem 2.5 in Filipović and Svindland ([14]). Finally, we extend previous result from L^1 space to general r.i. spaces in Chapter 4.

1.2 Model spaces

Measuring the risk of a financial position has to be evaluated based on the model space \mathcal{X} . A *function space* is a topological space whose points are functions. There are many different kinds of function spaces. We say a function space over a fixed nonatomic probability space $(\Omega, \mathcal{F}, \mathbb{P})$ is an *order ideal* of $L^0 := L^0(\Omega, \mathcal{F}, \mathbb{P})$, i.e., a subspace of L^0 such that if $X \in \mathcal{X}$ and Y is a random variable such that $|Y| \leq |X|$, then $Y \in \mathcal{X}$. There is an ongoing debate on the right model space for financial risk measures, i.e. about what an ideal domain of definition for risk measures would be. Typically, as risk occurs in face of randomness, the risks which are to be measured are identified with real-valued random variables on some measurable space (Ω, \mathcal{F}) . The question which causes debate, however, is which space of random variables one should use as model space. Since risk is often referred to uncertainty about the underlying probabilistic mechanism, many scholars argue that model spaces should be robust in the sense of not depending too heavily on some specific probabilistic model [22]. In the past decades, the literature usually suggests one of the following model spaces:

1. L^0 , the spaces of all \mathbb{P} -almost sure (\mathbb{P} -a.s.) equivalence classes of random variables for some probability measure \mathbb{P} on (Ω, \mathcal{F}) .
2. L^∞ , the spaces of all bounded random variables or \mathbb{P} -a.s. equivalence classes of bounded random variables, i.e.

$$L^\infty := \{f \in L^0 \mid |f| \leq c, \text{ a.e. for some } c \in \mathbb{R}\}.$$

3. L^p ($1 \leq p < \infty$), the space of \mathbb{P} -a.s. equivalence classes of random variables with finite p -th moment, i.e.

$$L^p := \{f \in L^0 \mid \int |f|^p \mathbf{d}\mathbb{P} < \infty\}.$$

The L^0 and L^∞ spaces satisfy minimal model dependence since in fact they only depend on the null sets of the probability measure \mathbb{P} . The problem with choosing L^0 , however, is that these spaces are in general too large to reasonably define aggregation based risk measures on them. Moreover, if (Ω, \mathcal{F}) is not finite, L^0 do not allow for a locally convex topology which make it no longer suitable for optimization problem. On the other hand, the L^∞ space that has been used as model space in the early stage of axiomatic theory of risk measures is a Banach space, so in particular a locally convex space. However, it is obviously a very limited space, for example, the random variables with log-normal distribution (under some pricing probability measure) involved in Black-Scholes market models cannot be incorporated in L^∞ . Assuming frictionless markets (i.e. financial markets without transaction costs), there is no upper bound on the volumes and thus value of financial positions. Hence unbounded distributions appear quite naturally. In applications, risks with unbounded support and potentially heavy-tailed distributions are also commonly employed in the insurance business. From this point of view model spaces should at least be sufficiently large to include these standard unbounded models, and the L^p spaces have been proposed to resolve this issue. Problematic though is the strong dependence of L^p on the probability measure \mathbb{P} in that it is not invariant under equivalent changes of measure anymore. When different agents in the market base their respective evaluations on applying different references probability measures

respectively or when optimizing the preferences of some agent under constraints given by some pricing rules, it is in general not clear which reference probability one should choose and on which (locally convex) model space the analysis should be carried out. Further more, L^p and L^∞ spaces are not everything we can work on. There are many other model spaces should be interested in our researches, such as Orlicz space and Orlicz heart. Thus, when measuring the risks, we may want our model space includes as many random variables as possible to reflect the fact of randomness of risk, and it should not have too much constraints as well. This led to more general model spaces, so-called rearrangement invariant(r.i) spaces, whose norms are measured only depends on the distribution of random variables in it. These spaces have a considerably richer structure than L^p spaces. The following definition of r.i. space is given in Braverman ([7]).

Definition 1.1 (Rearrangement invariant (r.i.) space). *A Banach space $\mathcal{X} \neq \{0\}$ over a fixed probability space $(\Omega, \mathcal{F}, \mathbb{P})$ is said to be an rearrangement invariant (r.i.) space, if the following conditions hold:*

- (1) *If $|X| \leq |Y|$ and $Y \in \mathcal{X}$, then $X \in \mathcal{X}$ and $\|X\| \leq \|Y\|$.*
- (2) *if $\mathbb{P}(X \leq t) = \mathbb{P}(Y \leq t)$ for all $t \in \mathbb{R}$, i.e., X, Y have same distributions (written as $X \sim Y$), and $Y \in \mathcal{X}$, then $X \in \mathcal{X}$ and $\|X\| = \|Y\|$.*

The main property of r.i. space is that the norm of any element X only depends on the distribution of X . This is why the probabilistic methods have been used successfully.

Example 1.2. (r.i. spaces)

- (1) L^p spaces for $1 \leq p \leq \infty$;
- (2) Orlicz space L^Φ and Orlicz Heart H^Φ (see chapter 4).

Remark 1.3. For an r.i. space \mathcal{X} , it holds that $L^\infty \subset \mathcal{X} \subset L^1$ (see Corollary 6.7 in [4]).

As we mentioned that convexity, continuity and law-invariance are the desirable properties of risk measure we are interested in optimization problems. In particular, the continuity

properties are useful for applications and are responsible for corresponding representation. Also, convex duality techniques are desirable and are available as soon as the risk measures involved admit tractable dual representations. Thus, when applying the inf-convolution methods in the risk sharing problems, we are interested in whether the inf-convolution of risk measures preserves these desirable properties on r.i. spaces. In the following chapter, we start discuss some axioms and Fatou-type properties of risk measures.

1.3 Axioms and Fatou-type properties of risk measures

The axiomatic approach to monetary risk measures was introduced by Artzner et al. ([2]). The space \mathcal{X} is going to describe all possible financial positions $X : \Omega \rightarrow \mathbb{R}$, where $X(\omega)$ is the discounted net worth of the position at the end of the time period if the scenario $\omega \in \Omega$ is realized. Given a financial position X in \mathcal{X} , a risk measure, we refer to Artzner et al. ([2]), is defined as

$$\rho(X) = \inf \{m \in \mathbb{R} \mid X + m\mathbf{1} \in \mathcal{A}_\rho\},$$

where $\mathcal{A}_\rho := \{X \in \mathcal{X} \mid \rho(X) \leq 0\}$ is the set of all acceptable financial positions (related to the preference system). In a point of view, $\rho(X)$ is the minimum capital need to added to the financial position makes it acceptable.

A map $\rho : \mathcal{X} \rightarrow (-\infty, \infty]$ is called a monetary risk measure if it satisfies the following properties:

- (1) *Grounded property:* If $\rho(\mathbf{0}) = 0$ (so ρ is proper, i.e. it does not coincide with ∞).
- (2) *Monotonicity:* If $X \leq Y$, then $\rho(X) \geq \rho(Y)$, for all $X, Y \in \mathcal{X}$.
- (3) *Cash invariance:* If $\rho(X + m) = \rho(X) - m$. for all $m \in \mathbb{R}$ and all $X \in \mathcal{X}$.

A monetary risk measure ρ is called a convex risk measure if it satisfies

- (4) *Convexity:* If $\rho(\lambda X + (1 - \lambda)Y) \leq \lambda\rho(X) + (1 - \lambda)\rho(Y)$, for any $\lambda \in [0, 1]$ and all $X, Y \in \mathcal{X}$.

ρ is called a coherent risk measure if it also satisfies

- (5) *Positive homogeneity*: If $\rho(\lambda X) = \lambda\rho(X)$, for all $\lambda \geq 0$ and all $X \in \mathcal{X}$.

Continuity property is an important tool for approximation arguments, and in particular it implies qualitative robustness properties of risk measures which are important for their application. Continuity property of convex risk measures on L^p and consequences for its representation have been investigated in detail in Föllmer and Schied ([15]). The Fatou-type properties are continuity property of risk measures that describes in a rigorous way how close are our risk evaluations for different financial positions that are expected to have similar performance. The tractable dual representation is ensured on L^p spaces if the risk measures have the Fatou property. However, the Fatou property no longer guarantees the risk measure admits such representation in the general r.i. spaces, such as Orlicz spaces (see chapter 4). In order to overcome this obstacle, Gao and Xanthos introduced the strong Fatou property and highlight the importance of strong Fatou property in ([19]), which turns out to be the right continuity adjustment in the Orlicz space framework.

Let \mathcal{X} be an r.i. space. We say that $\rho : \mathcal{X} \rightarrow (-\infty, \infty]$ has the

1. *Fatou property* if $\rho(X) \leq \liminf_n \rho(X_n)$ whenever $(X_n) \subseteq \mathcal{X}$ and $X \in \mathcal{X}$ satisfy

$$X_n \xrightarrow{a.s.} X \text{ and } |X_n| \leq X_0$$

for some $X_0 \in \mathcal{X}$ and all $n \in \mathbb{N}$.

2. *Strong Fatou property* if $\rho(X) \leq \liminf_n \rho(X_n)$ whenever $(X_n) \subseteq \mathcal{X}$ and $X \in \mathcal{X}$ satisfy

$$X_n \xrightarrow{a.s.} X \text{ and } \sup \|X_n\| < \infty.$$

It is clear that the Fatou property is weaker than the strong Fatou property, also it is not difficult to verify that they coincide on L^∞ .

Let's consider a common used risk measure Expected Shortfall as an example. For $\alpha \in (0, 1)$, we define **Value-at-Risk** at level α by

$$\text{VaR}_\alpha(X) := \inf \{m \in \mathbb{R} : \mathbb{P}(X + m < 0) \leq \alpha\}, \quad X \in L^0.$$

For $\alpha \in (0, 1]$, define **Expected Shortfall** at level α by

$$\text{ES}_\alpha(X) = \frac{1}{\alpha} \int_0^\alpha \text{Var}_\beta(X) d\beta, \quad X \in L^1.$$

By this definition, Expected Shortfall make sense for any $X \in L^1$ and it is well known that ES_α is a coherent risk measure. Note that $\text{ES}_1(X) = \mathbb{E}[X]$, so it has the Fatou property but not the strong Fatou property on L^1 . Indeed, in view of nonatomicity, take a decreasing sequence of measurable sets (A_n) such that $\mathbb{P}(A_n) = \frac{1}{n}$ for any $n \in \mathbb{N}$. Set $X_n = -n\mathbf{1}_{A_n}$ for every $n \in \mathbb{N}$. Then $\|X_n\| = 1$ for all $n \in \mathbb{N}$, $X_n \xrightarrow{a.s.} 0$, but $\liminf_n \mathbb{E}[X_n] = -1 < 0 = \mathbb{E}[0]$. However, when $\alpha \in (0, 1)$, the Expected Shortfall does have the strong Fatou property on any r.i space ([9]).

Proposition 1.4. *For $\alpha \in (0, 1)$, ES_α has the strong Fatou property on any r.i space \mathcal{X} .*

Proof. Fix any $\varepsilon \in (0, 1 - \alpha)$. By Egorov's Theorem, there exist a measurable set B and $n_0 \in \mathbb{N}$ such that

$$\mathbb{P}(B) < \varepsilon, \quad \text{and} \quad |X_n - X| < \varepsilon \quad \text{on } B^c \quad \text{for all } n \geq n_0.$$

Pick any $\beta \in (0, \alpha)$, and take any $n \geq n_0$. Let $m := \text{Var}_\beta(X_n)$ and $m' := m + \varepsilon$. It follows from $\{X + m' < 0\} \subseteq (\{X + m' < 0\} \cap B) \cup (\{X + m' < 0\} \cap \{X_n < X + \varepsilon\}) \subseteq B \cup \{X_n + m < 0\}$ that $\mathbb{P}(X + m' < 0) \leq \mathbb{P}(B) + \mathbb{P}(X_n + m < 0) \leq \varepsilon + \beta$, and consequently,

$$\text{Var}_{\beta+\varepsilon}(X) \leq m' = \text{Var}_\beta(X_n) + \varepsilon.$$

Since this holds for any $\beta \in (0, \alpha)$ and any $n \geq n_0$, integrating with respect to β over $(0, \alpha)$ implies $\frac{1}{\alpha} \int_\varepsilon^{\alpha+\varepsilon} \text{Var}_\beta(X) d\beta = \frac{1}{\alpha} \int_0^\alpha \text{Var}_{\beta+\varepsilon}(X) d\beta \leq \frac{1}{\alpha} \int_0^\alpha \text{Var}_\beta(X_n) d\beta + \varepsilon = \text{ES}_\alpha(X_n) + \varepsilon$ for all $n \geq n_0$. Taking infimum over $n \geq n_0$, we have

$$\frac{1}{\alpha} \int_\varepsilon^{\alpha+\varepsilon} \text{Var}_\beta(X) d\beta \leq \inf_{n \geq n_0} \text{ES}_\alpha(X_n) + \varepsilon \leq \liminf_n \text{ES}_\alpha(X_n) + \varepsilon.$$

Now, since $\text{Var}_\bullet(X) \in L^1(0, 1]$, letting $\varepsilon \rightarrow 0$, we have $\text{ES}_\alpha(X) \leq \liminf_n \text{ES}_\alpha(X_n)$. □

Chapter 2

Fenchel-Moreau Duality theorem

The convex analysis and, especially, the duality theory has surprisingly found in applications of risk measures in the field of convex optimization. Convex duality methods often lead to new insight, computational techniques and optimality conditions. In this section we will study some background theory in convexity, optimization, and convex duality needed to understand how duality methods are used in mathematical finance. we first recall some basic notions of convexity theory, such as convex sets, convex functions and properties of these, based on the presentation of convexity in Rochafellar ([24]). Then we study conjugate functions and Fenchel-Moreau dual representation theorem refer to Brezis ([8]), which is the most fundamental theorem in duality theory. Finally we discuss the applications of Fenchel-Moreau theorem for convex risk measures in L^p spaces.

2.1 Convex and lower semicontinuous functions

A topological space (\mathcal{X}, τ) is called a locally convex topological space, if every neighbourhood of zero includes a convex neighbourhood of zero.

Example 2.1. (Locally convex topological spaces)

- (1) $(\mathcal{X}, \|\cdot\|)$.
- (2) $(\mathcal{X}, \sigma(\mathcal{X}, \mathcal{Y}))$, where $\sigma(\mathcal{X}, \mathcal{Y})$ is weak topology.

Definition 2.2 (Convex set). *A subset $C \subseteq \mathcal{X}$ is said to be a convex set if*

$$tX + (1 - t)Y \in C, \text{ for all } X, Y \in C \text{ and } t \in [0, 1].$$

Moreover, C is said to be a cone if

$$tX \in C, \text{ for all } X \in C \text{ and } t > 0.$$

The following are some basic properties of convex sets in a topological vector space:

1. The sum of two convex sets is convex.
2. Scalar multiples of convex sets is convex.
3. The intersection of an arbitrary family of convex sets is convex.
4. Both the interior and the closure of a convex set are convex.

In the following context, \mathcal{X} denoted as a locally convex topological space. Let $\varphi : \mathcal{X} \rightarrow (-\infty, \infty]$ be a function on \mathcal{X} .

Definition 2.3 (Convex function). *We say φ is convex if the inequality*

$$\varphi(tX + (1 - t)Y) \leq t\varphi(X) + (1 - t)\varphi(Y)$$

holds for all $X, Y \in \mathcal{X}$ and all $t \in [0, 1]$.

Definition 2.4 ((Effective) domain). *The effective domain of φ denoted by $D(\varphi)$ is defined as*

$$D(\varphi) = \{X \in \mathcal{X} \mid \varphi(X) < +\infty\}.$$

In what follows, we denote $\{X \in \mathcal{X} \mid \varphi(X) = \alpha\}$ by $\{\varphi = \alpha\}$ and denote $\{X \in \mathcal{X} \mid \varphi(X) \geq \alpha\}$ by $\{\varphi \geq \alpha\}$, etc.

