Mathematisches Institut
Justus-Liebig-Universität Gießen

Capital Allocation and Systemic Risk

Dissertation

Submitted by

Florian Schindler

in fulfillment of the requirements for the degree of
„Doctor rerum naturalium“ (Dr. rer. nat.)

Supervisor: Prof. Dr. Ludger Overbeck

April 2024

Abstract

We study (scalar) systemic risk measures in a general framework. This framework embeds the existing approaches of [KOZ16] (first aggregate or axiomatic approach) and [BFFMB20] (first inject capital). It turns out, that even in non trivial situations systemic risk measures of the first inject capital approach can be described by a first aggregate systemic risk measure. Moreover, we study capital allocation rules. We introduce two types, one is connected to the dual representation of the underlying systemic risk measure and the other one is in the spirit of Aumann-Shapley. The canonical ways to this task in the frameworks of [KOZ16] and [BFFMB20] are special cases of our general dual capital allocation rule. In addition, we are able to extend the so called *scenario dependent allocation* to arbitrary configurations of the system.

Zusammenfassung

Diese Arbeit befasst sich mit (skalarwertigen) systemischen Risikomaßen auf allgemeinen Räumen. Die Allgemeinheit der Rahmenbedingungen ermöglicht es die bestehenden Ansätze von [KOZ16] (erst aggregieren oder axiomatischer Ansatz) und [BFFMB20] (erst Kapital zuführen) einzubetten. Trotz der offensichtlichen Unterschiede gelingt es auch in nicht trivialen Fällen systemische Risikomaße vom „erst Kapital zuführen“ Ansatz durch ein systemisches Risikomaß des „erst aggregieren“ Ansatzes zu beschreiben. Des Weiteren werden Allokationsmechanismen untersucht. Einer dieser Mechanismen steht im Zusammenhang mit der dualen Darstellung systemischer Risikomaße, der andere orientiert sich an Aumann-Shapley. Die kanonischen Mechanismen der Ansätze aus [KOZ16] und [BFFMB20] stellen sich als Spezialfälle heraus. Zusätzlich werden sog. *Szenario abhängige Allokationen* auf beliebige Teilsysteme erweitert.

Für Marina, Hannah und Nala.

Contents

Listings

Introduction

The financial crises of the past decades showed that measuring systemic risk in a suitable way is an urgency. If we consider a firm consisting of $n \in \mathbb{N}$ business units or a portfolio consisting of $n \in \mathbb{N}$ assets, allocating the overall risk to its constituent parts plays a crucial role. In [Kal05], [Tas07] and [Tsa09] the authors presented some helpful tools for this task. However, financial systems do not fit in the framework of these tools. First, the idea that the constituent parts subsidize each other, i.e. to simply sum up all profits and losses, is not a realistic scenario for a financial system $\bar{X} = (X_1, \ldots, X_n)$. More general aggregation rules Λ which reflect the structure of the system appropriately need to be considered. The authors in [CIM13], later extended to general measurable spaces by [KOZ16], and [BFFMB19] developed this idea in two directions. The difference between these two approaches is the order of aggregating and allocating. Starting with some axioms - partially motivated by economic or risk management reasoning and partially by mathematical structures - it is shown in [CIM13] and [KOZ16] that the risk of a financial system should have the form

$$\rho\left(\bar{X}\right) = \inf\left\{m \in \mathbb{R} \mid \Lambda\left(\bar{X}\right) - m \in \mathcal{A}_0\right\} = \rho_0\left(\Lambda\left(\bar{X}\right)\right). \tag{0.1}$$

In other words, a mapping ρ fulfills the given axioms if an only if it is of the form „$\rho_0 \circ \Lambda$“ for a suitable *single-firm* risk measure ρ_0 with corresponding acceptance set $\mathcal{A}_0$ and a suitable *aggregation rule* Λ. In particular, this representation implies that the aggregation takes place *before* the capital is injected to the system. The scalar $\rho\left(\bar{X}\right)$ can be interpreted - in analogy to single-firm risk measures - as the minimal amount of capital which has to be injected to the aggregated system to make it acceptable, i.e. saves the system of a collapse. This approach produces meaningful results, like dual representation theorems, in a general framework.

In [BFFMB19] an alternative procedure is presented. They argue that aggregating *after* injecting capital on the level of the single institutions could affect the overall systemic risk in a positive way. Starting with a mapping as in (0.1), these risk measures appear as

$$D\left(\bar{X}\right) = \inf\left\{\sum_{i=1}^{n} m_i \mid \bar{m} \in \mathbb{R}^n, \Lambda\left(\bar{X} - \bar{m}\right) \in \mathcal{A}_0\right\}. \tag{0.2}$$

Again, the set $\mathcal{A}_0$ is the acceptance set of a single-firm risk measure, the mapping $\Lambda \colon \mathbb{R}^n \to \mathbb{R}$ is an aggregation rule, and the overall systemic risk is described by a scalar. It can be interpreted as a valuation of the injected capital. Here, the valuation is given by adding up the injected capital, but more general valuations π are possible. An extension of (0.2) occurs if we allow for random capital injections, i.e.

$$R\left(\bar{X}\right) = \inf\left\{\sum_{i=1}^{n} Y_i \mid \bar{Y} \in \mathcal{C}, \Lambda\left(\bar{X} - \bar{Y}\right) \in \mathcal{A}_0\right\}, \tag{0.3}$$

where $\mathcal{C}$ is some set consisting of vectors of $\mathcal{F}$-measurable functions $Y_i\colon \Omega \to \mathbb{R}$ on $(\Omega, \mathcal{F}, \mathbb{P})$ (denoted by $(L^0)^n = (L^0(\Omega, \mathcal{F}, \mathbb{P}))^n$). The set $\mathcal{C}$ contains additional restrictions for the (possibly random) allocation $\bar{Y}$. In all studies of these systemic risk measures, we have

$$\mathcal{C} \subseteq \mathcal{C}(\mathbb{R}) := \left\{\bar{Y} \in \left(L^0\right)^n \mid \sum_{i=1}^{n} Y_i \in \mathbb{R}\right\}. \tag{0.4}$$

The interpretation of this constraint is that the overall capital which is needed to save the system is determined today, but the allocation depends on the occurring scenario. Under some restrictive assumptions, further studies are possible and dual representation results can be derived. At first sight, it seems that this approach yields a completely new, more flexible type of systemic risk measures which are only related to the type (0.1) in some trivial cases (for example if $\Lambda(\bar{x}) = \sum_{i=1}^{n} x_i$). However, we are able to show in Theorem 2.70 that there are many non trivial cases where systemic risk measures of type (0.3) can be embedded to the axiomatic approach. This is one of the main results of this work. This means that we are able to represent them via (0.1) for some suitable $\tilde{\rho}_0$ and $\tilde{\Lambda}$. This identification is demonstrated in Example 2.72. It refines the observations in [BFFMB19] and [DFG24] in the situation where dual representations are possible.

Besides [CIM13, KOZ16] and [BFFMB20] there are many other contributions to the theory of (scalar) systemic risk measures. We already mentioned [AKMM21] and [AR20]. Additionally, [AR20] captures dual representations for set-values systemic risk measures. We would like to mention that we do not consider set-valued systemic risk measures in this work and refer the interested reader to [FRW17]. A completely different approach appears for example in [RV13]. Initiated by the seminal work of [EN01], the financial system is represented by a stochastic network with specific structure. One then tries to analyze the weaknesses of the network and how it might fail. The object of interest is a so called clearing vector which settles all liabilities in the network within a simultaneously clearing mechanism. Existence results and efficient algorithms to compute such a clearing vector are studied. A more detailed review of the existing literature with a focus on scalar systemic risk measures is given in [DF21].

The structure of this monograph is as follows. We start our analysis by presenting a general framework to define systemic risk measures in Definition 2.1. The underlying spaces and our definition are general enough to embed both approaches. We formulate general representation results in Proposition 2.3 and Theorem 2.8. The existing results in the former approaches are then special cases. The additional structure yields a refinement of our general result. This yields a simplified proof for the dual representation result given in [KOZ16]. The authors in [AKMM21] and [AR20] also present an alternative way to proof the dual representation results for both approaches. However, they use different techniques and work with less general spaces. The remainder of Section 2 is devoted to the representation of the main results of the first aggregate and first inject capital approaches. Especially for systemic risk measures of type (0.1), some results can be generalized in some way. As mentioned before, the main result of this section is Theorem 2.70.

The second part of our analysis is dedicated to allocation rules. The systemic risk measures of [KOZ16] and [BFFMB20] suggest a natural way to allocate the risk to the participants of the system. In Section 3, we investigate a general natural capital allocation rule based on the general dual representation result. It turns out, that the proposed allocation principles from [KOZ16] and [BFFMB20] are exactly of this type. Moreover, we can extend the so called *scenario dependent capital allocation rule* or *random capital allocation rule* from [BFFMB20] to arbitrary subsystems $\bar{Z}$. In Example 3.6 we give explicit formulas for this type of capital allocation rule. In subsection 3.2, we describe the procedure to deduce a random allocation for an arbitrary subsystem. This procedure heavily relies on the fact that the allocation mechanism proposed in [BFFMB20] is a special case of our general dual systemic capital allocation rule. Additionally, we present an allocation rule which is in the spirit of the Aumann-Shapley allocation rule ([Den01, Tsa09]) for single-firm risk measures. Moreover, for Gâteaux differentiable systemic risk measures we can guarantee some type of full allocation. We close the section with some examples and a numerical case study. For the entropic systemic risk measure, the allocation and risk measurement principles derived theoretically are implemented in Python to compute the different values.

1. Model and Notation

Initiated by the seminal work of [ADEH99], the study of (scalar) risk measures is well developed in many different directions. From an economic and financial standpoint, risk is a tribute to the uncertainty of the future. The possibility of unfavorable events in the future should affect the evaluation today. The beginning of the mathematical task behind measuring risk is to model the financial positions in a correct way. To this end, let $(\Omega, \mathcal{F})$ be a general measurable space. By $L^0 = L^0(\Omega, \mathcal{F})$ we denote the vector space of $\mathcal{F}$-measurable functions $X\colon \Omega \to \mathbb{R}$. $\mathcal{M}_1 = \mathcal{M}_1(\Omega, \mathcal{F})$ denotes the set of all probability measures on $(\Omega, \mathcal{F})$. In situations, where the existence of base probability measure $\mathbb{P}$ on $(\Omega, \mathcal{F})$ is assumed, $\mathcal{M}_1(\mathbb{P}) = \mathcal{M}_1(\Omega, \mathcal{F}, \mathbb{P})$ denotes the set of all probability measures on $(\Omega, \mathcal{F})$ which are absolutely continuous w.r.t. $\mathbb{P}$. With L^0_+ we denote the cone of all nonnegative functions, i.e.

$$L^0_+ \coloneqq \left\{ X \in L^0 \mid X(\omega) \geq 0 \, \forall \, \omega \in \Omega \right\}.$$

The relation

$$X \succcurlyeq Y \Leftrightarrow X - Y \in L^0_+$$

then defines a partial order on L^0 and L^0 becomes an ordered vector space. The spaces under our consideration are always vector subspaces of L^0 containing the constants. We typically denote them by $\mathcal{X}$ or $\mathcal{X}_i$, $i \in \{1, \ldots, n\}$ if we consider more than just one space at a time. The ordering of L^0 transfers to its vector subspaces through

$$\mathcal{X}_+ \coloneqq L^0_+ \cap \mathcal{X} = \{X \in \mathcal{X} \mid X(\omega) \geq 0 \, \forall \, \omega \in \Omega\}.$$

Since we are interested in dual representations, we work with dual pairs (see A.2 for a detailed introduction of this concept). We denote this pair by $\langle \mathcal{X}, \mathcal{X}' \rangle$, its pairing by $\langle \cdot, \cdot \rangle$ and the elements of $\mathcal{X}'$ by ξ. We equip both spaces with consistent locally convex Hausdorff topologies. Furthermore, as in [RS06], we need the following technical assumption:

(C) If $\xi \notin \mathcal{X}'_+$, then there exists $X \in \mathcal{X}_+$ such that $\langle X, \xi \rangle < 0$.

[RS06] pointed out that this is a very mild requirement which ensures that the cone $\mathcal{X}'_+$ is dual to the cone $\mathcal{X}_+$, i.e. $\mathcal{X}'_+ = \{\xi \in \mathcal{X}' : \langle X, \xi \rangle \geq 0 \; \forall \; X \in \mathcal{X}_+\}$. A sufficient condition to guarantee (C) is that the space $\mathcal{X}$ contains all indicator functions $\mathbb{1}_A, A \in \mathcal{F}$.

We consider a model with one time period. Over this time period we are interested in measuring the risk of a financial position. Now, $X \in \mathcal{X}$ is understood as the future net loss of a single finanical position. A typical example of such a position is a firm, so we will call it *firm* from now on. For a financial system, consisting of a finite set of $n \in \mathbb{N}$ firms, representing n nodes in a financial network, the vector $\bar{X} = (X_1, \ldots, X_n) \in \bigtimes_{i=1}^{n} \mathcal{X}_i$ describes the net losses of these nodes, i.e. $X_i \in \mathcal{X}_i$ is the net loss of firm i. By abuse of notation, we write $\bar{X} \succcurlyeq \bar{Y}$ for $\bar{X}, \bar{Y} \in \bigtimes_{i=1}^n \mathcal{X}_i$ if $X_i \succcurlyeq Y_i$ for all $i = 1, \dots, n$. This means that the partial order is induced by

$$\left(\bigtimes_{i=1}^n \mathcal{X}_i\right)_+ \coloneqq \left\{\bar{X} \in \bigtimes_{i=1}^n \mathcal{X}_i \mid X_i(\omega) \geq 0 \ \forall \ \omega \in \Omega \text{ and } i \in \{1, \dots, n\}\right\} = \bigtimes_{i=1}^n (\mathcal{X}_i)_+ .$$

We make the following strctural assumptions on $\bigtimes_{i=1}^n \mathcal{X}_i$: We suppose that for each $i \in \{1, \dots, n\}$ $\langle \mathcal{X}_i, \mathcal{X}_i' \rangle$ forms a dual pair, where both spaces are equipped with consistent locally convex Hausdorff topolgies. We denote the elements of $\bigtimes_{i=1}^n \mathcal{X}_i'$ by $\Xi = (\xi_1, \dots, \xi_n)$ and the pairing by

$$\langle \bar{X}, \Xi \rangle_n \coloneqq \sum_{i=1}^n \langle X_i, \xi_i \rangle, \tag{1.1}$$

where $\langle \cdot, \cdot \rangle$ is the pairing of the dual pair $\langle \mathcal{X}_i, \mathcal{X}_i' \rangle$. This pairing induces the dual pair $\langle \bigtimes_{i=1}^n \mathcal{X}_i, \bigtimes_{i=1}^n \mathcal{X}_i' \rangle$. $\bigtimes_{i=1}^n \mathcal{X}_i$ and $\bigtimes_{i=1}^n \mathcal{X}_i'$ are equipped with their respective product topologies. Thus, they are locally convex Hausdorff spaces. Moreover, with (1.1) we obtain $\left(\bigtimes_{i=1}^n \mathcal{X}_i\right)' = \bigtimes_{i=1}^n \mathcal{X}_i'$ and vice versa which means the product topologies are consistent with the pairing. To simplyfy the notation, we will write $\bar{\mathcal{X}} \coloneqq \bigtimes_{i=1}^n \mathcal{X}_i$ and $\bar{\mathcal{X}}' \coloneqq \left(\bigtimes_{i=1}^n \mathcal{X}_i\right)'$. Typical examples for the underlying spaces are L^p-spaces and Orlicz spaces (see Example 2.26 (iii)). Endowed with their norm topologies, these spaces are Banach lattices. In these situations, the pairing is given by

$$\langle \bar{X}, \Xi \rangle_n \coloneqq \sum_{i=1}^n \mathbb{E}\left[X_i \xi_i\right]. \tag{1.2}$$

A special treatment is neccessary for $\langle L^\infty, (L^\infty)' \rangle$. There the expectaion above is interpreted as the integral of $X \in L^\infty$ w.r.t. the finitely additive set function corresponding to ξ. For more details on this special case, we refer to [FS08] Section A.6. Additionally, we will use the notation

$$\mathcal{C}(\mathbb{R}) \coloneqq \left\{\bar{Y} \in \left(L^0\right)^n \mid \sum_{i=1}^n Y_i \in \mathbb{R}\right\}, \tag{1.3}$$

e_i for the i-th unit vector in $\mathbb{R}^n$, $1_n = (1, \dots, 1) \in \mathbb{R}^n$ and set $\inf \emptyset \coloneqq \infty$.

2. Systemic Risk Measures

Once the model is specified, it is necessary to translate preferable economic properties of a risk measurement tool into mathematical properties. Scalar risk measures ρ address this task. They operate on a given space of finanical positions and return a scalar, i.e. a real number. Positions with positive risk can be interpreted as some kind of unfavorable or risky and the negative ones as favorable or safe. With this in mind, the value $-\infty$ is unrealistic. There is no position which is completely risk free. It is therefore excluded completely. We shall also exclude the trivial case where $\rho \equiv \infty$ in our analysis. However, it is possible that there are positions which carry that much risk that they are intolerable. In other words, risk measures have to be *proper* (see Definition A.1 for a rigourous definition) mappings. In the common literature, risk measures for single firms - called *single-firm risk measures* - and risk measures for systems - called *systemic risk measures* - are treated separately. We will see that both situations have distinctive features. But if we consider a single firm as a system, consisting of eaxtly one firm, a systemic risk measure should act as a single-firm risk measure. In this sense a systemic risk measure should extend the ideas of single-firm risk measure. The development of the literature worked that way. However, we will start from the other direction and give a formal definition of a systemic risk measure and identify a single-firm risk measure as the special case $n = 1$. All the results presented for systemic risk measures in this section include results for single-firm risk measures if one sets $n = 1$. The idea of finanical network behind a system of firms, where the firms interact with each other, yields us to use the alternative term *participants* for the firms. Let us start with the basic definition.

Definition 2.1. *Consider the following properties for a proper mapping* $\rho\colon \bar{\mathcal{X}} \to \mathbb{R} \cup \{\infty\}$*:*

(S1) Monotonicity: For all $\bar{X}, \bar{Y} \in \bar{\mathcal{X}}$ *with* $\bar{X} \succcurlyeq \bar{Y}$*,* ρ *satisfies* $\rho\left(\bar{X}\right) \geq \rho\left(\bar{Y}\right)$*.*

(S2) Convexity:

- *(S2a) Outcome convexity: For all* $\bar{X}, \bar{Y} \in \bar{\mathcal{X}}$ *and all* $\alpha \in [0, 1]$*,* ρ *satisfies* $\rho\left(\alpha\bar{X} + (1-\alpha)\bar{Y}\right) \leq \alpha\rho\left(\bar{X}\right) + (1-\alpha)\rho\left(\bar{Y}\right)$*.*
- *(S2b) Risk convexity: If* $\rho\left(\bar{Z}(\omega)\right) = \alpha\rho\left(\bar{X}(\omega)\right) + (1-\alpha)\rho\left(\bar{Y}(\omega)\right)$ *for a given* $\alpha \in [0, 1]$ *and for all* $\omega \in \Omega$ *holds, then* ρ *satisfies* $\rho\left(\bar{Z}\right) \leq \alpha\rho\left(\bar{X}\right) + (1-\alpha)\rho\left(\bar{Y}\right)$*.*

(S3) Positive homogeneity: For all $\bar{X} \in \bar{\mathcal{X}}$ *and all* $\alpha \in \mathbb{R}_+ := [0, \infty)$*,* ρ *satisfies* $\rho\left(\alpha\bar{X}\right) = \alpha\rho\left(\bar{X}\right)$*.*

(S4) Preference consistency: For all $\bar{X}, \bar{Y} \in \bar{\mathcal{X}}$ *with* $\rho\left(\bar{X}(\omega)\right) \geq \rho\left(\bar{Y}(\omega)\right)$ *for all* $\omega \in \Omega$*,* ρ *satisfies* $\rho\left(\bar{X}\right) \geq \rho\left(\bar{Y}\right)$*.*

(S5) $\mathcal{R}$*-Surjectivity:* $\rho\left(\mathbb{R}^n\right) = \mathcal{R}$ *for* $\mathcal{R} \in \{\mathbb{R}, \mathbb{R}_+\}$*.*

(S6) Normalization: $\rho\left(1_n\right) = n$*.*

(S7) Subadditivity: For all $\bar{X}, \bar{Y} \in \bar{\mathcal{X}}$, ρ satisfies $\rho\left(\bar{X}+\bar{Y}\right) \leq \rho\left(\bar{X}\right)+\rho\left(\bar{Y}\right)$.

(S8) Translation property (on $\mathcal{B}$): For all $\bar{Y} \in \mathcal{B} \subseteq \mathcal{C}(\mathbb{R}) \cap \bar{\mathcal{X}}$, ρ satisfies $\rho\left(\bar{X}+\bar{Y}\right) = \rho\left(\bar{X}\right)+\sum_{i=1}^{n} Y_i$.

If a mapping satisfies the properties (S1) and (S2a), it is called systemic risk measure. If a systemic risk measure additionally satisfies (S3), it is called positively homogeneous systemic risk measure.

As mentioned in the introduction to this section, these properties have an economic background. Their interpretation is also discussed in [CIM13] and [KOZ16]. To guarantee a self contained character, we will have a look at their motivation in the following. Monotonicity (S1) captures the simple fact that bigger losses are considered to be riskier than smaller losses. Outcome convexity (S2a) is linked to diversification: The risk of a diversified portfolio should be no greater than the mixture of the stand alone risks. Risk convexity (S2b) is connected to risk aversion: Consider a situation with two stages of randomness. In each scenario we will carry the scenariowise risk of the system $\bar{X}$ with probability α and that of the system $\bar{Y}$ with probability $1-\alpha$. Now the risk of the average system $\bar{Z}$, i.e. a system with scenariowise risk defined as the average of the scenariowise risks of $\bar{X}$ and $\bar{Y}$, should not exceed the average of the risks of $\bar{X}$ and $\bar{Y}$. Positive homogeneity (S3) allows for a linear relation between the risk of a system $\bar{X}$ and its size. For a positively homogeneous systemic risk measure normalization (S6) can be interpreted as some sort of scaling. Preference consistency (S4) is again connected to scenariowise risks. If the scenariowise risk of the system $\bar{X}$ is at most the scenariowise risk of the system $\bar{Y}$, then the risk of $\bar{X}$ should not exceed the risk of $\bar{Y}$. Subadditivity (S7) is an other way to think of the benefits occuring form diversification: If a system can be divided into two subsystems, the standalone evaluation should not reduce the risk. Finally, the translation property (S8) describes the sensitivity of the systemic risk measure to certain changes of the position $\bar{X}$. Usually, we have $\mathbb{R}^n \subseteq \mathcal{B}$. In this situation, the systemic risk can be interpreted as a capital requirement. For $n = 1$ (i.e. for single-firm risk measures) this property is also know under the name *cash additivity*. Roughly speaking, deterministic capital injections to the system yield a reduction of the systemic risk by the overall injected amount. Note, however, that only the overall injected amount has to be deterministic. As soon as $n > 1$, the set of possible injections may be enlarged and one may have $\mathbb{R}^n \subset \mathcal{B}$. The $\mathcal{R}$-surjectivity (S5) is of technical nature.

Remark 2.2. *Of course it is possible for a systemic risk measure to fulfill far more properties. However, this set of axioms reflects the most common features of a risk measurement tool from an economic standpoint. It is worth mentioning that the axioms are not independent of each other. More precisely, a given subset of axioms may already imply one of the others. The most popular implication chain* *is the following one: For a systemic risk measure with $\rho(0) = 0$, every pair of the axioms (S2), (S2a) and (S7) already implies the remaining third one. Other implications are presented in Remark 2.25 and Theorem 2.70. In each situation, the implications hold due to additional structural properties of the mapping ρ which are of mathematical nature.*

2.1. Representations of Systemic Risk Measures

Once the set of axioms is determined, one wishes to get a general notion of mappings which fit with these. There are two common ways to represent such mappings. One is the so called *primal* representation. For convex mappings, this is simply the representation through the epigraph $\operatorname{epi}(\rho) = \left\{\left(m, \bar{X}\right) \in \mathbb{R} \times \bar{\mathcal{X}} \mid m \geq \rho\left(\bar{X}\right)\right\}$. As it is common for the risk measure literature, we will use the notation $\mathcal{A}_\rho = \operatorname{epi}(\rho)$ for the epigraph of ρ.

Proposition 2.3. *Suppose that $\rho : \bar{\mathcal{X}} \to \mathbb{R} \cup \{\infty\}$ is a systemic risk measure.*

(i) For all $\bar{X} \in \bar{\mathcal{X}}$, ρ admits the primal representation

$$\rho\left(\bar{X}\right) = \inf\left\{m \in \mathbb{R} \mid \left(m, \bar{X}\right) \in \mathcal{A}_\rho\right\}, \tag{2.4}$$

with $\inf \emptyset := \infty$ as set earlier.

(ii) For the set $\mathcal{A}_\rho$ we have:

(a) $\mathcal{A}_\rho$ is nonempty.

(b) For all $\bar{X} \in \bar{\mathcal{X}}$, we have $\inf\left\{m \in \mathbb{R} \mid \left(m, \bar{X}\right) \in \mathcal{A}_\rho\right\} > -\infty$.

(c) $\mathcal{A}_\rho$ satisfies the monotonicity property, i.e. if $\left(m, \bar{X}_1\right) \in \mathcal{A}_\rho$ and for $\bar{X}_2 \in \bar{\mathcal{X}}$ we have $\bar{X}_1 \succcurlyeq \bar{X}_2$, then also $\left(m, \bar{X}_2\right) \in \mathcal{A}_\rho$.

(d) $\mathcal{A}_\rho$ satisfies the epigraph property, i.e. if $\left(m_1, \bar{X}\right) \in \mathcal{A}_\rho$ and for $m_2 \in \mathbb{R}$ we have $m_2 \geq m_1$, then also $\left(m_2, \bar{X}\right) \in \mathcal{A}_\rho$.

(e) $\mathcal{A}_\rho$ is convex.

(f) If ρ satisfies (S3), then $\mathcal{A}_\rho$ is a cone .

(g) If ρ satisfies (S8) on $\mathcal{B}$, then for all $\bar{Y} \in \mathcal{B} \subseteq \mathcal{C}(\mathbb{R}) \cap \bar{\mathcal{X}}$ and all $m \in \mathbb{R}$

$$\left(m - \sum_{i=1}^{n} Y_i, \bar{X}\right) \in \mathcal{A}_\rho \Leftrightarrow \left(m, \bar{X} + \bar{Y}\right) \in \mathcal{A}_\rho.$$

Proof. For (i) we have

$$\begin{aligned}\inf\left\{m \in \mathbb{R} \mid \left(m, \bar{X}\right) \in \mathcal{A}_\rho\right\} &= \inf\left\{m \in \mathbb{R} \mid m \geq \rho\left(\bar{X}\right)\right\} \\ &= \rho\left(\bar{X}\right).\end{aligned}$$

Since ρ is assumed to be proper, there exists some $\bar{X} \in \bar{\mathcal{X}}$ such that $\rho\left(\bar{X}\right) < \infty$.