Definition 2.5 (Epigraph). *The epigraph of φ denoted by $\text{epi}(\varphi)$ is defined as*

$$\text{epi}(\varphi) = \{[X, \lambda] \in \mathcal{X} \times \mathbb{R} \mid \varphi(X) \leq \lambda\}.$$

The following are some basic properties of the convex function φ :

1. If φ has a local minimum M , then M is also a global minimum for φ .
2. $\lambda\varphi$ is convex for $\lambda > 0$.
3. (*Effective*) domain $D(\varphi)$ is a convex set.
4. φ is convex if and only if $\text{epi}(\varphi)$ is convex in $\mathcal{X} \times \mathbb{R}$.
5. φ is convex, then for every $\lambda \in \mathbb{R}$ the set $\{\varphi(X) \leq \lambda\}$ is convex; but the converse is not true.
6. The sum of convex functions is convex, e.g. if φ_1 and φ_2 is convex, then $\varphi_1 + \varphi_2$ is convex.
7. For $X_1, \dots, X_n \in \mathcal{X}$, $a_1, \dots, a_n \in \mathbb{R}^+$, and $\sum_{i=1}^n a_i = 1$, the following inequality holds:

$$\varphi\left(\sum_{i=1}^n a_i X_i\right) \leq \sum_{i=1}^n a_i \varphi(X_i).$$

8. If $(\varphi_i)_{i \in I}$ is a family of (finite) convex function, then their superior envelope is also convex, that is $\varphi(X) = \sup_{i \in I} \varphi_i(X)$ is convex.

Definition 2.6 (Lower semicontinuous (l.s.c.)). *We say φ is τ -l.s.c. at a point X_0 , if for each $\lambda \in \mathbb{R}$ such that $\varphi(X_0) > \lambda$ there exists a neighbourhood U of X_0 such that $\varphi(X) > \lambda$ for all $X \in U$.*

We say φ is l.s.c on \mathcal{X} if it is l.s.c at every point $X \in \mathcal{X}$.

Lemma 2.7. *φ is τ -l.s.c. if and only if $\{\varphi \leq \lambda\}$ is τ -closed for every $\lambda \in \mathbb{R}$.*

Proof. The sublevel sets $\{\varphi \leq \lambda\}$ are closed for all $\lambda \in \mathbb{R}$ if and only if $\{\varphi \leq \lambda\}^c$ are open for all λ . But this happens if and only if all $Y \in \{\varphi \leq \lambda\}^c$ are interior points, which is equivalent with that for each $Y \in \{\varphi \leq \lambda\}^c$ there is a neighbourhood U such that $U \subseteq \{\varphi \leq \lambda\}^c$, i.e. $\varphi(U) > \lambda$. But this is the definition of φ being lower semicontinuous at the point Y . Since this argument holds for all $Y \in \mathcal{X}$ (by choosing different λ), φ is lower semicontinuous. \square

The following are some basic properties of l.s.c functions:

1. φ is l.s.c. if and only if $\text{epi}(\varphi)$ is closed in $\mathcal{X} \times \mathbb{R}$.
2. φ is l.s.c., then for every sequence (X_n) in \mathcal{X} such that $X_n \xrightarrow{\tau} X$, we have $\liminf_n \varphi(X_n) \geq \varphi(X)$, and conversely is true if \mathcal{X} is a metric space.
3. If φ_1 and φ_2 is l.s.c., then $\varphi_1 + \varphi_2$ is l.s.c..
4. If $(\varphi_i)_{i \in I}$ is a family of l.s.c function, then their superior envelope is also l.s.c., that is $\varphi(X) = \sup_{i \in I} \varphi_i(X)$ is l.s.c..

2.2 Fenchel-Moreau theorem and its applications

Another advantage with convex functions in optimization is that one can exploit the duality properties in order to solve problems. Dual pairs are an extremely useful way of obtaining locally convex topological spaces. The following definition of dual pair is given in Aliprantis and Border ([1]).

Definition 2.8 (Dual pair). *A dual pair is a pair $(\mathcal{X}, \mathcal{Y})$ of vector spaces together with a bilinear functional $(X, Y) \mapsto \langle X, Y \rangle$, from $\mathcal{X} \times \mathcal{Y} \rightarrow \mathbb{R}$, that separates the points of \mathcal{X} and \mathcal{Y} . That is*

- (1) *The mapping $Y \mapsto \langle X, Y \rangle$ is linear for each $X \in \mathcal{X}$.*
- (2) *The mapping $X \mapsto \langle X, Y \rangle$ is linear for each $Y \in \mathcal{Y}$.*
- (3) *If $\langle X, Y \rangle = 0$ for each $Y \in \mathcal{Y}$, then $X = 0$.*
- (4) *If $\langle X, Y \rangle = 0$ for each $X \in \mathcal{X}$, then $Y = 0$.*

Each space of the dual pair $(\mathcal{X}, \mathcal{Y})$ can be interpreted as a set of linear function on the other. for instance, each $X \in \mathcal{X}$, defines the linear functional $Y \mapsto \langle X, Y \rangle$. Conditions (1) and (2) are the ones required for the definition of a bilinear functional. The bilinear functional $(X, Y) \mapsto \langle X, Y \rangle$ is also called the duality (or the bilinearity) of the dual pair.

Example 2.9. (Examples of dual pair)

(1) $(\mathbb{R}^n, \mathbb{R}^n)$ under duality $\langle X, Y \rangle = \sum_{i=1}^n x_i y_i$.

(2) (L^p, L^q) under duality $\langle X, Y \rangle = \int XY \mathbf{d}\mathbb{P}$ for $1 \leq p \leq \infty$ and $\frac{1}{p} + \frac{1}{q} = 1$.

For a topological vector space (\mathcal{X}, τ) , the topological dual space of \mathcal{X} is the space of all linear continuous function on \mathcal{X} , defined as

$$(\mathcal{X}, \tau)' = \{f : \mathcal{X} \rightarrow \mathbb{R} \mid f \text{ is linear and } \tau\text{-continuous}\}.$$

If the τ is clear from the context, we will denote $(\mathcal{X}, \tau)'$ by \mathcal{X}' . If \mathcal{X} is a normed vector space, the norm dual space of $(\mathcal{X}, \|\cdot\|)'$ is denoted by \mathcal{X}^* , and is defined as

$$\mathcal{X}^* = \{f : \mathcal{X} \rightarrow \mathbb{R} \mid f \text{ is linear and norm continuous}\}.$$

Remark 2.10. The topological dual of $(\mathcal{X}, \sigma(\mathcal{X}, \mathcal{Y}))$ is \mathcal{Y} .

The topological dual of $\mathcal{X} \times \mathbb{R}$ is $\mathcal{X}' \times \mathbb{R}$ in the sense of the next lemma.

Lemma 2.11. For any function $\Phi \in (\mathcal{X} \times \mathbb{R})'$, there exists $(f, k) \in \mathcal{X}' \times \mathbb{R}$ such that

$$\Phi([X, \lambda]) = \langle f, X \rangle + k\lambda, \quad \text{for all } X \in \mathcal{X} \text{ and } \lambda \in \mathbb{R}.$$

Proof. Let us define $\langle f, X \rangle = \Phi([X, 0])$ for all $X \in \mathcal{X}$, then f is continuous linear function since $\Phi \in (\mathcal{X} \times \mathbb{R})^*$, thus $f \in \mathcal{X}'$. let $k = \Phi([0, 1])$, by linearity of Φ we have

$$\Phi([X, \lambda]) = \Phi([X, 0] + \lambda[0, 1]) = \Phi([X, 0]) + \lambda\Phi([0, 1]) = \langle f, X \rangle + k\lambda.$$

□

A Hyperplane is a set of the form $\{f = \alpha\}$, where f is a non-zero linear function on \mathcal{X} and $\alpha \in \mathbb{R}$. A hyperplane defines two strict half-spaces, $\{f > \alpha\}$ and $\{f < \alpha\}$, and two weak half-spaces, $\{f \geq \alpha\}$ and $\{f \leq \alpha\}$. The following Hahn-Banach separation theorem provides a fundamental understanding of the proof of Fenchel-Moreau duality theorem.

Theorem 2.12 (Hahn-Banach second geometric form). *Let \mathcal{X} be a locally convex topological space, A, B be two non-empty convex subsets of \mathcal{X} with $A \cap B = \emptyset$. Assume that A is closed and B is compact. Then there exists a closed hyperplane $\{f = \alpha\}$ with $f \in \mathcal{X}'$ that strictly separates A and B , that is there exists some $\varepsilon > 0$ such that*

$$f(X) + \varepsilon \leq \alpha \leq f(Y) - \varepsilon$$

for all $X \in A$ and all $Y \in B$.

We now come to a central notion of convex duality, the conjugate of a function.

Definition 2.13 (Conjugate function). *Let \mathcal{X} be a locally convex topological space, and $\varphi : \mathcal{X} \rightarrow (-\infty, +\infty]$ be a function such that $\varphi \not\equiv +\infty$ (i.e., $D(\varphi) \neq \emptyset$). We define the conjugate function $\varphi^* : \mathcal{X}' \rightarrow (-\infty, +\infty]$ to be*

$$\varphi^*(f) = \sup_{X \in \mathcal{X}} \{\langle f, X \rangle - \varphi(X)\} \quad (f \in \mathcal{X}')$$

Note that φ^* is convex and l.s.c. on $(\mathcal{X}', \sigma(\mathcal{X}', \mathcal{X}))$. Indeed, for each fixed $X \in \mathcal{X}$ the function $f \mapsto \langle f, X \rangle - \varphi(X)$ is convex and continuous (and thus l.s.c.) on \mathcal{X}' , it follows that the superior envelope of these functions (as X runs through \mathcal{X}) is convex and l.s.c.

Definition 2.14 (Biconjugate function). *Let \mathcal{X} be a locally convex topological space. For a function $\varphi : \mathcal{X} \rightarrow (-\infty, \infty]$, define the biconjugate of φ , φ^{**} , to be the conjugate of φ^* , so*

$$\varphi^{**}(X) = \sup_{f \in \mathcal{X}'} \{\langle f, X \rangle - \varphi^*(f)\} \quad (X \in \mathcal{X})$$

Proposition 2.15. *Let \mathcal{X} be a locally convex topological space, assume that $\varphi : \mathcal{X} \rightarrow (-\infty, \infty]$ is a l.s.c. convex function and $\varphi \not\equiv \infty$. Then $\varphi^* \not\equiv \infty$, and in particular, φ is bounded below by an affine continuous function.*

Proof. since $D(\varphi) \neq \emptyset$, let $X_0 \in D(\varphi)$ and let $\lambda_0 < \varphi(X_0)$. We apply Hahn-Banach second geometric form theorem in the space $\mathcal{X} \times \mathbb{R}$ with $A = \text{epi}(\varphi)$ and $B = [X_0, \lambda_0]$ (Since φ is convex and l.s.c., then A is closed and convex, B is compact since it is a singleton). So,

there exists a closed hyperplane $H = [\Phi = a]$ in $\mathcal{X} \times \mathbb{R}$ that strictly separates A and B, and function $\Phi \in (\mathcal{X} \times \mathbb{R})'$. By lemma 2.11 , we have

$$\Phi([X, \lambda]) = \Phi([X, 0] + \lambda[0, 1]) = \langle f, X \rangle + k\lambda$$

Writing that $\Phi > a$ on A and $\Phi < a$ on B, we have

$$\langle f, X \rangle + k\lambda > a \quad \text{for all } [X, \lambda] \in \text{epi}(\varphi)$$

and

$$\langle f, X_0 \rangle + k\lambda_0 < a$$

in particular, we have

$$\langle f, X \rangle + k\varphi(X) > a \quad \forall X \in D(\varphi) \tag{2.1}$$

and thus

$$\langle f, X_0 \rangle + k\varphi(X_0) > a > \langle f, X_0 \rangle + k\lambda_0$$

it follows that $k > 0$ since $\varphi(X_0) > \lambda_0$. Divided by $-k$ in both side of (2.1), we have

$$\langle -\frac{1}{k}f, X \rangle - \varphi(X) < -\frac{a}{k} \quad \text{for all } X \in D(\varphi)$$

therefore $\varphi^*(-\frac{1}{k}f) < +\infty$, and φ is bounded below, that is

$$\varphi(X) > \langle -\frac{1}{k}f, X \rangle + \frac{a}{k} \quad \text{for all } X \in \mathcal{X}$$

□

Theorem 2.16 (Fenchel-Moreau dual representation). *Let \mathcal{X} be a locally convex topological space, and $\varphi : \mathcal{X} \rightarrow (-\infty, \infty]$ be a convex function. Then the following are equivalent*

(1) φ is lower semicontinuous

(2) φ admits the representation as: $\varphi(X) = \sup_{f \in \mathcal{X}'} \{\langle f, X \rangle - \varphi^*(f)\}$,

where $\varphi^*(f) = \sup_{X \in \mathcal{X}} \{\langle f, X \rangle - \varphi(X)\}$

Proof. (2) \implies (1) is clear. Indeed, since for each $f \in \mathcal{X}'$ the $\varphi^*(f)$ is convex and l.s.c. then $\langle f, X \rangle - \varphi^*(f)$ is convex and l.s.c. on \mathcal{X} , thus it follows that φ is convex and l.s.c since the superior envelope of these functions (as f runs through \mathcal{X}') is convex and l.s.c.

(1) \implies (2). Let $\varphi^{**}(X) = \sup_{f \in \mathcal{X}'} \{\langle f, X \rangle - \varphi^*(f)\}$, we want to show that $\varphi = \varphi^{**}$.

First we prove that the result holds for $\varphi \geq 0$.

Note that $\varphi^{**} \leq \varphi$, since $\langle f, X \rangle - \varphi^*(f) \leq \varphi(X)$ for all $X \in \mathcal{X}$ and for all $f \in \mathcal{X}'$. In order to show that $\varphi^{**} = \varphi$ we prove by contradiction. We assume that $\varphi^{**}(X_0) < \varphi(X_0)$ for some $X_0 \in \mathcal{X}$, then $\varphi^{**}(X_0)$ is always finite. We apply Hahn-Banach second geometric form theorem in the space $\mathcal{X} \times \mathbb{R}$ with $A = \text{epi}(\varphi)$ and $B = [X_0, \varphi^{**}(X_0)]$. So, there exists linear continuous function $f \in \mathcal{X}'$, $k \in \mathbb{R}$, and $a \in \mathbb{R}$ such that

$$\langle f, X \rangle + k\lambda > a \quad \text{for all } [X, \lambda] \in \text{epi}(\varphi),$$

and

$$\langle f, X_0 \rangle + k\varphi^{**}(X_0) < a. \tag{2.2}$$

It follows that $k \geq 0$ (fix some $X \in D(\varphi)$ and let $\lambda \rightarrow \infty$), here we could possibly have $k = 0$.

Let $\varepsilon > 0$, since $\varphi \geq 0$, we have

$$\langle f, X \rangle + (k + \varepsilon)\varphi(X) \geq a \quad \text{for all } X \in D(\varphi).$$

Therefore

$$\varphi^*\left(-\frac{f}{k + \varepsilon}\right) \leq -\frac{a}{k + \varepsilon}.$$

It follows from the definition of $\varphi^{**}(X_0)$ that

$$\varphi^{**}(X_0) \geq \left\langle -\frac{f}{k + \varepsilon}, X_0 \right\rangle - \varphi^*\left(-\frac{f}{k + \varepsilon}\right) \geq \left\langle -\frac{f}{k + \varepsilon}, X_0 \right\rangle + \frac{a}{k + \varepsilon}.$$

Thus we have

$$\langle f, X_0 \rangle + (k + \varepsilon)\varphi^{**}(X_0) \geq a \quad \text{for all } \varepsilon > 0.$$

which contradicts with (2.2).

Now let us prove the general case.

Fix some $f_0 \in D(\varphi^*)$ (by proposition 2.15 we know that $D(\varphi^*) \neq \emptyset$). We define

$$\bar{\varphi}(X) = \varphi(X) - \langle f_0, X \rangle + \varphi^*(f_0).$$

Since f_0 is linear continuous then f_0 is convex l.s.c, thus $\bar{\varphi}$ is convex and l.s.c., $\bar{\varphi} \not\equiv \infty$ (since $\varphi \not\equiv \infty$ so there is some $X \in \mathcal{X}$ such that $\varphi(X) < \infty$, and $\varphi^*(f_0) < \infty$ since $f_0 \in D(\varphi^*)$), and $\bar{\varphi} \geq 0$ (since $\varphi^*(f_0) \geq \langle f_0, X \rangle - \varphi(X)$). We know from previous that $\bar{\varphi}^{**} = \bar{\varphi}$.

Now let us now compute $(\bar{\varphi})^*$ and $(\bar{\varphi})^{**}$. We have

$$\begin{aligned} (\bar{\varphi})^*(f) &= \sup_{X \in \mathcal{X}} \{ \langle f, X \rangle - \bar{\varphi}(X) \} \\ &= \sup_{X \in \mathcal{X}} \{ \langle f, X \rangle - \varphi(X) + \langle f_0, X \rangle + \varphi^*(f_0) \} \\ &= \sup_{X \in \mathcal{X}} \{ \langle f + f_0, X \rangle - \varphi(X) \} - \varphi^*(f_0) \\ &= \varphi^*(f + f_0) - \varphi^*(f_0) \end{aligned}$$

and

$$\begin{aligned} (\bar{\varphi})^{**}(X) &= \sup_{f \in \mathcal{X}'} \{ \langle f, X \rangle - \bar{\varphi}^*(f) \} \\ &= \sup_{f \in \mathcal{X}'} \{ \langle f, X \rangle - \varphi^*(f + f_0) + \varphi^*(f_0) \} \\ &= \sup_{f \in \mathcal{X}'} \{ \langle f + f_0, X \rangle - \varphi^*(f + f_0) - \langle f_0, X \rangle + \varphi^*(f_0) \} \\ &= \sup_{f \in \mathcal{X}'} \{ \langle f + f_0, X \rangle - \varphi^*(f + f_0) \} - \langle f_0, X \rangle + \varphi^*(f_0) \\ &= \sup_{g=f+f_0 \in \mathcal{X}'} \{ \langle g, X \rangle - \varphi^*(g) \} - \langle f_0, X \rangle + \varphi^*(f_0) \\ &= \varphi^{**}(X) - \langle f_0, X \rangle + \varphi^*(f_0) \end{aligned}$$

Writing that $(\bar{\varphi})^{**} = \bar{\varphi}$, we obtain $\varphi^{**} = \varphi$. □

To apply Fenchel-Moreau representation theorem to convex risk measures, the following extension of a classical theorem of Namioka stating that positive linear functionals on a Banach lattice are continuous is most useful. A Banach lattice is a real Banach space \mathcal{X}

endowed with an ordering \leq such that (\mathcal{X}, \leq) is a vector lattice, and the norm on \mathcal{X} is a lattice norm, that is, if $|X| \leq |Y|$ implies $\|X\| \leq \|Y\|$ for all $X, Y \in \mathcal{X}$.

Theorem 2.17 (Extended Namioka-Klee theorem [6]). *Any proper convex and monotone increasing functional $\pi : \mathcal{X} \rightarrow (-\infty, \infty]$ on a Banach lattice $(\mathcal{X}, \|\cdot\|)$ is continuous on $I_\pi := \text{int}(D(\pi))$ (the interior of $D(\pi)$). Moreover, it admits a dual representation as*

$$\pi(X) = \max_{Y \in \mathcal{X}_+^*} \{\langle Y, X \rangle - \pi^*(Y)\}, \text{ for all } X \in \text{int}(D(\pi))$$

Corollary 2.18. *Let $\rho : L^p \rightarrow (-\infty, \infty]$, $1 \leq p \leq \infty$, be a proper convex risk measure.*

Then

(1) ρ is continuous on I_ρ w.r.t. relative norm topology.

(2) Any finite convex risk measure on L^p , $1 \leq p \leq \infty$, is continuous.

Theorem 2.19 (Radon-Nikodym theorem). *Let μ, ν be two finite measures on (Ω, \mathcal{F}) such that $\nu \ll \mu$ (i.e. ν is absolutely continuous with respect to μ , that is, if $\mu(A) = 0$ for any $A \in \mathcal{F}$, then $\nu(A) = 0$). Then there exists a measurable function $f \geq 0$ such that*

$$\nu(A) = \int_A f \mathbf{d}\mu.$$

The function f is called Radon-Nikodym derivative and is denoted by $\frac{\mathbf{d}\nu}{\mathbf{d}\mu}$. Equivalently, we have $\int_\Omega g \mathbf{d}\nu = \int_\Omega f g \mathbf{d}\mu$ for any integrable and measurable function g .

In the following context, the function spaces \mathcal{X} we work on are r.i. spaces.

Definition 2.20 (Weak convergent). *Let (X_n) be a sequence in \mathcal{X} , we write $X_n \xrightarrow{\sigma(\mathcal{X}, \mathcal{X}')} X$ in \mathcal{X} if and only if $f(X_n) \rightarrow f(X)$ for every $f \in \mathcal{X}'$.*

Definition 2.21 (Order convergent). *A sequence $(X_n) \subseteq \mathcal{X}$ is said order converge to X in \mathcal{X} , written as $X_n \xrightarrow{o} X$, if $X_n \xrightarrow{a.s.} X \in \mathcal{X}$, and $|X_n| \leq X_0$ for some $X_0 \in \mathcal{X}$ and for all $n \in \mathbb{N}$.*

Definition 2.22 (Order continuous). *A linear function f on \mathcal{X} is said to be order continuous if $f(X_n) \rightarrow f(X)$ whenever $(X_n) \subseteq \mathcal{X}$ and $X_n \xrightarrow{o} X$.*

The *order continuous dual* of a r.i. space \mathcal{X} is denoted by \mathcal{X}_n^\sim and defined as

$$\mathcal{X}_n^\sim = \{f : \mathcal{X} \rightarrow \mathbb{R} \mid f \text{ is linear and order continuous}\}$$

In general, we have $\mathcal{X}_n^\sim \subseteq \mathcal{X}^*$, but they coincide if \mathcal{X} has *order continuous norm*, i.e., \mathcal{X} has a norm $\|\cdot\|$ such that $\|X_n\| \rightarrow 0$ whenever (X_n) is a sequence in \mathcal{X} and $X_n \downarrow 0$.

Lemma 2.23. $\mathcal{X}_n^\sim \subseteq \mathcal{X}^*$.

Proof. Let $f \in \mathcal{X}_n^\sim$ and $(X_n) \subseteq \mathcal{X}$ such that $X_n \xrightarrow{\|\cdot\|} 0$. Suppose that $f(X_n) \not\rightarrow 0$. Then there exists some $\varepsilon > 0$, and a subsequence (X_{n_k}) of (X_n) such that for all $k \in \mathbb{N}$, we have $\|X_{n_k}\| \leq \frac{1}{2^k}$ but $|f(X_{n_k})| > \varepsilon$. Put $Y_k = \sum_{i=k}^{\infty} |X_{n_i}|$, then $Y_k \in \mathcal{X}$ and $|X_{n_k}| \leq Y_k \downarrow 0$. Thus $X_{n_k} \xrightarrow{o} 0$. It follows $f(X_{n_k}) \rightarrow 0$, a contradiction. Therefore $f \in \mathcal{X}^*$. \square

Theorem 2.24. Let \mathcal{X} be a r.i. space, then the following are equivalent:

- (1) $\mathcal{X}_n^\sim = \mathcal{X}^*$.
- (2) \mathcal{X} has order continuous norm.
- (3) if $X_n \xrightarrow{o} X$ in \mathcal{X} , then $\|X_n - X\| \rightarrow 0$.

Proof. (1) \Leftrightarrow (2) see proof of theorem 2.4.2 in [23].

(2) \implies (3). Let $X_n \xrightarrow{o} X$ in \mathcal{X} , then there exists $X_0 \in \mathcal{X}$ such that $|X_n| \leq X_0$ for all n , and $X_n \xrightarrow{a.s.} X$. Then $|X_n - X| \xrightarrow{a.s.} 0$, and $|X_n - X| \leq X_0 + |X| \in \mathcal{X}$. Now, put

$$Y_n = \sup_{m \geq n} |X_m - X| \leq X_0 + |X|,$$

then $|X_n - X| \leq Y_n$ for all n , and $Y_n \downarrow 0$ in \mathcal{X} . By (2) we have

$$\|X_n - X\| \leq \|Y_n\| \rightarrow 0.$$

(3) \implies (1). By lemma 2.23, we only need to show that $\mathcal{X}^* \subseteq \mathcal{X}_n^\sim$. Let $f \in \mathcal{X}^*$, and let $X_n \xrightarrow{o} X$ in \mathcal{X} , by (3) we have $\|X_n - X\| \rightarrow 0$, this follows

$$|f(X_n) - f(X)| = |f(X_n - X)| \leq \|f\| \|X_n - X\| \rightarrow 0.$$

Thus $f \in \mathcal{X}_n^\sim$. \square

Example 2.25. (1). For L^p ($1 \leq p < \infty$) space, we have $(L^p)_n^\sim = (L^p)^*$ since L^p has order continuous norm. Indeed, let $X_n \downarrow 0$, then $|X_n|^p \downarrow 0$. By dominated convergent theorem, we have

$$\|X_n\| = \left(\int_{\Omega} |X_n|^p \mathbf{d}\mathbb{P} \right)^{\frac{1}{p}} \rightarrow 0.$$

(2). L^∞ in general does not has order continuous norm. Consider $([0, 1], \mathcal{B}_{[0,1]}, m)$, where m is a Lebesgue measure. Let $X_n = \mathbf{1}_{(0, \frac{1}{n}]}$ $\downarrow 0$, but $\|X_n\| = 1$ for all n .

Theorem 2.26 (Riesz representation). *Let $1 \leq p \leq \infty$ and $\frac{1}{p} + \frac{1}{q} = 1$. Let $\phi \in (L^p)_n^\sim$, then there exists a unique function $Y \in L^q$ such that*

$$\phi(X) = \int_{\Omega} XY \mathbf{d}\mathbb{P}, \quad \text{for all } X \in L^p,$$

Proof. By knowing the fact that any $\phi \in (L^p)_n^\sim$ can be expressed by $\phi = \phi_1 - \phi_2$, where ϕ_1, ϕ_2 are positive linear order continuous functions. W.l.o.g. we assume $\phi \geq 0$.

Step 1. For any $A \in \mathcal{F}$, $\mathbf{1}_A \in L^p$. Put

$$\mu(\mathbf{1}_A) = \phi(\mathbf{1}_A) \geq 0.$$

Claim that μ is a measure and $\mu \ll \mathbb{P}$. Indeed,

1. if $\mathbb{P}(A) = 0$, then $\mathbf{1}_A = 0$ a.s. in L^p . Thus $\mu(A) = \phi(0) = 0$. In particular, $\mu(\emptyset) = 0$.
2. $\mu(\Omega) = \phi(\mathbf{1}_\Omega) < \infty$.
3. Let $(A_n) \subseteq \mathcal{F}$ be disjointed. Then we have

$$\mathbf{1}_{\bigcup_{k=1}^n A_k}(\omega) \rightarrow \mathbf{1}_{\bigcup_{k=1}^\infty A_k}(\omega), \quad \text{and } \mathbf{1}_{\bigcup_{k=1}^n A_k}(\omega) \leq \mathbf{1}.$$

Thus, $\mathbf{1}_{\bigcup_{k=1}^n A_k} \xrightarrow{o} \mathbf{1}_{\bigcup_{k=1}^\infty A_k}$ in L^p . It follows $\phi(\mathbf{1}_{\bigcup_{k=1}^n A_k}) \rightarrow \phi(\mathbf{1}_{\bigcup_{k=1}^\infty A_k})$. By linearity of ϕ we have

$$\phi(\mathbf{1}_{\bigcup_{k=1}^n A_k}) = \phi(\mathbf{1}_{A_1}) + \cdots + \phi(\mathbf{1}_{A_n}) = \sum_{k=1}^n \mu(A_k).$$

Then $\sum_{k=1}^\infty \mu(A_k) = \mu(\bigcup_{k=1}^\infty A_k)$. Thus, μ is a finite measure and $\mu \ll \mathbb{P}$.

By Theorem 2.19, there exists $Y \geq 0$ and measurable such that $\phi(\mathbf{1}_A) = \mu(A) = \int_A Y \mathbf{dP}$.

Step 2. Let $X \in L^p$ and $X \geq 0$, then there exists simple function sequence $(X_n) \subseteq L^p$ such that $0 \leq X_n(\omega) \uparrow X(\omega)$ for all $\omega \in \Omega$. Thus $X_n \xrightarrow{o} X$ and $\phi(X_n) \rightarrow \phi(X)$. On the other hand, by step 1 and monotone convergent theorem, we have $\phi(X_n) = \int_\Omega X_n Y \mathbf{dP} \rightarrow \int_\Omega XY \mathbf{dP}$. Thus $\phi(X) = \int_\Omega XY \mathbf{dP}$.

Step 3. For general $X \in L^p$, we have $X = X^+ - X^-$, then

$$\int_\Omega X^+ Y \mathbf{dP} = \phi(X^+) < \infty, \text{ and } \int_\Omega X^- Y \mathbf{dP} = \phi(X^-) < \infty.$$

Thus $\int_\Omega XY \mathbf{dP} < \infty$. So, we have $\phi(X) = \phi(X^+) - \phi(X^-) = \int_\Omega XY \mathbf{dP} < \infty$.

Step 4. Claim that such Y is unique. Indeed, suppose Y is not unique. Let $Y' \in L^0$ be also such that $\int_\Omega |XY| \mathbf{dP} < \infty$ and $\phi(X) = \int_\Omega XY \mathbf{dP}$ for all $X \in L^p$. So,

$$\int_\Omega |X(Y - Y')| \mathbf{dP} < \infty \text{ and } \int_\Omega X(Y - Y') \mathbf{dP} = 0, \text{ for all } X \in L^p.$$

In particular, by taking $X = \mathbf{1}$, we have $Y - Y' \in L^1$. Moreover, we have

$$\int_A (Y - Y') \mathbf{dP} = 0, \text{ for all } A \in \mathcal{F}.$$

Thus $Y = Y'$ a.s., a contradiction.

Step 5. Claim that $Y \in L^p$. Indeed,

$$\text{Case (1), for } p = \infty, \text{ take } X = \text{Sgn}(Y) := \begin{cases} 1, & \text{if } Y(\omega) > 0 \\ -1, & \text{if } Y(\omega) < 0 \\ 0, & \text{if } Y(\omega) = 0. \end{cases}$$

Then $\|X\|_\infty \leq 1$, and

$$\|Y\|_1 = \int_\Omega |Y| \mathbf{dP} = \int_\Omega \text{Sgn}(Y) Y \mathbf{dP} = \int_\Omega XY \mathbf{dP} = \phi(X) \leq M \|X\|_\infty < \infty,$$

for some $M \in \mathbb{R}$, thus $Y \in L^1$.