But this implies $\left(\rho\left(\bar{X}\right), \bar{X}\right) \in \mathcal{A}_\rho$ and therefore (ii) (a) holds. (ii) (b) is again a direct consequence of the properness. To see (ii) (c) let $\left(m, \bar{X}_1\right) \in \mathcal{A}_\rho$ and $\bar{X}_2 \in \bar{\mathcal{X}}$ with $\bar{X}_1 \succcurlyeq \bar{X}_2$. Since ρ is monotone (S1) we also have $\rho\left(\bar{X}_1\right) \geq \rho\left(\bar{X}_2\right)$ and therefore $m \geq \rho\left(\bar{X}_1\right) \geq \rho\left(\bar{X}_2\right)$ which yields $\left(m, \bar{X}_2\right) \in \mathcal{A}_\rho$. Obviously, the epigraph of a mapping satisfies the epigraph property. So, (ii) (d) follows directly from this fact. Now, let $\left(m_1, \bar{X}_1\right), \left(m_2, \bar{X}_2\right) \in \mathcal{A}_\rho$. For all $\alpha \in [0, 1]$, the outcome convexity (S2a) of ρ implies

$$\begin{aligned}\rho\Big(\alpha \bar{X}_1 + (1-\alpha)\bar{X}_2\Big) &\leq \alpha\rho\left(\bar{X}_1\right) + (1-\alpha)\rho\left(\bar{X}_2\right)\\ &\leq \alpha m_1 + (1-\alpha) m_2,\end{aligned}$$

which proves (ii) (e). In a similar fashion, suppose that ρ is positively homogeneous (S3) and let $\left(m, \bar{X}\right) \in \mathcal{A}_\rho$. For all $\alpha \in \mathbb{R}_+$ we have

$$\begin{aligned}\rho\left(\alpha \bar{X}\right) &= \alpha\rho\left(\bar{X}\right)\\ &\leq \alpha m,\end{aligned}$$

which proves (ii) (f). Finally, suppose that ρ satisfies the translation property (S8) on $\mathcal{B}$ and let $\left(\sum_{i=1}^n Y_i, \bar{X}\right) \in \mathcal{A}_\rho$ for some $\bar{Y} \in \mathcal{B}$. We have

$$\begin{aligned} & \left(m - \sum_{i=1}^n Y_i, \bar{X}\right) \in \mathcal{A}_\rho\\ \Leftrightarrow\quad & m - \sum_{i=1}^n Y_i \geq \rho\left(\bar{X}\right)\\ \Leftrightarrow\quad & m \geq \rho\left(\bar{X}\right) + \sum_{i=1}^n Y_i\\ \Leftrightarrow\quad & m \geq \rho\left(\bar{X} + \bar{Y}\right)\\ \Leftrightarrow\quad & \left(m, \bar{X} + \bar{Y}\right) \in \mathcal{A}_\rho.\end{aligned}$$

□

We know that a mapping is convex if and only if its epigraph is a convex set. As seen above, besides convexity, the epigraph also inherits other properties. It is then natural to ask which properties a given set $\mathcal{A}$ has to fulfill, in order to obtain a systemic risk measure via

$$\rho_{\mathcal{A}} = \inf\left\{m \in \mathbb{R} \mid \left(m, \bar{X}\right) \in \mathcal{A}\right\}. \tag{2.5}$$

Proposition 2.6. *Suppose that* $\emptyset \neq \mathcal{A} \subset \mathbb{R} \times \bar{\mathcal{X}}$ *satisfies the monotonicity and epigraph property. Additionally, for all* $\bar{X} \in \bar{\mathcal{X}}$, *we have* $\inf\left\{m \in \mathbb{R} \mid \left(m, \bar{X}\right) \in \mathcal{A}\right\} > -\infty$. *Then for* $\rho_{\mathcal{A}} \colon \bar{\mathcal{X}} \to \mathbb{R} \cup \{-\infty, \infty\}$ *given by 2.5 the following holds:*
(a) $\rho_{\mathcal{A}}$ *is proper.*
(b) $\rho_{\mathcal{A}}$ *satisfies (S1).*
(c) If $\mathcal{A}$ *is a convex set, then* $\rho_{\mathcal{A}}$ *satisfies (S2a).*
(d) If $\mathcal{A}$ *is a cone, then* $\rho_{\mathcal{A}}$ *satisfies (S3).*
(e) If for all $\bar{Y} \in \mathcal{B} \subseteq \mathcal{C}(\mathbb{R}) \cap \bar{\mathcal{X}}$ *and all* $m \in \mathbb{R}$

$$\left(m - \sum_{i=1}^{n} Y_i, \bar{X}\right) \in \mathcal{A} \Leftrightarrow \left(m, \bar{X} + \bar{Y}\right) \in \mathcal{A},$$

then $\rho_{\mathcal{A}}$ *satisfies (S8) on* $\mathcal{B}$.
(f) $\mathcal{A}_{\rho_{\mathcal{A}}} = \mathcal{A}$ *if and only if for all* $\bar{X} \in dom(\rho_{\mathcal{A}})$ *we have* $\left\{m \in \mathbb{R} \mid \left\{m, \bar{X}\right\} \in \mathcal{A}\right\} = \left[\rho_{\mathcal{A}}\left(\bar{X}\right), \infty\right)$.

Proof. By the definition of the set $\mathcal{A}$, we directly obtain $\rho_{\mathcal{A}}\left(\bar{X}\right) > -\infty$ for all $\bar{X} \in \bar{\mathcal{X}}$. Additionally, since $\mathcal{A} \neq \emptyset$ is assumed there is some $\left(m, \bar{X}\right) \in \mathcal{A}$. But this means $\rho_{\mathcal{A}}\left(\bar{X}\right) \leq m < \infty$ which proves statement (a) that $\rho_{\mathcal{A}}$ is proper. Now, since $\mathcal{A}$ satisfies the monotonicity property, for $\bar{X}, \bar{Y} \in \bar{\mathcal{X}}$ with $\bar{X} \succcurlyeq \bar{Y}$, we have $\left\{m \in \mathbb{R} \mid \left(m, \bar{X}\right) \in \mathcal{A}\right\} \subseteq \left\{m \in \mathbb{R} \mid \left(m, \bar{Y}\right) \in \mathcal{A}\right\}$ and therefore

$$\begin{aligned}
\rho_{\mathcal{A}}\left(\bar{X}\right) &= \inf\left\{m \in \mathbb{R} \mid \left(m, \bar{X}\right) \in \mathcal{A}\right\} \\
&\geq \inf\left\{m \in \mathbb{R} \mid \left(m, \bar{Y}\right) \in \mathcal{A}\right\} \\
&= \rho_{\mathcal{A}}\left(\bar{Y}\right).
\end{aligned}$$

This proves (b). For (c), the set $\mathcal{A}$ is assumed to be convex. But this means for all $\alpha \in [0,1]$ and all $\bar{X}, \bar{Y} \in \bar{\mathcal{X}}$

$$\left\{\alpha m_1 + (1-\alpha) m_2 \in \mathbb{R} \mid \left(m_1, \bar{X}\right), \left(m_2, \bar{Y}\right) \in \mathcal{A}\right\} \subseteq \left\{m \in \mathbb{R} \mid \left(m, \alpha\bar{X} + (1-\alpha)\bar{Y}\right) \in \mathcal{A}\right\}.$$

Therefore we have

$$\begin{aligned}
\rho_{\mathcal{A}}\left(\alpha\bar{X} + (1-\alpha)\bar{Y}\right) &= \inf\left\{m \in \mathbb{R} \mid \left(m, \alpha\bar{X} + (1-\alpha)\bar{Y}\right) \in \mathcal{A}\right\} \\
&\leq \inf\left\{\alpha m_1 + (1-\alpha) m_2 \in \mathbb{R} \mid \left(m_1, \bar{X}\right), \left(m_2, \bar{Y}\right) \in \mathcal{A}\right\} \\
&= \alpha\rho_{\mathcal{A}}\left(\bar{X}\right) + (1-\alpha)\rho_{\mathcal{A}}\left(\bar{Y}\right).
\end{aligned}$$

Now, let $\mathcal{A}$ be a cone. This yields for all $\lambda \in \mathbb{R}_+$ and all $\bar{X} \in \bar{\mathcal{X}}$ that

$$\left\{m \in \mathbb{R} \mid \lambda\left(m, \bar{X}\right) \in \mathcal{A}\right\} = \left\{m \in \mathbb{R} \mid \left(m, \bar{X}\right) \in \mathcal{A}\right\}.$$

With this identity, we obtain

$$\begin{aligned}
\lambda\rho_{\mathcal{A}}\left(\bar{X}\right) &= \lambda \inf\left\{m \in \mathbb{R} \mid \left(m, \bar{X}\right) \in \mathcal{A}\right\} \\
&= \inf\left\{\lambda m \in \mathbb{R} \mid \left(m, \bar{X}\right) \in \mathcal{A}\right\} \\
&= \inf\left\{m \in \mathbb{R} \mid \left(\frac{1}{\lambda}m, \bar{X}\right) \in \mathcal{A}\right\} \\
&= \inf\left\{m \in \mathbb{R} \mid \frac{1}{\lambda}\left(m, \lambda\bar{X}\right) \in \mathcal{A}\right\} \\
&= \inf\left\{m \in \mathbb{R} \mid \left(m, \lambda\bar{X}\right) \in \mathcal{A}\right\} \\
&= \rho_{\mathcal{A}}\left(\lambda\bar{X}\right).
\end{aligned}$$

This proves (d). For (e), we have

$$\begin{aligned}
\rho_{\mathcal{A}}\left(\bar{X} + \bar{Y}\right) &= \inf\left\{m \in \mathbb{R} \mid \left(m, \bar{X} + \bar{Y}\right) \in \mathcal{A}\right\} \\
&= \inf\left\{m \in \mathbb{R} \mid \left(m - \sum_{i=1}^{n} Y_i, \bar{X}\right) \in \mathcal{A}\right\} \\
&= \inf\left\{z + \sum_{i=1}^{n} Y_i \in \mathbb{R} \mid \left(z, \bar{X}\right) \in \mathcal{A}\right\} \\
&= \inf\left\{z \in \mathbb{R} \mid \left(z, \bar{X}\right) \in \mathcal{A}\right\} + \sum_{i=1}^{n} Y_i \\
&= \rho_{\mathcal{A}}\left(\bar{X}\right) + \sum_{i=1}^{n} Y_i.
\end{aligned}$$

Finally (f) is obtained by the observation that for each proper mapping $f \colon \bar{\mathcal{X}} \to \mathbb{R} \cup \{\infty\}$, we have

$$\mathcal{A}_f = \bigcup_{\bar{X} \in \mathrm{dom}(f)} \left[f\left(\bar{X}\right), \infty\right) \times \left\{\bar{X}\right\}.$$

Note that by the definition of $\rho_{\mathcal{A}}$, for $\bar{X} \notin \mathrm{dom}\left(\rho_{\mathcal{A}}\right)$ the set $\left\{m \in \mathbb{R} \mid \left(m, \bar{X}\right) \in \mathcal{A}\right\}$ is empty. □

Remark 2.7. *In the literature, there are many authors that present similar results to the ones in Proposition 2.3 and Proposition 2.6. See for example [FS08] Proposition 4.6 and 4.7 and [KR09] Section 1. The results are presented for single-firm risk measures, but the proofs apply directly with some slight changes. However, one needs to be careful which exact set of axioms is assumed to be fulfilled by* ρ *or* $\mathcal{A}$

respectively. For example, in [FS08] the axioms for $\mathcal{A}$ are chosen in a way that $\rho_{\mathcal{A}}$ has the translation property (S8) on $\mathbb{R}$ (or $\mathbb{R}^n$ respectively) by definition. The remainder of the proof heavily relies on this fact. However, in our situation the proof does not rely on the translation property (S8).

The second way of representing systemic risk measures is called *dual* representation. Under additional continuity assumptions, namely *lower semicontinuity (l.s.c.)* (see Definition A.19), the famous Fenchel-Moreau Theorem A.32 applies. It provides the possibility to describe the systemic risk measure through objects from the topological dual space of $\bar{\mathcal{X}}$. We use the notation $e_i\mathbb{R} := \{xe_i \mid x \in \mathbb{R}\}$.

Theorem 2.8. *Suppose that $\rho\colon \bar{\mathcal{X}} \to \mathbb{R} \cup \{\infty\}$ is a systemic risk measure which is l.s.c.. Then ρ admits the representation*

$$\rho\left(\bar{X}\right) = \sup_{\Xi \in \mathcal{D}'} \left\{\langle \bar{X}, \Xi\rangle_n - \alpha^\rho(\Xi)\right\} \qquad \forall \bar{X} \in \bar{\mathcal{X}}, \tag{2.9}$$

where $\alpha^\rho\colon \bar{\mathcal{X}}' \to \mathbb{R} \cup \{\infty\}$ is the (convex) conjugate of ρ given by

$$\alpha^\rho(\Xi) = \rho^*(\Xi) := \sup_{\bar{X} \in \bar{\mathcal{X}}} \left\{\langle \bar{X}, \Xi\rangle_n - \rho\left(\bar{X}\right)\right\} \tag{2.10}$$

and $\mathcal{D}' = dom(\alpha^\rho) := \left\{\Xi \in \bar{\mathcal{X}}' \mid \alpha^\rho(\Xi) < \infty\right\}$. Moreover, for the set $\mathcal{D}'$ the following holds true:

(i) $\mathcal{D}' \subseteq \bar{\mathcal{X}}'_+$.
(ii) ρ satisfies (S8) on $e_i\mathbb{R}$ if and only if $\langle e_i, \Xi\rangle_n = \langle 1, \xi_i\rangle = 1$, for all $\Xi \in \mathcal{D}'$.
(iii) ρ satisfies (S3) if and only if ρ admits the representation

$$\rho\left(\bar{X}\right) = \sup_{\Xi \in \mathcal{D}'} \left\{\langle \bar{X}, \Xi\rangle_n\right\} \qquad \forall \bar{X} \in \bar{\mathcal{X}}. \tag{2.11}$$

Proof. See for example [RS06] Theorem 2.2. Their proof is formulated for single-firm risk measures. However, only the structural properties of the domain of the considered mappings are relevant. Indeed, the spaces they consider and the spaces in our setup share all relevant properties. To maintain the self contained character, let us present the proof in full detail. The representation (2.9) is, as stated above, a direct consequence of the Fenchel-Moreau Theorem A.32. To prove (i), consider some $\Xi \notin \bar{\mathcal{X}}'_+$. By Assumption (C), there exists some $\bar{Z} \in \left(\bar{\mathcal{X}}\right)_+$ with $\langle \bar{Z}, \Xi\rangle_n < 0$. Now, for all $\bar{X} \in \mathrm{dom}(\rho)$ and all $t \in \mathbb{R}_+$, we can define $\bar{X}_t := \bar{X} - t\bar{Z}$. Now we have

$\bar{X} \succcurlyeq \bar{X}_t$ and hence $\rho\left(\bar{X}\right) \geq \rho\left(\bar{X}_t\right)$. But this means

$$\begin{aligned}\alpha^\rho(\Xi) &\geq \sup_{t\in\mathbb{R}_+}\left\{\langle\bar{X}_t,\Xi\rangle_n - \rho\left(\bar{X}_t\right)\right\}\\ &\geq \sup_{t\in\mathbb{R}_+}\left\{\langle\bar{X},\Xi\rangle_n - t\langle\bar{Z},\Xi\rangle_n - \rho\left(\bar{X}\right)\right\}\\ &= \infty.\end{aligned}$$

Now suppose that ρ satisfies (S8) on $e_i\mathbb{R}$. Then for $\bar{X} \in \operatorname{dom}(\rho)$, it holds that

$$\begin{aligned}\alpha^\rho(\Xi) &\geq \sup_{a\in\mathbb{R}}\left\{\langle\bar{X}+ae_i,\Xi\rangle_n - \rho\left(\bar{X}+ae_i\right)\right\}\\ &= \sup_{a\in\mathbb{R}}\left\{\langle\bar{X},\Xi\rangle_n + a\left(\langle 1,\xi_i\rangle - 1\right) - \rho\left(\bar{X}\right)\right\}.\end{aligned}$$

But this means $\alpha^\rho(\Xi) = \infty$ whenever $\langle 1,\xi_i\rangle \neq 1$. Conversly if $\langle 1,\xi_i\rangle = 1$, we have $\langle\bar{X}+ae_i,\Xi\rangle_n = \langle\bar{X},\Xi\rangle_n + a$ for all $a \in \mathbb{R}$. Since the supremum is translation invariant, (2.9) yields that ρ satisfies (S8). This means (ii) holds true. For (iii), we first observe that a mapping given by (2.11) satisfies (S3). Finally, suppose that ρ satisfies (S3). Since $\bar{\mathcal{X}}$ is a vector space, we have for all $t \in \mathbb{R}_+$

$$\begin{aligned}\alpha^\rho(\Xi) &= \sup_{\bar{X}\in\bar{\mathcal{X}}}\left\{\langle\bar{X},\Xi\rangle_n - \rho\left(\bar{X}\right)\right\}\\ &= t\sup_{\bar{X}\in\bar{\mathcal{X}}}\left\{\langle\frac{1}{t}\bar{X},\Xi\rangle_n - \rho\left(\frac{1}{t}\bar{X}\right)\right\}\\ &= t\alpha^\rho(\Xi).\end{aligned}$$

This implies $\alpha^\rho \in \{0,\infty\}$ and (2.9) reduces to (2.11). □

Theorem 2.8 provides an elegant way to represent systemic risk measures via dual variables. However, to guarantee a representation, one needs to verify the l.s.c.. In general, verifying continuity assumptions poses a challenging problem. If we reduce our attention to locally convex Fréchet lattices, any finite systemic risk measure is continuous and sub-differentiable on the interior of its domain and hence admits the representation (2.9). This result is an extension of the Namioka-Klee Theorem and was studied in detail in [BF09] (see also Theorems A.38 and A.47). For finite systemic risk measures this result ensures a dual representation. Note, however, that for non finite systemic risk measures additional assumptions are needed, in order guarantee the dual representation. But even if we know that there is a dual representation, the question which arises subsequently, is: How can we interpret the objects in $\mathcal{D}'$?

Measuring risk targets the problems caused by the uncertainty of the future. A probability measure is the suitable object to model the uncertainty. Indeed if the objects in $\mathcal{D}'$ are vectors consisting of probability measures (or more precisely probability densities) the economic interpretation of the dual representation (2.9) would be reasonable: The set $\mathcal{D}'$ consists of probabilistic models of the future. To compute the systemic risk, we first add up the expected losses in each model. The choice of a model comes with a *penalization* by α^ρ. The risk is now given as the worst *prediction* of the future.

Example 2.12. *Let us consider the case n=1, i.e. the system consists of a single firm. In this case, we use the term single-firm risk measure. We will have a closer look at this concept in Section 2.3. Now, a typical example for the space $\mathcal{X}$ is L^∞. A single-firm risk measure which also satisfies (S8) takes only finite values on L^∞, is Lipschitz continuous and hence admits a representation via* (2.9) *(see [FS08] Section 4). We also know that only positive and normalized elements of $(L^\infty)'$ are relevant to consider. However, the set $\mathcal{D}'$ still contains objects which are not in $\mathcal{M}_1$, namely the finitely additive set functions. To overcome this issue, $\sigma(L^\infty, L^1)$-l.s.c is needed. In the case L^∞, this is implied by the so called Fatou property. First introduced by [Del02] for single-firm risk measures on L^∞, it also appears to be the right instrument for convex single-firm risk measures on L^p to guarantee l.s.c on the whole space (even in the non finite case). It requires that for a given dominated sequence $(X_n)_{n\in\mathbb{N}}$ in the underlying L^p-space, $p \in [1,\infty]$, which converges $\mathbb{P}$-a.s. to X, we have*

$$\rho_0(X) \leq \liminf_{n\to\infty} \rho_0(X_n).$$

For convex monetary risk measures on L^p, $p \in [1,\infty]$, the Fatou property provides the preferred representation. In a more general setting, i.e. if the space $\mathcal{X}$ is a (or all the spaces $\mathcal{X}_i$ are) locally convex Fréchet lattice, this property translates to order l.s.c.. First envisoned by [BF09], under addtional assumptions one can ignore the singular objects in $\mathcal{D}'$ for the dual representation. In Section A.4 this situation is studied in more detail.

2.2. Differentiability and Optimal Solutions

Once the dual problem (2.9) is formulated, the natural question is: Are there optimal solutions, i.e. are there objects $\Xi \in \bar{\mathcal{X}}'$ for which the supremum is actually attained. There is a connection between optimal solutions to the dual problem (2.9) (even to their existence) and some notion of differentiability. The right notion in our context is the one of directional derivatives and Gâteaux differentiability (see A.5 for the respective definitions and more details on this concept).

Theorem 2.13. *Suppose that $\rho\colon \bar{\mathcal{X}} \to \mathbb{R} \cup \{\infty\}$ is a systemic risk measure which is l.s.c. and for some $\bar{X} \in dom(\rho)$, ρ is Gâteaux differentiable at $\bar{X}$ with derivative $\nabla\rho\left(\bar{X}\right)$. Then $\nabla\rho\left(\bar{X}\right)$ is an optimal solution to the dual problem* (2.9) *of ρ at $\bar{X}$,* *i.e. the supremum in* (2.9) *is attained at* $\nabla\rho\left(\bar{X}\right)$

$$\rho\left(\bar{X}\right) = \langle \bar{X}, \nabla\rho\left(\bar{X}\right)\rangle_n - \alpha^{\rho}\left(\nabla\rho\left(\bar{X}\right)\right).$$

Proof. If ρ is Gâteaux differentiable at $\bar{X} \in \operatorname{int}(\operatorname{dom}(\rho))$ with derivative $\nabla\rho\left(\bar{X}\right)$, we can rearrange (A.44) in the following way:

$$\begin{aligned} & \langle \bar{V}, \nabla\rho\left(\bar{X}\right)\rangle_n \le \rho\left(\bar{X}+\bar{V}\right) - \rho\left(\bar{X}\right) \quad \forall \bar{V} \in \bar{\mathcal{X}} \\ \Leftrightarrow \quad & \langle \bar{V}-\bar{X}, \nabla\rho\left(\bar{X}\right)\rangle_n \le \rho\left(\bar{V}\right) - \rho\left(\bar{X}\right) \quad \forall \bar{V} \in \bar{\mathcal{X}} \\ \Leftrightarrow \quad & \rho\left(\bar{X}\right) \le \langle \bar{X}, \nabla\rho\left(\bar{X}\right)\rangle_n - \rho^{*}\left(\nabla\rho\left(\bar{X}\right)\right). \end{aligned}$$

Since $\alpha^{\rho} := \rho^{*}$, Theorem 2.8 yields the claim. □

Even if ρ is not Gâteaux differentiable at some $\bar{X} \in \operatorname{int}(\operatorname{dom}(\rho))$ with derivative $\nabla\rho\left(\bar{X}\right)$, the last inequality in the previous proof presents a necessary and sufficient condition for an arbitrary $\Xi \in \bar{\mathcal{X}}'$ to be an optimal solution to the dual problem (2.9). This means that optimal solutions are elements of the subgradient $\partial\rho\left(\bar{X}\right)$ (see Section A.5). The following Corollary collects some more connections of optimal solutions and notions of differentiablity for systemic risk measures.

Corollary 2.14. *Suppose that* $\rho\colon \bar{\mathcal{X}} \to \mathbb{R}\cup\{\infty\}$ *is a systemic risk measure which is l.s.c. and fix* $\bar{X} \in \bar{\mathcal{X}}$*. Then the following holds:*

(i) There exists an optimal solution to the dual problem (2.9) *at* $\bar{X}$ *if and only if* $\partial\rho\left(\bar{X}\right) \neq \emptyset$*. Moreover, the solution is unique if and only if* $|\partial\rho\left(\bar{X}\right)| = 1$*. In these situations, we have*

$$\rho\left(\bar{X}\right) = \langle \bar{X}, \Xi\rangle_n - \alpha^{\rho}(\Xi), \qquad \Xi \in \partial\rho\left(\bar{X}\right).$$

(ii) If $\bar{X} \in \operatorname{int}(\operatorname{dom}(\rho))$ *and* ρ *is Gâteaux differentiable at* $\bar{X}$ *with derivative* $\nabla\rho\left(\bar{X}\right)$*, then* $\partial\rho\left(\bar{X}\right) = \left\{\nabla\rho\left(\bar{X}\right)\right\}$*.*

If we can additionally assume that $\bar{X} \in \operatorname{dom}(\rho)$ *and* ρ *is continuous at* $\bar{X}$*, we have:*

(iii) $\partial\rho\left(\bar{X}\right) \neq \emptyset$*,* $\delta_{+}\rho\left(\bar{X},\cdot\right)$ *is continuous and*

$$\delta_{+}\rho\left(\bar{X},\cdot\right) = \max_{\Xi\in\partial\rho\left(\bar{X}\right)} \left\{\langle \cdot, \Xi\rangle_n\right\}.$$

(iv) ρ *is Gâteaux differentiable at* $\bar{X}$ *with derivative* $\nabla\rho\left(\bar{X}\right)$ *if and only if* $\partial\rho\left(\bar{X}\right) = \left\{\nabla\rho\left(\bar{X}\right)\right\}$*.*

Proof. (i) follows directly from the definition of the subdifferential. In Theorem 2.13, we have already seen that $\nabla\rho\left(\bar{X}\right) \in \partial\rho\left(\bar{X}\right)$. From A.46, it follows now that it is indeed the only element of $\partial\rho\left(\bar{X}\right)$, which shows (ii). Part (iii) and (iv) are also direct consequences of A.46. □

2.3. Aggregation Rules and Single-Firm Risk Measures

Systemic risk measures of type (0.1) and (0.3) fit in the general framework presented in the previous section. Since these systemic risk measures come in with more structural properties, it is possible to refine the representation results. In the following subsections we will briefly review some of the fundamental results for systemic risk measures of type (0.1) and (0.3).

At first stage, systemic risk measures of type (0.1) and (0.3) have in common that they use a general *aggregation rule* to aggregate the system into a univariate risk factor. The risk of this univariate factor is then measured with a *single-firm risk measure*. Let us start by defining a general aggregation rule.

Definition 2.15. *Consider the following properties for a mapping* $\Lambda\colon \mathbb{R}^n \to \mathbb{R}$*:*

(A1) Monotonicity: For all $\bar{x}, \bar{y} \in \mathbb{R}^n$ *with* $\bar{x} \geq \bar{y}$*,* Λ *satisfies* $\Lambda(\bar{x}) \geq \Lambda(\bar{y})$*.*

(A2) Convexity: For all $\bar{x}, \bar{y} \in \mathbb{R}^n$ *and all* $\alpha \in [0,1]$*,* Λ *satisfies* $\Lambda(\alpha\bar{x} + (1-\alpha)\bar{y}) \leq \alpha\Lambda(\bar{x}) + (1-\alpha)\Lambda(\bar{y})$*.*

(A3) $\mathcal{R}$*-Surjectivity:* $\Lambda(\mathbb{R}^n) = \mathcal{R}$ *for* $\mathcal{R} \in \{\mathbb{R}, \mathbb{R}_+\}$*.*

(A4) Positive homogeneity: For all $\bar{x}, \bar{y} \in \mathbb{R}^n$ *and all* $\alpha \in \mathbb{R}_+$*,* Λ *satisfies* $\Lambda(\alpha\bar{x}) = \alpha\Lambda(\bar{x})$*.*

(A5) Normalization: $\Lambda(1_n) = n$*.*

If a mapping satisfies (A1)- (A3), it is called convex aggregation rule. If it additionally satisfies the property (A4), it is called positively homogeneous aggregation rule.

The interpretation of these properties is straight forward except for (A3). Let us give a motivation. A mapping from $\mathbb{R}^n$ to $\mathbb{R}$ with property (A2) is continuous (see [Roc97] Corollary 10.1.1). In addition if the mapping is not constant, it is not bounded from above. However, a lower bound may exist. So the $\mathcal{R}$-Surjectivity pays tribute to this fact. Of course, other lower bounds than 0 are possible, but there is no loss of generality by our restriction. The aggregation rule is a nice instrument to include the strucutre of the system, while measruing its risk. Let us continue with some examples.

Example 2.16. *(i) In [KOZ16], the authors already figured out that, obviously, simple summation*

$$\Lambda^{sum}(\bar{x}) = \sum_{i=1}^{n} x_i \tag{2.17}$$

is an aggregation rule with all properties form Definition 2.15. If a shift with some parameter $c \in \mathbb{R}$ is allowed, we obtain

$$\Lambda^{sum,c}(\bar{x}) = \sum_{i=1}^{n} x_i - c. \tag{2.18}$$

These choices are only reasonable if the participants cross-subsidize each other. An easy way to avoid cross-subsidization is to add up the losses only, i.e.

$$\Lambda^{loss}(\bar{x}) = \sum_{i=1}^{n} x_i^+. \tag{2.19}$$

A simple modification occurs if we again allow a shift. Then, only losses beyond a given threshold level $b > 0$ are considered. The corresponding aggregation rule is then given by

$$\Lambda^{loss,b}(\bar{x}) = \sum_{i=1}^{n} (x_i - b)^+, \qquad b \in \mathbb{R}_+. \tag{2.20}$$

(ii) In the setting of [BFFMB20], the main analysis is restricted to systemic risk measure which use the following aggregation:

$$\Lambda^{ut}(\bar{x}) = \sum_{i=1}^{n} l_i(x_i), \tag{2.21}$$

where each $l_i : \mathbb{R} \to \mathbb{R}$ is a loss function, i.e. an increasing and convex function which also satisfies $\lim_{x\to\infty} \frac{l_i(x)}{x} = \infty$. If we choose the exponential loss functions $l_i(x) = \frac{1}{\alpha_i} \exp(\alpha_i x)$, $\alpha_i > 0$, the corresponding mapping Λ^{exut} is indeed a convex aggregation rule.

(iii) [CIM13] and [KOZ16] also presented examples with more sensitivity to the structure. One aggregation rule is motivated by the famous structural contagion model of [EN01]. In this setup, the liabilities of firm i to firm j are captured within the so called relative liabilities matrix $\Pi = (\Pi_{ij})_{i,j=1,\dots,n}$, where Π_{ij} represents the proportion of the total liabilities of firm i received by firm j. Moreover we assume that an external regulator has the possibility to inject capital in the system. If the vector $\bar{x} \in \mathbb{R}^n$ captures the realized losses of all

participants of the system, then each of them has two possibilities to cover these losses: One way is to receive money from the regulator and the other way is to reduce the payments to other firms by y_i. If firm i decides to reduce their payments to the other firms, then firm j faces new losses $\Pi_{ij} y_i$.

$$\Lambda^{CM}(\bar{x}) = \min_{\substack{b_i + y_i \geq x_i + \sum_{j=1}^{n} \Pi_{ji} y_j \\ \forall i=1,\ldots,n, \bar{b}, \bar{y} \in \mathbb{R}_+^n}} \left\{ \sum_{i=1}^{n} (y_i + \gamma b_i) \right\}, \qquad \gamma > 1. \tag{2.22}$$

(iv) *Suppose that there are some nodes, classified in a set $A \subset \{1, \ldots, n\}$, which would be dangerous for the stability of the system in case of a default. Then the aggregation rule*

$$\Lambda^{crit}(\bar{x}) = \exp\left(\gamma \sum_{i \in A} x_i^+\right) - 1 + \sum_{i \in N \setminus A} x_i^+, \qquad \gamma > 0. \tag{2.23}$$

penalizes large losses of the critical nodes on an exponential scale. Since for small x we have $e^x - 1 \approx x$, small losses of the critical nodes are treated the same way as the losses of the less relevant nodes. This aggregation rule is also due to [KOZ16].

Now, let us move on to single-firm risk measures. Roughly speaking, a single-firm risk measure is simply a systemic risk measure, where the system consists of only one single participant, i.e. $n = 1$. To guarantee a clear distinction in our analysis, we give a separate definition.

Definition 2.24. *Consider the following properties for a proper mapping $\rho_0 \colon \mathcal{X} \to \mathbb{R} \cup \{\infty\}$:*

(R1) *Monotonicity: For all $X, Y \in \mathcal{X}$ with $X \succcurlyeq Y$, ρ_0 satisfies $\rho_0(X) \geq \rho_0(Y)$.*

(R2) *Convexity: For all $X, Y \in \mathcal{X}$ and all $\alpha \in [0,1]$, ρ_0 satisfies $\rho_0(\alpha X + (1-\alpha) Y) \leq \alpha \rho_0(X) + (1-\alpha)\rho_0(Y)$.*

(R3) *Positive homogeneity: For all $X \in \mathcal{X}$ and all $\alpha \in \mathbb{R}_+$, ρ_0 satisfies $\rho_0(\alpha X) = \alpha \rho_0(X)$.*

(R4) *Constancy on $\mathcal{R} \subset \mathbb{R}$: For all $m \in \mathcal{R} \subset \mathbb{R}$, ρ_0 satisfies $\rho_0(m) = m$.*

(R5) *Subadditivity: For all $X, Y \in \mathcal{X}$, ρ_0 satisfies $\rho(X+Y) \leq \rho(X) + \rho(Y)$.*

(R6) *Translation property: For all $X \in \mathcal{X}$ and all $m \in \mathbb{R}$, ρ_0 satisfies $\rho_0(X+m) = \rho_0(X) + m$.*

If a mapping satisfies the property (R1) and at least one of the other properties we will use the generic term single-firm risk measure. More specific, a convex single-firm risk measure is a mapping which satisfies (R1) and (R2). A positively homogeneous single-firm risk measure is a convex single-firm risk measure that additionally satisfies the property (R3). A coherent single-firm risk measure is a positively homogeneous single-firm risk measure that additionally satisfies the property (R6).

The interpretation of these properties is exactly the same as given earlier. Compared to the general systemic case with $n \geq 2$, the preference consistency (S4) follows directly from the monotonicity (R1). Normalization (S6) translates to constancy on $\{1\}$ and at least $\mathbb{R}$-surjectivity (S5) can be explained through constancy on whole $\mathbb{R}$. The only non standard property is (R4). This property is of technical nature and plays an important role for the decomposition Theorem 2.33 presented in the next section. It was introduced by [FG02]. In the single-firm risk literature, single-firm risk measures with the translation property (R6) are often called monetary risk measures. This label pays tribute to the already mentioned interpretation as a capital requirement.

Remark 2.25. *Constancy on whole $\mathbb{R}$ follows from the translation property (R6) together with the property $\rho_0(0) = 0$. For l.s.c. convex single-firm risk measures with $\rho_0(0) = 0$, the translation property (R6) is indeed equivalent to constancy on whole $\mathbb{R}$.*

Let us continue with some examples of single-firm risk measures.

Example 2.26. *(i) One of the most popular single-firm risk measures is Value at Risk. It is defined as the $(1-\alpha)$ quantile (for small α) of the loss variable X, i.e.*

$$V@R_{1-\alpha}(X) = F_X^{-1}(1-\alpha) = \inf\{m \in \mathbb{R} \mid F_X(m) \geq 1-\alpha\}. \tag{2.27}$$

Since it only penalizes the losses, it yields a better fit for measuring risk as variance. It is easy to see that it is a positively homogeneous monetary risk measure. Unfortunately, Value at Risk in general does not reward diversification if it is defined on the whole space L^0. This has to be seen as a major drawback. Additionally, it is not able to detect extreme events in the upper tale.

(ii) A possible way to overcome the drawbacks of Value at Risk is to consider an average over all the values above a certain level $(1-\alpha)$, i.e.

$$AV@R_{1-\alpha}(X) = \frac{1}{\alpha}\int_{1-\alpha}^{1} V@R_\gamma(X)\, d\gamma. \tag{2.28}$$

This single-firm risk measure is called Average Value at Risk or Expected Shortfall. Indeed, it is a coherent single-firm risk measure. For continuously distributed X, Average Value at Risk can be described via

$$\begin{aligned} AV@R_{1-\alpha}(X) &= \mathbb{E}\left[X \mid X > V@R_{1-\alpha}(X)\right] \\ &= V@R_{1-\alpha}(X) + \mathbb{E}\left[X - V@R_{1-\alpha}(X) \mid X > V@R_{1-\alpha}(X)\right]. \end{aligned}$$

Obviously, it captures extrem events in the upper tails and is therefore an upgrade to Value at Risk. Since Average Value at Risk is a coherent single-firm risk measure, it also rewards diversification and it admits a dual representation given by

$$AV@R_{1-\alpha}(X) = \sup_{\mathbb{Q}\in\mathcal{M}_1^{\alpha}(\mathbb{P})} \left\{\mathbb{E}_{\mathbb{Q}}[X]\right\}, \tag{2.29}$$

where

$$\mathcal{M}_1^{\alpha}(\mathbb{P}) := \left\{\mathbb{Q} \in \mathcal{M}_1(\mathbb{P}) \mid \frac{d\mathbb{Q}}{d\mathbb{P}} \leq \frac{1}{\alpha}\right\}.$$

(iii) Consider a setting in which the regulator uses the exponential loss function to express his preference. The suitable single firm risk measure ρ_0 in this setup is clearly the entropic risk measure given by

$$\rho_0^{entr}(X) = \frac{1}{\theta}\ln \mathbb{E}\left[\exp(\theta X)\right], \qquad \theta > 0. \tag{2.30}$$

The suitable domain for this risk measure (independent of θ) is the Orlicz heart

$$M^{\phi^{\exp}} = M^{\exp} := \left\{X \in L^0 \mid \mathbb{E}\left[\phi^{\exp}(\alpha X)\right] < \infty \text{ for all } \alpha > 0\right\},$$

with the Young function $\phi^{\exp}(x) := \exp(|x|) - 1$ (see [FW15]). Note that in this example the existence of a base probability measure $\mathbb{P}$ is assumed. The parameter θ reflects the risk aversion of the operator, smaller θ means lower risk aversion. ρ_0^{entr} admits a dual representation via the relative entropy

$$H(\mathbb{Q} \mid \mathbb{P}) = \mathbb{E}_{\mathbb{Q}}\left[\log\left(\frac{d\mathbb{Q}}{d\mathbb{P}}\right)\right],$$

given by

$$\rho_0^{entr}(X) = \sup_{\mathbb{Q}\in\mathcal{M}_1^{\phi^*}(\mathbb{P})} \left\{\mathbb{E}_{\mathbb{Q}}[X] - \frac{1}{\alpha}H(\mathbb{Q} \mid \mathbb{P})\right\}, \tag{2.31}$$

where

$$\mathcal{M}_1^{\phi^*}(\mathbb{P}) \coloneqq \left\{\mathbb{Q} \in \mathcal{M}_1(\mathbb{P}) \mid \frac{d\mathbb{Q}}{d\mathbb{P}} \in L^{\phi^*}\right\}.$$

(iv) The single-firm risk measure, used in [BFFMB20], is given as follows: For some constant $B \in \mathbb{R}$, consider the set

$$\mathcal{A} = \left\{Z \in L^1(\Omega, \mathcal{F}, \mathbb{P}) \mid \mathbb{E}[Z] \leq B\right\}.$$

Now,

$$\rho_0^{\mathcal{A}}(X) = \rho_{\mathcal{A}}(X) = \inf\{m \in \mathbb{R} \mid X - m \in \mathcal{A}\} \tag{2.32}$$

defines a convex single-firm risk measure which can be represented as

$$\rho_0^{\mathcal{A}}(X) = \mathbb{E}[X] - B.$$

If $B = 0$, it is obviously a coherent single-firm risk measure.

2.4. First Aggregate

To formulate the main result which justifies a deeper analysis of systemic risk measures of type (0.1), we need a technical requirement on ρ. Consider a mapping $\rho\colon \bar{\mathcal{X}} \to \mathbb{R} \cup \{\infty\}$. Its restriction to $\mathbb{R}^n$ is the mapping $\rho\mid_{\mathbb{R}^n}\colon \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$. The $\mathcal{R}$-surjectivity (S5) ensures that the restriction is finite and in this situation $\rho\mid_{\mathbb{R}^n}\colon \mathbb{R}^n \to \mathbb{R}$ satisfies (A3). We also have that the properties (S1) and (S2a) of ρ translate to (A1) and (A2) for $\rho\mid_{\mathbb{R}^n}$. A mapping with these three properties is always continuous and hence measurable. Now

$$\rho|_{\mathbb{R}^n}\left(\bar{\mathcal{X}}\right) \coloneqq \left\{\rho|_{\mathbb{R}^n} \circ \bar{X}\colon \Omega \to \mathbb{R} \mid \bar{X} \in \bar{\mathcal{X}}\right\} \subseteq L^0(\Omega, \mathcal{F})$$

always holds true. In the same fashion we think of $\Lambda\left(\bar{\mathcal{X}}\right)$.

Theorem 2.33 ([KOZ16])**.** *Let $\mathcal{X}$ be an arbitrary locally convex Hausdorff space and $\mathcal{C} \subseteq \mathcal{X}$ an arbitrary convex set. Then $\rho\colon \bar{\mathcal{X}} \to \mathbb{R} \cup \{\infty\}$ is a systemic risk measure which additionally satisfies the properties (S2b), (S4), (S5) and $\rho|_{\mathbb{R}^n}\left(\bar{\mathcal{X}}\right) = \mathcal{C}$ if and only if there exists a convex aggregation rule $\Lambda : \mathbb{R}^n \to \mathbb{R}$ which additionally satisfies $\Lambda\left(\bar{\mathcal{X}}\right) = \mathcal{C}$ and a proper convex mapping $\rho_0\colon \mathcal{X} \to \mathbb{R} \cup \{\infty\}$ which is monotone on $\Lambda\left(\bar{\mathcal{X}}\right) = \mathcal{C}$ and satisfies the constancy property (R4) on $\Lambda(\mathbb{R}^n)$ such that ρ is fully described by ρ_0 and Λ for all $\bar{X} \in \bar{\mathcal{X}}$ via*

$$\rho\left(\bar{X}\right) = \rho_0\left(\Lambda\left(\bar{X}\right)\right) \coloneqq \rho_0\left(\Lambda \circ \bar{X}\right). \tag{2.34}$$

In the above situation, we also have the following equivalences:

(i) ρ additionally satisfies (S3) if and only if Λ and ρ_0 additionally satisfy (A4) and (R3), respectively.

(ii) ρ additionally satisfies (S6) if and only if Λ additionally satisfies (A5).

Moreover if $\mathcal{R} = \mathbb{R}_+$, we have $\mathcal{C} \subseteq \mathcal{X}_+$.

Proof. We set

$$\Lambda\left(\bar{x}\right) \coloneqq \rho\left(\bar{x}\right), \tag{2.35}$$

for all $\bar{x} \in \mathbb{R}^n$, or equivalently $\Lambda \coloneqq \rho \mid_{\mathbb{R}^n}$. As already mentioned, the properties (S1), (S2a) and (S5) of ρ, translate to the properties (A1), (A2) and (A3) for $\Lambda \coloneqq \rho \mid_{\mathbb{R}^n}$. We also have that (S3) and (S6) translate to (A4) and (A5). The additional property for $\Lambda\left(\bar{\mathcal{X}}\right)$ also follows directly from the definition of Λ. So, in each situation this construction yields an aggregation rule with all desired properties.

In the next step, we consider $\rho_0 \colon \mathcal{X} \to \mathbb{R} \cup \{\infty\}$, defined by

$$\rho_0\left(X\right) \coloneqq \rho\left(\bar{X}\right), \tag{2.36}$$

for $\bar{X} \in \bar{\mathcal{X}}$ with $\Lambda\left(\bar{X}\right) \coloneqq \Lambda \circ \bar{X} = X$ and

$$\rho_0\left(X\right) = \infty$$

for $X \notin \Lambda\left(\bar{\mathcal{X}}\right)$. We first note that ρ_0 is well-defined. Indeed, for $\bar{X}, \bar{Y} \in \bar{\mathcal{X}}$ with $X = \Lambda\left(\bar{X}\right) = \Lambda\left(\bar{Y}\right) = Y$, we have

$$\rho\left(\bar{X}\left(\omega\right)\right) = \Lambda\left(\bar{X}\left(\omega\right)\right) = \Lambda\left(\bar{X}\right)\left(\omega\right) = X\left(\omega\right) \geq Y\left(\omega\right) = \Lambda\left(\bar{Y}\right)\left(\omega\right) = \Lambda\left(\bar{Y}\left(\omega\right)\right) = \rho\left(\bar{Y}\left(\omega\right)\right)$$

and

$$\rho\left(\bar{X}\left(\omega\right)\right) = \Lambda\left(\bar{X}\left(\omega\right)\right) = \Lambda\left(\bar{X}\right)\left(\omega\right) = X\left(\omega\right) \leq Y\left(\omega\right) = \Lambda\left(\bar{Y}\right)\left(\omega\right) = \Lambda\left(\bar{Y}\left(\omega\right)\right) = \rho\left(\bar{Y}\left(\omega\right)\right)$$

for all $\omega \in \Omega$. The preference consistency (S4) of ρ now implies

$$\rho_0\left(X\right) = \rho\left(\bar{X}\right) = \rho\left(\bar{Y}\right) = \rho_0\left(Y\right).$$

Since ρ is a proper mapping, we obviously have that $\rho_0\left(X\right) \neq -\infty$ for all $X \in \mathcal{X}$. ρ also satisfies $\mathcal{R}$-surjectivity (S5). So for all reals $a \in \mathcal{R}$ there is some $\bar{x} \in \mathbb{R}^n$ with $\rho\left(\bar{x}\right) = \Lambda\left(\bar{x}\right) = a$. But this means $\rho_0\left(a\right) = a$ for all $a \in \mathcal{R}$, so ρ_0 is constant on $\mathcal{R}$ and hence proper. Next, consider $X, Y \in \Lambda\left(\bar{\mathcal{X}}\right)$ with $X \succcurlyeq Y$. Then there are $\bar{X}, \bar{Y} \in \bar{\mathcal{X}}$ with $\Lambda\left(\bar{X}\right) = X$ and $\Lambda\left(\bar{Y}\right) = Y$, respectively. For all $\omega \in \Omega$, we have

$$\rho\left(\bar{X}\left(\omega\right)\right) = \Lambda\left(\bar{X}\right)\left(\omega\right) = X\left(\omega\right) \geq Y\left(\omega\right) = \Lambda\left(\bar{Y}\right)\left(\omega\right) = \rho\left(\bar{Y}\left(\omega\right)\right),$$

which together with the preference consistency (S4) of ρ implies

$$\rho_0(X) = \rho\left(\bar{X}\right) \geq \rho\left(\bar{Y}\right) = \rho_0(Y).$$

This means that ρ_0 is monotone on $\Lambda\left(\bar{\mathcal{X}}\right)$. Now, we show that ρ_0 is a convex mapping. The inequality (A.4) obviously holds if either $X \notin \Lambda\left(\bar{\mathcal{X}}\right)$, $Y \notin \Lambda\left(\bar{\mathcal{X}}\right)$ or $X, Y \notin \Lambda\left(\bar{\mathcal{X}}\right)$. So, consider $X, Y \in \Lambda\left(\bar{\mathcal{X}}\right)$ and $\alpha \in [0,1]$. Since $\Lambda\left(\bar{\mathcal{X}}\right)$ is a convex set, we have $Z \coloneqq \alpha X + (1-\alpha) Y \in \Lambda\left(\bar{\mathcal{X}}\right)$ and there are $\bar{X}, \bar{Y}, \bar{Z} \in \bar{\mathcal{X}}$ with $\Lambda\left(\bar{X}\right) = X$, $\Lambda\left(\bar{Y}\right) = Y$ and $\Lambda\left(\bar{Z}\right) = Z$, respectively. Now, for all $\omega \in \Omega$

$$\begin{aligned}
\rho\left(\bar{Z}(\omega)\right) &= \Lambda\left(\bar{Z}\right)(\omega) \\
&= Z(\omega) \\
&= \alpha X(\omega) + (1-\alpha) Y(\omega) \\
&= \alpha \Lambda\left(\bar{X}\right)(\omega) + (1-\alpha) \Lambda\left(\bar{Y}\right)(\omega) \\
&= \alpha \rho\left(\bar{X}(\omega)\right) + (1-\alpha) \rho\left(\bar{Y}(\omega)\right).
\end{aligned}$$

The risk convexity (S2b) of ρ implies

$$\rho_0(Z) = \rho\left(\bar{Z}\right) \leq \alpha\rho\left(\bar{X}\right) + (1-\alpha)\rho\left(\bar{Y}\right) = \alpha\rho_0(X) + (1-\alpha)\rho_0(Y).$$

Moreover, ρ_0 is positively homogeneous if ρ is positively homogeneous. Finally, for all $\bar{X} \in \bar{\mathcal{X}}$ and in every situation formulated, we have $\rho\left(\bar{X}\right) = \rho_0\left(\Lambda\left(\bar{X}\right)\right)$, as desired.

To prove the opposite implication, consider a mapping $\Lambda\colon \mathbb{R}^n \to \mathbb{R}$ such that $\Lambda\left(\bar{\mathcal{X}}\right) = \mathcal{C} \subseteq \mathcal{X}$ is convex, a mapping $\rho_0\colon \mathcal{X} \to \mathbb{R} \cup \{\infty\}$ and define $\rho\colon \bar{\mathcal{X}} \to \mathbb{R} \cup \{\infty\}$ by

$$\rho\left(\bar{X}\right) \coloneqq \rho_0\left(\Lambda\left(\bar{X}\right)\right).$$

It immediately follows that $\rho\left(\bar{X}\right) \neq -\infty$ for all $\bar{X} \in \bar{\mathcal{X}}$. Now if Λ satisfies the $\mathcal{R}$-surjectivity (A3) and ρ is constant on $\Lambda(\mathbb{R}^n) = \mathcal{R}$, ρ itself satisfies the $\mathcal{R}$-surjectivity (S5) and hence is proper. By definition, we now have $\rho \mid_{\mathbb{R}^n} = \Lambda$. If both Λ and ρ_0 are convex, ρ satisfies the outcome convexity (S2a). If Λ is monotone and ρ_0 is monotone on $\Lambda\left(\bar{\mathcal{X}}\right)$, ρ satisfies monotonicity (S1). If all the above mentioned properties for Λ and ρ_0 hold at the same time, ρ is indeed a systemic risk measure with the additional property (S5). It remains to show that ρ also satisfies (S4) and (S2b). To this end, consider $\bar{X}, \bar{Y} \in \bar{\mathcal{X}}$ with

$$\rho_0\left(\Lambda\left(\bar{X}\right)(\omega)\right) = \rho\left(\bar{X}(\omega)\right) \geq \rho\left(\bar{Y}(\omega)\right) = \rho_0\left(\Lambda\left(\bar{Y}\right)(\omega)\right),$$

for all $\omega \in \Omega$. Since ρ_0 is constant on $\Lambda(\mathbb{R}^n)$ and $\Lambda\left(\bar{X}\right)(\omega) = \Lambda\left(\bar{X}(\omega)\right)$, this directly implies $\Lambda\left(\bar{X}\right) \succcurlyeq \Lambda\left(\bar{Y}\right)$. Now, ρ_0 is monotone on $\Lambda\left(\bar{\mathcal{X}}\right)$, which finally implies

$$\rho\left(\bar{X}\right) = \rho_0\left(\Lambda\left(\bar{X}\right)\right) \geq \rho_0\left(\Lambda\left(\bar{Y}\right)\right) = \rho\left(\bar{Y}\right).$$

Thus, ρ satisfies the preference consistency (S4). For the risk convexity (S2b), let $\alpha \in [0,1]$, $\bar{X}, \bar{Y}, \bar{Z} \in \bar{\mathcal{X}}$ and suppose that for all $\omega \in \Omega$

$$\rho_0\left(\Lambda\left(\bar{Z}(\omega)\right)\right) = \alpha\rho_0\left(\Lambda\left(\bar{X}(\omega)\right)\right) + (1-\alpha)\rho_0\left(\Lambda\left(\bar{Y}(\omega)\right)\right).$$

Now, since we have $\Lambda(\mathbb{R}^n) = \mathcal{R}$ by (A3) and $\rho_0(a) = a$ for $a \in \mathcal{R}$ by (R4), this means

$$\Lambda\left(\bar{Z}(\omega)\right) = \alpha\Lambda\left(\bar{X}(\omega)\right) + (1-\alpha)\Lambda\left(\bar{Y}(\omega)\right).$$

The convexity of ρ_0 implies

$$\rho\left(\bar{X}\right) = \rho_0\left(\Lambda\left(\bar{Z}\right)\right) \leq \alpha\rho_0\left(\Lambda\left(\bar{X}\right)\right) + (1-\alpha)\rho_0\left(\Lambda\left(\bar{Y}\right)\right) = \alpha\rho\left(\bar{X}\right) + (1-\alpha)\rho\left(\bar{Y}\right),$$

as desired. Now if Λ and ρ_0 additionally satisfy (A4) and (R3), respectively, ρ inherits this property and therefore additionally satisfies (S3). If Λ satisfies (A5), we have by (R4)

$$\rho(1_n) = \rho_0(\Lambda(1_n)) = \rho_0(n) = n.$$

So, ρ satisfies (S6). If $\mathcal{R} = \mathbb{R}_+$, it is obvious that $\Lambda\left(\bar{\mathcal{X}}\right) = \rho\mid_{\mathbb{R}^n}\left(\bar{\mathcal{X}}\right) = \mathcal{C} \subseteq \mathcal{X}_+$. Finally, in every situation formulated, ρ satisfies the claimed properties. $\square$

Theorem 2.33 can be seen as a blueprint for the construction of systemic risk measures of type (0.1). In particular, every convex single-firm risk measure which is constant on $\mathbb{R}$ can serve as a building block. Additionally if a given mapping satisfies the desired properties, the proof yields an instruction to find ρ_0 and Λ. Note that the presented extension of the mapping ρ_0 from $\mathcal{C}$ to the whole $\mathcal{X}$ is not unique. However, in combination with Λ this fact does not propose any problem at this stage. When it comes to the dual representation of such systemic risk measures, we will see a different extension in Remark 2.50 for a special situation. Before we move on to some examples, let us focus on the technical requirements in the previous Theorem. First of all, there is the space $\bar{\mathcal{X}}$. In many situations, the spaces $\mathcal{X}_i$ are quiet similar or even identical. For example we have $\mathcal{X}_i = L^{p_i}(\Omega, \mathcal{F}, \mathbb{P})$ for different or identical p_i. There, we have some $j \in \{1, \ldots, n\}$ with $\mathcal{X}_i \subseteq \mathcal{X}_j$ for all $i \in \{1, \ldots, n\}$. Simple aggregation rules like $\Lambda^{\text{sum,c}}$ or Λ^{loss} yield $\Lambda\left(\bar{\mathcal{X}}\right) = \mathcal{X}_j$. But clearly, the explicit structure of $\mathcal{X}$ heavily depends on Λ. In [KOZ16] the authors presented some conditions for Λ to guarantee $\Lambda(\mathcal{X}^n) = \mathcal{X}$ in the situation where $\mathcal{X} = L^p(\Omega, \mathcal{F}, \mathbb{P})$. We can extend this result to the case, where $\mathcal{X}$ is a Banach lattice. The Lemma makes use of the function $f_\Lambda\colon \mathbb{R} \to \mathbb{R}$ defined by

$$f_\Lambda(a) := \Lambda(a \cdot 1_n)\,. \tag{2.37}$$

Lemma 2.38. *Let $\mathcal{X}$ be a Banach lattice with norm $\|\cdot\|$ and $\mathcal{R} = \mathbb{R}$. If any of the assumptions below hold then $\Lambda(\mathcal{X}^n) = \mathcal{X}$.*

(a) $\Lambda\colon \mathbb{R}^n \to \mathbb{R}$ satisfies the properties (A1)-(A3) and the function f_Λ satisfies $\|f_\Lambda(X)\| < \infty$ and $\|f_\Lambda^{-1}(X)\| < \infty$, for all $X \in \mathcal{X}$.

(b) $\Lambda\colon \mathbb{R}^n \to \mathbb{R}$ is measurable, satisfies the property (A1), the function f_Λ is bijective and satisfies $\|f_\Lambda(X)\| < \infty$ and $\|f_\Lambda^{-1}(X)\| < \infty$, for all $X \in \mathcal{X}$.

(c) $\Lambda\colon \mathbb{R}^n \to \mathbb{R}$ satisfies the property (A3), is strictly increasing, continuous and the function f_Λ satisfies $\|f_\Lambda(X)\| < \infty$ and $\|f_\Lambda^{-1}(X)\| < \infty$, for all $X \in \mathcal{X}$.

Proof. For all $\bar{X} \in \bar{\mathcal{X}}$, we can define

$$Z_{\bar{X}} := \sup_{i\in\{1,\ldots,n\}} \{X_i\}\, \mathbb{1}_A + \inf_{i\in\{1,\ldots,n\}} \{X_j\}\, \mathbb{1}_{A^C},$$

where $A = \left\{\omega \in \Omega \mid \Lambda\left(\bar{X}(\omega)\right) \geq 0\right\}$. In each situation, the mapping Λ is measurable which implies $A \in \mathcal{F}$ and hence $Z_{\bar{X}} \in L^0(\Omega, \mathcal{F})$. Since $\mathcal{X}$ is a Banach lattice, we have $Z_{\bar{X}} \in \mathcal{X}$. Λ is also monotone increasing in each situation, which now implies

$$\begin{aligned} 0 &\leq \Lambda\left(\bar{X}(\omega)\right) \leq \Lambda\left(Z_{\bar{X}}(\omega)\, 1_n\right) && \text{for } \omega \in A, \\ 0 &> \Lambda\left(\bar{X}(\omega)\right) \geq \Lambda\left(Z_{\bar{X}}(\omega)\, 1_n\right) && \text{for } \omega \in A^C. \end{aligned}$$

Combining these observations yields to

$$|\Lambda\left(\bar{X}\right)| \leq |\Lambda\left(Z_{\bar{X}} 1_n\right)| = |f_\Lambda\left(Z_{\bar{X}}\right)|.$$

From $\|f_\Lambda(X)\| < \infty$ and the fact that $\|\cdot\|$ is a lattice norm, we obtain

$$\|\Lambda\left(\bar{X}\right)\| \leq \|f_\Lambda(X)\| < \infty,$$

which means $\Lambda\left(\bar{X}\right) \in \mathcal{X}$ and hence $\Lambda(\mathcal{X}^n) \subseteq \mathcal{X}$. To see the opposite inclusion, first note that the possible collection of properties for Λ always imply that it is measurable and that f_Λ is a measurable, bijective and strictly increasing mapping. For f_Λ^{-1}, we also have that it is a measurable, bijective and strictly increasing mapping So, for an arbitrary $X \in \mathcal{X}$ we can define $Y\colon \Omega \to \mathbb{R}$ by $Y := f_\Lambda^{-1}(X)$. Obviously, $Y \in L^0(\Omega, \mathcal{F})$ and by assumption $Y \in \mathcal{X}$, since we always have $\|Y\| = \|f_\Lambda^{-1}(X)\| < \infty$, for all $X \in \mathcal{X}$. But this means $Y1_n \in \mathcal{X}^n$ and

$$\Lambda(Y1_n) = f_\Lambda(Y) = X$$

and therefore $X \in \Lambda(\mathcal{X}^n)$. Now, since $X \in \mathcal{X}$ was choosen arbitrarily, $\mathcal{X} \subseteq \Lambda(\mathcal{X}^n)$ and finally $\mathcal{X} = \Lambda(\mathcal{X}^n)$. □

Note that even in the general setting of Theorem 2.33, $\Lambda\left(\bar{\mathcal{X}}\right) = \rho \mid_{\mathbb{R}^n} \left(\bar{\mathcal{X}}\right) = \mathcal{X}$ always implies that ρ_0 is a single-firm risk measure. We continue with some examples.