Case (2), for $1 < p < \infty$, take

$$X_n = \frac{|Y|^{q-1} \mathbf{1}_{\{|Y| \leq n\}} \text{Sgn}(Y)}{\| |Y| \mathbf{1}_{\{|Y| \leq n\}} \|_q^{\frac{q}{p}}}.$$

Then $\|X_n\|_p = 1$. and

$$\int_{\Omega} X_n Y \mathbf{d}\mathbb{P} = \int_{\Omega} \frac{|Y|^q \mathbf{1}_{\{|Y| \leq n\}}}{\left\| |Y| \mathbf{1}_{\{|Y| \leq n\}} \right\|_q^{\frac{q}{p}}} \mathbf{d}\mathbb{P} = \frac{\left\| |Y| \mathbf{1}_{\{|Y| \leq n\}} \right\|_q^q}{\left\| |Y| \mathbf{1}_{\{|Y| \leq n\}} \right\|_q^{\frac{q}{p}}} = \left\| |Y| \mathbf{1}_{\{|Y| \leq n\}} \right\|_q^{q - \frac{q}{p}} = \left\| |Y| \mathbf{1}_{\{|Y| \leq n\}} \right\|_q.$$

Since $\int_{\Omega} X_n Y \mathbf{d}\mathbb{P} = \phi(X_n) \leq M \|X_n\|_p \leq M$ for some $M \in \mathbb{R}$, then $\left\| |Y| \mathbf{1}_{\{|Y| \leq n\}} \right\|_q \leq M$ for all $n \in \mathbb{N}$. Now letting $n \rightarrow \infty$, we have $\|Y\|_q \leq M$, thus $Y \in L^q$.

Case (3), for $p = 1$, we want $\|Y\|_{\infty} \leq M$ for some $0 \leq M \in \mathbb{R}$, in other words, we want $\mathbb{P}(|Y| > M) = 0$. Now, suppose $\mathbb{P}(|Y| > M) > 0$, so there exists $k \in \mathbb{N}$ such that $\mathbb{P}(|Y| > M + \frac{1}{k}) > 0$. Let

$$X = \frac{\mathbf{1}_{\{|Y| > M + \frac{1}{k}\}} \text{Sgn}(Y)}{\mathbb{P}(|Y| > M + \frac{1}{k})}.$$

Then $\|X\|_1 = 1$, and

$$\int_{\Omega} XY \mathbf{d}\mathbb{P} = \int_{\Omega} \frac{\mathbf{1}_{\{|Y| > M + \frac{1}{k}\}} |Y|}{\mathbb{P}(|Y| > M + \frac{1}{k})} \mathbf{d}\mathbb{P} \geq \frac{\int_{\Omega} \mathbf{1}_{\{|Y| > M + \frac{1}{k}\}} (M + \frac{1}{k}) \mathbf{d}\mathbb{P}}{\mathbb{P}(|Y| > M + \frac{1}{k})} = M + \frac{1}{k}.$$

But $\int_{\Omega} XY \mathbf{d}\mathbb{P} = \phi(X) \leq M \|X\|_1 = M$, this is a contradiction. So, $Y \in L^{\infty}$. The proof is completed. \square

Let C be a subset of \mathcal{X} , we say C is closed if its complement C^c is open in \mathcal{X} . Equivalently, C is closed if and only if it contains all of its limit points.

Remark 2.27. C is *norm closed* if the limit of every norm convergent sequence is also in it. C is $\sigma(\mathcal{X}, \mathcal{Y})$ *closed* if the limit of every weak convergent sequence is also in it. C is *order closed* if the limits of every order convergent sequences is also in it.

Proposition 2.28. *If a subset C of \mathcal{X} is convex, then the following are equivalent*

(1) C is norm closed.

(2) C is $\sigma(\mathcal{X}, \mathcal{X}^*)$ closed.

Proof. (1) \implies (2). Suppose C is not $\sigma(\mathcal{X}, \mathcal{X}^*)$ closed. Then there exists $X \in \overline{C}^{\sigma(\mathcal{X}, \mathcal{X}^*)} \setminus C$. Since C is convex and norm closed, by theorem 2.12 there exists a hyperplane strictly separate X and C . So we can find some function $f \in \mathcal{X}^*$ such that

$$f(X) + \varepsilon < f(Y) \text{ for all } Y \in C.$$

Since $X \in \overline{C}^{\sigma(\mathcal{X}, \mathcal{X}^*)}$, then there exist a net $(X_a) \subseteq C$ such that $X_a \xrightarrow{\sigma(\mathcal{X}, \mathcal{X}^*)} X$. Thus $f(X_a) \rightarrow f(X)$, this is a contradiction since $\lim_a f(X_a) > f(X)$. Therefore C is $\sigma(\mathcal{X}, \mathcal{X}^*)$ closed.

(2) \implies (1). Let $(X_n) \subseteq C$ such that $X_n \xrightarrow{\|\cdot\|} X$, then for every $f \in \mathcal{X}^*$ we have $f(X_n) \rightarrow f(X)$, thus by definition of weak convergent we have $X_n \xrightarrow{\sigma(\mathcal{X}, \mathcal{X}^*)} X$. Since we are assuming that C is $\sigma(\mathcal{X}, \mathcal{X}^*)$ closed, it follows that $X \in C$. \square

Lemma 2.29. *For every convex risk measure $\rho : \mathcal{X} \rightarrow (-\infty, \infty]$, the following are equivalent:*

(1) ρ has the Fatou property.

(2) $\{\rho \leq \lambda\}$ is order closed for any $\lambda \in \mathbb{R}$.

Proof. (1) \implies (2) Take $(X_n) \subseteq \{\rho \leq \lambda\}$ such that $X_n \xrightarrow{o} X$. By (1), we have $\rho(X) \leq \liminf_n \rho(X_n) \leq \lambda$, thus $X \in \{\rho \leq \lambda\}$.

(2) \implies (1). Take $(X_n) \subseteq \mathcal{X}$ such that $X_n \xrightarrow{o} X$. If $\liminf_n \rho(X_n) = \infty$, the statement is clear, otherwise put $\lambda = \liminf_n \rho(X_n)$, by passing to a subsequence we assume that $\rho(X_n) \rightarrow \lambda$, then for every $\varepsilon > 0$, there exists $N \in \mathbb{N}$ such that $|\rho(X_n) - \lambda| \leq \varepsilon$ for all $n \geq N$, thus $\rho(X_n) \leq \lambda + \varepsilon$. Since $X_n \xrightarrow{o} X$, we have $\rho(X) \leq \lambda + \varepsilon$. Therefore $\rho(X) \leq \liminf_n \rho(X_n)$ since ε is arbitrary small. \square

An important topic in the theory of risk measures is to determine when a risk measure on \mathcal{X} admits a representation with respect to some duality involving \mathcal{X} . The first major result in this direction was obtained by Delbaen ([12]), who used the space L^∞ of all bounded random variables as model space and considered the dual (L^∞, L^1) .

Theorem 2.30 (Delbaen). *For every convex risk measure $\rho : L^\infty \rightarrow (-\infty, \infty]$, the following are equivalent:*

(1) ρ is $\sigma(L^\infty, L^1)$ -l.s.c.

(2) ρ has the Fatou property.

In this case, ρ admits representation via dual terms as follows:

$$\rho(X) = \sup_{Y \in L^1} \{\mathbb{E}[XY] - \rho^*(Y)\}, \quad \text{for all } X \in L^\infty.$$

If in addition ρ is a coherent risk measure, then ρ admits a more simplified representation as

$$\rho(X) = \sup_{Y \in \mathcal{Q}} \mathbb{E}[-XY] \quad \text{for all } X \in L^\infty,$$

Where \mathcal{Q} is a set of nonnegative random variables with expectation 1.

In above theorem, the set \mathcal{Q} can be interpreted as a set of probability distributions (scenarios) and the risk measure of X is obtained as the worst expected loss over the set of scenarios (stress tests). In general, such dual representations play an important role in optimization problems and portfolio selection. The representation in theorem 2.30 is connected with condition (1) $\sigma(L^\infty, L^1)$ -l.s.c. of ρ via the Fenchel-Moreau duality theorem in convex analysis. Here, $\sigma(L^\infty, L^1)$ -l.s.c. of ρ refers to the property that the sets

$$\{\rho \leq \lambda\} = \{X \in L^\infty \mid \rho(X) \leq \lambda\}$$

are $\sigma(L^\infty, L^1)$ closed for any $\lambda \in \mathbb{R}$. On the other hand, condition (2) in Theorem 2.30 is equivalent to the fact that the sets $\{\rho \leq \lambda\}$ are closed with respect to dominated convergence of sequences.

As consequence of the Fenchel-Moreau theorem one obtains the following representation result of convex risk measures on L^p spaces (see Kaina and Rüschendorf ([21])).

Theorem 2.31 (Representation of convex risk measures on L^p). *Let $\rho : L^p \rightarrow (-\infty, \infty]$, $1 \leq p < \infty$, be a proper convex risk measure. Then the following are equivalent:*

- (1) ρ is $\sigma(L^p, L^q)$ -l.s.c..
- (2) ρ is $\|\cdot\|$ -l.s.c..
- (3) $\rho(X) = \sup_{Y \in L^q} \{\mathbb{E}[XY] - \rho^*(Y)\}$, for all $X \in L^p$.
- (4) ρ has Fatou property, i.e. $\rho(X) \leq \liminf_n \rho(X_n)$ whenever $X_n \xrightarrow{o} X$.

Proof. (1) \iff (2). By lemma 2.7 it is equivalent to show that $\{\rho \leq \lambda\}$ is $\|\cdot\|$ closed $\iff \{\rho \leq \lambda\}$ is $\sigma(L^p, L^q)$ closed for every $\lambda \in R$, and this immediately follows by proposition 2.28.

(1) \iff (3). This is a consequence of Fenchel-Moreau theorem.

(2) \implies (4). Take $(X_n) \subseteq L^p$ such that $X_n \xrightarrow{o} X$, since L^p ($1 \leq p < \infty$) has order continuous norm, by theorem 2.24 we have $X_n \xrightarrow{\|\cdot\|} X$. Then by (2) we have $\rho(X) \leq \liminf_n \rho(X_n)$.

(4) \implies (2). Let $(X_n) \subseteq \{\rho \leq \lambda\}$ such that $X_n \xrightarrow{\|\cdot\|} X$, then there exists a subsequence X_{n_k} such that $\|X_{n_k} - X\| \leq \frac{1}{2^k}$ for all $k \in \mathbb{N}$. Put $Y_k = \sum_{i=k}^{\infty} |X_{n_i} - X|$, then $Y_k \in L^p$ and $|X_{n_k} - X| \leq Y_k \downarrow 0$. Thus $X_{n_k} \xrightarrow{o} X$. By (4), we have $X \in \{\rho \leq \lambda\}$. \square

A coherent risk measure on L^p have a more simplified representation as in the following theorem.

Theorem 2.32 (Representation of coherent risk measures on L^p). *Let $\rho : L^p \rightarrow (-\infty, \infty]$, $1 \leq p < \infty$, be a proper coherent risk measure and lower semicontinuous (w.r.t. $\|\cdot\|_p$), Then ρ admits the following representation:*

$$\rho(X) = \sup_{Y \in \mathcal{Q}_p} \mathbb{E}[-XY], \quad X \in L^p.$$

where $\mathcal{Q}_p := \{Y \in L^q_+ \mid \rho^*(-Y) = 0 \text{ and } \mathbb{E}[Y] = 1\}$.

Proof. Step 1, by theorem 2.31 we have $\rho(X) = \sup_{Y \in L^q} \{\mathbb{E}[XY] - \rho^*(Y)\}$, $X \in L^p$. Choose $Z \in L^p$ such that $\rho(Z) < \infty$. From positive homogeneity of ρ , then for any $\lambda > 0$ we have

$$\begin{aligned} \rho^*(Y) &= \sup_{X \in L^p} \{\mathbb{E}[XY] - \rho(X)\} \\ &= \sup_{\lambda X \in L^p} \{\mathbb{E}[\lambda XY] - \rho(\lambda X)\} \\ &= \sup_{X \in L^p} \{\lambda \mathbb{E}[XY] - \lambda \rho(X)\} \\ &= \lambda \rho^*(Y) \end{aligned}$$

Since λ is an arbitrary positive number, we must have $\rho^*(Y) = 0$, or ∞ .

Step 2, suppose $W \in L^q$ and $\rho^*(W) = 0$. Let $A = \{\omega \mid W(\omega) > 0\}$ and set $X = \mathbf{1}_A$, then $0 \leq X \in L^p$ and $\mathbb{E}[XW] \geq 0$. For any $\lambda > 0$, we have $\lambda X + Z \geq Z$, then by monotonicity of ρ we have

$$\begin{aligned} \lambda \mathbb{E}[XW] + \mathbb{E}[ZW] &= \mathbb{E}[(\lambda X + Z)W] \\ &\leq \rho(\lambda X + Z) + \rho^*(W) \\ &\leq \rho(Z) + \rho^*(W) < \infty \end{aligned}$$

Since λ is arbitrary positive number, we have $\mathbb{E}[XW] = \mathbb{E}[W\mathbf{1}_A] = 0$. Thus $W \leq 0$ a.s.. Define

$$\mathcal{Q} := \{Y \in L^q \mid \rho^*(-Y) = 0\}.$$

By the preceding argument we have that $Y \geq 0$ a.s.. In fact $\rho(-Y) = \infty$ for any $Y \in L^q \setminus \mathcal{Q}$. Thus we have $\rho(X) = \sup_{Y \in \mathcal{Q}} \mathbb{E}[-XY]$.

Step 3, for any $m \in \mathbb{R}$, by cash invariance of ρ we have

$$\rho(Z) - m = \rho(Z + m\mathbf{1}) = \sup_{Y \in \mathcal{Q}} \{-\mathbb{E}[ZY] - m\mathbb{E}[Y]\},$$

hence, for any $Y \in \mathcal{Q}$ and any $m \in \mathbb{R}$ we have

$$\rho(Z) \geq -\mathbb{E}[ZY] + m(1 - \mathbb{E}[Y]).$$

Therefore, $\mathbb{E}[Y] = 1$. This complete the proof. \square

Example 2.33 (Representation of the Expected Shortfall on L^1). For $\alpha \in (0, 1]$, the Expected Shortfall risk measure $ES_\alpha : L^1 \rightarrow \mathbb{R}$ has the following representation

$$ES_\alpha = \sup_{Q \in \mathcal{Z}_\alpha} \mathbb{E}_Q[-X]$$

where $\mathcal{Z}_\alpha := \{Q \mid Q \ll \mathbb{P}, \text{ and } \frac{dQ}{d\mathbb{P}} \leq \frac{1}{\alpha}\}$ (see proof of theorem 4.39 in [15] pg.184).

We can see from above that $ES_\alpha(X)$ is equal to the worst case of the expected losses of the position X over a given set of probability models.

Chapter 3

Inf-convolution of risk measures on L^1

In general, the risk sharing problem refers to a problem that allocating a risky position in an optimal way among several agents. It can be interpreted as a problem that not only to minimize the total risk but also to determine the optimal allocation. In this chapter, we study the inf-convolution of risk measure on L^1 space.

In mathematical language, the risk sharing problem can be described as following:

- Let \mathcal{X} be the set of possible discounted financial net worth at a fixed future date.
- n is the number of agents.
- ρ_i is the risk measure of agent i .
- The set of allocation of a position $X \in \mathcal{X}$ is defined as follows

$$\mathbb{A}(X) = \left\{ (X_1, \dots, X_n) \mid X_i \in \mathcal{X} \text{ and } \sum_{i=1}^n X_i = X \right\}.$$

- The risk sharing problem is to solve the optimization problem

$$\arg \min_{(X_1, \dots, X_n) \in \mathbb{A}(X)} \rho_1(X_1) + \dots + \rho_n(X_n). \quad (*)$$

where *arg min* stands for arguments of the minimum, and it is a collection of the points of the domain of some function at which the function values are minimized.

Definition 3.1. An allocation $(X_i) \in \mathbb{A}(X)$ is called *Pareto optimal allocation (POA)* if for any other allocation $(Y_i) \in \mathbb{A}(X)$ such that

$$\rho_i(Y_i) \leq \rho_i(X_i), \text{ it holds that } \rho_i(Y_i) = \rho_i(X_i), \text{ for } 1 \leq i \leq n.$$

We say an allocation $(X_i) \in \mathbb{A}(X)$ is *comonotone* if there exist increasing functions $f_1, \dots, f_n : \mathbb{R} \rightarrow \mathbb{R}$ such that $\sum_{i=1}^n f_i = \mathbf{Id}_{\mathbb{R}}$ and $X_i = f_i(X)$ for all i . These functions f_i are obviously 1-Lipschitz-continuous.

It is generally not easy to find the optimal allocation explicitly. The inf-convolutions of risk measure introduced in Barrieu and El Karoui (2005) has provided a new mathematical approach to the risk sharing problems.

Definition 3.2 (Inf-convolution). Given the functionals $\rho_i : \mathcal{X} \rightarrow (-\infty, \infty]$, $i = 1, \dots, n$, their inf-convolution is defined by

$$\square_{i=1}^n \rho_i(X) = \inf \left\{ \sum_{i=1}^n \rho_i(X_i) : X_i \in \mathcal{X}, i = 1, \dots, n, \text{ and } \sum_{i=1}^n X_i = X \right\}, X \in \mathcal{X}.$$

The inf-convolution is said to be *exact* at X if there exists an allocation (X_1, \dots, X_n) of X such that $\square_{i=1}^n \rho_i(X) = \sum_{i=1}^n \rho_i(X_i)$. Such a minimizing allocation is called an *optimal* allocation of X . The inf-convolution is said to be *exact* if it is exact at every $X \in \mathcal{X}$. Hence, the following proposition gives the connection of $(*)$ and definition of Pareto optimal allocation in risk sharing problem.

Proposition 3.3. Consider the risk allocation problem with finite convex risk measures $\rho_i : L^1 \rightarrow \mathbb{R}$, $i = 1, \dots, n$, and assume that $\bigcap_{i=1}^n D(\rho_i^*) \neq \emptyset$. Then for an allocation $(Z_i) \in \mathbb{A}(X)$ the following are equivalent:

- (1) (Z_i) is a Pareto optimal allocation.
- (2) $\square_{i=1}^n \rho_i(X) = \sum_{i=1}^n \rho_i(Z_i)$.
- (3) There exists $Y \in L^\infty$ such that $\rho_i(Z_i) = \mathbb{E}[Z_i Y] - \rho_i^*(Y)$, $1 \leq i \leq n$.

Proof. (2) \implies (1) is obvious. Indeed, by definition of $\square_{i=1}^n \rho$ we have $\square_{i=1}^n \rho(X) = \sum_{i=1}^n \rho_i(Z_i) \leq \sum_{i=1}^n \rho_i(W_i)$ for any other allocation $(W_i) \in \mathbb{A}(X)$. If $\rho_i(W_i) \leq \rho_i(Z_i)$ for all $1 \leq i \leq n$, we must have $\rho_i(W_i) = \rho_i(Z_i)$. Thus (1) holds.

(1) \implies (2). Suppose that (Z_i) is a Pareto optimal allocation in $\mathbb{A}(X)$. Let $m_i \geq 0$ for all $1 \leq i \leq n$ such that $\sum_{i=1}^n m_i = m \in \mathbb{R}$. Then $Z_i + m_i \geq Z_i$ for all i , by monotonicity of ρ_i we have $\rho_i(Z_i + m_i) \leq \rho_i(Z_i)$ for all i . But (Z_i) is a Pareto optimal allocation then $\rho_i(Z_i + m_i) = \rho_i(Z_i)$ for all i , thus

$$\sum_{i=1}^n \rho_i(Z_i) \geq \square_{i=1}^n \rho_i(X) \geq \sum_{i=1}^n \rho_i(Z_i) - m = \sum_{i=1}^n \rho_i(Z_i + m_i) = \sum_{i=1}^n \rho_i(Z_i).$$

(2) \implies (3). Suppose (2) holds, this implies that $\square_{i=1}^n \rho_i(X)$ is finite because each ρ_i is finite, by corollary 2.18 $\square_{i=1}^n \rho_i$ is norm continuous on L^1 . Thus by theorem 2.17, there exists $Y \in L^\infty$ such that

$$\square_{i=1}^n \rho_i(X) = \mathbb{E}[XY] - (\square_{i=1}^n \rho_i)^*(Y).$$

Since $(\square_{i=1}^n \rho_i)^*(Y) = \sum_{i=1}^n \rho_i^*(Y)$ (see lemma 3.8), we have

$$\sum_{i=1}^n \rho_i(Z_i) = \square_{i=1}^n \rho_i(X) = \mathbb{E}[XY] - \sum_{i=1}^n \rho_i^*(Y) = \sum_{i=1}^n (\mathbb{E}[Z_i Y] - \rho_i^*(Y)).$$

Since for all $1 \leq i \leq n$, $\rho_i(Z_i) \geq \mathbb{E}[Z_i Y] - \rho_i^*(Y)$, then we must have $\rho_i(Z_i) = \mathbb{E}[Z_i Y] - \rho_i^*(Y)$ for all i .

(3) \implies (2). Suppose (3) holds. Since ρ_i is finite, then $\square_{i=1}^n \rho_i$ is finite and thus is norm continuous on L^1 . By theorem 2.31 and definition of $\square_{i=1}^n \rho_i$, we have

$$\square_{i=1}^n \rho_i(X) \geq \mathbb{E}[XY] - (\square_{i=1}^n \rho_i)^*(Y) = \sum_{i=1}^n (\mathbb{E}[Z_i Y] - \rho_i^*(Y)) = \sum_{i=1}^n \rho_i(Z_i) \geq \square_{i=1}^n \rho_i(X),$$

thus, $\square_{i=1}^n \rho_i(X) = \sum_{i=1}^n \rho_i(Z_i)$. □

In optimization of capital and risk allocation problem, the existence of an optimal allocation is always to be interested. The following theorem in Filipović and Svindland ([14]) states that the capital and risk allocation problem always admits a solution via contracts whose payoffs are defined as (increasing Lipschitz-continuous) functions of the aggregate risk X . We note that this extremely useful fact is often assumed in economic contract theory.

Theorem 3.4 (Filipović and Svindland). *Let $\rho_i : L^1 \rightarrow (-\infty, \infty]$, $i = 1, \dots, n$, be l.s.c. convex, cash-additive, law-invariant functionals. Then $\square_{i=1}^n \rho_i : \mathcal{X} \rightarrow (-\infty, \infty]$ is l.s.c. convex, cash-additive, law-invariant. Moreover, for each $X \in L^1$ there exist increasing functions $f_i : \mathbb{R} \rightarrow \mathbb{R}$, $i = 1, \dots, n$, such that $\sum_{i=1}^n f_i(x) = x$ for every $x \in \mathbb{R}$ and*

$$\square_{i=1}^n \rho_i(X) = \sum_{i=1}^n \rho_i(f_i(X)).$$

In other words, $\square_{i=1}^n \rho_i : L^1 \rightarrow (-\infty, \infty]$, $i = 1, \dots, n$ is exact, and amongst the optimal allocations of any $X \in L^1$, there always is a comonotone one.

The proof of Theorem 3.4 relies on the following Lemmas and proposition.

Lemma 3.5. *For all $X \in L^1$, and $Z \in L^\infty$ we have that*

$$\int_0^1 q_X(s)q_Z(s)ds = \sup_{\hat{X} \sim X} \mathbb{E}[\hat{X}Z] = \sup_{\hat{Z} \sim Z} \mathbb{E}[X\hat{Z}], \quad (3.1)$$

where $q_X : (0, 1) \rightarrow \mathbb{R}$, and $q_X(s) = \inf\{x \in \mathbb{R} \mid \mathbb{P}(X \leq x) \geq s\}$ (see lemma 4.55 in [15]).

We denote the convex order by \succeq on L^1 , that is, $X \succeq Y \iff \mathbb{E}[u(X)] \geq \mathbb{E}[u(Y)]$, for every convex function $u : \mathbb{R} \rightarrow \mathbb{R}$.

Lemma 3.6. *Let $X, Y \in L^1$, then*

$$X \succeq Y \iff \int_0^1 q_X(s)g(s)ds \geq \int_0^1 q_Y(s)g(s)ds \quad (3.2)$$

for all increasing $g : (0, 1) \rightarrow \mathbb{R}$ such that both integrals exist (see lemma 2.2 in [11]).

We say ρ is *law-invariant* if $\rho(X) = \rho(Y)$ whenever $X, Y \in \mathcal{X}$ have the same distribution (written as $X \sim Y$). A subset $C \subset \mathcal{X}$ is said to be *law-invariant* if it contains all the random variables that have the same distribution as some element in it.

Lemma 3.7. *Let $\rho : L^1 \rightarrow (-\infty, \infty]$ be a proper l.s.c. convex function; then the following are equivalent:*

(1) ρ is law-invariant.

(2) ρ^* is law-invariant, where $\rho^*(Z) := \sup_{X \in L^1} \{\mathbb{E}[XZ] - \rho(X)\}$ for $Z \in L^\infty$.

Proof. (1) \implies (2) By definition of ρ^* and (3.1) we have that

$$\begin{aligned}\rho^*(Z) &= \sup_{X \in L^1} \{\mathbb{E}[XZ] - \rho(X)\} \\ &= \sup_{X \in L^1} \left\{ \sup_{\hat{X} \sim X} (\mathbb{E}[\hat{X}Z] - \rho(\hat{X})) \right\} \\ &= \sup_{X \in L^1} \left\{ \int_0^1 q_X(s)q_Z(s)ds - \rho(X) \right\}\end{aligned}$$

since the last expression depends on the distribution of Z only, so if $W \sim Z$ on L^∞ , then $\rho^*(W) = \rho^*(Z)$, thus ρ^* is law-invariant.

(2) \implies (1), since ρ is l.s.c. convex function, by Fenchel-Moreau theorem and (3.1), for $X \sim Y$ we have that

$$\begin{aligned}\rho(X) = \rho^{**}(X) &= \sup_{Z \in L^\infty} \{\mathbb{E}[XZ] - \rho^*(Z)\} \\ &= \sup_{Z \in L^\infty} \left\{ \sup_{\hat{Z} \sim Z} \mathbb{E}[X\hat{Z}] - \rho^*(\hat{Z}) \right\} \\ &= \sup_{Z \in L^\infty} \left\{ \int_0^1 q_X(s)q_Z(s)ds - \rho^*(Z) \right\} \\ &= \sup_{Z \in L^\infty} \left\{ \int_0^1 q_Y(s)q_Z(s)ds - \rho^*(Z) \right\} \\ &= \rho(Y)\end{aligned}$$

□

Lemma 3.8. For $\rho_i : L^1 \rightarrow (-\infty, \infty]$, $i = 1, \dots, n$, be proper convex function, then

$$(\square_{i=1}^n \rho_i)^*(Y) = \sum_{i=1}^n \rho_i^*(Y).$$

for all $Y \in L^\infty$. Moreover, if ρ_i is law-invariant, then $(\square_{i=1}^n \rho_i)^*$ is also law-invariant.

Proof. It is enough to prove when $n = 2$, the rest follows by induction.

by definition of convex conjugate, we have

$$\begin{aligned}
(\rho_1 \square \rho_2)^*(Y) &= \sup_{X \in L^1} \{\mathbb{E}[XY] - \rho_1 \square \rho_2(X)\} \\
&= \sup_{X \in L^1} \{\mathbb{E}[XY] - \inf_{X_1+X_2=X} \{\rho_1(X_1) + \rho_2(X_2)\}\} \\
&= \sup_{X \in L^1} \{\mathbb{E}[X_1Y] + \mathbb{E}[X_2Y] - \sup_{X_1+X_2=X} \{-\rho_1(X_1) - \rho_2(X_2)\}\} \\
&= \sup_{X_1+X_2=X \in L^1} \{\mathbb{E}[X_1Y] - \rho_1(X_1) + \mathbb{E}[X_2Y] - \rho_2(X_2)\} \\
&= \rho_1^*(Y) + \rho_2^*(Y),
\end{aligned}$$

for all $Y \in L^\infty$. If $Z \sim Y$ in L^∞ , then by Lemma 3.7 we have that

$$(\rho_1 \square \rho_2)^*(Y) = \rho_1^*(Y) + \rho_2^*(Y) = \rho_1^*(Z) + \rho_2^*(Z) = (\rho_1 \square \rho_2)^*(Z).$$

Thus $(\rho_1 \square \rho_2)^*$ is law-invariant. \square

Lemma 3.9. *Let $\rho : L^1 \rightarrow (-\infty, \infty]$ be a proper l.s.c., law-invariant convex function. If $X \succeq Y$, then $\rho(X) \geq \rho(Y)$. Moreover, ρ is dilatation monotone, i.e. $\rho(\mathbb{E}[X|\mathcal{G}]) \leq \rho(X)$ for all $X \in L^1$ and all sub- σ -algebras $\mathcal{G} \subset \mathcal{F}$*

Proof. Since ρ is l.s.c. convex function, by Fenchel-Moreau theorem and (3.2), for $X \succeq Y$ we have that

$$\begin{aligned}
\rho(X) = \rho^{**}(X) &= \sup_{Z \in L^\infty} \{\mathbb{E}[XZ] - \rho^*(Z)\} \\
&= \sup_{Z \in L^\infty} \left\{ \int_0^1 q_X(s)q_Z(s)ds - \rho^*(Z) \right\} \\
&\geq \sup_{Z \in L^\infty} \left\{ \int_0^1 q_Y(s)q_Z(s)ds - \rho^*(Z) \right\} \\
&= \rho(Y)
\end{aligned}$$

Moreover, we have $\mathbb{E}[X|\mathcal{G}] \preceq X$. Indeed, by Jensen's inequality, for any convex function u , we have $u(\mathbb{E}[X|\mathcal{G}]) \leq \mathbb{E}[u(X)|\mathcal{G}]$ for all $X \in L^1$ and all sub- σ -algebras $\mathcal{G} \subseteq \mathcal{F}$. Thus $\mathbb{E}[u(\mathbb{E}[X|\mathcal{G}])] \leq \mathbb{E}[E[u(X)|\mathcal{G}]] = \mathbb{E}[u(X)]$. By definition of convex order, we have $\mathbb{E}[X|\mathcal{G}] \preceq X$. By previous result we have that $\rho(\mathbb{E}[X|\mathcal{G}]) \leq \rho(X)$. \square

Lemma 3.10. Let $\mathcal{A} := \{(f, g) \mid f, g : \mathbb{R} \rightarrow \mathbb{R} \text{ are increasing, } f + g = \mathbf{Id}_{\mathbb{R}}\}$, then f, g are 1-Lipschitz-continuous functions for $(f, g) \in \mathcal{A}$.

Lemma 3.11. Let $f_n : \mathbb{R} \rightarrow \mathbb{R}$, $n \in \mathbb{N}$, be a sequence of increasing 1-Lipschitz-continuous functions such that $f_n(0) \in [-K, K]$ for all $n \in \mathbb{N}$, where $K \geq 0$ is a constant. Then there is a subsequence $(f_{n_k})_{k \in \mathbb{N}}$ of $(f_n)_{n \in \mathbb{N}}$ and an increasing 1-Lipschitz-continuous functions $f : \mathbb{R} \rightarrow \mathbb{R}$ such that $\lim_{k \rightarrow \infty} f_{n_k}(x) = f(x)$ for all $x \in \mathbb{R}$.