Example 2.39. *(i) Let us construct different versions of the systemic entropic risk measure. As a base single-firm risk measure, we will always use the single-firm entropic risk measure ρ_0^{entr} presented in Example 2.26 (i). In this setup, the parameter θ can be understood as a systemic risk aversion parameter. It is a global parameter which serves for the whole system and gives the regulator the opportunity to adjust the risk measure. A suitable way to derive such a systemic risk aversion parameter is to simply assign an individual risk aversion parameter α_i to each participant of the system. Note that α_i reflects the risk aversion from the point of view of the regulator, i.e. risk seeking participants will receive bigger α. Now the systemic risk aversion parameter has to act contrary to the risk aversion of the system and its participants. Risk averse participants of the system give the regulator the opportunity to be more flexible with the systemic risk measure, whereas risk friendly participants yield a conservative systemic risk measurement. Therefore, we set*

$$\theta = \frac{1}{\sum_{i=1}^{n} \frac{1}{\alpha_i}}.$$

Now we use different aggregation rules from Example 2.16 to obtain different versions. For Λ^{sum}

$$\rho^{ses}\left(\bar{X}\right) = \left(\rho_0^{entr} \circ \Lambda^{sum}\right)\left(\bar{X}\right) = \frac{1}{\theta} \ln \mathbb{E}\left[\exp\left(\theta \sum_{i=1}^{n} X_i\right)\right], \tag{2.40}$$

for $\Lambda^{sum,c}$

$$\begin{aligned} \rho^{ses,c}\left(\bar{X}\right) &= \left(\rho_0^{entr} \circ \Lambda^{sum,c}\right)\left(\bar{X}\right) = \frac{1}{\theta} \ln \mathbb{E}\left[\exp\left(\theta\left(\sum_{i=1}^{n} X_i - c\right)\right)\right] \\ &= \frac{1}{\theta} \ln \mathbb{E}\left[\exp\left(\theta \sum_{i=1}^{n} X_i\right)\right] - c = \left(\rho_0^{entr} \circ \Lambda^{sum}\right)\left(\bar{X}\right) - c, \end{aligned} \tag{2.41}$$

and for Λ^{loss}

$$\rho^{sel}\left(\bar{X}\right) = \left(\rho_0^{entr} \circ \Lambda^{loss}\right)\left(\bar{X}\right) = \frac{1}{\theta}\ln \mathbb{E}\left[\exp\left(\theta\sum_{i=1}^{n} X_i^+\right)\right]. \tag{2.42}$$

(ii) Now we use the aggregation rule Λ^{ut} *presented in Example 2.16 and the single-firm risk measure* $\rho_0^{\mathcal{A}}$ *presented in Example 2.26 from [BFFMB20], to construct a first aggregate systemic risk measure. It is given as follows:*

$$\rho\left(\bar{X}\right) = \sum_{i=1}^{n} \frac{1}{\alpha_i}\mathbb{E}\left[\exp\left(\alpha_i X_i\right)\right] - B \tag{2.43}$$

The explicit structure of systemic risk measures of type (0.1) has consequences for their primal and dual representation. As mentioned rearlier, by $\mathcal{A}_{\rho_0}$ we denote the corresponding epigraph. By abuse of notation, for the mapping $\Lambda \circ \cdot : \bar{\mathcal{X}} \to \mathcal{X}$ with $\Lambda \circ \cdot \left(\bar{X}\right) = \Lambda \circ \bar{X} = \Lambda\left(\bar{X}\right)$, we write

$$\mathcal{A}_{\Lambda\circ\cdot} := \left\{\left(Y, \bar{Z}\right) \in \mathcal{X} \times \bar{\mathcal{X}} \mid Y \succcurlyeq \Lambda\left(\bar{Z}\right)\right\}.$$

Proposition 2.44 ([KOZ16]). *Let* $\mathcal{X}$ *be an arbitrary locally convex Hausdorff space and* $\mathcal{C} \subseteq \mathcal{X}$ *an arbitrary convex set. Suppose that* $\rho = \rho_0 \circ \Lambda$ *is a systemic risk measure with a convex aggregation rule* $\Lambda : \mathbb{R}^n \to \mathbb{R}$ *that satisfies* $\Lambda\left(\bar{\mathcal{X}}\right) = \mathcal{C}$*, and a convex mapping* $\rho_0 : \mathcal{X} \to \mathbb{R} \cup \{\infty\}$ *which is monotone on* $\Lambda\left(\bar{\mathcal{X}}\right) = \mathcal{C}$ *and satisfies the constancy property (R4) on* $\mathcal{R} = \Lambda(\mathbb{R}^n)$*. Then, for all* $\bar{X} \in \bar{\mathcal{X}}$*,* ρ *admits the primal representation*

$$\rho\left(\bar{X}\right) = \inf\left\{m \in \mathbb{R} \mid (m, Y) \in \mathcal{A}_{\rho_0}, \left(Y, \bar{X}\right) \in \mathcal{A}_{\Lambda\circ\cdot}\right\} \tag{2.45}$$

where we set as usual $\inf \emptyset := \infty$.

Proof. First, note that the proof of our general primal representation result 2.3 never relied on the fact that the considered mapping is monotone. If we drop this assumption, we still have a represenation and the parts (ii)(a), (ii)(b), (ii)(d) and (ii)(e) still hold. So for ρ_0, we have

$$\rho_0(X) = \inf\left\{m \in \mathbb{R} \mid (m, X) \in \mathcal{A}_{\rho_0}\right\}.$$

For ρ, we already know

$$\rho\left(\bar{X}\right) = \inf\left\{m \in \mathbb{R} \mid \left(m, \bar{X}\right) \in \mathcal{A}_{\rho}\right\}.$$

But

$$\begin{aligned}\left(m,\bar{X}\right)\in\mathcal{A}_{\rho}&\Leftrightarrow m\geq\rho\left(\bar{X}\right)\\&\Leftrightarrow m\geq\rho_0\left(\Lambda\left(\bar{X}\right)\right)\\&\Leftrightarrow\left(m,\Lambda\left(\bar{X}\right)\right)\in\mathcal{A}_{\rho_0}.\end{aligned}$$

We also have that ρ_0 is monotone on $\Lambda\left(\bar{\mathcal{X}}\right)=\mathcal{C}$. This implies

$$\left\{m\in\mathbb{R}\mid\left(m,\Lambda\left(\bar{X}\right)\right)\in\mathcal{A}_{\rho_0}\right\}=\left\{m\in\mathbb{R}\mid(m,Y)\in\mathcal{A}_{\rho_0},\left(Y,\bar{X}\right)\in\mathcal{A}_{\Lambda\circ\cdot}\right\},$$

as desired. □

In some sense, the decomposition of ρ into ρ_0 and Λ transfers to the primal representation. This is also true for the dual representation. Here, α^{ρ_0} denotes the corresponding convex conjugate. By abuse of notation, for the mapping $\Lambda\circ\cdot\colon\bar{\mathcal{X}}\to\mathcal{X}$ with $\Lambda\circ\cdot\left(\bar{X}\right)=\Lambda\circ\bar{X}=\Lambda\left(\bar{X}\right)$, the mapping $\alpha^{\Lambda\circ\cdot}\colon\mathcal{X}'\times\bar{\mathcal{X}}'\to\mathbb{R}$ is given by

$$\alpha^{\Lambda\circ\cdot}\left(\xi,\Xi\right)\coloneqq\sup_{\bar{Z}\in\bar{\mathcal{X}}}\left\{\langle\bar{Z},\Xi\rangle_n-\langle\Lambda\left(\bar{Z}\right),\xi\rangle\right\}.$$

Theorem 2.46 ([KOZ16]). *Let $\mathcal{X}$ be an arbitrary locally convex Hausdorff space and $\mathcal{C}\subseteq\mathcal{X}$ an arbitrary convex set. Suppose that $\rho=\rho_0\circ\Lambda$ is a systemic risk measure with a convex aggregation rule $\Lambda:\mathbb{R}^n\to\mathbb{R}$ that is continuous on $\bar{\mathcal{X}}$ and satisfies $\Lambda\left(\bar{\mathcal{X}}\right)=\mathcal{C}$, and a l.s.c. convex single-firm risk measure $\rho_0:\mathcal{X}\to\mathbb{R}\cup\{\infty\}$ which satisfies the constancy property (R4) on $\mathcal{R}=\Lambda(\mathbb{R}^n)$. Then, for all $\bar{X}\in\bar{\mathcal{X}}$,*

$$\rho\left(\bar{X}\right)=\sup_{(\xi,\Xi)\in\mathcal{D}^{\#}}\left\{\langle\bar{X},\Xi\rangle_n-\alpha\left(\xi,\Xi\right)\right\},\tag{2.47}$$

where $\alpha\colon\mathcal{X}'\times\bar{\mathcal{X}}'\to\mathbb{R}\cup\{\infty\}$ is given by

$$\alpha\left(\xi,\Xi\right)=\alpha^{\rho_0}\left(\xi\right)+\alpha^{\Lambda\circ\cdot}\left(\xi,\Xi\right)\tag{2.48}$$

and

$$\mathcal{D}^{\#}\coloneqq dom\left(\alpha\right)=dom\left(\alpha^{\rho_0}\right)\times dom\left(\alpha^{\Lambda\circ\cdot}\right)\subseteq\mathcal{X}'_+\times\bar{\mathcal{X}}'_+.\tag{2.49}$$

In addition, the ξ-component of a feasible solution satisfies $\langle 1,\xi\rangle=1$ in case of $\rho\left(\mathbb{R}^n\right)=\mathbb{R}$ and $\langle 1,\xi\rangle\leq 1$ in case of $\rho\left(\mathbb{R}^n\right)=\mathbb{R}_+$.

Proof. For $\xi\in\mathcal{X}'$, consider the mapping $f_{\xi,\Lambda}\colon\bar{\mathcal{X}}\to\mathbb{R}$ with $f_{\xi,\Lambda}\left(\bar{X}\right)\coloneqq\langle\Lambda\left(\bar{X}\right),\xi\rangle$. This mapping is finite valued and continuous. If $\langle\cdot,\xi\rangle$ is monotone (at least) on

$\mathcal{C} = \Lambda\left(\bar{\mathcal{X}}\right)$, this mapping is also convex and monotone. Hence in this situation, by Theorem A.32, it admits a dual representation, given by

$$f_{\xi,\Lambda}\left(\bar{X}\right) = \sup_{\Xi\in\mathcal{D}'}\left\{\langle\bar{X},\Xi\rangle_n - \alpha^{f_{\xi,\Lambda}}\left(\Xi\right)\right\},$$

where $\mathcal{D}' = \operatorname{dom}\left(\alpha^{f_{\xi,\Lambda}}\right) \subseteq \bar{\mathcal{X}}'_+$. Moreover, we have

$$\alpha^{f_{\xi,\Lambda}}\left(\Xi\right) = \sup_{\bar{Z}\in\bar{\mathcal{X}}}\left\{\langle\bar{Z},\Xi\rangle_n - \langle\Lambda\left(\bar{Z}\right),\xi\rangle\right\} = \alpha^{\Lambda\circ\cdot}\left(\xi,\Xi\right)$$

and therefore $\operatorname{dom}\left(\alpha^{f_{\xi,\Lambda}}\right) = \operatorname{dom}\left(\alpha^{\Lambda\circ\cdot}\right)$. Since ρ_0 is also a convex and l.s.c. mapping, we have

$$\rho_0\left(X\right) = \sup_{\xi\in\mathcal{D}'_0}\left\{\langle X,\xi\rangle - \alpha^{\rho_0}\left(\xi\right)\right\},$$

where $\mathcal{D}'_0 = \operatorname{dom}\left(\alpha^{\rho_0}\right)$. Theorem 2.8 also tells us that $\mathcal{D}'_0 = \operatorname{dom}\left(\alpha^{\rho_0}\right) \subseteq \mathcal{X}'_+$. Property (C) yields that every $\xi \in \mathcal{D}'_0$ induces a convex, montone and continuous mapping $f_{\xi,\Lambda}$. Now, for all $\bar{X} \in \bar{\mathcal{X}}$ we have

$$\begin{aligned}
\rho\left(\bar{X}\right) &= \rho_0\left(\Lambda\left(\bar{X}\right)\right)\\
&= \sup_{\xi\in\mathcal{D}'_0}\left\{\langle\Lambda\left(\bar{X}\right),\xi\rangle - \alpha^{\rho_0}\left(\xi\right)\right\}\\
&= \sup_{\xi\in\mathcal{D}'_0}\left\{\sup_{\Xi\in\mathcal{D}'}\left\{\langle\bar{X},\Xi\rangle_n - \alpha^{\Lambda\circ\cdot}\left(\xi,\Xi\right)\right\} - \alpha^{\rho_0}\left(\xi\right)\right\}\\
&= \sup_{(\xi,\Xi)\in\mathcal{D}^{\#}}\left\{\langle\bar{X},\Xi\rangle_n - \left(\alpha^{\rho_0}\left(\xi\right) + \alpha^{\Lambda\circ\cdot}\left(\xi,\Xi\right)\right)\right\}.
\end{aligned}$$

Hence, ρ admits the desired representation and $\operatorname{dom}\left(\alpha\right)$ has the claimed structure.

Now consider the ξ-component of an arbitrary feasible solution $(\xi,\Xi) \in \mathcal{D}^{\#}$. The proof of Theorem 2.33 tells us that $\Lambda\left(\mathbb{R}^n\right) = \rho\left(\mathbb{R}^n\right) = \mathcal{R} \in \{\mathbb{R}_+,\mathbb{R}\}$. Since we assumed that ρ_0 is constant on $\mathcal{R}$, we have $(\lambda,\lambda) \in \mathcal{A}_{\rho_0}$ for all $\lambda \in \mathcal{R}$. This yields

$$\alpha^{\rho_0}\left(\xi\right) \geq \sup_{\lambda\in\mathcal{R}}\left\{\lambda\left(\langle 1,\xi\rangle - 1\right)\right\}.$$

But this means $\alpha^{\rho_0}\left(\xi\right) = \infty$ if $\langle 1,\xi\rangle \neq 1$ for $\mathcal{R} = \mathbb{R}$ or if $\langle 1,\xi\rangle > 1$ for $\mathcal{R} = \mathbb{R}_+$. □

Remark 2.50. *1. Consider the situation, where $\mathcal{X}$ is a Riesz space and we have $\Lambda\left(\bar{\mathcal{X}}\right) = \mathcal{X}_+$. The mapping ρ_0 constructed in Theorem 2.33 is only monotone on $\Lambda\left(\bar{\mathcal{X}}\right) = \mathcal{X}_+$, and hence seems not to fit in the framework of the previous Theorem. However, in this situation, we can extend ρ_0 from $\Lambda\left(\bar{\mathcal{X}}\right) = \mathcal{X}_+$ to the whole $\mathcal{X}$ by setting $\rho_0\left(X\right) = \rho_0\left(X^+\right)$ due to the fact that $\mathcal{X}$ is a Riesz* *space. Using the properties*

$$(X+Y)^+ \leq X^+ + Y^+,$$
$$X \leq Y \Rightarrow X^+ \leq Y^+$$

yields that ρ_0 is monotone on the whole $\mathcal{X}$, and hence a single-firm risk measure with all additional properties stated in Theorem 2.33. This means that the alternative proof presented above indeed proves the full Theorem from [KOZ16].

2. *The original representation result in [KOZ16] states that*

$$\begin{aligned}\alpha(\xi,\Xi) &= \sup_{(m,Y)\in\mathcal{A}_{\rho_0}} \{-m+\langle Y,\xi\rangle\} + \sup_{\left(V,\bar{Z}\right)\in\mathcal{A}_{\Lambda\circ\cdot}} \left\{-\langle V,\xi\rangle + \langle \bar{Z},\Xi\rangle_n\right\} \\ &= \sup_{(m,Y)\in\mathcal{A}_{\rho_0},\left(V,\bar{Z}\right)\in\mathcal{A}_{\Lambda\circ\cdot}} \left\{-m+\langle Y-V,\xi\rangle + \langle \bar{Z},\Xi\rangle_n\right\}.\end{aligned}$$

Note, however, that by monotonicity of the supremum and $\xi \in \mathcal{X}'_+$, relevant pairs to consider are of the form $(\rho_0(Y),Y)$ and $\left(\Lambda\left(\bar{Z}\right),\bar{Z}\right)$, respectively. This observation transforms the above representation for α to the one presented in (2.48)*.*

From Theorem 2.8, we already know that the penalty term of a positively homogeneous systemic risk measure reduces to an indicator function in the sense of convex analysis. If we can decompose the systemic risk measure, the domain of α changes if at least one of the components, ρ_0 or Λ, is positively homogeneous. In this situation, the sign changed dual cones to $\mathcal{A}_{\rho_0}$ and $\mathcal{A}_{\Lambda\circ\cdot}$, given by

$$\begin{aligned}\mathcal{A}'_{\rho_0} &\coloneqq \left\{(x,\xi)\in\mathbb{R}\times\mathcal{X}' \mid mx - \langle Y,\xi\rangle \geq 0 \;\forall\; (m,Y)\in\mathcal{A}_{\rho_0}\right\}, \\ \mathcal{A}'_{\Lambda\circ\cdot} &\coloneqq \left\{(\xi,\Xi)\in\mathcal{X}'\times\bar{\mathcal{X}}' \mid \langle Y,\xi\rangle - \langle\bar{Z},\Xi\rangle_n \geq 0 \;\forall\; \left(Y,\bar{Z}\right)\in\mathcal{A}_{\Lambda\circ\cdot}\right\},\end{aligned}$$

play a crucial role.

Lemma 2.51. *Consider the situation of Theorem 2.46. If ρ_0 is additionally positively homogeneous we have*

$$\alpha^{\rho_0}(\xi) = \begin{cases} 0, & (1,\xi)\in\mathcal{A}'_{\rho_0}, \\ \infty, & (1,\xi)\notin\mathcal{A}'_{\rho_0}, \end{cases}$$

and if Λ is additionally positively homogeneous we have

$$\alpha^{\Lambda\circ\cdot}(\xi,\Xi) = \begin{cases} 0, & (\xi,\Xi)\in\mathcal{A}'_{\Lambda\circ\cdot}, \\ \infty, & (\xi,\Xi)\notin\mathcal{A}'_{\Lambda\circ\cdot}. \end{cases}$$

Proof. Note that Theorem 2.8 already ensures that the respective penalty functions are indicators in the sense of convex analysis. The important thing here is that we can determine the domain. Let us start with ρ_0. We always have

$$\begin{aligned}\alpha^{\rho_0}(\xi) &= \sup_{X\in\mathcal{X}}\{\langle X,\xi\rangle - \rho_0(X)\}\\ &\geq \langle 0,\xi\rangle - \rho_0(0)\\ &= 0.\end{aligned}$$

Now, suppose that $(1,\xi)\in\mathcal{A}'_{\rho_0}$. Then the supremum is taken over elements which are non-positive. Hence $\alpha^{\rho_0}(\xi)=0$. If $(1,\xi)\notin\mathcal{A}'_{\rho_0}$, there exists some $(m,Y)\in\mathcal{A}_{\rho_0}$ with

$$\langle Y,\xi\rangle - \rho_0(Y) \geq \langle Y,\xi\rangle - m > 0.$$

Since $\mathcal{A}_{\rho_0}$ is a cone, we have $t(\rho_0(Y),Y)\in\mathcal{A}_{\rho_0}$, for all $t\in\mathbb{R}_+$. But this yields

$$\begin{aligned}\alpha^{\rho_0}(\xi) &= \sup_{X\in\mathcal{X}}\{\langle X,\xi\rangle - \rho_0(X)\}\\ &\geq \sup_{t\in\mathbb{R}_+}\{t(\langle Y,\xi\rangle - \rho_0(Y))\}\\ &= \infty.\end{aligned}$$

The properties of the domain of $\alpha^{\Lambda\circ\cdot}(\xi,\Xi)$ follow with the same arguments. □

Finally, the previous Lemma allows us to identify the domain of α presented in Theorem 2.46 if both, Λ and ρ_0, are positively homogeneous case.

Theorem 2.52 ([KOZ16])**.** *Let $\mathcal{X}$ be an arbitrary locally convex Hausdorff space and $\mathcal{C}\subseteq\mathcal{X}$ an arbitrary convex set. Suppose that $\rho=\rho_0\circ\Lambda$ is a positively homogeneous systemic risk measure with a positively homogeneous aggregation rule $\Lambda\colon\mathbb{R}^n\to\mathbb{R}$ that is continuous on $\bar{\mathcal{X}}$ and satisfies $\Lambda\left(\bar{\mathcal{X}}\right)=\mathcal{C}$ and a l.s.c. positively homogeneous single-firm risk measure $\rho_0\colon\mathcal{X}\to\mathbb{R}\cup\{\infty\}$ that satisfies the constancy property (R4) on $\Lambda(\mathbb{R}^n)$. Then, for all $\bar{X}\in\bar{\mathcal{X}}$, the representation* (2.47) *reduces to*

$$\rho\left(\bar{X}\right)=\sup_{\mathcal{V}^{\#}}\langle\bar{X},\Xi\rangle_n \tag{2.53}$$

where the set $\mathcal{V}^{\#}$ is defined by

$$\mathcal{V}^{\#}:=\left\{(\xi,\Xi)\in\mathcal{D}^{\#}\mid(1,\xi)\in\mathcal{A}'_{\rho_0},(\xi,\Xi)\in\mathcal{A}'_{\Lambda\circ\cdot}\right\}. \tag{2.54}$$

If Λ also satisfies the normalization property (A5), the Ξ-component of a feasible solution (ξ,Ξ) satisfies $\langle 1_n,\Xi\rangle_n\leq n$.

Proof. We only need to show the additional property of the Ξ-component of a feasible solution (ξ,Ξ). Since in this setup by Theorem 2.33, Λ satisfies (A5) if and only ρ

satisfies (S6). The representation now gives us

$$\begin{aligned} n &= \rho\left(1_n\right) \\ &\geq \langle 1_n, \Xi \rangle_n \end{aligned}$$

for all Ξ-components of $(\xi, \Xi) \in \mathcal{V}^{\#}$. $\square$

Example 2.55. *Consider the systemic risk measure ρ^{ses} presented in Example 2.39. Obviously Λ^{sum} is Gâteaux differentiable. Moreover, we have that for all $X, V \in M^{\exp}$*

$$\begin{aligned} \delta\rho_0^{entr}\left(X, V\right) &= \frac{\mathbb{E}\left[V \exp\left(\theta X\right)\right]}{\mathbb{E}\left[\exp\left(\theta X\right)\right]} \\ &= \langle V, \frac{\exp\left(\theta X\right)}{\mathbb{E}\left[\exp\left(\theta X\right)\right]} \rangle. \end{aligned}$$

But this means ρ_0^{entr} is Gâteaux differentiable at $X \in M^{\exp}$ with derivative, denoted by

$$\nabla\rho_0^{entr}\left(X\right) = \frac{\exp\left(\theta X\right)}{\mathbb{E}\left[\exp\left(\theta X\right)\right]}$$

Now, the chain rule for Gâteaux differentials (see Proposition A.43 (iii)) yields that ρ^{ses} is also Gâteaux differentiable with derivative for all $X \in \left(M^{\exp}\right)^n$, given by

$$\nabla\rho^{ses}\left(\bar{X}\right) = 1_n \frac{\exp\left(\theta \sum_{i=1}^n X_i\right)}{\mathbb{E}\left[\exp\left(\theta \sum_{i=1}^n X_i\right)\right]} = 1_n \nabla\rho_0^{entr}\left(\sum_{i=1}^n X_i\right).$$

Since ρ^{ses} is continuous on $\left(M^{\exp}\right)^n$, we have uniqueness in the second component of optimal solutions. But the same holds true for ρ_0^{entr} in the respective setting and therefore the unique optimal solution is given by

$$\left(\xi^{\bar{X}}, \Xi^{\bar{X}}\right) := \left(\nabla\rho_0^{entr}\left(\sum_{i=1}^n X_i\right), 1_n \nabla\rho_0^{entr}\left(\sum_{i=1}^n X_i\right)\right).$$

Note, that this is also the unique optimal solution to the dual problem of $\rho^{ses,c}$ also presented in Example 2.39. For ρ^{sel}, we are also able to present an optimal solution to the dual problem (2.47)*. To this end, consider*

$$1_{>0}\left(\bar{X}\right) := \left(\mathbb{1}_{\{X_1>0\}}, \ldots, \mathbb{1}_{\{X_n>0\}}\right)$$

and

$$\left(\xi, \Xi\right) = \left(\nabla\rho_0^{entr}\left(\sum_{i=1}^n X_i^+\right), 1_{>0}\left(\bar{X}\right) \nabla\rho_0^{entr}\left(\sum_{i=1}^n X_i^+\right)\right) \in \mathcal{A}'_{\Lambda^{loss}\circ}.$$

Λ^{loss} *is positively homogeneous and hence* $\alpha^{\Lambda^{loss}\circ\cdot}(\xi,\Xi)=0$. *Moreover,*

$$\begin{aligned}
\rho^{sel}\left(\bar{X}\right) &= \left(\rho_0^{entr}\circ\Lambda^{loss}\right)\left(\bar{X}\right)\\
&= \frac{1}{\theta}\ln\mathbb{E}\left[\exp\left(\theta\sum_{i=1}^{n}X_i^+\right)\right]\\
&= \sup_{(\xi,\Xi)\in\mathcal{D}^{\#}}\left\{\langle\bar{X},\Xi\rangle_n-\left(\alpha^{\rho_0^{entr}}(\xi)+\alpha^{\Lambda^{loss}\circ\cdot}(\xi,\Xi)\right)\right\}\\
&\geq \langle\bar{X},1_{>0}\left(\bar{X}\right)\nabla\rho_0^{entr}\left(\sum_{i=1}^{n}X_i^+\right)\rangle_n-\alpha^{\rho_0^{entr}}\left(\nabla\rho_0^{entr}\left(\sum_{i=1}^{n}X_i^+\right)\right)\\
&= \frac{1}{\theta}\ln\mathbb{E}\left[\exp\left(\theta\sum_{i=1}^{n}X_i^+\right)\right].
\end{aligned}$$

2.5. First Inject Capital

To give dual representation results for systemic risk measures of type (0.3), the authors in [BFFMB20] presented a couple of assumptions on the ingredients. The natural environment is then an Orlicz setup connected with a probability space $(\Omega,\mathcal{F},\mathbb{P})$. So, for a strict Young function ϕ, i.e. a finite valued, even and convex function on $\mathbb{R}$ with $\phi(0)=0$ and $\lim_{x\to\infty}\frac{\phi(x)}{x}=\infty$, the Orlicz space is given by

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

The subspace of L^{ϕ} given by

$$M^{\phi}:=\left\{X\in L^0(\Omega,\mathcal{F},\mathbb{P})\mid\mathbb{E}[\phi(\alpha X)]<\infty\text{ for all }\alpha>0\right\}$$

is called Orlicz heart. We equip these spaces with the Luxemburg norm corresponding to ϕ, given by

$$\|X\|_{\phi}=\inf\left\{\alpha\in(0,\infty)\mid\mathbb{E}\left[\phi\left(\frac{X}{\alpha}\right)\right]\leq 1\right\}.$$

With this norm, both spaces become Banach spaces. Moreover, the Luxemburg norm is in fact a lattice norm (see A.4) which makes them Banach lattices. The spaces are then endowed with the corresponding norm topology. We obtain $\left(M^{\phi}\right)'=L^{\phi^*}$. Together with the pairing

$$\langle X,\xi\rangle=\mathbb{E}[X\xi],$$

these spaces fit in our general framework, as mentioned in Section 1. Now, for strict Young functions ϕ_i, $i\in\{1,\dots,n\}$, we write $M^{\Phi}:=\bigtimes_{i=1}^{n}M^{\phi_i}$ and equip M^{Φ} with the product topology. Then we obtain $\left(M^{\Phi}\right)'=L^{\Phi^*}=\bigtimes_{i=1}^{n}L^{\phi_i^*}$ and the pairing is given by (1.2). The assumptions are now given as follows:

Assumption. *(i)* $\Lambda\colon \mathbb{R}^n \to \mathbb{R}$ *is given by* $\Lambda(\bar{x}) = \sum_{i=1}^n l_i(x_i)$, *where* $l_i : \mathbb{R} \to \mathbb{R}$ *is increasing, strictly convex, differentiable and satisfies the Inada conditions*

$$l_i'(\infty) := \lim_{x\to\infty} l_i'(x) = \infty, \quad l_i'(-\infty) := \lim_{x\to-\infty} l_i'(x) = 0.$$

(ii) *We set* $\phi_i(x) := l_i(|x|) - l_i(0)$. *Then for all* $i \in \{1,\dots,n\}$, ϕ_i *is a strict Young function. Now,* $\mathcal{C}_0$ *is subset of* $\mathcal{C}(\mathbb{R})$ *such that* $\mathcal{C} = \mathcal{C}_0 \cap M^\Phi$ *is a convex cone which satisfies* $\mathbb{R}^n \subseteq \mathcal{C} \subseteq \mathcal{C}(\mathbb{R})$.
(iii) $B > \Lambda(-\infty)$, *i.e. there exists* $\bar{m} \in \mathbb{R}^n$ *such that* $\Lambda(\bar{m}) = \sum_{i=1}^n l_i(m_i) \le B$.
(iv) $\mathcal{A}_0 := \{Z \in L^1(\Omega, \mathcal{F}, \mathbb{P}) \mid \mathbb{E}[Z] \le B\}$.
(v) *For all* $i \in \{1,\dots,n\}$, *it holds that for all probability measures* $\mathbb{Q} \ll \mathbb{P}$ *with density* $\frac{d\mathbb{Q}}{d\mathbb{P}} := \xi$

$$\mathbb{E}[l_i^*(\xi)] < \infty \quad \textit{if and only if} \quad \mathbb{E}[l_i^*(\lambda\xi)] < \infty, \textit{ for all } \lambda > 0.$$

Under these assumptions mappings $R\colon M^\Phi \to \mathbb{R} \cup \{-\infty\} \cup \{\infty\}$ given by

$$\begin{aligned} R\left(\bar{X}\right) &= \inf\left\{\sum_{i=1}^n Y_i \mid \bar{Y} \in \mathcal{C}, \Lambda\left(\bar{X} - \bar{Y}\right) \in \mathcal{A}_0\right\} \\ &= \inf\left\{\sum_{i=1}^n Y_i \mid \bar{Y} \in \mathcal{C}, \mathbb{E}\left[\sum_{i=1}^n l_i(X_i - Y_i)\right] \le B\right\} \end{aligned} \tag{2.56}$$

are analyzed. As already mentioned, this particular form of R is motivated by the primal representation of systemic risk measures of type (0.1). The aim is to change the order of aggregating and injecting capital. To this end, the scalar m has to be replaced by a (random) vector.