Proof. Since f_n is increasing 1-Lipschitz-continuous function, and $f_n(0) \in [-K, K]$ for all $n \in \mathbb{N}$, then we have that

$$|f_n(x)| \leq |f_n(x) - f_n(0)| + |f_n(0)| \leq |x| + |f_n(0)|.$$

this guarantees that $f_n(x) \in [-K, K + x]$ if $x \geq 0$ and $f_n(x) \in [-K + x, K]$ if $x \leq 0$. Hence, by a procedure familiar from the standard proof of the Arzelà-Ascoli theorem, we are able to extract a subsequence $(f_{n_k})_{k \in \mathbb{N}}$ of $(f_n)_{n \in \mathbb{N}}$ such that $\lim_{k \rightarrow \infty} f_{n_k}(q)$ exists for all $q \in \mathbb{Q}$. In fact, we can show that sequences $(f_{n_k}(x))_{k \in \mathbb{N}}$ must converge for all $x \in \mathbb{R}$. Indeed, for $\varepsilon > 0$, and choose $q \in \mathbb{Q}$ and $N_0 \in \mathbb{N}$ such that $|q - x| < \varepsilon/3$ and $|f_{n_k}(q) - f_{n_l}(q)| \leq \varepsilon/3$ for all $k, l \geq N_0$. Then we have

$$\begin{aligned} |f_{n_k}(x) - f_{n_l}(x)| &\leq |f_{n_k}(x) - f_{n_k}(q)| + |f_{n_k}(q) - f_{n_l}(q)| + |f_{n_l}(q) - f_{n_l}(x)| \\ &\leq 2|x - q| + |f_{n_k}(q) - f_{n_l}(q)| \\ &< \varepsilon, \end{aligned}$$

then $f_{n_k}(x)$ is Cauchy sequence in \mathbb{R} , thus it must converge to some $f(x) := \lim_{k \rightarrow \infty} f_{n_k}(x)$.

It is easy to verify that f is an 1-Lipschitz-continuous increasing function due to the fact

$$|f(x) - f(y)| = \lim_{k \rightarrow \infty} |f_{n_k}(x) - f_{n_k}(y)| \leq |x - y|$$

□

Proposition 3.12. For any allocation (Y, Z) of $X \in L^1$, there is $(f, g) \in \mathcal{A}$ such that $f(X) \preceq Y$ and $g(X) \preceq Z$, where

$$\mathcal{A} =: \{(f, g) \mid f, g : \mathbb{R} \rightarrow \mathbb{R} \text{ are increasing, } f + g = \mathbf{Id}_{\mathbb{R}}\}.$$

Proof. Step 1; By Jensen's inequality and lemma 3.9, we have that $(\mathbb{E}[Y|X], \mathbb{E}[Z|X])$ is an allocation of X which is at least as good as (Y, Z) , that is, $\mathbb{E}[Y|X] \preceq Y$ and $\mathbb{E}[Z|X] \preceq Z$. Let $h_1, h_2 : \mathbb{R} \rightarrow \mathbb{R}$ be measurable functions such that $h_1(X) = \mathbb{E}[Y|X]$, $h_2(X) = \mathbb{E}[Z|X]$. Clearly, we may assume that $h_1 + h_2 = \mathbf{Id}_{\mathbb{R}}$. If h_1 and h_2 are increasing, we are done; if not, we will improve this allocation as showed in step 2. During the remainder of this proof we may restrict ourselves to improve allocations (Y, Z) of type $Y = h_1(X)$ and $Z = h_2(X)$ for some measurable functions $h_1, h_2 : \mathbb{R} \rightarrow \mathbb{R}$ such that $h_1 + h_2 = \mathbf{Id}_{\mathbb{R}}$.

Step 2: Suppose that X is a simple random variable, i.e., $X = \sum_{i=1}^n x_i \mathbf{1}_{A_i}$ for a partition of A_1, \dots, A_n of Ω and real numbers x_i such that $x_i \neq x_j$ for $i \neq j$. Let $y_i := h_1(x_i)$ and $z_i := h_2(x_i)$. Then $h_1(X) = \sum_{i=1}^n y_i \mathbf{1}_{A_i}$ and $h_2(X) = \sum_{i=1}^n z_i \mathbf{1}_{A_i}$. We set $x := (x_1, \dots, x_n)$, $y := (y_1, \dots, y_n)$, $z := (z_1, \dots, z_n)$, and $p_k := \mathbb{P}(A_k)$, $k = 1, \dots, n$. Let π be a permutation of $\{1, \dots, n\}$ such that

$$x_\pi := (x_{\pi(1)}, \dots, x_{\pi(n)}) \in \mathcal{D} := \{\tilde{x} \in \mathbb{R}^n \mid \tilde{x}_1 \leq \tilde{x}_2 \leq \dots \leq \tilde{x}_n\}.$$

Observe that $(h_1(X), h_2(X))$ is comonotone if and only if $y_\pi, z_\pi \in \mathcal{D}$. For the sake of brevity, we may and do assume w.l.o.g. that $x \in \mathcal{D}$ already. Supposing that (y, z) is not comonotone, i.e., $y \notin \mathcal{D}$ or $x \notin \mathcal{D}$ or both, the following algorithm by M. Landsberger and I. Meilijson transfers (y, z) into a comonotone allocation:

Since (y, z) is not comonotone, there must exists an i such that $y_1 \leq \dots \leq y_i, z_1 \leq \dots, \leq z_i$ but either $y_{i+1} < y_i$ or $z_{i+1} < z_i$. W.l.o.g. let us assume that $z_{i+1} < z_i$. Then there is a smallest j such that $z_{i+1} < z_j$. For $k = j, \dots, i$, we set

$$y_k^{new} = y_k + \frac{p_{i+1}}{\sum_{l=j}^{i+1} p_l} (z_j - z_{i+1}) \quad \text{and} \quad z_k^{new} = z_k - \frac{p_{i+1}}{\sum_{l=j}^{i+1} p_l} (z_j - z_{i+1}).$$

whereas

$$y_{i+1}^{new} = y_{i+1} - \frac{\sum_{l=j}^i p_l}{\sum_{l=j}^{i+1} p_l} (z_j - z_{i+1}) \quad \text{and} \quad z_{i+1}^{new} = z_{i+1} + \frac{\sum_{l=j}^i p_l}{\sum_{l=j}^{i+1} p_l} (z_j - z_{i+1}).$$

The other coordinates of y and z are left unchanged. Finally, set $y := y^{new}$ and $z := z^{new}$ and repeat the procedure in case the output is not comonotone.

Let $(Y^{new}, Z^{new}) := (\sum_{i=1}^n y_i^{new} \mathbf{1}_{A_i}, \sum_{i=1}^n z_i^{new} \mathbf{1}_{A_i})$. Firstly, (Y^{new}, Z^{new}) is obviously an allocation of X ; secondly, we claim that $Y^{new} \preceq Y$ and $Z^{new} \preceq Z$, i.e., each cycle of the algorithm improves the allocation. In order to show that $Y^{new} \preceq Y$ and $Z^{new} \preceq Z$, Let $u : \mathbb{R} \rightarrow \mathbb{R}$ be any convex function. We put

$$\alpha := \frac{p_{i+1}}{\sum_{l=j}^{i+1} p_l} \in (0, 1) \quad \text{and} \quad \lambda_k = \frac{z_j - z_{i+1}}{z_k - z_{i+1}} \in (0, 1].$$

and recall that convexity is equivalent to

$$\frac{u(b) - u(a)}{b - a} \leq \frac{u(c) - u(a)}{c - a} \leq \frac{u(c) - u(b)}{c - b}, \quad \forall a < b < c.$$

We compute

$$\begin{aligned} \sum_{k=j}^{i+1} u(z_k^{new}) p_k &= \sum_{k=j}^i u(z_k^{new}) p_k + u(z_{i+1}^{new}) p_{i+1} \\ &= \sum_{k=j}^i u(z_k - \alpha(z_j - z_{i+1})) p_k + u(z_{i+1} + (1 - \alpha)(z_j - z_{i+1})) p_{i+1} \\ &= \sum_{k=j}^i u(z_k - \alpha \lambda_k (z_k - z_{i+1})) p_k + u(z_{i+1} + (1 - \alpha)(z_j - z_{i+1})) p_{i+1} \\ &= \sum_{k=j}^i u((1 - \alpha \lambda_k) z_k + \alpha \lambda_k z_{i+1}) p_k + u((1 - \alpha) z_j + \alpha z_{i+1}) p_{i+1} \\ &\leq \sum_{k=j}^i [(1 - \alpha \lambda_k) u(z_k) + \alpha \lambda_k u(z_{i+1})] p_k + [(1 - \alpha) u(z_j) + \alpha u(z_{i+1})] p_{i+1} \\ &= \sum_{k=j}^{i+1} u(z_k) p_k + (1 - \alpha)(u(z_j) - u(z_{i+1})) p_{i+1} - \alpha \sum_{k=j}^i \lambda_k (u(z_k) - u(z_{i+1})) p_k \\ &\leq \sum_{k=j}^{i+1} u(z_k) p_k + (1 - \alpha)(u(z_j) - u(z_{i+1})) p_{i+1} - \alpha \sum_{k=j}^i (u(z_j) - u(z_{i+1})) p_k \\ &= \sum_{k=j}^{i+1} u(z_k) p_k, \end{aligned}$$

the last inequality above is due to the fact that $\lambda_k(u(z_k) - u(z_{i+1})) \geq u(z_j) - u(z_{i+1})$, Thus we have $\mathbb{E}[u(z^{new})] \leq \mathbb{E}[u(z)]$, and it implies $Z^{new} \preceq Z$. A similar computation for Y^{new} shows that $Y^{new} \preceq Y$.

Step 3: Let X be any integrable random variable in L^1 . Let $(Y_n)_{n \in \mathbb{N}}$ and $(Z_n)_{n \in \mathbb{N}}$ be sequences of simple random variables such that $|Y_n| \leq |Y|$ and $|Z_n| \leq |Z|$ for all $n \in \mathbb{N}$, and $Y_n \xrightarrow{a.s.} Y$ and $Z_n \xrightarrow{a.s.} Z$ in L^1 . Then $X_n = (Y_n + Z_n) \xrightarrow{a.s.} X$ in L^1 . By step 2, for each $n \in \mathbb{N}$, there exists a comonotone improvement $(f_n(X_n), g_n(X_n))$ of (Y_n, Z_n) . choose $N \in \mathbb{N}$ such that $\|Y_n\|_1 \leq \|Y\|_1 + 1$, $\|Z_n\|_1 \leq \|Z\|_1 + 1$, and $\|X_n\|_1 \leq \|X\|_1 + 1$, for all $n \geq N$. Since all f_n (and g_n) are 1-Lipschitz-continuous, we have that $|f_n(0)| \leq |X_n| + |f_n(X_n)|$. Taking the expectations on both sides yields

$$|f_n(0)| \leq \mathbb{E}[|X_n|] + \mathbb{E}[|f_n(X_n)|] \leq \mathbb{E}[|X_n|] + \mathbb{E}[|Y_n|],$$

because $f_n(X_n) \preceq Y_n$ and function $x \mapsto |x|$ is convex. Hence, if $n \geq N$, we get $|f_n(0)| \leq \|X\|_1 + \|Y\|_1 + 2 := K_1$ and similarly $|g_n(0)| \leq \|X\|_1 + \|Z\|_1 + 2 := K_2$, and thus $f_n(0), g_n(0) \in [-K, K]$ for $K := \max\{K_1, K_2\}$. Therefore by Lemma 3.11, there is a subsequence $(f_{n_k})_{k \in \mathbb{N}}$ of $(f_n)_{n \in \mathbb{N}}$ and a 1-Lipschitz-continuous increasing function $f : \mathbb{R} \rightarrow \mathbb{R}$ such that $f(a) = \lim_{k \rightarrow \infty} f_{n_k}(a)$, $a \in \mathbb{R}$. Now it is easily verified that $(g_{n_k})_{k \in \mathbb{N}}$ converges pointwise to the 1-Lipschitz-continuous increasing function $g := \text{Id}_{\mathbb{R}} - f$. Hence, the sequence $f_{n_k}(X_{n_k}) \xrightarrow{a.s.} f(X)$, and $g_{n_k}(X_{n_k}) = (X_{n_k} - f_{n_k}(X_{n_k})) \xrightarrow{a.s.} g(X)$. Since

$$|f_{n_k}(X_{n_k})| \leq |X_{n_k}| + K \leq |Y| + |Z| + K$$

for large enough $K \in \mathbb{N}$, we can apply the dominated convergence theorem, which yields $f(X), g(X) \in L^1$ and $\|f(X) - f_{n_k}(X_{n_k})\|_1 \rightarrow 0$, $\|g(X) - g_{n_k}(X_{n_k})\|_1 \rightarrow 0$, for $k \rightarrow \infty$. Moreover, we have that

$$\mathbb{E}[|f(X)|] = \lim_{k \rightarrow \infty} \mathbb{E}[|f_{n_k}(X_{n_k})|] \leq \lim_{k \rightarrow \infty} \mathbb{E}[|Y_{n_k}|] = \mathbb{E}[|Y|],$$

and similar proof for g . Hence, $(f(X), g(X))$ is a comonotone allocation of X satisfying $f(X) \preceq Y$ and $g(X) \preceq Z$. This complete the proof. \square

Now based on the lemmas and proposition established above, we can start the proof of Theorem 3.4 as follows.

Proof. First note that it is enough to prove for $n = 2$, the rest follows by induction.

Step 1. $\rho_1 \square \rho_2$ is proper, convex, and cash-invariant.

1. $\rho_1 \square \rho_2$ is convex.

Indeed, since ρ_1, ρ_2 is convex, then for $\lambda \in [0, 1]$, and $X_1, X_2, Y_1, Y_2 \in L^1$, such that $X_1 + X_2 = X \in L^1$, and $Y_1 + Y_2 = Y \in L^1$, we have that

$$\begin{aligned} & \rho_1(\lambda X_1 + (1 - \lambda)Y_1) + \rho_2(\lambda X_2 + (1 - \lambda)Y_2) \\ & \leq \lambda \rho_1(X_1) + (1 - \lambda)\rho_1(X_2) + \lambda \rho_2(Y_1) + (1 - \lambda)\rho_2(Y_2) \\ & \leq \lambda(\rho_1(X_1) + \rho_2(X_2)) + (1 - \lambda)(\rho_1(Y_1) + \rho_2(Y_2)). \end{aligned}$$

By taking infimum over $X_1 + X_2 = X$ first and then over $Y_1 + Y_2 = Y$ we have that

$$\begin{aligned} & \rho_1 \square \rho_2(\lambda X + (1 - \lambda)Y) \\ & = \inf_{X_1 + X_2 = X, Y_1 + Y_2 = Y} \{\rho_1(\lambda X_1 + (1 - \lambda)Y_1) + \rho_2(\lambda X_2 + (1 - \lambda)Y_2)\} \\ & \leq \lambda \inf_{X_1 + X_2 = X} \{\rho_1(X_1) + \rho_2(X_2)\} + (1 - \lambda) \inf_{Y_1 + Y_2 = Y} \{\rho_1(Y_1) + \rho_2(Y_2)\} \\ & = \lambda \rho_1 \square \rho_2(X) + (1 - \lambda) \rho_1 \square \rho_2(Y). \end{aligned}$$

2. $\rho_1 \square \rho_2$ is proper and cash-invariant. Indeed, since ρ_i for $i = 1, 2$ is convex and law-invariant, by Lemma 3.9 then ρ_i is dilatation monotone, i.e, $\rho_i(\mathbb{E}[X|G]) \leq \rho_i(X)$, for $X \in L^1$. By cash-invariance, we have that $\rho_i(X) \geq \rho_i(0 + \mathbb{E}[X]) = \rho_i(0) - \mathbb{E}[X]$ for all $X \in L^1$. Hence,

$$\begin{aligned} \rho_1 \square \rho_2(X) & = \inf_{X_1 + X_2 = X} \{\rho_1(X_1) + \rho_2(X_2)\} \\ & \geq \inf_{X_1 + X_2 = X} \{-\mathbb{E}[X_1] - \mathbb{E}[X_2] + \rho_1(0) + \rho_2(0)\} \\ & = -\mathbb{E}[X] + \rho_1(0) + \rho_2(0). \end{aligned}$$

So $\rho_1 \square \rho_2$ is proper and $\rho_1 \square \rho_2(0) = \rho_1(0) + \rho_2(0) < \infty$. Furthermore, for all $m \in \mathbb{R}$, we have that

$$\rho_1 \square \rho_2(X + m) = \inf_{Y \in L^1} \{\rho_1(X_1 + m - Y) + \rho_2(Y)\} = \rho_1 \square \rho_2(X) - m$$

due to the cash-invariance of ρ_1 .