Remark 2.57. *A reasonable choice for* B *is given by* $\sum_{i=1}^n l_i(0)$. *It ensures that* $R(0) = 0$ *which is a desirable property in some situations, as we will see later.*

The following Proposition collects some properties. First and foremost, it shows that these mappings are systemic risk measures according to our Definition 2.1.

Proposition 2.58 ([BFFMB20])**.** *Consider the mapping in* (2.56)*.*
(i) *For all* $\bar{X} \in M^\Phi$ $R\left(\bar{X}\right) > -\infty$.
(ii) *R is a systemic risk measure on* M^Φ *according to Definition 2.1.*
(iii) $dom(R) = M^\Phi$, *i.e.,* $R\colon M^\Phi \to \mathbb{R}$.
(iv) *R is continuous and sub-differentiable on* $dom(R) = M^\Phi$.
(v) *R satisfies (S8) on* $\mathcal{C}$.

Before we move on to the proof of the previous Proposition, we need the following Lemma which provides some estimates for the loss functions l_i.

Lemma 2.59 ([BFFMB20]). *The loss functions connected to the aggregation rule have the following properties under the standing assumption of this paragraph: There exist constants $b \in \mathbb{R}_+$ and $c, d \in \mathbb{R}$ such that for all $i \in \{1, \dots, n\}$*

(i) $l_i(x) \geq bx + c$ for all $x \leq 0$,

(ii) $l_i(x) \geq 2bx + d$ for all $x \geq 0$.

Proof. First of all, since all the functions are strictly convex they are not constant and hence unbounded. Thus, by the rule of de l'Hospital and differentiability we obtain for all $i \in \{1, \dots, n\}$

$$\lim_{x \to \infty} \frac{l_i(x)}{x} = \lim_{x \to \infty} l_i'(x) = \infty.$$

From the convexity of each l_i there exist constants $c_i \in \mathbb{R}$ such that

$$l_i(x) \geq l_i'(0)\, x + c_i$$

for all $x \in \mathbb{R}$. If we set $b := \min_{i \in \{1,\dots,n\}} \{l_i(0)\}$ and $c := \min_{i \in \{1,\dots,n\}} \{c_i\}$ we obtain (i). Now the assumption $\lim_{x \to \infty} \frac{l_i(x)}{x} \to \infty$, for all $i \in \{1, \dots, n\}$, implies that for each $M \in \mathbb{R}$ there exists a constant $K > 0$, depending on M, such that

$$l_i(x) \geq Mx,$$

for all $x \geq K$. Therefore $l_i(x) - Mx \geq 0$ for all $x \in [K, \infty)$. Since $l_i(x) - Mx$ is continuous on $[0, K]$, we can subtract a constant $d \in \mathbb{R}$ such that $l_i(x) - Mx - d \geq 0$ for all $x \in [0, \infty)$. Choosing $M = 2b$ concludes the proof. □

Proof of Proposition 2.58. The proofs of (i)-(iv) are stated in [BFFMB20] Proposition 2.4. Let us reproduce their statements in our framework. In order to prove (i) we assume that for some $\bar{X} \in M^\Phi$ we have $R\left(\bar{X}\right) = -\infty$. This readily implies the existence of a monotone decreasing sequence $\left(\bar{Y}^m\right)_{m \in \mathbb{N}} \subseteq \mathcal{C}$ with $\sum_{i=1}^n Y_i^m \to -\infty$. As a direct consequence, we also have $\sum_{i=1}^n \mathbb{E}\left[Y_i^m\right] \to -\infty$. Now, by assumption and Jensens inequality, for all $m \in \mathbb{N}$, we have

$$B \geq \mathbb{E}\left[\sum_{i=1}^n l_i(X_i - Y_i)\right] \geq \sum_{i=1}^n l_i\left(\mathbb{E}\left[X_i\right] - \mathbb{E}\left[Y_i^m\right]\right).$$

Let us denote $x_i^m := \mathbb{E}\left[Y_i^m\right]$. Then $\sum_{i=1}^n x_i^m \to -\infty$ implies the existence of an index $n_0 \in \{1, \dots, n\}$ and a subsequence $\left(\bar{x}^{m_k}\right)_{k \in \mathbb{N}}$ with $x_{n_0}^{m_k} \to -\infty$, where the convergence is monotone. If there is an other index $j \in \{1, \dots, n\} \setminus \{n_0\}$ with $\liminf x_j^{m_k} = -\infty$, we move on to another subsequence of $(x^{m_k})_{k\in\mathbb{N}}$, where the coordinate sequences for n_0 and j converge monotone to $-\infty$. This procedure is realized at most $n-1$ times. Afterwards, we obtain a subsequence $\left(\bar{y}^l\right)$ of our original sequence $(\bar{x}^m)_{m\in\mathbb{N}}$, and three disjoint sets of coordinates N_-, N_+ and $N_\star$ such that

$$
\begin{aligned}
&y_i^l \downarrow -\infty \text{ if } i \in N_- \subseteq \{1,\dots,n\}\,,\\
&\limsup y_i^l = \infty \text{ if } i \in N_+ \subset \{1,\dots,n\}\,,\\
&|y_i^l| \leq K \text{ for all } l \in \mathbb{N} \text{ and all } i \in N_\star = \{1,\dots,n\} \setminus (N_- \cup N_+)\,,
\end{aligned}
$$

where $K \in \mathbb{R}$ is a constant independent of l. Possibly, we have $N_+ = \emptyset$ or/and $N_\star = \emptyset$, but we definitely have $N_- \neq \emptyset$. Since by assumption $\sum_{i=1}^n y_i^m \to -\infty$, for sufficiently large $m \in \mathbb{N}$, we have $\sum_{i=1}^n y_i^m \leq 0$. So, for fixed large m, we obtain

$$
\sum_{i\in N_+} y_i^m \leq -\sum_{i\in N_-} y_i^m - \sum_{i\in N_\star} y_i^m \leq -\sum_{i\in N_-} y_i^m + nK. \tag{2.60}
$$

This fact, together with Lemma 2.59, finally allows the following computations for sufficiently large m:

$$
\begin{aligned}
\sum_{i=1}^n l_i\left(\mathbb{E}\left[X_i\right] - y_i^m\right) &= \sum_{i\in N_+} l_i\left(\mathbb{E}\left[X_i\right] - y_i^m\right) + \sum_{i\in N_-} l_i\left(\mathbb{E}\left[X_i\right] - y_i^m\right) + \sum_{i\in N_\star} l_i\left(\mathbb{E}\left[X_i\right] - y_i^m\right)\\
&\geq C - \sum_{i\in N_+} b y_i^m - \sum_{i\in N_-} 2b y_i^m - \sum_{i\in N_\star} l_i\left(K\right)\\
&\geq C + \sum_{i\in N_-} b y_i^m - \sum_{i\in N_-} 2b y_i^m - bnK\\
&= C - b\sum_{i\in N_-} y_i^m,
\end{aligned}
$$

where the constant C changes over lines but stays independent of m. Since $\sum_{i=1}^n y_i^m \to -\infty$, $\sum_{i=1}^n l_i\left(\mathbb{E}\left[X_i\right] - y_i^m\right)$ is unbounded which is a contradiction, and hence $R\left(\bar{X}\right) \neq -\infty$ for all $\bar{X} \in M^\Phi$.

Now, let us move on to the properness of R. We show (iii) directly, i.e. $\operatorname{dom}(R) = M^\Phi$. So let $\bar{X} \in M^\Phi$. Then for $m \to \infty$ we have $\bar{X} - m1_n \downarrow -\infty$ $\mathbb{P}$-almost surely. Now, we have $\Lambda\left(\bar{X}\right) - \Lambda\left(\bar{X} - m1_n\right) \geq 0$ and $\Lambda\left(\bar{X}\right) - \Lambda\left(\bar{X} - m1_n\right) \uparrow \Lambda\left(\bar{X}\right) - \Lambda\left(-\infty\right)$. So, the monotone convergence theorem implies $\mathbb{E}\left[\Lambda\left(\bar{X} - m1_n\right)\right] \downarrow \Lambda\left(-\infty\right) < B$. Since $m1_n \in \mathbb{R}^n \subseteq \mathcal{C}$, the set $\left\{\bar{Y} \in \mathcal{C} \mid \mathbb{E}\left[\sum_{i=1}^n l_i\left(X_i - Y_i\right)\right] \leq B\right\}$ is nonempty which directly implies $R\left(\bar{X}\right) < \infty$. Monotonicity (S1) and outcome convexity (S2a) are now obvious, which completes (ii).

(iv) is exactly the result of the extended version of the Namioka-Klee Theorem presented in Theorem A.38 and Theorem A.47. (S8), and hence (v), immediately follows from the fact that $\mathcal{C}$ is a convex cone and the translation property of the infimum. $\square$

The set $\mathcal{C}$ of admissible (random) allocations is a key feature of these systemic risk measures. It gives the regulator another instrument to implement the structure of the system or to simply restrict the possible operations.

Example 2.61. *Suppose that the participants of the financial system are divided into $h \in \{1, \dots, n\}$ groups. This means if we set $\bar{n} = (n_1, \dots, n_h) \in \mathbb{N}^h$ with $n_{j-1} < n_j$, $j = 1 \dots, h$, $n_0 := 0$ and $n_h := n$, group j consists of the firms $I_j := \{n_{j-1} + 1, \dots, n_j\}$ for $j = 1 \dots, h$. Consider the set $\mathcal{C}^{(\bar{n})} = \mathcal{C}_0^{(\bar{n})} \cap M^\Phi$, where*

$$\mathcal{C}_0^{(\bar{n})} = \left\{ \bar{Y} \in \left(L^0\right)^n \mid \exists\, \bar{d} \in \mathbb{R}^h : \sum_{i \in I_j} Y_i = d_j \text{ for } j = 1, \dots, h \right\} \subseteq \mathcal{C}(\mathbb{R}) . \tag{2.62}$$

The random vectors in $\mathcal{C}^{(\bar{n})}$ set a deterministic value for the allocation to each of the h groups. Inside each group the allocation of this fixed value is dependent on the occurrent scenario. Obviously, there are two extreme cases, i.e. $h = 1$ and $h = n$. The case $h = 1$ leads to arbitrary random allocations with the only constraint $\bar{Y} \in \mathcal{C}(\mathbb{R})$. The case $h = n$ leads to fully deterministic allocations.

Of course, systemic risk measures given by (2.56) admit a refined dual representation.

Theorem 2.63 ([BFFMB20])**.** *For all $\bar{X} \in M^\Phi$,*

$$R\left(\bar{X}\right) = \max_{\Xi \in \mathcal{D}'} \left\{ \langle \bar{X}, \Xi \rangle_n - \alpha^R(\Xi) \right\}, \tag{2.64}$$

where the penalty function is given by

$$\alpha^R(\Xi) = \sup_{\bar{Z} \in \mathcal{A}} \left\{ \langle \bar{Z}, \Xi \rangle_n \right\}, \quad \Xi \in \partial R\left(\bar{X}\right), \tag{2.65}$$

with $\mathcal{A} := \left\{ \bar{Z} \in M^\Phi \mid \sum_{i=1}^n \mathbb{E}\left[l_i(Z_i)\right] \leq B \right\}$. The maximum in (2.64) is attained for $\Xi \in \partial R\left(\bar{X}\right)$. Moreover, all $\Xi \in \partial R\left(\bar{X}\right)$ satisfies $\Xi \in L_+^{\Phi^*}$, $\langle 1, \xi_i \rangle = 1$ for all $i \in \{1, \dots, n\}$ and $\langle \bar{Y}, \Xi \rangle_n - \sum_{i=1}^n Y_i \leq 0$ for all $\bar{Y} \in \mathcal{C}$.

(i) If we additionally have $\pm\left(e_i \mathbb{1}_A - e_j \mathbb{1}_A\right) \in \mathcal{C}$ for some $i, j \in \{1, \dots, n\}$, $i \neq j$ and all $A \in \mathcal{F}$, then all $\Xi \in \partial R\left(\bar{X}\right)$ satisfies $\xi_i = \xi_j$.

(ii) If we additionally have $\pm\left(e_i \mathbb{1}_A - e_j \mathbb{1}_A\right) \in \mathcal{C}$ for all $i, j \in \{1, \dots, n\}$ and all $A \in \mathcal{F}$, then all $\Xi \in \partial R\left(\bar{X}\right)$ satisfies $\xi_i = \xi_j$ for all $i, j \in \{1, \dots, n\}$ and $\langle \bar{Y}, \Xi \rangle_n = \sum_{i=1}^n Y_i$ for all $\bar{Y} \in \mathcal{C}$.

Proof. A proof is given in [BFFMB20] Proposition 3.1. However, let us give an alternative proof, which only relies on statements presented earlier in this paper.

First of all, by Proposition 2.58 R is a systemic risk measure which is continuous and therefore l.s.c. Hence, Theorem 2.8 yields the claimed representation with a supremum instead of a maximum. But Proposition 2.58 also yields that R is sub-differentiable, i.e. $\partial R\left(\bar{X}\right) \neq \emptyset$. So, the supremum is attained for every $\Xi \in \partial R\left(\bar{X}\right)$ which also shows that it is indeed a maximum. Moreover, Theorem 2.8 also yields that all feasible solutions Ξ satisfy $\Xi \in L_+^{\Phi^*}$ due to monotonicity and $\langle 1, \xi_i \rangle = 1$, for all $i \in \{1, \dots, n\}$, due to the translation property. Now, for $\Xi \in \partial R\left(\bar{X}\right)$ and an arbitrary $\bar{Y} \in \mathcal{C}$, we also have

$$\begin{aligned}
\alpha^R(\Xi) &\geq \langle \bar{X} + \bar{Y}, \Xi \rangle_n - R\left(\bar{X} + \bar{Y}\right) \\
&= \langle \bar{Y}, \Xi \rangle_n - \sum_{i=1}^n Y_i + \langle \bar{X}, \Xi \rangle_n - R\left(\bar{X}\right) \\
&\geq \langle \bar{Y}, \Xi \rangle_n - \sum_{i=1}^n Y_i + \alpha^R(\Xi),
\end{aligned}$$

where the last inequality follows from $\Xi \in \partial R\left(\bar{X}\right)$. Since $\Xi \in \partial R\left(\bar{X}\right)$ and $\bar{Y} \in \mathcal{C}$ were chosen arbitrarily, $\langle \bar{Y}, \Xi \rangle_n - \sum_{i=1}^n Y_i \leq 0$ holds for all $\Xi \in \partial R\left(\bar{X}\right)$ and all $\bar{Y} \in \mathcal{C}$. Now, this property together with $\pm\left(e_i \mathbb{1}_A - e_j \mathbb{1}_A\right) \in \mathcal{C}$ for some $i, j \in \{1, \dots, n\}$, $i \neq j$, and all $A \in \mathcal{F}$ immediately implies $\xi_i = \xi_j$ for $\Xi \in \partial R\left(\bar{X}\right)$. This shows (i). (ii) is then obvious. It remains to show the explicit form of the penalty function. To this end, we first notice that

$$\begin{aligned}
\alpha^R(\Xi) &= \sup_{\bar{X} \in M^\Phi} \left\{ \langle \bar{X}, \Xi \rangle_n - R\left(\bar{X}\right) \right\} \\
&\geq \sup_{\bar{Z} \in \mathcal{A}} \left\{ \langle \bar{Z}, \Xi \rangle_n - R\left(\bar{Z}\right) \right\} \\
&\geq \sup_{\bar{Z} \in \mathcal{A}} \left\{ \langle \bar{Z}, \Xi \rangle_n \right\},
\end{aligned}$$

where the last inequality follows from the fact that for $\bar{Z} \in \mathcal{A}$ we have $R\left(\bar{Z}\right) \leq 0$. To establish the reverse inequality, let $\Xi \in \partial R\left(\bar{X}\right)$ and observe that for all $\bar{X} \in M^\Phi$ and $\bar{Y}^{\bar{X}} \in \left\{ \bar{Y} \in \mathcal{C} \mid \Lambda\left(\bar{X} - \bar{Y}\right) \in \mathcal{A}_0 \right\} \neq \emptyset$ we have $\bar{X} - \bar{Y}^{\bar{X}} \in \mathcal{A}$. Therefore, we have

$$\begin{aligned}
\sup_{\bar{Z} \in \mathcal{A}} \left\{ \langle \bar{Z}, \Xi \rangle_n \right\} &\geq \sup_{\bar{X} \in M^\Phi} \left\{ \langle \bar{X} - \bar{Y}^{\bar{X}}, \Xi \rangle_n \right\} \\
&\geq \sup_{\bar{X} \in M^\Phi} \left\{ \langle \bar{X}, \Xi \rangle_n - \sum_{i=1}^n Y_i^{\bar{X}} \right\},
\end{aligned}$$

where the last inequality follows from $\langle \bar{Y}, \Xi\rangle_n - \sum_{i=1}^n Y_i \leq 0$. Taking the supremum over all $\bar{Y}^{\bar{X}} \in \left\{\bar{Y} \in \mathcal{C} \mid \Lambda\left(\bar{X} - \bar{Y}\right) \in \mathcal{A}_0\right\}$ finally yields the desired representation. □

The explicit structure allows us to give an estimate for the value of the penalty function.

Proposition 2.66 ([BFFMB20]). *Consider the penalty function α^R from (2.65). It holds*

$$\alpha^R(\Xi) := \sup_{\bar{Z} \in \mathcal{A}} \left\{\langle \bar{Z}, \Xi\rangle_n\right\} \leq \inf_{\lambda > 0} \left\{\frac{1}{\lambda} B + \frac{1}{\lambda} \sum_{i=1}^n \mathbb{E}\left[l_i^*\left(\lambda \xi_i\right)\right]\right\}. \tag{2.67}$$

Proof. To keep the notation simple, we will use the pairing notation throughout this proof. This means $\mathbb{E}[X] = \langle X, 1\rangle$. First of all, we notice that

$$\bar{Z} \in \mathcal{A} \quad \Leftrightarrow \quad f\left(\bar{Z}\right) := \sum_{i=1}^n \langle l_i(Z_i), 1\rangle - B \leq 0.$$

The mapping f is convex and increasing. By the Young-Fenchel inequality (A.31), we obtain for all $\bar{Z} \in \mathcal{A}$ and all $\lambda > 0$

$$\begin{aligned}\langle \bar{Z}, \Xi\rangle_n &= \frac{1}{\lambda}\langle \bar{Z}, \lambda\Xi\rangle_n \\ &\leq \frac{1}{\lambda}\left(f\left(\bar{Z}\right) + f^*(\lambda\Xi)\right) \\ &\leq \frac{1}{\lambda} f^*(\lambda\Xi).\end{aligned}$$

This yields

$$\alpha^R(\Xi) := \sup_{\bar{Z} \in \mathcal{A}} \left\{\langle \bar{Z}, \Xi\rangle_n\right\} \leq \inf_{\lambda > 0}\left\{\frac{1}{\lambda} f^*(\lambda\Xi)\right\}.$$

For f^* we make the following observation:

$$\begin{aligned}f^*(\Xi) &= \sup_{\bar{Z} \in M^\Phi} \left\{\langle \bar{Z}, \Xi\rangle_n - f\left(\bar{Z}\right)\right\} \\ &= B + \sup_{\bar{Z} \in M^\Phi} \left\{\sum_{i=1}^n \langle Z_i, \xi_i\rangle - \langle l_i(Z_i), 1\rangle\right\} \\ &= B + \sum_{i=1}^n \sup_{Z \in M^{\phi_i}} \left\{\langle Z, \xi_i\rangle - \langle l_i(Z), 1\rangle\right\}.\end{aligned}$$

Now, for all $i \in \{1, \ldots, n\}$ $y \in \mathbb{R}$, we have

$$l_i^*(y) \geq xy - l_i(x) \qquad \text{for all } x \in \mathbb{R}.$$

So, for all $\xi \in L^{\phi_i^*}$, $Z \in M^{\phi_i}$ and all $\omega \in \Omega$,

$$l_i^*(\xi(\omega)) \geq Z(\omega)\,\xi(\omega) - l_i(Z(\omega))\,.$$

This readily implies

$$B + \sum_{i=1}^{n} \langle 1, l_i^*(\xi_i)\rangle \geq f^*(\Xi)$$

and completes the proof. $\square$

Remark 2.68. *The original Proposition from [BFFMB20] (Proposition 3.4) states even more. For dual objects* Ξ *with* $\alpha^R(\Xi) < \infty$*, the inequality in* (2.67) *becomes an equality. Moreover* $\mathbb{E}[l_i^*(\lambda\xi_i)] < \infty$ *for all* i *and all* $\lambda > 0$ *and the infimum is attained, i.e.,*

$$\begin{aligned} \alpha^R(\Xi) &= \sum_{i=1}^{n} \langle (l_i^*)'(\lambda^*\xi_i), \xi_i\rangle \\ &= \sum_{i=1}^{n} \mathbb{E}\left[\xi_i (l_i^*)'(\lambda^*\xi_i)\right], \end{aligned} \tag{2.69}$$

where $\lambda^* > 0$ *is the unique solution of the equation*

$$B + \sum_{i=1}^{n} \mathbb{E}[l_i^*(\lambda\xi_i)] - \lambda \sum_{i=1}^{n} \mathbb{E}\left[\xi_i (l_i^*)'(\lambda\xi_i)\right] = 0.$$

For our purposes, we only need the above stated and proven inequality. The full proof can be found in [BFFMB20]. The situation $n = 1$ *was already discussed in [FS08] Theorem 4.115.*

The main benefit of the refined dual representation is that it enables us to verify more properties of systemic risk measures of type (0.3). The following results states more commonalities between systemic risk measures of type (0.1) and (0.3).

Theorem 2.70. *Suppose that for all* $\bar{X} \in M^{\Phi}$ *every* $\Xi \in \partial R\left(\bar{X}\right)$ *satisfies* $\xi_i = \xi_j$ *for all* $i, j \in \{1, \ldots, n\}$*.*

(i) The systemic risk measure R*, defined in* (2.56)*, satisfies the properties (S2b), (S4) and (S5).*

(ii) If $R\,|_{\mathbb{R}^n}\left(M^{\Phi}\right) = \mathcal{X}$ *for some locally convex Hausdorff space* $\mathcal{X}$*, we have*

$$R\left(\bar{X}\right) = \rho_0 \circ \Lambda^{sum,c},$$

for a convex single-firm risk measure $\rho_0 \colon \mathcal{X} \to \mathbb{R}$ *that satisfies the constancy property (R4) on* $\mathbb{R}$*.*

Proof. **(i):** By assumption, we have $\Xi = 1_n \xi$ for some suitable probability density ξ. We denote the corresponding probability measure by $\mathbb{Q}$. Suppose that

$$R\left(\bar{Z}(\omega)\right) = \alpha R\left(\bar{X}(\omega)\right) + (1-\alpha) R\left(\bar{Y}(\omega)\right), \qquad \alpha \in [0,1]$$

for almost all $\omega \in \Omega$. Since R admits a dual representation given by (2.64) we obtain

$$\begin{aligned} R\left(\bar{Z}(\omega)\right) &= \max_{\Xi \in \mathcal{D}} \left\{ \sum_{i=1}^{n} Z_i(\omega) - \alpha^R(\Xi) \right\} \\ &= \sum_{i=1}^{n} Z_i(\omega) + \max_{\Xi \in \mathcal{D}} \left\{ -\alpha^R(\Xi) \right\} \end{aligned}$$

and

$$\begin{aligned} \alpha R\left(\bar{X}(\omega)\right) + (1-\alpha) R\left(\bar{Y}(\omega)\right) &= \alpha \max_{\Xi \in \mathcal{D}} \left\{ \sum_{i=1}^{n} X_i(\omega) - \alpha^R(\Xi) \right\} \\ &\quad + (1-\alpha) \max_{\Xi \in \mathcal{D}} \left\{ \sum_{i=1}^{n} Y_i(\omega) - \alpha^R(\Xi) \right\} \\ &= \alpha \sum_{i=1}^{n} X_i(\omega) + (1-\alpha) \sum_{i=1}^{n} Y_i(\omega) + \max_{\Xi \in \mathcal{D}} \left\{ -\alpha^R(\Xi) \right\}. \end{aligned}$$

But this means

$$\sum_{i=1}^{n} Z_i = \alpha \sum_{i=1}^{n} X_i + (1-\alpha) \sum_{i=1}^{n} Y_i$$

$\mathbb{P}-$almost surely and since $\mathbb{Q} \ll \mathbb{P}$ also $\mathbb{Q}-$almost surely. Therefore,

$$\begin{aligned} R\left(\bar{Z}\right) &= \max_{\Xi \in \mathcal{D}} \left\{ \langle \sum_{i=1}^{n} Z_i, \xi \rangle - \alpha^R(\Xi) \right\} \\ &= \max_{\Xi \in \mathcal{D}} \left\{ \langle \alpha \sum_{i=1}^{n} X_i + (1-\alpha) \sum_{i=1}^{n} Y_i, \xi \rangle - \alpha^R(\Xi) \right\} \\ &= R\left(\alpha \bar{X} + (1-\alpha) \bar{Y}\right) \\ &\leq \alpha R\left(\bar{X}\right) + (1-\alpha) R\left(\bar{Y}\right), \end{aligned}$$

where the last inequality follows from (S2a). But this means R satisfies (S2b). To show preference constancy (S4) suppose that

$$R\left(\bar{X}(\omega)\right) \geq R\left(\bar{Y}(\omega)\right)$$

for almost all $\omega \in \Omega$. With the same arguments as stated above we obtain

$$\sum_{i=1}^{n} X_i \geq \sum_{i=1}^{n} Y_i$$

$\mathbb{P}-$almost surely and since $\mathbb{Q} \ll \mathbb{P}$ also $\mathbb{Q}-$almost surely. But this means

$$\begin{aligned} R\left(\bar{X}\right) &= \max_{\Xi \in \mathcal{D}} \left\{ \langle \sum_{i=1}^{n} X_i, \xi \rangle - \alpha^R\left(\Xi\right) \right\} \\ &\geq \max_{\Xi \in \mathcal{D}} \left\{ \langle \sum_{i=1}^{n} Y_i, \xi \rangle - \alpha^R\left(\Xi\right) \right\} \\ &= R\left(\bar{Y}\right). \end{aligned}$$

The $\mathcal{R}$-surjectivity property (S5) follows with the same arguments.
(ii): Together with part (i), R satisfies all the assumptions of Theorem 2.33. Theorem 2.63 yields that for all $\bar{x} \in \mathbb{R}^n$ $R(\bar{x}) = \Lambda^{\text{sum,c}}$, where $c = \min_{\Xi \in \mathcal{D}} \left\{ \alpha^R\left(\Xi\right) \right\}$. Now, the constancy property (R4) is fulfilled on $\Lambda^{\text{sum,c}}\left(\mathbb{R}^n\right) = \mathbb{R}$. ☐

Remark 2.71.

1. *The property $R|_{\mathbb{R}^n}\left(M^{\Phi}\right) = \mathcal{X}$ for some locally convex Hausdorff space $\mathcal{X}$ holds for example if $\phi_i \succ \phi_j$ for all $i \neq j$, where $\phi_i \succ \phi_j$ if*

$$\phi_j(x) \leq b\phi_i(ax) \qquad \forall x \geq 0$$

for some $b > 0$ and $a > 0$. In this case, we have $M^{\phi_i} \subseteq M^{\phi_j}$ (see [RGMP16] Theorem 16.2.1.) and therefore $R|_{\mathbb{R}^n}\left(M^{\Phi}\right) = M^{\phi_j}$.

2. *$\pm\left(e_i \mathbb{1}_A - e_j \mathbb{1}_A\right) \in \mathcal{C}$ for all $i, j \in \{1, \ldots, n\}$ and all $A \in \mathcal{F}$ presents a sufficient condition for the crucial assumption in Theorem 2.70. The economic interpretation of this property is that capital transfers between the firms is an accepted instrument for the regulator to ensure the stability of the system. Obviously, the task of capital allocation exactly works that way. A situation where the regulator does not have this instrument would imply that the subgroups are regulated in an isolated manner, e.g. insurances are regulated ignoring what is happening in the banking sector. Obviously, the assumption is always fulfilled if we set $\mathcal{C} = \mathcal{C}(\mathbb{R})$. According to the definition of the systemic risk measure R, in the least conservative or least restrictive situation we are always able to apply Theorem 2.70.*

3. *In [BFFMB19] Theorem 4.2, the authors already pointed out that even in the more general setup without the running Assumption of this section mappings of type* (0.3) *are compositions of Λ^{sum} and a general quasi-convex mapping if $\mathcal{C} = \mathcal{C}(\mathbb{R})$. In [DFG24] this result is studied for the corresponding shortfall type risk measures. They also provide an explicit formula for the decomposition*

and cover more cases than the ones captured by Theorem 2.70. However, the interesting observation in this Theorem is not the fact that a decomposition exists. It is the fact that the components fit in the framework of [KOZ16].