Step 2. $\rho_1 \square \rho_2(X) = \inf_{(f,g) \in \mathcal{A}} \{\rho_1(f(X)) + \rho_2(g(X))\}$, $X \in L^1$, where

$$\mathcal{A} := \{(f, g) \mid f, g : \mathbb{R} \rightarrow \mathbb{R} \text{ are increasing, } f + g = \mathbb{I}_{\mathbb{R}}\}.$$

Proof of Step 2 is immediate consequence of Proposition 3.12 and Lemma 3.9.

Step 3. $\rho_1 \square \rho_2$ is exact, and for each $X \in L^1$, there exists a comonotone optimal allocation.

Suppose that $X \in L^1$ is such that $\rho_1 \square \rho_2(X) = \infty$. Then every allocation $(f(X), g(X))$, $(f, g) \in \mathcal{A}$, is optimal. Now let $X \in D(\rho_1 \square \rho_2)$ (i.e. $\rho_1 \square \rho_2(X) < \infty$) and choose a sequence $(f_n, g_n) \in \mathcal{A}$, $n \in \mathbb{N}$, such that $\rho_1 \square \rho_2(X) = \lim_{n \rightarrow \infty} (\rho_1(f_n(X)) + \rho_2(g_n(X)))$. By the cash-invariance we may assume that $f_n(0) = g_n(0) = 0$ for all $n \in \mathbb{N}$. Hence, by Lemma 3.11 there is a subsequence $(f_{n_k})_{k \in \mathbb{N}}$ of $(f_n)_{n \in \mathbb{N}}$ and a 1-Lipschitz-continuous and increasing function $f : \mathbb{R} \rightarrow \mathbb{R}$ such that $f(a) = \lim_{k \rightarrow \infty} f_{n_k}(a)$ for all $a \in \mathbb{R}$. Clearly, the sequence $f_{n_k}(X) \xrightarrow{a.s.} f(X)$ and $g_{n_k}(X) = (X - f_{n_k}(X)) \xrightarrow{a.s.} g(X)$, where $g := \mathbf{Id}_{\mathbb{R}} - f$ is a 1-Lipschitz-continuous increasing function. Since $|f_{n_k}(X)| \leq |X|$ and $|g_{n_k}(X)| \leq |X|$ for all $k \in \mathbb{N}$, we may apply the dominated convergence theorem and get $f(X), g(X) \in L^1$ and $\|f(X) - f_{n_k}(X)\|_1 \rightarrow 0$, $\|g(X) - g_{n_k}(X)\|_1 \rightarrow 0$ for all $k \rightarrow \infty$. On the one hand, by l.s.c. of ρ_1 and ρ_2 , we have

$$\begin{aligned} \rho_1 \square \rho_2(X) &= \lim_{k \rightarrow \infty} (\rho_1(f_{n_k}(X)) + \rho_2(g_{n_k}(X))) \\ &\geq \liminf_{k \rightarrow \infty} \rho_1(f_{n_k}(X)) + \liminf_{k \rightarrow \infty} \rho_2(g_{n_k}(X)) \\ &\geq \rho_1(f(X)) + \rho_2(g(X)). \end{aligned}$$

On the other hand, by the definition of inf-convolution, we have that $\rho_1 \square \rho_2(X) \leq \rho_1(f(X)) + \rho_2(g(X))$. Thus $\rho_1 \square \rho_2(X) = \rho_1(f(X)) + \rho_2(g(X))$ and the $(f(X), g(X))$ is an optimal allocation of X .

Step 4. $\rho_1 \square \rho_2$ is l.s.c. and law-invariant.

1. $\rho_1 \square \rho_2$ is l.s.c..

To show that $\rho_1 \square \rho_2$ is l.s.c. is enough to show that $\mathcal{A}_{\rho_1 \square \rho_2}$ is norm closed, where

$$\mathcal{A}_{\rho_1 \square \rho_2} := \{X \in L^1 \mid \rho_1 \square \rho_2(X) \leq 0\}.$$

Let $(X_n) \subseteq \mathcal{A}_{\rho_1 \square \rho_2}$ such that $X_n \xrightarrow{\|\cdot\|} X$ in L^1 . We will show that $X \in \mathcal{A}_{\rho_1 \square \rho_2}$. According to Step 3, there are $(f_n(X), g_n(X)) \in \mathcal{A}$, $n \in \mathbb{N}$, such that $0 \geq \rho_1 \square \rho_2(X_n) = \rho_1(f_n(X_n)) + \rho_2(g_n(X_n))$. By the cash-invariance again, we may assume that $f_n(0) = g_n(0) = 0$ for all $n \in \mathbb{N}$. Similarly to Step 3, by applying Lemma 3.11, we find a subsequence $(f_{n_k}, g_{n_k})_{k \in \mathbb{N}}$ of $(f_n, g_n)_{n \in \mathbb{N}}$ and $(f, g) \in \mathcal{A}$ such that $f_{n_k}(X_{n_k})$ converges to $f(X)$ in L^1 and $g_{n_k}(X_{n_k}) = X_{n_k} - f_{n_k}(X_{n_k})$ converges to $g(X)$ in L^1 . By l.s.c. of ρ_1 and ρ_2 we have

$$\begin{aligned}
\rho_1 \square \rho_2(X) &\leq \rho_1(f(X)) + \rho_2(g(X)) \\
&\leq \liminf_{k \rightarrow \infty} \rho_1(f_{n_k}(X_{n_k})) + \liminf_{k \rightarrow \infty} \rho_2(g_{n_k}(X_{n_k})) \\
&\leq \liminf_{k \rightarrow \infty} \{\rho_1(f_{n_k}(X_{n_k})) + \rho_2(g_{n_k}(X_{n_k}))\} \\
&\leq 0.
\end{aligned}$$

Thus $X \in \mathcal{A}_{\rho_1 \square \rho_2}$, and therefore $\mathcal{A}_{\rho_1 \square \rho_2}$ is norm closed. By cash-invariant of $\rho_1 \square \rho_2$, we have that $\{\rho_1 \square \rho_2 \leq m\}$ is norm closed for all $m \in \mathbb{R}$. Hence $\rho_1 \square \rho_2$ is l.s.c..

2. $\rho_1 \square \rho_2$ is law-invariant follows Lemma 3.7 and Lemma 3.8. Indeed, by Lemma 3.8 we have $(\rho_1 \square \rho_2)^*$ is law-invariant since ρ_1 and ρ_2 are law-invariant, and by Lemma 3.7 we have $\rho_1 \square \rho_2$ is law-invariant.

□

Chapter 4

Inf-convolution of risk measures on rearrangement invariant spaces

In this chapter, we will extend the results of inf-convolution of risk measures in Filipović and Svindland theorem from L^1 space to more general r.i. spaces, e.g. Orlicz spaces. We also show that the strong Fatou property is preserved by inf-convolution of risk measures on r.i. spaces. This ensures that the inf-convolution of risk measures also admits a tractable dual representation.

4.1 Orlicz space

We give below some basic results of Orlicz spaces. For more detail, see Edgar and Sucheston ([13]) and Cheridito and Li ([10]).

Definition 4.1 (Orlicz function). *A function $\Phi : [0, \infty) \rightarrow [0, \infty)$ is called an Orlicz function if it satisfies:*

- (1) *it is convex, i.e. $\Phi(\lambda t_1 + (1 - \lambda)t_2) \leq \lambda\Phi(t_1) + (1 - \lambda)\Phi(t_2)$, for any $\lambda \in [0, 1]$.*
- (2) *it is increasing, i.e. if $t_1 \leq t_2$, then $\Phi(t_1) \leq \Phi(t_2)$.*
- (3) *$\Phi(0) = 0$ and $\lim_{t \rightarrow \infty} \frac{\Phi(t)}{t} = \infty$.*

Example 4.2. (Examples of Orlicz function)

(1) $\Phi(t) = \frac{t^p}{p}$ for $p > 1$

(2) $\Phi(t) = e^t - 1$, and $\Phi(t) = e^t - t - 1$

Convex conjugate of Orlicz function will appear to be linked to the duality of Orlicz spaces. We define the conjugate function Ψ of Φ as

$$\Psi(s) = \sup\{ts - \Phi(t) : t \geq 0\}, \quad s \geq 0.$$

From the restriction (3) of definition of Orlicz function, Ψ is finite-valued and is also an Orlicz function; moreover, its conjugate is Φ .

The Orlicz space of Φ is the collection of all the measurable random variables $X : \Omega \rightarrow \mathbb{R}$ such that $\mathbb{E}[\Phi(\frac{X}{a})] < \infty$ for some $a > 0$, is denoted by L^Φ , i.e.

$$L^\Phi := L^\Phi(\Omega, \mathcal{F}, \mathbb{P}) = \{X \in L^0 \mid \mathbb{E}[\Phi(\frac{X}{a})] < \infty, \text{ for some } a > 0\}.$$

For any $X \in L^\Phi$, we define the *Luxemburg norm* of X as

$$\|X\|_\Phi := \inf\{a > 0 \mid \mathbb{E}[\Phi(\frac{X}{a})] \leq 1\}.$$

Equivalently, L^Φ can also be described as

$$L^\Phi = \{X \in L^0 \mid \|X\|_\Phi < \infty\}.$$

Indeed, if $X \in L^\Phi$, then $\mathbb{E}[\Phi(\frac{X}{n})] < \infty$ for some integer n . But $\frac{|X|}{n} \rightarrow 0$ a.s., then $\Phi(\frac{X}{n}) \rightarrow 0$ a.s. (here we assume that $\Phi(t) < \infty$ for some $t > 0$). By dominated convergence theorem we have $\mathbb{E}[\Phi(\frac{X}{n})] \leq 1$ for some n , and thus $\|X\|_\Phi < \infty$. Conversely, if $\|X\|_\Phi < \infty$, it is clear that $X \in L^\Phi$.

The subspace of L^Φ consisting of all $X \in L^\Phi$ such that $\mathbb{E}[\Phi(\frac{X}{a})] < \infty$ for all $a > 0$ is called the *Orlicz heart* of L^Φ and is denoted by H^Φ . It is clearly that $H^\Phi \subseteq L^\Phi$. In fact, we have that $L^\infty \subset H^\Phi \subseteq L^\Phi \subset L^1$. The restriction on Φ eliminate the case where L^Φ coincides with L^∞ or L^1 .

Proposition 4.3. *In general, we have $L^\infty \subset H^\Phi \subseteq L^\Phi \subset L^1$*

Proof. We first show that $L^\infty \subset L^\Phi$. Let $X \in L^\infty$, then there exists $0 < c \in \mathbb{R}$ such that $|X| \leq c$ a.e. Let $E \subseteq \Omega$ such that $|X| > c$ on E^c , then

$$\int_{\Omega} \Phi\left(\frac{|X|}{a}\right) d\mathbb{P} \leq \int_E \Phi\left(\frac{|c|}{a}\right) d\mathbb{P} = \Phi\left(\frac{|c|}{a}\right) \mathbb{P}(E).$$

We can choose $a \in \mathbb{R}$ such that $\Phi\left(\frac{|c|}{a}\right) < \infty$, then we have $\int_{\Omega} \Phi\left(\frac{|X|}{a}\right) d\mathbb{P} < \infty$, thus $X \in L^\Phi$

Now we show that $L^\Phi \subset L^1$. Let $X \in L^\Phi$, then we have $\int \Phi\left(\frac{|X|}{\|X\|_\Phi}\right) d\mathbb{P} \leq 1$. Since Φ is convex, there exists $\lambda > 0$ such that $\Phi\left(\frac{|X|}{\|X\|_\Phi}\right) \geq \lambda \frac{|X|}{\|X\|_\Phi}$ (for increasing convex function Φ , there exists $\lambda > 0$ and $x_0 > 0$ such that $\Phi(x) \leq \lambda x$ for all $x \geq x_0$). Thus we have

$$\int \lambda \left(\frac{|X|}{\|X\|_\Phi}\right) d\mathbb{P} \leq \int \Phi\left(\frac{|X|}{\|X\|_\Phi}\right) d\mathbb{P} < \infty,$$

so we have $X \in L^1$, thus $L^\Phi \subset L^1$. □

Under the Luxemburg norm $\|\cdot\|$ the Orlicz space L^Φ is a Banach space. Moreover, the Orlicz Heart H^Φ is a norm closed subspace of L^Φ , that is, if sequence $(X_n) \subseteq H^\Phi$ and $X_n \xrightarrow{\|\cdot\|} X$, then $X \in H^\Phi$ (see Theorem 2.1.11 in [13]).

The following proposition shows that Orlicz space L^Φ is a rearrangement invariant function space. This means if X is a random variable that has same distribution as some $Y \in L^\Phi$, then $X \in L^\Phi$ and $\|X\| = \|Y\|$.

Proposition 4.4. *L^Φ is an r.i. space.*

Proof. Let X be a random variable that has the same distribution as $Y \in L^\Phi$, i.e. $\mathbb{P}(\{|X| > t\}) = \mathbb{P}(\{|Y| > t\})$ for all $t > 0$. It is enough to show that

$$\int \Phi(|X|) d\mathbb{P} = \int \Phi(|Y|) d\mathbb{P}.$$

Since Φ is strictly increasing (except possibly where it is 0), we have that $\mathbb{P}(\{\Phi(|X|) > t\}) = \mathbb{P}(\{|X| > \Phi^{-1}(t)\})$. Then

$$\begin{aligned}
\int \Phi(|X|)d\mathbb{P} &= \int_0^\infty \mathbb{P}(\{\Phi(|X|) > t\})dt \\
&= \int_0^\infty \mathbb{P}(\{|X| > \Phi^{-1}(t)\})dt \\
&= \int_0^\infty \mathbb{P}(\{|Y| > \Phi^{-1}(t)\})dt \\
&= \int_0^\infty \mathbb{P}(\{\Phi(|Y|) > t\})dt \\
&= \int \Phi(|Y|)d\mathbb{P}.
\end{aligned}$$

□

4.2 Dual representation of risk measures on Orlicz space

It is known since Delbaen ([12]) that a convex risk measure ρ is $\sigma(L^\infty, L^1)$ lower semicontinuous if and only if ρ has the Fatou property on L^∞ , and such ρ admits a nice dual representation via the Fenchel-Moreau duality theorem. By theorem 2.31, a convex risk measure ρ has Fatou property if and only if ρ is $\sigma(L^p, L^q)$ -lower semicontinuous for $1 \leq p \leq \infty$. It has been an open question since Biagini and Frittelli ([6]) whether these equivalent conditions for a convex risk measure continue to hold on the general Orlicz space L^Φ . The following theorem in Gao et al. ([17]) shows that a convex risk measure ρ on L^Φ has the strong Fatou property if and only if it is $\sigma(L^\Phi, H^\Psi)$ -lower semicontinuous.

Theorem 4.5 (Gao and Xanthos). *Let $\rho : L^\Phi \rightarrow (-\infty, \infty]$ be a proper convex risk measure, the following are equivalent:*

- (1) ρ is $\sigma(L^\Phi, H^\Psi)$ -lower semicontinuous
- (2) ρ admits the representation: $\rho(X) = \sup_{Y \in H^\Psi} \{\mathbb{E}[XY] - \rho^*(Y)\}$ for any $X \in L^\Phi$, where $\rho^*(Y) = \sup_{X \in L^\Phi} \{\mathbb{E}[XY] - \rho(X)\}$, for any $Y \in H^\Psi$.
- (3) ρ has strong Fatou property, i.e. $\rho(X) \leq \liminf_n \rho(X_n)$ whenever $\sup \|X_n\| < \infty$ and $X_n \xrightarrow{a.s.} X$.

The following theorem in Gao et al. ([18]) shows that under the assumption of law-invariance, a convex risk measure ρ has strong Fatou property is equivalent to the Fatou property and to $\sigma(L^\Phi, L^\Psi)$ (respectively $\sigma(L^\Phi, H^\Psi)$, $\sigma(L^\Phi, L^\infty)$)-lower semicontinuous on Orlicz space.