Let us continue with the master example of these systemic risk measures presented in [BFFMB20].

Example 2.72. *Set $\mathcal{C} = \mathcal{C}^{\bar{n}}$ and $l_i(x) = \frac{1}{\alpha_i}\exp(\alpha_i x)$, $\alpha_i > 0$, $i \in \{1, \dots, n\}$. This means $\phi_i(x) = \frac{1}{\alpha_i}(\exp(\alpha_i |x|) - 1)$. Now we have $\phi^{\exp} \succ \phi_i$ and $\phi_i \succ \phi^{\exp}$ with $\phi^{\exp}(x) = \exp(|x|) - 1$ for all $i \in \{1, \dots, n\}$ and therefore $M^\Phi \equiv (M^{\exp})^n$. Additionally, set $B > \sum_{i=1}^n l_i(-\infty) = 0$. The systemic risk measure $R^{ex} : (M^{\exp})^n \to \mathbb{R}$ becomes*

$$R^{ex}\left(\bar{X}\right) = \inf\left\{\sum_{i=1}^n Y_i \mid \bar{Y} \in \mathcal{C}^{\bar{n}}, \mathbb{E}\left[\sum_{i=1}^n \frac{1}{\alpha_i}\exp\left(\alpha_i\left(X_i - Y_i\right)\right)\right] = B\right\}. \tag{2.73}$$

Now, for all $\bar{X} \in (M^{\exp})^n$, R^{ex} is Gâteaux differentiable with derivative $\nabla R^{ex}\left(\bar{X}\right)$ at $\bar{X}$, where for $j \in \{1, \dots, h\}$ the components $l \in I_j$ are given by

$$\left(\nabla R^{ex}\left(\bar{X}\right)\right)_l = \frac{\exp\left(\theta_j \sum_{i \in I_j} X_i\right)}{\mathbb{E}\left[\exp\left(\theta_j \sum_{i \in I_j} X_i\right)\right]}, \tag{2.74}$$

with $\theta_j = \frac{1}{\sum_{i \in I_j} \frac{1}{\alpha_i}}$. So $\nabla R^{ex}\left(\bar{X}\right)$ is the unique optimal solution to the dual problem (2.64). On the other hand, Proposition 2.66 enables us to give an estimate for the penalty function. We have $l_i^*(y) = \frac{y}{\alpha_i}(\ln(y) - 1)$ and the infimum is attained for

$$\lambda^* = \theta B \tag{2.75}$$

Hence,

$$\begin{aligned}\alpha^R(\Xi) &\leq \langle \bar{Z}^*, \nabla R^{ex}\left(\bar{X}\right)\rangle_n \\ &:= \sum_{i=1}^n \langle \frac{1}{\alpha_i}\left(\ln\left(\left(\theta B \nabla R^{ex}\left(\bar{X}\right)\right)_i\right)\right), \left(\nabla R^{ex}\left(\bar{X}\right)\right)_i\rangle \\ &= \sum_{i=1}^n \frac{1}{\alpha_i}\left(H\left(\mathbb{Q}_i^{\bar{X}} \mid \mathbb{P}\right) + \ln(\theta B)\right)\end{aligned}$$

where $\mathbb{Q}_i^{\bar{X}}$ is the probability measure with density $\left(\nabla R^{ex}\left(\bar{X}\right)\right)_i$. But $\bar{Z}^* \in \mathcal{A}$, so the inequality is indeed an equality. Additionally if the system consists of exactly one group, the set $\mathcal{C}$ satisfies all the assumptions of Theorem 2.70. Since $R^{ex}\mid_{\mathbb{R}^n}((M^{\exp})^n) = M^{\exp}$ we are able to apply Theorem 2.33. But the previous observations *immediately yield*

$$R^{ex}\left(\bar{X}\right)=\frac{1}{\theta}\ln\mathbb{E}\left[\exp\left(\theta\sum_{i=1}^{n}X_i\right)\right]-c=\left(\rho_0^{entr}\circ\Lambda^{sum,c}\right)\left(\bar{X}\right)=\rho^{ses,c}\left(\bar{X}\right)$$

for $c=\frac{1}{\theta}\ln(\theta B)$. *In other words, the systemic risk measure* R^{ex} *is equivalent to a version of the entropic systemic risk measure presented in Example 2.39 (i). In the situation where more groups are relevant, we use the structure of the penalty function to simply split the system into these groups. Then, we compute the risk of each group and add the risk up to the risk of the whole system. More precisely, we set* $B_j=\frac{\theta}{\theta_j}B$ *and* $\mathcal{C}^{I_j}:=\pi_{I_j}\circ\mathcal{C}^{\bar{n}}=\left\{\bar{Y}\in\bigtimes_{i\in I_j}M^{\phi_i}\mid\sum_{i\in I_j}Y_i\in\mathbb{R}\right\}$ *for* $j\in\{1,\ldots,h\}$. *Now, we consider the systemic risk measures*

$$R_j^{ex}\left(\bar{Z}\right)=\inf\left\{\sum_{i\in I_j}Y_i\mid\bar{Y}\in\mathcal{C}^{I_j},\mathbb{E}\left[\sum_{i\in I_j}\frac{1}{\alpha_i}\exp\left(\alpha_i\left(Z_i-Y_i\right)\right)\right]=B_j\right\},$$

for $\bar{Z}\in\bigtimes_{i\in I_j}M^{\phi_i}$. *Each systemic risk measure* R_j^{ex} *for subgroup* j *uses the set* $\mathcal{C}^{I_j}$ *which again satisfies all the assumptions of Theorem 2.70 and the additional property highlighted in Remark 2.71. Therefore, we are able to decompose all* R_j^{ex}. *So, for* $\bar{X}\in M^{\Phi}$, *let* $\pi_{I_j}\circ\bar{X}=\left(X_{\tilde{n}_{j-1}},\ldots,X_{\tilde{n}_j}\right)$. *Then, we have*

$$R^{ex}\left(\bar{X}\right)=\sum_{j=1}^{h}R_j^{ex}\left(\pi_{I_j}\circ\bar{X}\right)=\sum_{j=1}^{h}\rho_j^{ses,c}\left(\pi_{I_j}\circ\bar{X}\right).$$

The interpretation of this property is that the impact between the groups is fully determined through their risk aversion coefficients. Additionally, the system-wide threshold B effects every group. In other words, the potential outcomes of each group does not affect the risk of the other groups.

3. Capital Allocation Rules for Systemic Risk Measures

Once the systemic risk is computed, the subsequent challenging task for the regulator is to *fairly* allocate the risk to the constituent parts of the system. For systemic risk measures of type (0.1), the allocation problem is broadly studied in the specific case where $\Lambda = \Lambda^{sum}$. For coherent single-firm risk measures [Kal05] and [Tas07] provided solutions to the problem. More precisely, [Kal05] proved the existence of capital allocation rules with desirable properties for coherent single-firm risk measures. In addition, he pointed out under which additional conditions the capital allocation rule is a generalized version of the gradient allocation also studied in [Tas07] under the name *Euler allocation*. From a game theoretic viewpoint, this allocation is derived by [Den01]. In the convex case [Tsa09] used some of the ideas presented in [Den01] and proposed a capital allocation rule in the spirit of the Aumann-Shapley value. This capital allocation rule reduces to the gradient allocation if the single-firm risk measure is coherent. Therefore it is a generalization of the very. An other interesting approach in the convex case is given in [CG18]. Their capital allocation rules are connected to the dual representation of the underlying convex single-firm risk measure. For more general aggregation rules [CIM13] and [KOZ16] address this problem. For systemic risk measures of type (0.3) solutions are presented in [BFFMB20]. We will discuss these ideas and present a general approach for systemic capital allocation rules. This approach contains the previously mentioned contributions to the theory and enables us to link certain special cases. Let us start by giving a formal definition of a capital allocation rule for systemic risk measures.

Definition 3.1. *Given a systemic risk measure ρ on $\bar{\mathcal{X}}$, a systemic capital allocation rule is a mapping $CS : \bar{\mathcal{X}} \times \bar{\mathcal{X}} \to \mathbb{R}$ such that $CS(\bar{X}; \bar{X}) = \rho(\bar{X})$ for all $\bar{X} \in \bar{\mathcal{X}}$. For a fixed system $\bar{X} \in \bar{\mathcal{X}}$, a systemic capital allocation rule for ρ at $\bar{X}$ (or local systemic capital allocation rule) is a mapping $CS_{\bar{X}} : \bar{\mathcal{X}} \to \mathbb{R}$ with $CS_{\bar{X}}\left(\bar{X}\right) = \rho\left(\bar{X}\right)$.*

With this in mind, one can think of a systemic capital allocation rule as family of mappings $(CS_{\bar{X}})_{\bar{X} \in \bar{\mathcal{X}}}$ where each mapping is a systemic capital allocation rule for ρ at $\bar{X}$ and for all $\bar{X}, \bar{Z} \in \bar{\mathcal{X}}$ we simply set

$$CS\left(\bar{Z}; \bar{X}\right) = CS_{\bar{X}}\left(\bar{Z}\right).$$

Now, $CS(\bar{Z}; \bar{X})$ describes the portion of risk carried by $\bar{Z}$ considered as an arbitrary sub-system of $\bar{X}$. The condition $CS(\bar{X}; \bar{X}) = \rho(\bar{X})$ means that the capital allocated to the whole system $\bar{X}$ is exactly the risk capital of $\bar{X}$, i.e. $\rho(\bar{X})$.

Obviously if we can allocate the risk to any arbitrary subsystem $\bar{Z}$, we can also allocate it to $\bar{X} - \bar{Z}$. In this spirit, we can divide the system in $d \in \mathbb{N}$ subsystems and the following question arises: Does our systemic capital allocation rule allocate the whole risk $\rho\left(\bar{X}\right)$? We say a systemic capital allocation rule fulfills the *full allocation property* if the following holds:

(C1) Full allocation: $CS\left(\bar{X};\bar{X}\right) = \sum_{j=1}^{d} CS\left(\bar{Z}_j;\bar{X}\right)$, *for all* $\bar{X}, \bar{Z}_1, \ldots, \bar{Z}_d \in \bar{\mathcal{X}}$, *such that* $\bar{X} = \sum_{i=1}^{d} \bar{Z}_i$.

The full allocation property (C1) represents some kind of fairness condition. On the one hand, the regulator of the system can guarantee that the risk of the system is completely allocated. On the other hand, it is not over conservative in the sense that the overall allocated risk is no greater than the risk that actually occurs. From a mathematical viewpoint, it is just additivity w.r.t. the first component. Note that the full allocation property (C1) already implies

(C1R) Restricted Full allocation: $CS\left(\bar{X};\bar{X}\right) = \sum_{i=1}^{n} CS\left(e_i X_i;\bar{X}\right)$.

This condition represents the most natural demand of full allocation. The systemic risk should be allocated completely to its constituent parts. However, other configurations than the canonical one (i.e. the components of the vector representing the system) may be reasonable to consider while addressing the allocation problem. Unfortunate configurations may lead to stronger penalized risk profiles. Especially if the systemic risk measure is positively homogeneous, we will have full linearity (in the first component)

(C2) Linearity: $CS\left(\alpha_1\bar{Z}_1 + \alpha_2\bar{Z}_2;\bar{X}\right) = \alpha_1 CS\left(\bar{Z}_1;\bar{X}\right) + \alpha_2 CS\left(\bar{Z}_2;\bar{X}\right)$, *for all* $\alpha_1, \alpha_2 \in \mathbb{R}$ *and all* $\bar{X}, \bar{Z}_1, \bar{Z}_2 \in \bar{\mathcal{X}}$.

Another fairness indicator is the so called no-undercut property.

(C3) No-undercut: $CS\left(\bar{Z};\bar{X}\right) \leq \rho\left(\bar{Z}\right)$ *for all* $\bar{X}, \bar{Z} \in \bar{\mathcal{X}}$.

It is some kind of reward for the system $\bar{Z}$ to be part of $\bar{X}$, i.e. the allocated risk is no greater than the standalone risk. This property has a direct link to the subadditivity property (S7) of the underlying systemic risk measure and seems to be natural in this case. Note, that a positively homogeneous systemic risk measure is always subadditive. However, for purely convex systemic risk measures, it is no longer valid that the combination of positions comes in hand with a risk reduction. Therefore, this property is only reasonable if the systemic risk measure is subadditive. This property also appears in [Kal05] under the name *diversifying* and in [Tsa09] under the name *non-split*. All of the above properties can be localized if we fix some system $\bar{X} \in \bar{\mathcal{X}}$. In these situations, we simply say that the localized capital allocation rule fulfills the respective property. The last property, introduced by [Kal05], is a continuity property.

(C4) Continuity (at $\bar{X}$*):* $\lim_{h\to 0} CS\left(\bar{Z};\bar{X}+h\bar{Z}\right) = CS\left(\bar{Z};\bar{X}\right)$ *for fixed* $\bar{X} \in \bar{\mathcal{X}}$ *and all* $\bar{Z} \in \bar{\mathcal{X}}$.

It simply states that the effect of small changes in the system on the allocated risk to its subsystems is also small.

3.1. Dual Representation Capital Allocation Rules

The dual representation of a systemic risk measure gives an idea how the risk is assembled. For the sake of simplicity, let us assume that the risk is measured with a positively homogeneous systemic risk measure. Now, fix some $\bar{X} \in \bar{\mathcal{X}}$ and assume that ρ is Gâteaux differentiable at $\bar{X} \in \bar{\mathcal{X}}$. The main theorem from [Kal05] (Theorem 4.3) tells us that the unique capital allocation rule for ρ at $\bar{X}$ which satisfies Linearity (C2) and No-undercut (C3) is given by

$$CS^G_{\bar{X}}\left(\bar{Z}\right) = \delta\left(\bar{X}, \bar{Z}\right). \tag{3.2}$$

We call this systemic capital allocation rule *gradient* capital allocation rule. Now if ρ is in addition continuous at $\bar{X}$, Corollary 2.14 tells us that $\partial\rho\left(\bar{X}\right) \neq \emptyset$ which means that there are optimal solutions to the dual problem. Additionally, the Gâteaux differential at $\bar{X}$ is continuous and hence ρ is Gâteaux differentiable at $\bar{X}$. Its derivative $\nabla\rho\left(\bar{X}\right)$ is the only element of the subgradient and therefore the unique optimal solution to the dual problem. Now, the unique capital allocation rule for ρ at $\bar{X}$ with Linearity (C2) and No-undercut (C3) takes the form

$$CS^G_{\bar{X}}\left(\bar{Z}\right) = \langle \bar{Z}, \nabla\rho\left(\bar{X}\right)\rangle_n.$$

In the spirit of this result, it is quiet natural to define a systemic capital allocation rule which is connected to optimal solutions of the dual problem. Even if the systemic risk measure is not continuous and not Gâteaux differentiable, the main properties are preserved. However if the systemic risk measure is not positively homogeneous, we have to take care of the additional penalty term in order to guarantee some notion of full allocation. The following proposition subsumes these ideas. It generalizes the contributions of [KOZ16] and [CG18].

Proposition 3.3. *Suppose that $\rho\colon \bar{\mathcal{X}} \to \mathbb{R}\cup\{\infty\}$ is a systemic risk measure which is l.s.c.. Fix some $\bar{X} \in dom(\rho)$ and assume that the corresponding dual problem* (2.9) *has a unique solution $\Xi^{\bar{X}}$ at $\bar{X}$. Then*

$$CS^{\Xi}_{\bar{X}}\left(\bar{Z}\right) = \langle \bar{Z}, \Xi^{\bar{X}}\rangle_n - \mu_{\bar{X}}\left(\bar{Z}\right)\alpha^{\rho}\left(\Xi^{\bar{X}}\right), \tag{3.4}$$

where $\mu_{\bar{X}}\colon \bar{\mathcal{X}} \to \mathbb{R}$ is a weight function with $\mu_{\bar{X}}\left(\bar{X}\right) = 1$, is a systemic capital allocation rule for ρ at $\bar{X}$. If $\mu_{\bar{X}}$ is in addition fully additive w.r.t. the first component, it also fulfills the full allocation property (C1). If $\mu_{\bar{X}}$ is only additive w.r.t. the components of the vector $\bar{X}$, it fulfills the restricted full allocation property (C1R). Moreover if the systemic risk measure ρ is in addition positively homogeneous, (3.4) *indeed reduces to*

$$CS^{\Xi}_{\bar{X}}\left(\bar{Z}\right) = \langle \bar{Z}, \Xi^{\bar{X}}\rangle_n. \tag{3.5}$$

Proof. All claims follow directly from the assumptions on $\mu_{\bar{X}}$. $\square$

We will refer to this type of (local) systemic capital allocation rule as *dual capital allocation rule*. It is worth mentioning, that the uniqueness of the optimal solution is not necessary. However, in general we have different capital allocation rules for different optimal solutions. At this point, one needs reasonable and fair criteria to choose the right capital allocation rule.

As already pointed out after the definition, a systemic capital allocation rule for ρ can be seen as a collection of local versions $CS_{\bar{X}}$ for all $\bar{X}$. Of course, it is possible to have different weight functions for different systems. But it seems to be reasonable, that the same type of weight function is used. With this in mind, we can consider the weight function as a mapping $\mu \colon \bar{\mathcal{X}} \times \bar{\mathcal{X}} \to \mathbb{R}$. For a fixed second component, we obtain the localized version from above. Let us continue by presenting some possible choices for the weight function μ.

$$\mu^1\left(\bar{Z},\bar{X}\right) = 1$$

$$\mu^2\left(\bar{Z},\bar{X}\right) = \frac{\rho\left(\bar{Z}\right)}{\rho\left(\bar{X}\right)}$$

$$\mu^3\left(\bar{Z},\bar{X}\right) = \frac{\rho\left(\bar{X}\right) - \rho\left(\bar{X}-\bar{Z}\right)}{\rho\left(\bar{X}\right)}$$

Note, that $\rho\left(\bar{X}\right) \neq 0$ is required. However, for μ^3 we additionally need $\rho(0) = 0$ to obtain $\mu\left(\bar{X},\bar{X}\right) = 1$. So, for μ^2 and μ^3 the situations where $\rho\left(\bar{X}\right) = 0$ have to be handled separately. A possible way to overcome this miscue is to use an exponential version of these rules.

$$\tilde{\mu}^2\left(\bar{Z},\bar{X}\right) = \frac{\exp\left(\gamma\rho\left(\bar{Z}\right)\right)}{\exp\left(\gamma\rho\left(\bar{X}\right)\right)}$$

$$\tilde{\mu}^3\left(\bar{Z},\bar{X}\right) = \frac{\exp\left(\gamma\left(\rho\left(\bar{X}\right) - \rho\left(\bar{X}-\bar{Z}\right)\right)\right)}{\exp\left(\gamma\rho\left(\bar{X}\right)\right)}$$

In this situation, the exogenous parameter $\gamma > 0$ gives the regulator an additional instrument. It controls the sensitivity of the systemic capital allocation rule for deviations from $\rho\left(\bar{X}\right)$ in a global manner. As we will see later, it is also possible to introduce this type of additional instrument on the level of the participants of the system. The parameters may be motivated by economic reasoning. However, none of these weight functions is additive in its first component in a non trivial case. Therefore, capital allocation rules derived with these weight functions lack any type of full allocation. Some of the following examples for μ were presented in [KOZ16] and are motivated by other works on the allocation task (see [CK11] for more details). Again, it is necessary to guarantee that the denominator is not equal to 0. But they all produce capital allocation rules which at least satisfy the restricted full allocation property (C1R).

$$\mu^4\left(\bar{Z},\bar{X}\right) = \frac{\sum_{i=1}^n \rho\left(e_i Z_i\right)}{\sum_{i=1}^n \rho\left(e_i X_i\right)}$$

$$\mu^5\left(\bar{Z},\bar{X}\right) = \frac{\sum_{i=1}^n \rho\left(\bar{X}\right) - \rho\left(\bar{X} - e_i Z_i\right)}{\sum_{i=1}^n \rho\left(\bar{X}\right) - \rho\left(\bar{X} - e_i X_i\right)}$$

$$\mu^6\left(\bar{Z},\bar{X}\right) = \frac{\sum_{i=1}^n \delta\rho\left(\bar{X}, e_i Z_i\right)}{\sum_{i=1}^n \delta\rho\left(\bar{X}, e_i X_i\right)}$$

$$\mu^7\left(\bar{Z},\bar{X}\right) = \frac{\langle \bar{Z}, \Xi^{\bar{X}} \rangle_n}{\langle \bar{X}, \Xi^{\bar{X}} \rangle_n}$$

If we want to introduce exogenous parameters on the level of the institutions, we can work with a two step procedure. In the first step, the weights for the participants of the system are set as the exogenous parameters, i.e.

$$\gamma_i = \mu^\gamma\left(e_i X_i, \bar{X}\right), \quad \sum_{i=1}^n \gamma_i = 1.$$

In the second step, the weight function for an arbitrary subsystem $\bar{Z}$ can be derived with a comparison of the components of $\bar{Z}$ and $\bar{X}$. In general, we have an additional vector of functions $\nu_i \colon \mathcal{X}_i \times \mathcal{X}_i \to \mathbb{R}$ with $\nu_i\left(X_i, X_i\right) = 1$ for all $i \in \{1, \ldots, n\}$. The weight function now takes the form

$$\mu^\gamma\left(\bar{Z},\bar{X}\right) = \sum_{i=1}^n \gamma_i \nu_i\left(Z_i, X_i\right).$$

The one dimensional versions of μ^4, μ^5, μ^6 and μ^7 present suitable choices for the mappings ν_i, i.e. for $\bar{X}, \bar{Z} \in \bar{\mathcal{X}}$ we set

$$\nu_i^1 (Z_i, X_i) = \frac{\rho (e_i Z_i)}{\rho (e_i X_i)}$$

$$\nu_i^2 (Z_i, X_i) = \frac{\rho \left(\bar{X}\right) - \rho \left(\bar{X} - e_i Z_i\right)}{\rho \left(\bar{X}\right) - \rho \left(\bar{X} - e_i X_i\right)}$$

$$\nu_i^3 (Z_i, X_i) = \frac{\delta\rho \left(\bar{X}, e_i Z_i\right)}{\delta\rho \left(\bar{X}, e_i X_i\right)}$$

$$\nu_i^4 (Z_i, X_i) = \frac{\langle Z_i, \xi_i^{\bar{X}} \rangle}{\langle X_i, \xi_i^{\bar{X}} \rangle}$$

Possible choices for γ_i are studied in the next example for the systemic entropic risk measure $\rho^{ses,c}$ presented in example 2.55. As shown in Example 2.72, it is equivalent to the master example in the setting of the *first inject then aggregate* approach.

Example 3.6. *We have already seen that the systemic entropic risk measure $\rho^{ses,c}$ has a unique optimal solution for every $X \in (M^{\exp})^n$ to the dual problem given by*

$$\begin{aligned}\left(\xi^{\bar{X}}, \Xi^{\bar{X}}\right) &:= \left(\nabla \rho_0^{entr} \left(\Lambda^{sum,c} \left(\bar{X}\right)\right), 1_n \nabla \rho_0^{entr} \left(\Lambda^{sum,c} \left(\bar{X}\right)\right)\right) \\ &= \left(\frac{\exp\left(\theta \sum_{i=1}^n X_i\right)}{\mathbb{E}\left[\exp\left(\theta \sum_{i=1}^n X_i\right)\right]}, 1_n \frac{\exp\left(\theta \sum_{i=1}^n X_i\right)}{\mathbb{E}\left[\exp\left(\theta \sum_{i=1}^n X_i\right)\right]}\right)\end{aligned}$$

The penalty function is given by

$$\begin{aligned}\alpha\left(\xi^{\bar{X}}, \Xi^{\bar{X}}\right) &= \alpha^{\rho_0^{entr}}\left(\xi^{\bar{X}}\right) + \alpha^{\Lambda^{sum,c}}\left(\Xi^{\bar{X}}\right) \\ &= \frac{1}{\theta} H\left(\mathbb{Q}^{\bar{X}} \mid \mathbb{P}\right) + c \\ &= \frac{1}{\theta}\mathbb{E}\left[\xi^{\bar{X}} \ln\left(\xi^{\bar{X}}\right)\right] + c \\ &= \mathbb{E}\left[\xi^{\bar{X}} \sum_{i=1}^n X_i\right] - \frac{1}{\theta} \ln\left(\mathbb{E}\left[\exp\left(\theta \sum_{i=1}^n X_i\right)\right]\right) + c,\end{aligned}$$

where $\mathbb{Q}^{\bar{X}}$ is the probability measure with density $\xi^{\bar{X}}$. A natural economic way to apportion the penalty term to the participants of the system is to use the individual risk aversion parameter, i.e.

$$\tilde{\gamma}_i = \mu^{\tilde{\gamma}}\left(e_i X_i, \bar{X}\right) = \frac{\theta}{\alpha_i}.$$

γ_i is simply the contribution of participant i to the systemic risk aversion parameter. Therefore, risk friendly participants are penalized more than risk averse participants.

The final step to obtain a systemic capital allocation rule is to extend the weights of the participants X_i to arbitrary configurations $\bar{Z}$ which are then rated in the presence of $\bar{X}$. This means that $\bar{Z}$ is considered as a subsystem of $\bar{X}$. We can now use the weight functions ν^1, ν^2, ν^3 and ν^4 to construct a systemic capital allocation rule for the systemic entropic risk measure. As already mentioned, all weight functions at least yield a systemic capital allocation rule which fulfills the restricted full allocation property (C1R). ν^3 and ν^4 also induce the full allocation property (C1) in this specific situation. So we have, for $\mu^t = \mu^{\tilde{\gamma},t} = \sum_{i=1}^n \tilde{\gamma}_i \nu_i^t$, $t = 1, 2, 3, 4$, that a systemic capital allocation rule for the systemic entropic risk measure $\rho^{ses,c}$ as in (3.4) *is given by*

$$\begin{aligned} &CS^{\Xi}\left(\bar{Z};\bar{X}\right) \\ &= \mathbb{E}\left[\xi^{\bar{X}}\left(\sum_{i=1}^n Z_i - \mu^t\left(\bar{Z},\bar{X}\right)\left(\sum_{i=1}^n X_i - \frac{1}{\theta}\ln\left(\mathbb{E}\left[\exp\left(\theta\sum_{i=1}^n X_i\right)\right]\right) + c\right)\right)\right]. \end{aligned} \tag{3.7}$$

By construction, we obtain for $\bar{Z} = e_j X_j$

$$\begin{aligned} &CS^{\Xi}\left(e_j X_j;\bar{X}\right) \\ &= \mathbb{E}\left[\xi^{\bar{X}}\left(X_j - \frac{\theta}{\alpha_j}\sum_{i=1}^n X_i + \frac{1}{\alpha_j}\ln\left(\mathbb{E}\left[\exp\left(\theta\sum_{i=1}^n X_i\right)\right]\right) - \frac{\theta}{\alpha_j}c\right)\right]. \end{aligned} \tag{3.8}$$

The same procedure can be done if we change the aggregation to Λ^{loss}, i.e. if we consider the systemic risk measure ρ^{sel}. An optimal solution to the dual problem is given by

$$(\xi,\Xi) = \left(\nabla\rho_0^{entr}\left(\sum_{i=1}^n X_i^+\right), 1_{>0}\left(\bar{X}\right)\nabla\rho_0^{entr}\left(\sum_{i=1}^n X_i^+\right)\right) \in \mathcal{A}'_{\Lambda^{loss}\circ}.$$

Now, we obtain

$$\begin{aligned} &CS^{\Xi}\left(\bar{Z};\bar{X}\right) \\ &= \mathbb{E}\left[\xi\left(\sum_{i=1}^n Z_i \mathbb{1}_{\{X_i>0\}} - \mu^t\left(\bar{Z},\bar{X}\right)\left(\sum_{i=1}^n X_i^+\right) - \frac{1}{\theta}\ln\left(\mathbb{E}\left[\exp\left(\theta\sum_{i=1}^n X_i^+\right)\right]\right)\right)\right] \end{aligned} \tag{3.9}$$

and for $\bar{Z} = e_j X_j$

$$CS^{\Xi}\left(e_j X_j;\bar{X}\right) = \mathbb{E}\left[\xi\left(X_j^+ - \frac{\theta}{\alpha_j}\sum_{i=1}^n X_i^+ + \frac{1}{\alpha_j}\ln\left(\mathbb{E}\left[\exp\left(\theta\sum_{i=1}^n X_i^+\right)\right]\right)\right)\right]. \tag{3.10}$$

3.2. Scenario Dependent Allocations

Since optimal solutions to the primal problem (2.56) of R are elements of $\mathcal{C}(\mathbb{R})$, they have a special interpretation. In every scenario the losses add up to $R\left(\bar{X}\right)$. But except for trivial cases, the components of the vector are random variables. Therefore, such an optimal solution is called *scenario dependent allocation* in [BFFMB20]. It is some sort of manual, how to allocate the risk capital $R\left(\bar{X}\right)$ to the participants of the system, according to the incurred scenario. For the systemic risk measure R^{ex} (h=1), the optimal solution was computed in [BFFMB20] (Section 6) and is explicitly given by

$$Y_i^{\bar{X}} = X_i - \frac{\theta}{\alpha_i}\sum_{i=1}^n X_i + \frac{1}{\alpha_i}\ln\left(\mathbb{E}\left[\exp\left(\theta\sum_{i=1}^n X_i\right)\right]\right) - \frac{\theta}{\alpha_i}c,$$

with $c = \frac{1}{\theta}\ln(\theta B)$. It was also shown that this solution is unique and *fair*. In Example 2.72, we have seen that $R^{ex} = \rho^{ses,c}$. Now, the previous example says that the dual systemic capital allocation rule for the participant j is given by

$$CS^{\Xi}\left(e_j X_j; \bar{X}\right) = \mathbb{E}\left[\xi^{\bar{X}} Y_j^{\bar{X}}\right].$$

So in other words, for the systemic risk measure R^{ex} one obtains the optimal solutions to the primal problem (i.e. the optimal scenario dependent allocation) if one computes the dual systemic capital allocation rule and ignores the expectation. In this sense, the systemic dual capital allocation rule can be seen as the expected scenario wise allocation. For an arbitrary R, the weights

$$\gamma_i = \frac{\langle (l_i^*)'\left(\lambda^*\Xi_i^{\bar{X}}\right), \Xi_i^{\bar{X}}\rangle}{\alpha^R\left(\Xi^{\bar{X}}\right)}$$

are the ones to choose in order to derive the optimal $\bar{Y}^{\bar{X}}$ with this procedure. If we extend the weights to a full weight function which is additive w.r.t. its first component, we can also give scenario dependent allocations for arbitrary subsystems of $\bar{X}$. However, these allocations obviously depend on the specific extension, whereas the scenario dependent allocation to the participants of the systems does not change. Note, that these weights also work if we have more than one group or even no group structure.

3.3. Aumann-Shapley Capital Allocation Rule and Weights for Systemic Risk Measures

The dual capital allocation rule for convex systemic risk measures presents a suitable way to generalize the ideas for positively homogeneous systemic risk measures. A different way is to define a local systemic capital allocation rule in the spirit of the Aumann-Shapley allocation for single-firm risk measures. If the systemic risk measure ρ is Gâteaux differentiable at $\gamma\bar{X} \in \operatorname{int}(\operatorname{dom}(\rho))$ for all $\gamma \in [0,1]$ we can define

$$CS^{AS}\left(\bar{Z};\bar{X}\right) = \int_0^1 \delta\rho\left(\gamma\bar{X},\bar{Z}\right)d\gamma. \tag{3.11}$$

If we can assume that $\rho(0) = 0$, we have

$$\begin{aligned}
\rho\left(\bar{X}\right) &= \rho\left(1\cdot\bar{X}\right) - \rho\left(0\cdot\bar{X}\right)\\
&= \int_0^1 \frac{d}{d\gamma}\left(\rho\left(\gamma\bar{X}\right)\right)d\gamma\\
&= \int_0^1 \lim_{\varepsilon\to 0}\frac{\rho\left((\gamma+\varepsilon)\bar{X}\right) - \rho\left(\gamma\bar{X}\right)}{\varepsilon}d\gamma\\
&= \int_0^1 \lim_{\varepsilon\to 0}\frac{\rho\left(\gamma\bar{X}+\varepsilon\bar{X}\right) - \rho\left(\gamma\bar{X}\right)}{\varepsilon}d\gamma\\
&= \int_0^1 \delta\rho\left(\gamma\bar{X},\bar{X}\right)d\gamma\\
&= CS^{AS}\left(\bar{X};\bar{X}\right).
\end{aligned}$$

The major benefit of this systemic capital allocation rule is that it provides an intrinsic way to allocate the penalty term. Additionally, the full allocation property (C1) is always satisfied, since the Gâteaux differential is always linear w.r.t. the second argument. If all the Gâteaux differentials are also continuous, i.e. ρ is Gâteaux differentiable at $\gamma\bar{X} \in \operatorname{int}(\operatorname{dom}(\rho))$ with derivative $\nabla\rho\left(\gamma\bar{X}\right)$ for all $\gamma \in [0,1]$ we have

$$CS^{AS}\left(\bar{Z};\bar{X}\right) = \int_0^1 \langle\bar{Z},\nabla\rho\left(\gamma\bar{X}\right)\rangle_n d\gamma.$$

Remark 3.12. *In many common situations, for example if the spaces $\mathcal{X}_i$ are L^p-spaces or even more general Orlicz spaces, the pairing appears to be an expectation as seen earlier. Now, for Gâteaux differentiable systemic risk measures ρ with $\rho(0) = 0$, the Aumann-Shapley capital allocation rule yields an alternative representation by interchanging the integrals. It is given by*

$$\rho\left(\bar{X}\right) = \sum_{i=1}^n \mathbb{E}\left[X_i L_i^{\bar{X}}\right],$$

where

$$L_i^{\bar{X}} = \int_0^1 \left(\nabla\rho\left(\gamma\bar{X}\right)\right)_i d\gamma.$$

This type of representation is studied in [KO14] for single-firm risk measures in a *dynamic setting.*

If the underlying systemic risk measure is positively homogeneous (S3), it also reduces to the gradient capital allocation rule. For systemic risk measures of type (0.1) the chain rule for Gâteaux differentials applies. We can rewrite (3.11) and obtain

$$CS^{AS}\left(\bar{Z};\bar{X}\right)=\int_0^1 \delta\rho_0\left(\Lambda\left(\gamma\bar{X}\right),\delta\Lambda\left(\gamma\bar{X},\bar{Z}\right)\right)d\gamma, \tag{3.13}$$

where $\delta\Lambda\left(\bar{X},\bar{Z}\right)$ is meant point wise, i.e. $\delta\Lambda\left(\bar{X},\bar{Z}\right):\Omega\to\mathbb{R}$ with

$$\delta\Lambda\left(\bar{X},\bar{Z}\right)(\omega)=\delta\Lambda\left(\bar{X}(\omega),\bar{Z}(\omega)\right).$$

In this situation if the corresponding single-firm risk measure ρ_0 is Gâteaux differentiable at $\Lambda\left(\gamma\bar{X}\right)$ for all $\gamma\in[0,1]$ with derivative $\nabla\rho_0\left(\Lambda\left(\gamma\bar{X}\right)\right)$, it reduces to

$$CS^{AS}\left(\bar{Z};\bar{X}\right)=\int_0^1 \langle\nabla\rho_0\left(\Lambda\left(\gamma\bar{X}\right)\right),\delta\Lambda\left(\gamma\bar{X},\bar{Z}\right)\rangle d\gamma. \tag{3.14}$$

However, the differentiability conditions on ρ and on ρ_0 and Λ to derive (3.11) and (3.14) are very restrictive. For example, on $\{X=0\}$ Λ^{loss} is not Gâteaux differentiable. An alternative way for systemic risk measures of type (0.1) to define a systemic capital allocation rule in the spirit of Aumann-Shapley is given by

$$\bar{CS}^{AS}\left(\bar{Z};\bar{X}\right)=\int_0^1 \delta\rho_0\left(\gamma\Lambda\left(\bar{X}\right),\Lambda\left(\bar{Z}\right)\right)d\gamma. \tag{3.15}$$

In this situation we only need the single-firm risk measure ρ_0 to be Gâteaux-differentiable at $\gamma\Lambda\left(\bar{X}\right)$ for all $\gamma\in[0,1]$. The price we have to pay is that full allocation properties are only satisfied if the aggregation rule Λ fulfills some additivity assumption. Moreover, for Λ^{sum} these systemic capital allocation rules coincide. An other interesting way to include the Aumann-Shapley ideas into the allocation process is to use the weight function

$$\mu^{AS}\left(\bar{Z},\bar{X}\right)=\frac{CS^{AS}\left(\bar{Z};\bar{X}\right)}{CS^{AS}\left(\bar{X};\bar{X}\right)}.$$

Note that if $\rho(0)\neq 0$, it no longer holds that $CS^{AS}\left(\bar{X};\bar{X}\right)=\rho\left(\bar{X}\right)$. This weight function always yields a dual systemic capital allocation rule which at least satisfies the restricted full allocation property (C1R). Of course, it is also possible to define a weight function μ which gives us $CS^{AS} = CS^{\Xi}$. It is given by

$$\mu^{8}\left(\bar{Z},\bar{X}\right) = \frac{\langle \bar{Z}, \Xi^{\bar{X}} \rangle_n - CS^{AS}\left(\bar{Z};\bar{X}\right)}{\alpha^{\rho}\left(\Xi^{\bar{X}}\right)}.$$

Let us continue by presenting some examples.

Example 3.16. *Consider the systemic risk measure ρ^{ses} defined in* (2.40)*. Since both, ρ_0^{entr} and Λ^{sum}, are (everywhere) Gâteaux differentiable with derivative and $\rho^{ses}(0) = 0$ we are able to apply* (3.11)*. It yields*

$$CS^{AS}\left(\bar{Z};\bar{X}\right) = \int_0^1 \mathbb{E}\left[\sum_{i=1}^{n} Z_i \frac{\exp\left(\theta\gamma \sum_{i=1}^{n} X_i\right)}{\mathbb{E}\left[\exp\left(\theta\gamma \sum_{i=1}^{n} X_i\right)\right]}\right] d\gamma.$$

As already seen in Example 2.72, we have $R^{ex} = \rho^{ses}$ for $h = 1$ and $B = \sum_{i=1}^{n} l_i(0) = \frac{1}{\theta}$. So for this systemic risk measure the Aumann-Shapley allocation is also given by the previous formula. If we have $h > 1$ groups the Aumann-Shapley capital allocation rule becomes

$$CS^{AS}\left(\bar{Z};\bar{X}\right) = \int_0^1 \mathbb{E}\left[\sum_{j=1}^{h} \sum_{i\in I_j} Z_i \frac{\exp\left(\theta_j\gamma \sum_{i\in I_j} X_i\right)}{\mathbb{E}\left[\exp\left(\theta_j\gamma \sum_{i\in I_j} X_i\right)\right]}\right] d\gamma.$$

Note that in both situations the full allocation property (C1) is satisfied. We are also able to compute a systemic capital allocation rule for ρ^{sel}, defined in (2.42)*. As already mentioned above, the aggregation rule $\Lambda^{loss}\left(\bar{X}\right)$ is not Gâteaux differentiable. Therefore, a computation of a systemic capital allocation rule via* (3.14) *fails. But all requirements for* (3.15) *are fulfilled and we obtain*

$$\bar{CS}^{AS}\left(\bar{Z};\bar{X}\right) = \int_0^1 \mathbb{E}\left[\Lambda^{loss}\left(\bar{Z}\right) \frac{\exp\left(\theta\gamma\Lambda^{loss}\left(\bar{X}\right)\right)}{\mathbb{E}\left[\exp\left(\theta\gamma\Lambda^{loss}\left(\bar{X}\right)\right)\right]}\right] d\gamma.$$

This systemic capital allocation rule only fulfills the restricted full allocation property (C1R).

Example 3.17. *Consider the systemic risk measure defined in Example 2.39 (ii). Since both, Λ^{exut} and ρ_0 are (everywhere) Gâteaux differentiable with derivative, we are able to compute the systemic capital allocation via* (3.14) *if we can guarantee that $\rho(0) = 0$. This condition is clearly satisfied if we set $B = \frac{1}{\theta}$. In this case, we* *obtain*

$$
\begin{aligned}
CS^{AS}\left(\bar{Z};\bar{X}\right) &= \sum_{i=1}^{n}\int_0^1 \mathbb{E}\left[Z_i \exp\left(\gamma\alpha_i X_i\right)\right] d\gamma. \\
&= \sum_{i=1}^{n}\mathbb{E}\left[\frac{Z_i}{\alpha_i X_i}\left(\exp\left(\gamma\alpha_i X_i\right)-1\right)\right]
\end{aligned}
$$

Obviously, the full allocation property (C1) is satisfied. We can also use (3.15) *to compute a systemic capital allocation rule for this risk measure. Again, we set* $B=\frac{1}{\theta}$ *and obtain*

$$
\begin{aligned}
\bar{CS}^{AS}\left(\bar{Z};\bar{X}\right) &= \int_0^1 \mathbb{E}\left[\sum_{i=1}^{n}\frac{1}{\alpha_i}\exp\left(\alpha_i Z_i\right)\right] d\gamma \\
&= \mathbb{E}\left[\sum_{i=1}^{n}\frac{1}{\alpha_i}\exp\left(\alpha_i Z_i\right)\right] \\
&= \rho\left(\bar{Z}\right)+B.
\end{aligned}
$$

This systemic capital allocation rule only fulfills the restricted full allocation property (C1R).

3.4. Numerical Example on a Finite Probability Space

We now want to study a financial system $\bar{X}$ on a finite probability space $\Omega = \{\omega_1,\ldots,\omega_4\}$ and compute different systemic capital allocation rules. The calculations can be derived with Python. The code is provided in 1. It extends the analysis of Example 7.1 from [BFFMB19]. So, let

$$
\bar{X} = \begin{pmatrix} X_1(\omega_1) & X_1(\omega_2) & X_1(\omega_3) & X_1(\omega_4) \\ X_2(\omega_1) & X_2(\omega_2) & X_2(\omega_3) & X_2(\omega_4) \\ X_3(\omega_1) & X_3(\omega_2) & X_3(\omega_3) & X_3(\omega_4) \\ X_4(\omega_1) & X_4(\omega_2) & X_4(\omega_3) & X_4(\omega_4) \end{pmatrix} = \begin{pmatrix} -100 & 50 & -100 & 50 \\ -50 & 25 & -50 & 25 \\ 25 & -50 & 25 & -50 \\ -50 & -50 & 25 & 25 \end{pmatrix}
$$

and

$$
\mathbb{P} = \begin{pmatrix} \mathbb{P}(\omega_1) \\ \mathbb{P}(\omega_2) \\ \mathbb{P}(\omega_3) \\ \mathbb{P}(\omega_4) \end{pmatrix} = \begin{pmatrix} 0.64 \\ 0.16 \\ 0.16 \\ 0.04 \end{pmatrix}.
$$

The first systemic risk measure we want to study is ρ^{ses} presented in Example 2.55. We assume that all participant of the system have an identical risk aversion coefficient $\alpha_i = 0.3$, $i = 1,\ldots,4$. In this situation, the overall systemic risk is given by

$$
\rho^{ses}\left(\bar{X}\right) = 7.273.
$$

Now, the corresponding Aumann-Shapley capital allocation rule in this setup yields the following values for the participants X_i, $i \in \{1, \ldots, 4\}$,

$$\begin{pmatrix} CS^{AS}\left(e_1X_1; \bar{X}\right) \\ CS^{AS}\left(e_2X_2; \bar{X}\right) \\ CS^{AS}\left(e_3X_3; \bar{X}\right) \\ CS^{AS}\left(e_4X_4; \bar{X}\right) \end{pmatrix} = \begin{pmatrix} 28.542 \\ 14.271 \\ -39.271 \\ 3.732 \end{pmatrix}.$$

For the dual capital allocation rule, the values are given as

$$\begin{pmatrix} CS^{\Xi}\left(e_1X_1; \bar{X}\right) \\ CS^{\Xi}\left(e_2X_2; \bar{X}\right) \\ CS^{\Xi}\left(e_3X_3; \bar{X}\right) \\ CS^{\Xi}\left(e_4X_4; \bar{X}\right) \end{pmatrix} = \begin{pmatrix} 39.579 \\ 14.583 \\ -60.409 \\ 13.520 \end{pmatrix}.$$

Finally, the scenario dependent allocation is given by

$$\begin{aligned} \bar{Y}^{\bar{X}} &= \begin{pmatrix} Y_1^{\bar{X}}(\omega_1) & Y_1^{\bar{X}}(\omega_2) & Y_1^{\bar{X}}(\omega_3) & Y_1^{\bar{X}}(\omega_4) \\ Y_2^{\bar{X}}(\omega_1) & Y_2^{\bar{X}}(\omega_2) & Y_2^{\bar{X}}(\omega_3) & Y_2^{\bar{X}}(\omega_4) \\ Y_3^{\bar{X}}(\omega_1) & Y_3^{\bar{X}}(\omega_2) & Y_3^{\bar{X}}(\omega_3) & Y_3^{\bar{X}}(\omega_4) \\ Y_4^{\bar{X}}(\omega_1) & Y_4^{\bar{X}}(\omega_2) & Y_4^{\bar{X}}(\omega_3) & Y_4^{\bar{X}}(\omega_4) \end{pmatrix} \\ &= \begin{pmatrix} -54.432 & 58.068 & -73.182 & 39.318 \\ -4.432 & 33.068 & -23.182 & 14.318 \\ 70.568 & -41.932 & 51.818 & -60.682 \\ -4.432 & -41.932 & 51.818 & 14.318 \end{pmatrix}. \end{aligned}$$

As already mentioned several times, this systemic risk measure coincides with R^{ex} for $B = \frac{1}{\theta}$ with one group. Next, we want to study the systemic risk measure ρ^{sel}. The overall systemic risk is given by

$$\rho^{sel}\left(\bar{X}\right) = 64.658.$$

To compute the allocation in the spirit of Aumann-Shapley we need to use the modified version (3.15). We obtain

$$\begin{pmatrix} \bar{CS}^{AS}\left(e_1X_1; \bar{X}\right) \\ \bar{CS}^{AS}\left(e_2X_2; \bar{X}\right) \\ \bar{CS}^{AS}\left(e_3X_3; \bar{X}\right) \\ \bar{CS}^{AS}\left(e_4X_4; \bar{X}\right) \end{pmatrix} = \begin{pmatrix} 29.740 \\ 14.870 \\ 10.130 \\ 9.919 \end{pmatrix}.$$

The dual and scenario dependent allocations are given as

$$\begin{pmatrix} CS^{\Xi}\left(e_1X_1;\bar{X}\right) \\ CS^{\Xi}\left(e_2X_2;\bar{X}\right) \\ CS^{\Xi}\left(e_3X_3;\bar{X}\right) \\ CS^{\Xi}\left(e_4X_4;\bar{X}\right) \end{pmatrix} = \begin{pmatrix} 40.317 \\ 17.466 \\ -3.235 \\ 10.111 \end{pmatrix}$$

and

$$\begin{aligned} \bar{Y}^{\bar{X}} &= \begin{pmatrix} Y_1^{\bar{X}}(\omega_1) & Y_1^{\bar{X}}(\omega_2) & Y_1^{\bar{X}}(\omega_3) & Y_1^{\bar{X}}(\omega_4) \\ Y_2^{\bar{X}}(\omega_1) & Y_2^{\bar{X}}(\omega_2) & Y_2^{\bar{X}}(\omega_3) & Y_2^{\bar{X}}(\omega_4) \\ Y_3^{\bar{X}}(\omega_1) & Y_3^{\bar{X}}(\omega_2) & Y_3^{\bar{X}}(\omega_3) & Y_3^{\bar{X}}(\omega_4) \\ Y_4^{\bar{X}}(\omega_1) & Y_4^{\bar{X}}(\omega_2) & Y_4^{\bar{X}}(\omega_3) & Y_4^{\bar{X}}(\omega_4) \end{pmatrix} \\ &= \begin{pmatrix} 9.915 & 47.415 & 3.665 & 41.165 \\ 9.915 & 22.415 & 3.665 & 16.165 \\ 34.915 & -2.585 & 28.665 & -8.835 \\ 9.915 & -2.585 & 28.665 & 16.165 \end{pmatrix}. \end{aligned}$$

If one ranks the participants of the system according to their expected loss and compares this ranking to a ranking according to the risk contributions of the participants measured with the different allocation rules, one observes a big difference. Since both informations are useful in order to rate the participants, the return on risk adjusted capital (RORAC) seems to be a suitable tool. Like [Tas07], we distinguish between the RORAC for the whole system, defined as

$$RORAC\left(\bar{X}\right) = \frac{\sum_{i=1}^{n}\mathbb{E}\left[X_i\right]}{\rho\left(\bar{X}\right)},$$

and for the participants, defined as

$$RORAC\left(e_iX_i;\bar{X}\right) = \frac{\mathbb{E}\left[X_i\right]}{CS\left(e_iX_i;\bar{X}\right)}.$$

For ρ^{ses}, the RORAC of the whole system is given by

$$RORAC\left(\bar{X}\right) = -17.874.$$

The individual RORAC measured with the dual systemic capital allocation rule of the participants is given by

$$\begin{pmatrix} RORAC^{\Xi}\left(e_1X_1;\bar{X}\right) \\ RORAC^{\Xi}\left(e_2X_2;\bar{X}\right) \\ RORAC^{\Xi}\left(e_3X_3;\bar{X}\right) \\ RORAC^{\Xi}\left(e_4X_4;\bar{X}\right) \end{pmatrix} = \begin{pmatrix} -1.769 \\ -2.400 \\ -0.166 \\ -2.589 \end{pmatrix}$$

and with the Aumann-Shapley systemic capital allocation rule by

$$\begin{pmatrix} RORAC^{AS}\left(e_1X_1;\bar{X}\right) \\ RORAC^{AS}\left(e_2X_2;\bar{X}\right) \\ RORAC^{AS}\left(e_3X_3;\bar{X}\right) \\ RORAC^{AS}\left(e_4X_4;\bar{X}\right) \end{pmatrix} = \begin{pmatrix} -2.453 \\ -2.453 \\ -0.255 \\ -9.379 \end{pmatrix}.$$

For ρ^{sel}, we obtain

$$RORAC\left(\bar{X}\right) = 0.619,$$

$$\begin{pmatrix} RORAC^{\Xi}\left(e_1X_1;\bar{X}\right) \\ RORAC^{\Xi}\left(e_2X_2;\bar{X}\right) \\ RORAC^{\Xi}\left(e_3X_3;\bar{X}\right) \\ RORAC^{\Xi}\left(e_4X_4;\bar{X}\right) \end{pmatrix} = \begin{pmatrix} 0.248 \\ 0.286 \\ -6.183 \\ 0.495 \end{pmatrix},$$

and

$$\begin{pmatrix} RORAC^{AS}\left(e_1X_1;\bar{X}\right) \\ RORAC^{AS}\left(e_2X_2;\bar{X}\right) \\ RORAC^{AS}\left(e_3X_3;\bar{X}\right) \\ RORAC^{AS}\left(e_4X_4;\bar{X}\right) \end{pmatrix} = \begin{pmatrix} 0.336 \\ 0.336 \\ 1.974 \\ 0.504 \end{pmatrix}.$$

```python
#%%
# import necessary packages

import numpy as np
from scipy.integrate import quad

# definition of the functions for the allocation
def optimal_density(x, theta, firms, prob):
    # function computes the optimal density for a weighted system
    # computing the numerator
    exp_scaled_systemic_loss = np.exp(theta * np.sum(x * firms, axis = 0))
    # computing the denominator
    norm_constant = np.dot(exp_scaled_systemic_loss, prob)
    return(exp_scaled_systemic_loss/norm_constant)

def cs_dual(alpha, B, theta, firms, prob):
    # function computes the (local) dual and scenario dependent allocation for the participants
    # computing the measure Q corresponding to the optimal density
    optimal_Q = optimal_density(1, theta, firms, prob) * prob
    # computing the optimal scenario dependent allocation
    optimal_Y = np.zeros([np.size(firms[:,0]), np.size(firms[0,:])])
    for row in range(np.size(firms[:,0])):
        for column in range(np.size(firms[0,:])):
            optimal_Y[row,column] = firms[row,column] - (1/alpha[row]) * np.log(optimal_density(1, theta, firms, prob)[column]) - (1/alpha[row]) * np.log(theta*B)
    return(optimal_Y, optimal_Q, optimal_Y @ optimal_Q, sum(optimal_Y @ optimal_Q))

def integrand(x, theta, firms, prob, firm):
    # function computes the integrand for the Aumann-Shapley allocation
    return(np.dot(firms[firm, :] * optimal_density(x, theta, firms, prob), prob))

def cs_aumann_shapely(theta, firms, prob):
    # function computes the Aumann-Shapley allocation
    allocation = np.zeros(np.size(firms[:,0]))
    for firm in range(np.size(firms[:,0])):
        allocation[firm] = quad(integrand, 0, 1, args=(theta, firms, prob, firm))[0]
    return(allocation)

def systemic_risk_allocations(alpha, B, theta, firms, prob):
    print("The overall systemic risk is given by",
          cs_dual(alpha, B, theta, firms, prob)[3], ".")
    print("The dual systemic capital allocation is given by",
          cs_dual(alpha, B, theta, firms, prob)[2], ".")
    print("The scenario dependent systemic capital allocation is given 
```

```python
        by",
            cs_dual(alpha, B, theta, firms, prob)[0], ".")
    print("The Aumann-Shapley systemic capital allocation is given by",
            cs_aumann_shapely(theta, firms, prob), ".")

# definitions of the functions for the RORAC
def total_systemic_rorac(alpha, B, theta, firms, prob):
    # function computes the total RORAC
    system_expected_loss = sum(firms @ prob)
    return(system_expected_loss/cs_dual(alpha, B, theta, firms, prob)[3])

def system_related_rorac_dual(alpha, B, theta, firms, prob):
    # function computes the system related RORAC of the participants of the system for the dual systemic capital allocation rule
    firms_expected_loss = firms @ prob
    return(firms_expected_loss/cs_dual(alpha, B, theta, firms, prob)[2])

def system_related_rorac_aumann_shapley(theta, firms, prob):
    # function computes the system related RORAC of the participants of the system for the Aumann-Shapley systemic capital allocation rule
    firms_expected_loss = firms @ prob
    return(firms_expected_loss/cs_aumann_shapely(theta, firms, prob))

def rorac(alpha, B, theta, firms, prob):
    for firm in range(np.size(firms[:,0])):
        print("The system-related RORAC of firm", firm+1, "measured with the dual systemic capital allocation rule is given by", system_related_rorac_dual(alpha, B, theta, firms, prob)[firm])
        print("The system-related RORAC of firm", firm+1, "measured with the Aumann-Shapley systemic capital allocation rule is given by", system_related_rorac_aumann_shapley(theta, firms, prob)[firm])
    print("The total system RORAC is given by", total_systemic_rorac(alpha, B, theta, firms, prob))
```

Listing 1: Python Code

A. Appendix

A.1. Convexity and Topology

Let us start by presenting some topological background. In the following, the space under consideration $\mathcal{V}$ is at least a (real) vector space. The objects of main interest are *convex sets* and *proper convex mappings*.

Definition A.1. *A set $\mathcal{C} \subseteq \mathcal{V}$ is called convex if, for all $X, Y \in \mathcal{C}$ and all $\alpha \in [0,1]$, we have*

$$\alpha X + (1-\alpha) Y \in \mathcal{C}.$$

An extended real valued mapping $f \colon \mathcal{V} \to \mathbb{R} \cup \{-\infty, \infty\}$ is called convex if its epigraph, defined by

$$epi(f) = \mathcal{A}_f = \{(m, X) \in \mathbb{R} \times \mathcal{V} \mid m \geq f(X)\}, \tag{A.2}$$

is a convex subset of $\mathbb{R} \times \mathcal{V}$. An extended real valued mapping $g \colon \mathcal{V} \to \mathbb{R} \cup \{-\infty, \infty\}$ is called concave if $-g$ is convex.

The effective domain of an extended real valued (not necessarily convex) mapping is given by

$$dom(f) := \{X \in \mathcal{V} \mid f(X) < \infty\}. \tag{A.3}$$

Finally, an extended real valued mapping is called proper if $dom(f) \neq \emptyset$ and, for all $X \in \mathcal{V}$, we have $f(X) \neq -\infty$. In this situation, we sometimes omit $-\infty$ from the notation, write $f \colon \mathcal{V} \to \mathbb{R} \cup \{\infty\}$ and call f proper.

For a real valued mapping defined on a convex subset $\mathcal{C}$ of $\mathcal{V}$ convexity is equivalent to the famous inequality

$$f(\alpha X + (1-\alpha) Y) \leq \alpha f(X) + (1-\alpha) f(Y), \tag{A.4}$$

for all $X, Y \in \mathcal{C}$ and all $\alpha \in [0,1]$ (see [AB06] Lemma 5.39). One may think of a proper convex mapping $\hat{f}$ as the extension of a real valued convex mapping f defined on a convex subset $\mathcal{C}$ of $\mathcal{V}$ with $\hat{f}(X) = \infty$ for $X \notin \mathcal{C}$. In this situation, we have $dom\left(\hat{f}\right) = \mathcal{C}$. More general, the effective domain of a convex mapping (defined on the whole space $\mathcal{V}$) is always convex. If we set the conventions $\infty + \infty = \infty$ and $0 \cdot \infty = \infty \cdot 0 = 0$, (A.4) (formulated for all $X, Y \in \mathcal{V}$) also characterizes proper convex mappings. For $\theta > 1$ and $X, Y \in \mathcal{V}$, we have

$$X = \frac{1}{\theta}(\theta X + (1-\theta) Y) + \frac{\theta - 1}{\theta} Y.$$

So, for all proper convex mappings f, $\theta > 1$ and $X, Y \in \operatorname{dom}(f)$ using (A.4) and rearranging the result yields the opposite inequality, i.e.

$$f(\theta X + (1-\theta) Y) \geq \theta f(X) + (1-\theta) f(Y). \tag{A.5}$$

If $\theta < 0$ we still have $1 - \theta > 1$. In this situation, the roles of X and Y are changed and (A.5) still holds. So, for a proper convex mapping equation (A.5) holds, for all $X, Y \in \operatorname{dom}(f)$ and all $\theta \in (-\infty, 0) \cup (1, \infty)$.