Theorem 4.6 (GLMX). *Let $\rho : L^\Phi \rightarrow (-\infty, \infty]$ be a convex risk measure and law-invariant. Then the following are equivalent:*

- (1) ρ has the Fatou property.
- (2) ρ has the strong Fatou property.
- (3) ρ is $\sigma(L^\Phi, L^\infty)$ -lower semicontinuous.
- (4) ρ is $\sigma(L^\Phi, L^\Psi)$ -lower semicontinuous.
- (5) ρ is $\sigma(L^\Phi, H^\Psi)$ -lower semicontinuous.

The following dual representation of functionals with the Fatou property is an immediate consequence of the above theorem and Fenchel-Moreau duality theorem.

Corollary 4.7. *Let $\rho : L^\Phi \rightarrow (-\infty, \infty]$ be a law-invariant, convex risk measure and has Fatou property. Then $\rho(X)$ admits the following representation:*

$$\rho(X) = \sup_{Y \in L^\infty} \{\mathbb{E}[XY] - \rho^*(Y)\}$$

where $\rho^*(Y) = \sup_{X \in L^\Phi} \{\mathbb{E}[XY] - \rho(X)\}$.

4.3 Inf-convolution of risk measures on r.i. spaces

The following theorem in ([9]) asserts that a law-invariant convex risk measure on an r.i. space that has the strong Fatou property can be extended uniquely to a law-invariant convex risk measure on L^1 with the Fatou property. This is an important result that allow us extend theorem 3.4 from L^1 to any r.i. spaces.

Theorem 4.8. *Let \mathcal{X} be an r.i. space, let $\rho : \mathcal{X} \rightarrow (-\infty, \infty]$ a law-invariant convex risk measures that has the strong Fatou property. Then ρ is $\sigma(\mathcal{X}, L^\infty)$ lower semicontinuous and it extends uniquely to a law-invariant convex risk measure on L^1 with the Fatou property.*

The result of ρ is $\sigma(\mathcal{X}, L^\infty)$ lower semicontinuous ensures ρ admitting a dual representation via L^∞ space. The following version of Komlos' Theorem is also useful.

Proposition 4.9. *Let (X_n) be a norm bounded sequence of random variables in an r.i. space \mathcal{X} , then there exists a random variable X (not necessarily in \mathcal{X}) and a subsequence (X_{n_k}) of (X_n) such that the arithmetic means of all subsequences of (X_{n_k}) converges to X almost surely, that is, $\frac{1}{k} \sum_{j=1}^k X_{n_j} \xrightarrow{a.s.} X$.*

With preceding theorem and proposition, we now can start to establish the main result of this paper, that is, inf-convolution of risk measures preserves the strong Fatou property on r.i. spaces, and the infimum in the definition of inf-convolution is attainable everywhere.

Theorem 4.10. *Let \mathcal{X} be an r.i. space over a fixed nonatomic probability space, and $\rho_i : \mathcal{X} \rightarrow (-\infty, \infty]$, $i = 1, \dots, n$, be law-invariant convex risk measure with the strong Fatou property. Then $\square_{i=1}^n \rho_i : \mathcal{X} \rightarrow (-\infty, \infty]$ is a law-invariant convex risk measure, and exact, and has the strong Fatou property. Moreover, for each $X \in \mathcal{X}$ there exist increasing functions $f_i : \mathbb{R} \rightarrow \mathbb{R}$, $i = 1, \dots, n$, such that $\sum_{i=1}^n f_i(x) = x$ for every $x \in \mathbb{R}$ and*

$$\square_{i=1}^n \rho_i(X) = \sum_{i=1}^n \rho_i(f_i(X)).$$

Proof. It is enough to prove when $n = 2$, the rest follows by induction. For $i = 1, 2$, ρ_i is law-invariant convex risk measure with the strong Fatou property. By theorem 4.8, each ρ_i extends to a functional $\bar{\rho}_i : L^1 \rightarrow (-\infty, \infty]$ that is law-invariant convex risk measure with Fatou property, thus is $\|\cdot\|_1$ lower semicontinuous on L^1 . Let $\bar{\rho}_1 \square \bar{\rho}_2 : L^1 \rightarrow (-\infty, \infty]$ be the inf-convolution of $\bar{\rho}_1$ and $\bar{\rho}_2$. Clearly,

$$\bar{\rho}_1 \square \bar{\rho}_2(X) \leq \rho_1 \square \rho_2(X) \text{ for any } X \in \mathcal{X}.$$

Now, pick any $X \in \mathcal{X}$. By Theorem 3.4, there exist increasing functions $f_1, f_2 : \mathbb{R} \rightarrow \mathbb{R}$ such that $f_1(x) + f_2(x) = x$ for each $x \in \mathbb{R}$ and $\bar{\rho}_1 \square \bar{\rho}_2(X) = \bar{\rho}_1(f_1(X)) + \bar{\rho}_2(f_2(X))$. Since

ρ_1, ρ_2 are cash-additive, without loss of generality we may assume that $f_1(0) = f_2(0) = 0$. By lemma 3.10 we have that f_1 and f_2 are 1-Lipschitz functions and thus $|f_i(X)| \leq |X|$ for $i = 1, 2$. Since \mathcal{X} is an order ideal of L^0 , we have that $f_i(X) \in \mathcal{X}$ for $i = 1, 2$. Therefore,

$$\overline{\rho_1} \square \overline{\rho_2}(X) = \overline{\rho_1}(f_1(X)) + \overline{\rho_2}(f_2(X)) = \rho_1(f_1(X)) + \rho_2(f_2(X)) \geq \rho_1 \square \rho_2(X).$$

It follows that $\overline{\rho_1} \square \overline{\rho_2}(X) = \rho_1(f_1(X)) + \rho_2(f_2(X)) = \rho_1 \square \rho_2(X)$, implying that $\rho_1 \square \rho_2$ is exact and $\overline{\rho_1} \square \overline{\rho_2}$ extends $\rho_1 \square \rho_2$. By Theorem 3.4, $\overline{\rho_1} \square \overline{\rho_2}$, and therefore, $\rho_1 \square \rho_2$, is law-invariant.

It remains to show that $\rho_1 \square \rho_2$ has the strong Fatou property. Pick an arbitrary $m \in \mathbb{R}$, and consider the sublevel set $\mathcal{C} := \{X \in \mathcal{X} : \rho_1 \square \rho_2(X) \leq m\}$. Let (X_n) be a norm bounded sequence in \mathcal{C} that a.s.-converges to $X \in \mathcal{X}$. It suffices to show that $X \in \mathcal{C}$. By the exact solution described above, we can find $Y_n, Z_n \in \mathcal{X}$ with $X_n = Y_n + Z_n$, $|Y_n| \leq |X_n|$, $|Z_n| \leq |X_n|$, and $\rho_1 \square \rho_2(X_n) = \rho_1(Y_n) + \rho_2(Z_n)$. Note that $(Y_n), (Z_n)$ are norm bounded sequences in \mathcal{X} . Applying Proposition 4.9 twice, we can find strictly increasing (n_j) and two random variables $Y, Z \in L^0$ such that $\frac{1}{k} \sum_{j=1}^k Y_{n_j} \xrightarrow{a.s.} Y$ and $\frac{1}{k} \sum_{j=1}^k Z_{n_j} \xrightarrow{a.s.} Z$. Since $|\frac{1}{k} \sum_{j=1}^k Y_{n_j}| \leq \frac{1}{k} \sum_{j=1}^k |X_{n_j}| \xrightarrow{a.s.} |X|$, we get that $|Y| \leq |X|$, so that $Y \in \mathcal{X}$. Similarly, we have $Z \in \mathcal{X}$. Note also that $Y + Z = X$ and that $(\frac{1}{k} \sum_{j=1}^k Y_{n_j})$ and $(\frac{1}{k} \sum_{j=1}^k Z_{n_j})$ are both norm bounded sequences in \mathcal{X} . Thus, applying the strong Fatou property and convexity of ρ_i 's, we get that

$$\begin{aligned} \rho_1 \square \rho_2(X) &\leq \rho_1(Y) + \rho_2(Z) \\ &\leq \liminf_k \rho_1\left(\frac{1}{k} \sum_{j=1}^k Y_{n_j}\right) + \liminf_k \rho_2\left(\frac{1}{k} \sum_{j=1}^k Z_{n_j}\right) \\ &\leq \liminf_k \frac{\sum_{j=1}^k \rho_1(Y_{n_j})}{k} + \liminf_k \frac{\sum_{j=1}^k \rho_2(Z_{n_j})}{k} \\ &\leq \liminf_k \left(\frac{\sum_{j=1}^k \rho_1(Y_{n_j})}{k} + \frac{\sum_{j=1}^k \rho_2(Z_{n_j})}{k} \right) \\ &= \liminf_k \left(\frac{\sum_{j=1}^k \rho_1 \square \rho_2(X_{n_j})}{k} \right) \leq m. \end{aligned}$$

This proves that $X \in \mathcal{C}$ and completes the proof of the proposition. \square

Proposition 4.11. *Let $\rho_i : \mathcal{X} \rightarrow (-\infty, \infty]$, $i = 1, \dots, n$, be coherent and law-invariant risk measure with the strong strong Fatou property. Then $\square_{i=1}^n \rho_i : \mathcal{X} \rightarrow (-\infty, \infty]$ is coherent,*

law-invariant, and exact, and has the strong Fatou property. In particular, $\square_{i=1}^n \rho_i$ admits the following representation:

$$\square_{i=1}^n \rho_i(X) = \sup_{Y \in \mathcal{Q}} \mathbb{E}[-XY]$$

where $\mathcal{Q} = \{Y \in L_+^\infty \mid \sum_{i=1}^n \rho_i^*(-Y) = 0, \text{ and } \mathbb{E}[Y] = 1\}$.

Proof. Let $\hat{\rho}(X) = \square_{i=1}^n \rho_i(X)$, by theorem 4.10 we have $\hat{\rho}$ is a law-invariant convex risk measure and has strong Fatou property on \mathcal{X} . Then by theorem 4.8, $\hat{\rho}$ is $\sigma(\mathcal{X}, L^\infty)$ -lower semicontinuous. Also, it is easy to verify that $\hat{\rho}$ has positive homogeneity thus a coherent risk measure. Applying Fenchel-Moreau theorem we have that $\hat{\rho}$ admits following representation:

$$\hat{\rho}(X) = \sup_{Y \in \mathcal{Q}} \mathbb{E}[-XY],$$

where $\mathcal{Q} = \{Y \in L_+^\infty \mid \sum_{i=1}^n \rho_i^*(-Y) = 0, \text{ and } \mathbb{E}[Y] = 1\}$. □

As was mentioned previously, for $\alpha \in (0, 1)$, the ES_α is a coherent risk measure with the strong Fatou property on any r.i. spaces, and ES_α is law-invariant by its definition. The following inf-convolution of two Expected shortfall risk measures provide a concrete example for solving an optimal allocation in risk sharing problem.

Example 4.12. For any $0 < \beta \leq \alpha < 1$, we have the following (see the proof of theorem 4.39 in [15]):

$$ES_\alpha^*(-Y) = \begin{cases} 0 & 0 < Y \leq \frac{1}{\alpha} \\ \infty & \text{otherwise} \end{cases}, \quad \text{and} \quad ES_\beta^*(-Y) = \begin{cases} 0 & 0 < Y \leq \frac{1}{\beta} \\ \infty & \text{otherwise} \end{cases},$$

Thus, by proposition 4.11, we have

$$(ES_\alpha \square ES_\beta)(X) = \sup_{Y \in L_+^\infty, ES_\alpha^*(-Y) + ES_\beta^*(-Y) = 0, \mathbb{E}[Y] = 1} \mathbb{E}[-XY] = ES_\alpha(X)$$

So, if we have two Expected shortfall risk measures ES_α and ES_β with $0 < \beta \leq \alpha < 1$, then their inf-convolution is exact, and it is equal to the ES_α . In this case, the allocation will be putting total position X only on the company with risk measure ES_α .

Bibliography

- [1] Aliprantis, C.D. and Border, K.C.: Infinite dimensional analysis: a hitchhiker's guide. *Studies in Economic Theory*, 4. (1999)
- [2] Artzner, P., Delbaen, F., Eber, J.M., and Heath, D.: Coherent measures of risk. *Mathematical finance*, 9(3), 203-228. (1999)
- [3] Barrieu, P. and El Karoui, N.: Inf-convolution of risk measures and optimal risk transfer. *Finance and stochastics*, 9(2), 269-298. (2005)
- [4] Bennett, C. and Sharpley, R. C.: Interpolation of operators. *Academic press*, Vol 129. (1988)
- [5] Biagini, S. and Frittelli, M.: On the extension of the Namioka-Klee theorem and on the Fatou property for risk measures. *In Optimality and risk-modern trends in mathematical finance*, Springer, Berlin, Heidelberg, 1-28. (2009)
- [6] Biagini, S. and Frittelli, M.: On the extension of the Namioka-Klee theorem and on the Fatou property for risk measures. *In Optimality and risk-modern trends in mathematical finance*, pp. 1-28. Springer, Berlin, Heidelberg. (2009)
- [7] Braverman, M. S. and Braverman, M. S.: Independent random variables and rearrangement invariant spaces. *Cambridge University Press*, Vol. 194. (1994)
- [8] Brezis, H.: Functional analysis, Sobolev spaces and partial differential equations. *Springer Science and Business Media*. (2010)

- [9] Chen, S., Gao, N., and Xanthos, F.: The strong Fatou property of risk measures. *Dependence Modeling*, 6: 183-196. (2018)
- [10] Cheridito, P. and Li, T.: Risk measures on Orlicz hearts. *Mathematical Finance: An International Journal of Mathematics, Statistics and Financial Economics*, 19(2), 189-214. (2009)
- [11] Dana, R. A.: A representation result for concave Schur concave functions. *Mathematical Finance: An International Journal of Mathematics, Statistics and Financial Economics*, 15(4), 613-634. (2005)
- [12] Delbaen, F.: Coherent risk measures on general probability spaces. *In Advances in finance and stochastics*, pp. 1-37. (2002)
- [13] Edgar, G.A., Sucheston, L.: Stopping Times and Directed Processes. *Cambridge University Press*, Cambridge. (1992)
- [14] Filipović, D. and Svindland, G.: Optimal capital and risk allocations for law- and cash-invariant convex functions. *Finance stochastics*, 12(3) 423-439. (2008)
- [15] Föllmer, H. and Schied, A.: Stochastic Finance: An Introduction in Discrete Time, 2nd edition. *de Gruyter Studies in Mathematics*, 27. (2002)
- [16] Frittelli, M. and Gianin, E.R.: Law-invariant convex risk measures. *In Advances in mathematical economics* 7, 33-46. (2005)
- [17] Gao, N., Leung, D.H., and Xanthos, F.: Closedness of convex sets in Orlicz spaces with applications to dual representation of risk measures. Preprint: arXiv:1610.08806. (2017)
- [18] Gao, N., Leung, D.H., Munari, C., and Xanthos, F.: Fatou property, representations, and extensions of law-invariant risk measures on general Orlicz spaces, *Finance and Stochastics*, 22(2), 395-415. (2018)

- [19] Gao, N., and Xanthos, F.: On the C-property and w^* -representations of risk measures, *Math finance*. 28(2), 748-754. (2018)
- [20] Jouini, E., Schachermayer, W., and Touzi, N.: Law invariant risk measures have the Fatou property. *In Advances in mathematical economics*, pp. 49-71. (2006)
- [21] Kaina, M. and Rüschendorf, L.: On convex risk measures on L^p -spaces. *Mathematical methods of operations research*, 69(3), 475-495. (2009)
- [22] Liebrich, F. B. and Svindland, G.: Model spaces for risk measures. *Insurance: Mathematics and Economics*, 77, 150-165. (2017)
- [23] Meyer-Nieberg, P.: Banach lattices. *Berlin-Heidelberg-New York*. (1991)
- [24] Rockafellar, R. T.: Convex Analysis. *Princeton landmarks in mathematics*. (1997)