One of the most important results in functional analysis is the Hahn-Banach Theorem. In its general form, it provides a condition under which real linear mappings defined on a vector subspace (also called linear *functionals*) admit an extension to real linear mappings defined on the whole vector space.

Theorem A.6 ([AB06] Theorem 5.53)**.** *Let $f\colon \mathcal{V} \to \mathbb{R}$ be any convex mapping and $\mathcal{W} \subset \mathcal{V}$ be a vector subspace of $\mathcal{V}$. Consider now a linear functional $l\colon \mathcal{W} \to \mathbb{R}$ which is dominated (on $\mathcal{W}$) by f, i.e.*

$$l(X) \leq f(X), \quad \forall X \in \mathcal{W}.$$

Then there exists a linear extension $\hat{l}\colon \mathcal{V} \to \mathbb{R}$ of l which is dominated (on $\mathcal{V}$) by f, i.e.

$$l(X) = \hat{l}(X), \quad \forall X \in \mathcal{W},$$

and

$$\hat{l}(X) \leq f(X), \quad \forall X \in \mathcal{V}.$$

Now, let us introduce more structure on a given space.

Definition A.7. *The pair $(\mathcal{T}, \tau)$, consisting of a set $\mathcal{T}$ together with a collection of subsets $\tau \subseteq \mathcal{P}(\mathcal{T})$, is called topological space if the following properties hold:*

1. *$\emptyset \in \tau$.*
2. *For an arbitrary index set $\mathcal{I} \neq \emptyset$ with $\mathcal{O}_i \in \tau$, we have $\bigcup_{i \in \mathcal{I}} \mathcal{O}_i \in \tau$.*
3. *For a finite index set $\{1, \ldots, n\}$ with $\mathcal{O}_i \in \tau$, we have $\bigcap_{i=1}^n \mathcal{O}_i \in \tau$.*

τ is called topology (on $\mathcal{T}$). For a subspace $\mathcal{S} \subseteq \mathcal{T}$ of the topological space $(\mathcal{T}, \tau)$ the set

$$\tau_{\mathcal{S}} \coloneqq \{\mathcal{O} \cap \mathcal{S} \mid \mathcal{O} \in \tau\}$$

defines a topology on $\mathcal{S}$. It is called the subspace topology (on $\mathcal{S}$).

We call the sets $\mathcal{O} \in \tau$ *open*. For an arbitrary subset $\mathcal{B} \subseteq \mathcal{T}$ there are two important open sets connected to $\mathcal{B}$. The first one is the largest open set contained in $\mathcal{B}$ called *interior* (of $\mathcal{B}$) and denoted by $\operatorname{int}(\mathcal{B})$. The second one is the largest open set disjoint from $\mathcal{B}$ called *exterior* (of $\mathcal{B}$) and denoted by $\operatorname{ext}(\mathcal{B})$. Elements of these open sets are called *interior* and *exterior* points. The complementary concept is closedness. A subset $\mathcal{A} \subseteq \mathcal{T}$ is called *closed* if its complement $\mathcal{A}^{\mathcal{C}} := \mathcal{T} \setminus A$ is open. For an arbitrary subset $\mathcal{B} \subseteq \mathcal{T}$ there are two important closed sets connected to $\mathcal{B}$. The smallest closed set containing $\mathcal{B}$ called *closure* (of $\mathcal{B}$) and denoted by $\operatorname{cl}(\mathcal{B})$, and the *boundary* of $\mathcal{B}$ denoted by $\partial\mathcal{B} := \operatorname{cl}(\mathcal{B}) \setminus \operatorname{int}(\mathcal{B})$. With the additional structure of a topology the notion of continuity comes in play. A mapping between two topological spaces is called *continuous* if the inverse image of an open set in the codomain is an open set in the domain. Depending on the underlying topology, the set of continuous mappings between two topological spaces may differ. Sometimes it is desirable that a given family of mappings $\mathcal{M}$ is continuous w.r.t. the considered topology. The *coarsest* (or smallest i.e. the topology with the fewest open sets) topology for which a given family of mappings is continuous is called *initial topology* denoted by $\tau(\mathcal{M})$. If the family consists of real valued mappings, we write $\sigma(\mathcal{T}, \mathcal{M}) := \tau(\mathcal{M})$. One prominent initial topology comes in play if we search for a natural topology on the cartesian product of topological spaces.

Definition A.8. *For an arbitrary index set $\mathcal{I} \neq \emptyset$ let $((\mathcal{T}_i, \tau_i))_{i\in\mathcal{I}}$ be a family of topological spaces. The initial topology of the projections $(\pi_i)_{i\in\mathcal{I}}$, where*

$$\pi_i : \bigtimes_{j\in I} \mathcal{T}_j \to \mathcal{T}_i, \quad (x_j)_{j\in\mathcal{I}} \mapsto x_i,$$

on $\bigtimes_{j\in\mathcal{I}} \mathcal{T}_j$ is called product topology.

A minimal continuity requirement on a topology on a vector space should be that the vector space operations, i.e. vector addition and scalar multiplication, are continuous. They are given as

$$+ : \mathcal{V} \times \mathcal{V} \to \mathcal{V}, \quad (X, Y) \mapsto X + Y$$

and

$$\cdot : \mathbb{R} \times \mathcal{V} \to \mathcal{V}, \quad (\alpha, X) \mapsto \alpha X.$$

The spaces $\mathcal{V} \times \mathcal{V}$ and $\mathbb{R} \times \mathcal{V}$ are equipped with their respective product topologies.

Definition A.9. *A topology τ on a vector space $\mathcal{V}$ is called linear topology if the vector space operations are τ-continuous mappings. The pair $(\mathcal{V}, \tau)$ is called topological vector space (t.v.s.) if the topology τ is linear.*

Vector subspaces of t.v.s. equipped with the respective subspace topology are again t.v.s.. The same holds true for the cartesian product of t.v.s. equipped with the product topology (see [AB06] Lemma 5.1 and Theorem 5.2).

Since the Hahn-Banach Theorem makes a statement about linear functionals, it is natural to ask if there exist continuous linear functionals on a given t.v.s. and to examine their further properties. The space of all continuous linear functionals on a t.v.s. $(\mathcal{V}, \tau)$ is denoted by $\mathcal{V}'$ and is called *(topological) dual space.* In general, there are many topologies on a given vector space which can be considered. As already pointed out, the question whether a mapping is continuous or not sticks close to the properties of the given topology. Hence, the dual space also depends on the topology of the underlying space. So only in non trivial cases, statements about dual spaces are interesting. A rich class of spaces with non trivial dual spaces are so called *locally convex* t.v.s.. They generalize the concept of normed spaces and hence include Banach spaces as important representatives.

Definition A.10. *A linear topology τ on a vector space $\mathcal{V}$ is called locally convex topology if every neighborhood of zero includes a convex neighborhood of zero. The pair $(\mathcal{V}, \tau)$ is called locally convex t.v.s (l.c.t.v.s.) if the topology τ is linear and locally convex.*

Remember, a *neighborhood* of a point $x \in \mathcal{T}$ in an arbitrary topological space $(\mathcal{T}, \tau)$ is a set $\mathcal{U}$ with the property $x \in \mathcal{O} \subseteq \mathcal{U}$ for some $\mathcal{O} \in \tau$. A *neighborhood basis* at a point x is a set $\mathcal{B}(x)$ of neighborhoods of x such that for every neighborhood $\mathcal{U}$ of x we have $x \in \mathcal{B} \subseteq \mathcal{U}$ for some $\mathcal{B} \in \mathcal{B}(x)$. The set of all neighborhoods of a point $x \in \mathcal{T}$, we will denote by $\mathcal{U}(x)$. Clearly, the set of all open neighborhoods of a point form a neighborhood basis. In a l.c.t.v.s., it is always possible to find a neighborhood basis of 0 consisting of open, convex and circled neighborhoods of 0 (see [AB06] Lemma 5.72). Recall, that a set $\mathcal{C}$ is *circled* if for all $X \in \mathcal{C}$ and all $\alpha \in \mathbb{R}$ with $|\alpha| \leq 1$ we have $\alpha X \in \mathcal{C}$. It is also possible to find a neighborhood basis $\mathcal{B}_c(0)$ of 0 consisting of closed, convex and circled neighborhoods of 0. With each $\mathcal{B} \in \mathcal{B}_c(0)$ comes a mapping $p_{\mathcal{B}} \colon \mathcal{V} \to \mathbb{R}_+$, given by

$$p_{\mathcal{B}}(X) = \inf \{\alpha \in \mathbb{R} \mid \alpha > 0,\ X \in \alpha\mathcal{B} := \{\alpha B \mid B \in \mathcal{B}\}\}. \tag{A.11}$$

This mapping is a *semi norm* in the sense of the following definition.

Definition A.12. *Consider the following properties for a real mapping $f \colon \mathcal{V} \to \mathbb{R}$:*

(SA) Subadditivity: For all $X, Y \in \mathcal{V}$, f satisfies $f(X+Y) \leq f(X) + f(Y)$.

(PH) Positive homogeneity: For all $X \in \mathcal{V}$ and all $\alpha \in \mathbb{R}_+$, f satisfies $f(\alpha X) = \alpha f(X)$.

(AH) Absolute homogeneity: For all $X \in \mathcal{V}$ and all $\alpha \in \mathbb{R}$, f satisfies $f(\alpha X) = |\alpha| f(X)$.

If a mapping satisfies (SA) and (PH), it is called sublinear. If a mapping satisfies (SA) and (AH), it is called seminorm.

The term semi is due to the fact that a seminorm which only takes the value 0 at $X = 0$ is a norm. As seen above, l.c.t.v.s. naturally carry seminorms. The connection is even stronger. The topology of a l.c.t.v.s. is uniquely determined by the family $(p_{\mathcal{B}})_{\mathcal{B}\in\mathcal{B}_c(0)}$. And conversely, for a given arbitrary family of seminorms $(p_i)_{i\in\mathcal{I}}$ on a vector space $\mathcal{V}$ there is a unique topology τ which makes $(\mathcal{V},\tau)$ to a l.c.t.v.s. (see [AB06] Theorems 5.6 and 5.73). This topology is called *generated* by the family of seminorms and the seminorms are continuous w.r.t. this topology.

Another interesting topological property is *separation*. The different forms of separation describe the capability of the topology to distinguish between different objects. One of the most common notions of separation is the Hausdorff property.

Definition A.13. *A topological space $(\mathcal{T},\tau)$ is called Hausdorff space (or simply Hausdorff) if for all distinct points $x, y \in \mathcal{T}$ there exist neighborhoods $\mathcal{U}_x$ of x and $\mathcal{U}_y$ of y such that $\mathcal{U}_x \cap \mathcal{U}_y = \emptyset$. We say that a topology τ on $\mathcal{T}$ is Hausdorff iff the topological space $(\mathcal{T},\tau)$ is a Hausdorff space.*

Obviously, every vector subspace of a Hausdorff t.v.s. endowed with the subspace topology is again a Hausdorff t.v.s.. If we consider the product of Hausdorff t.v.s. equipped with the product topology, it is again a Hausdorff t.v.s.. L.c.t.v.s. allow for more notions of separation. The continuous linear functionals on these spaces are a useful tool to characterize these notions.

Theorem A.14 ([AB06] Theorem 5.79)**.** *Let $(\mathcal{V},\tau)$ be a l.c.t.v.s.. For two convex sets $\mathcal{C}_1,\mathcal{C}_2 \subset \mathcal{V}$ with $\mathcal{C}_1 \cap \mathcal{C}_2 = \emptyset$ where one is compact and the other is closed there exists some nonzero $l \in \mathcal{V}'$ such that*

$$\sup_{X\in\mathcal{C}_1} l\left(X\right) < \inf_{Y\in\mathcal{C}_2} l\left(Y\right).$$

This form of separation is often referred to as *strong* separation. We say that the corresponding continuous linear functional *strongly* separates the two sets. As a direct consequence of the previous Theorem, we have the following Corollary.

Corollary A.15 ([AB06] Corollary 5.80)**.** *Let $(\mathcal{V},\tau)$ be a l.c.t.v.s.. For a closed convex set $\mathcal{C} \subset \mathcal{V}$ and a point $Z \notin \mathcal{C}$ there exists some nonzero $l \in \mathcal{V}'$ such that*

$$\sup_{X\in\mathcal{C}} l\left(X\right) < l\left(Z\right).$$

If we slightly weaken then assumptions, we obtain the following result.

Corollary A.16. *Let $(\mathcal{V},\tau)$ be a l.c.t.v.s.. For two convex sets $\mathcal{C}_1,\mathcal{C}_2 \subset \mathcal{V}$ with $\mathcal{C}_1 \cap \mathcal{C}_2 = \emptyset$ where $\mathcal{C}_2$ is open there exists some nonzero $l \in \mathcal{V}'$ such that for all $Y \in \mathcal{C}_2$*

$$\sup_{X\in\mathcal{C}_1} l\left(X\right) < l\left(Y\right).$$

Proof. It is a direct consequence of [AB06] Theorem 5.67. □

This situation is simply referred to as separation and we say that the corresponding continuous linear functional separates the two sets. If both sets are singletons, strong separation and separation describe the same behavior. In this context, we say that a family of functionals $\mathcal{F}$ on some space $\mathcal{V}$ separates points (of $\mathcal{V}$) if for all $X, Y \in \mathcal{V}$ with $X \neq Y$ there is some $f \in \mathcal{F}$ such that $f(X) \neq f(Y)$. It is now an easy Corollary to characterize the Hausdorff property for l.c.t.v.s. in terms of continuous linear functionals.

Corollary A.17 ([AB06] Corollary 5.82)**.** *A l.c.t.v.s. $(\mathcal{V}, \tau)$ is also Hausdorff if and only if its dual $\mathcal{V}'$ separates points.*

We have already mentioned that the Hausdorff property of a t.v.s. is preserved if one restricts the attention to a vector subspace of the given t.v.s. endowed with the subspace topology and if we consider products of Hausdorff t.v.s. equipped with the product topology. For l.c.t.v.s. these claims are also true. The question which arises in the situation of a vector subspace of a l.c.t.v.s. is the following: Are the dual spaces connected to each other in some way? For l.c.t.v.s. the answer is yes. The following Theorem holds:

Theorem A.18 ([AB06] Theorem 5.87)**.** *Let $(\mathcal{V}, \tau)$ be a l.c.t.v.s., $\mathcal{U} \subset \mathcal{V}$ be a vector subspace and $(\mathcal{U}, \tau_{\mathcal{U}})$ be the corresponding l.c.t.v.s.. Then every $l \in \mathcal{U}'$ has an extension (not necessarily unique) $\hat{l} \in \mathcal{V}'$. Moreover, the elements of $\mathcal{U}'$ are exactly the restrictions to $\mathcal{U}$ of the elements of $\mathcal{V}'$.*

If we analyze convex extended real valued mappings on Hausdorff l.c.t.v.s., continuity is often to strong. There are weaker forms of continuity which will be of interest.

Definition A.19. *An extended real valued mapping $f\colon \mathcal{T} \to \mathbb{R} \cup \{-\infty, \infty\}$ on a topological space $(\mathcal{T}, \tau)$ is called lower semicontinuous (l.s.c.) (more precisely τ-l.s.c.) if for each $c \in \mathbb{R}$ the set $\{x \in \mathcal{T} \mid f(x) \leq c\}$ is closed. f is called upper semicontinuous (u.s.c.) if $-f$ is l.s.c..*

The following Lemma captures some equivalent notions of and facts involving lower semicontinuity. If the underlying space is for example a metric space, one of these formulations can be stated with sequences. However, in more general topological spaces we need the more general concept of *nets* to capture such things correctly. So, a *net* in a topological space $(\mathcal{T}, \tau)$ is any mapping $f\colon \mathcal{J} \to \mathcal{T}$ where $\mathcal{J}$ is a upward directed set (with direction denoted by $\leq_{\mathcal{J}}$). We write $f(j) = x_j$ and $f = (x_j)_{j \in \mathcal{J}}$. In this sense, a net can be seen as a subset of $\mathcal{T}$ consisting of elements which are indexed by the elements of $\mathcal{J}$. Obviously, a sequence is a net where we have $\mathcal{J} = \mathbb{N}$. We say that a net $(x_j)_{j \in \mathcal{J}}$ is *τ-convergent* to some point $x \in \mathcal{T}$ when there is an index j_0 for every $\mathcal{U} \in \mathcal{U}(x)$ such that $x_j \in \mathcal{U}$ whenever $j_0 \leq_{\mathcal{J}} j$, write $x_j \xrightarrow{\tau} x$, and call x the *limit* of the net.

Lemma A.20 ([AB06] Lemma 2.41, Lemma 2.42, Corollary 2.60)**.** *Let* $f\colon \mathcal{T} \to \mathbb{R} \cup \{-\infty, \infty\}$ *be an extended real valued mapping on a topological space* $(\mathcal{T}, \tau)$ *and* $\mathcal{I} \neq \emptyset$ *an arbitrary index set. Then the following statements are equivalent:*

(i) f *is l.s.c..*

(ii) $\liminf_{j\in\mathcal{J}} f(x_j) := \sup_{j\in\mathcal{J}} \inf_{j\leq_{\mathcal{J}} k} \{f(x_k)\} \geq f(x)$ *for every* τ*-convergent net* $(x_j)_{j\in\mathcal{J}}$ *with limit* $x \in \mathcal{T}$.

(iii) *epi* (f) *is a closed subset of* $\mathcal{T} \times \mathbb{R}$.

Moreover, for an arbitrary family $(f_i)_{i\in\mathcal{I}}$ *of l.s.c. mappings,*

$$g(x) := \sup_{i\in I} f_i(x)$$

is l.s.c..

A.2. Dual Pairs and the Fenchel-Moreau Theorem

Hausdorff l.c.t.v.s. and their duals represent the most general framework to establish main the theoretical background of the meaningful (dual) representations results for single-firm and systemic risk measures presented earlier in this monograph. Let us start by introducing a sort of mechanism to obtain such spaces.

Definition A.21. *A dual pair* $\langle\mathcal{V}, \tilde{\mathcal{V}}\rangle$ *consists of a pair of vector spaces* $\mathcal{V}$ *and* $\tilde{\mathcal{V}}$ *together with a bilinear mapping, called pairing,*

$$\left(V, \tilde{V}\right) \mapsto \langle V, \tilde{V}\rangle \tag{A.22}$$

which sends elements from $\mathcal{V} \times \tilde{\mathcal{V}}$ *to* $\mathbb{R}$ *and satisfies the following two properties:*

(i) $\langle V, \tilde{V}\rangle = 0$ *for all* $\tilde{V} \in \tilde{\mathcal{V}} \Rightarrow V = 0$,

(ii) $\langle V, \tilde{V}\rangle = 0$ *for all* $V \in \mathcal{V} \Rightarrow \tilde{V} = 0$.

We say that the pairing induces the dual pair.

With this pairing the vector space $\mathcal{V}$ can be identified with a vector space of linear functionals on $\tilde{\mathcal{V}}$ and vice versa. More precisely, due to the separtion properties of the pairing, $\mathcal{V}$ can be identified with $\{\langle V, \cdot\rangle \mid V \in \mathcal{V}\}$ a subspace of all real mappings on $\tilde{\mathcal{V}}$ denoted by $\mathbb{R}^{\tilde{\mathcal{V}}}$. For this subspace, a natural choice for a topology is the subspace topology of the product topology on $\mathbb{R}^{\tilde{\mathcal{V}}}$. Since the product topology on $\mathbb{R}^{\tilde{\mathcal{V}}}$ is a locally convex Hausdorff topology (see [AB06] Lemma 5.74), all subspace topologies inherit these properties. Now, turning back to $\mathcal{V}$, the topology on $\mathcal{V}$ corresponding to this subspace topology on $\{\langle V, \cdot\rangle \mid V \in \mathcal{V}\}$ is exactly the inital topology $\sigma\left(\mathcal{V},\left(\langle\cdot,\tilde{V}\rangle\right)_{\tilde{V}\in\tilde{\mathcal{V}}}\right)$ called *weak* topology. If it is clear form the context that the vector spaces $\mathcal{V}$ and $\tilde{\mathcal{V}}$ form a dual pair, we simply write $\sigma\left(\mathcal{V},\tilde{\mathcal{V}}\right)$. So, with this topology the vector space $\mathcal{V}$ becomes a (or more precisely $\left(\mathcal{V},\sigma\left(\mathcal{V},\tilde{\mathcal{V}}\right)\right)$ is a) Hausdorff l.c.t.v.s.. If we interchange the roles of $\mathcal{V}$ and $\tilde{\mathcal{V}}$ we obtain that $\left(\tilde{\mathcal{V}},\sigma\left(\tilde{\mathcal{V}},\mathcal{V}\right)\right)$ is also a Hausdorff l.c.t.v.s.. In the following, we frequently use the above described identifications for dual pairs $\langle\mathcal{V},\tilde{\mathcal{V}}\rangle$ without explicit mention.

The term *dual pair* originates from the following important Theorem.

Theorem A.23 ([AB06] Theorem 5.93)**.** *If $\langle\mathcal{V},\tilde{\mathcal{V}}\rangle$ is a dual pair, then the topological dual of $\left(\mathcal{V},\sigma\left(\mathcal{V},\tilde{\mathcal{V}}\right)\right)$ is exactly $\tilde{\mathcal{V}}$.*

Dual pairs are an appropriate way to obtain Hausdorff l.c.t.v.s.. However, the above Theorem also states that every dual pair is obtained from a Hausdorff l.c.t.v.s. $(\mathcal{V},\tau)$ together with its dual space $\mathcal{V}'$. The pairing in this situation is simply the evaluation map. For this reason, we will use the notation $\langle\mathcal{V},\mathcal{V}'\rangle$ for duals pairs. However, this fact gives rise to the following question: Given a dual pair $\langle\mathcal{V},\mathcal{V}'\rangle$ equipped with the weak topologies, are there other topologies on the underlying spaces which preserve the duality? Let us give a definition which addresses this issue.

Definition A.24. *A locally convex topology τ on $\mathcal{V}$ is consistent (or compatible) with the dual pair $\langle\mathcal{V},\mathcal{V}'\rangle$ iff $(\mathcal{V},\tau)'=\mathcal{V}'$. Consistent topologies on $\mathcal{V}'$ are defined in the same fashion.*

Lemma A.25 ([AB06] Lemma 5.97)**.** *Every topology consistent with a dual pair is Hausdorff.*

If we change between consistent topologies, the dual space remains the same. There are other topological „invariant" objects for different consistent topologies.

Theorem A.26 ([AB06] Theorem 5.98, Lemma 5.99)**.** *All locally convex topologies consistent with a given dual pair have the same collection of closed convex sets and therefore the same collection of l.s.c. convex mappings.*

It is quiet natural to assume that the weak topology is the smallest consistent topology. In fact, this claim is true and there is always a largest consistent topology which is called *Mackey* topology denoted by $\tau\left(\mathcal{V},\mathcal{V}'\right)$. We refer the interested reader to [AB06] for a deeper analysis and detailed definition of this topology and close our our review on consistent topologies with the following Theorem.

Theorem A.27 ([AB06] Theorem 5.113)**.** *A locally convex topology τ on $\mathcal{V}$ is consistent with the dual pair $\langle\mathcal{V},\mathcal{V}'\rangle$ if and only if $\sigma\left(\mathcal{V},\mathcal{V}'\right)\subseteq\tau\subseteq\tau\left(\mathcal{V},\mathcal{V}'\right)$.*

Now, lets us consider a dual pair $\langle \mathcal{V}, \mathcal{V}' \rangle$. Each space is equipped with the respective weak topology. Then, for an extended real valued mapping $f\colon \mathcal{V} \to \mathbb{R} \cup \{-\infty, \infty\}$ we can define its *(convex) conjugate* $f^*\colon \mathcal{V}' \to \mathbb{R} \cup \{-\infty, \infty\}$ by

$$f^*(X') = \sup_{X \in \mathcal{V}} \{\langle X, X' \rangle - f(X)\} = \sup_{X \in dom(f)} \{\langle X, X' \rangle - f(X)\}. \tag{A.28}$$

Analogously, for an extended real valued mapping $h\colon \mathcal{V}' \to \mathbb{R} \cup \{-\infty, \infty\}$ its *(convex) conjugate* $h^*\colon \mathcal{V} \to \mathbb{R} \cup \{-\infty, \infty\}$ is defined by

$$h^*(X) = \sup_{X' \in \mathcal{V}'} \{\langle X, X' \rangle - h(X')\} = \sup_{X' \in dom(h)} \{\langle X, X' \rangle - h(X')\}. \tag{A.29}$$

The following Theorem collects some properties of conjugates.

Theorem A.30 ([Zal02] Theorem 2.3.1)**.** *Let $f, g\colon \mathcal{V} \to \mathbb{R} \cup \{-\infty, \infty\}$ and $h\colon \mathcal{V}' \to \mathbb{R} \cup \{-\infty, \infty\}$.*

(i) f^* is convex and $\sigma(\mathcal{V}', \mathcal{V})$-l.s.c., h^* is convex and $\sigma(\mathcal{V}, \mathcal{V}')$-l.s.c..

(ii) The Young-Fenchel inequality holds:

$$f(X) + f^*(X') \geq \langle X, X' \rangle, \quad \forall X \in \mathcal{V}, X' \in \mathcal{V}'. \tag{A.31}$$

(iii) $f \leq g$ implies $f^* \geq g^*$.

An interesting question in this context is the following: Is it true that $f^{**} = (f^*)^* = f$? Since conjugates are always convex and l.s.c. w.r.t. the respective weak topology, these properties represent the minimal requirements on a mapping f to guarantee this equality. Moreover, properness is needed since we have $f^* \equiv \infty$ if there is some $X \in \mathcal{V}$ with $f(X) = -\infty$. In fact, these assumptions are necessary and sufficient. Notice, if we consider convex mappings, weak-l.s.c. already implies l.s.c. w.r.t. every consistent topology. This yields us to the following fundamental Theorem.

Theorem A.32 ([Zal02] Theorem 2.3.3)**.** *Let $\langle \mathcal{V}, \mathcal{V}' \rangle$ be a dual pair where $\mathcal{V}$ is equipped with a consistent topology τ and $\mathcal{V}'$ is equipped with a consistent topology τ'. Then for a proper convex mapping $f\colon \mathcal{V} \to \mathbb{R} \cup \{-\infty, \infty\}$ which is τ-l.s.c. we have*

$$f^{**} = (f^*)^* = f.$$

A.3. Order Structure and Riesz Spaces

Besides the topological structure the vector spaces under our consideration carry an order structure. So, let us assume that there is a partial order $\leq$ defined on the vector space $\mathcal{V}$ which is compatible with the vector space operations. This means that for all $X, Y, Z \in \mathcal{V}$ and all $\alpha \in [0, \infty)$

(i) $X \leq Y \Rightarrow X + Z \leq Y + Z$,

(ii) $X \leq Y \Rightarrow \alpha X \leq \alpha Y$.

Then, we call the pair $(\mathcal{V}, \leq)$ (or simply $\mathcal{V}$ itself) an *ordered vector space*. For a nonempty set $\mathcal{U} \subseteq \mathcal{V}$ an element $X \in \mathcal{V}$ is called *upper bound* if $U \leq X$ for all $U \in \mathcal{U}$. Consequently, the set $\mathcal{U}$ is called *order bounded (from above)*. An element $S \in \mathcal{V}$ which is an upper bound of $\mathcal{U}$ and has the property that for every other upper bound $S \leq X$ holds true is called *least upper bound* or *supremum*. We denote it by $S = \sup\{U \mid U \in \mathcal{U}\}$. If S itself is an element of $\mathcal{U}$, then it is called *largest element* or *maximum*. In each situation, such an element S is unique if it exists. In the same spirit, one can define a *lower bound*, *order bounded (from below)*, *greatest lower bound* or *infimum* denoted by $\inf\{U \mid U \in \mathcal{U}\}$, and *smallest element* or *minimum*. If every subset of $\mathcal{V}$ consisting of two elements has a supremum and an infimum, $\mathcal{V}$ is called *vector lattice* or *Riesz space*. If the same holds true for every nonempty subset, it is called *order complete*. The partial order also gives rise to the notions of *monotonicity, positivity*, and *negativity*. For example, we call a net $(X_j)_{j \in \mathcal{J}}$ *monotone decreasing* if $j \leq_{\mathcal{J}} k$ implies $X_k \leq X_j$. The set

$$\mathcal{V}_+ \coloneqq \{X \in \mathcal{V} \mid X \geq 0\}$$

is called *positive cone* and the elements *positive elements*. The following properties hold for the positive cone in any ordered vector space:

(i) $X, Y \in \mathcal{V}_+ \Rightarrow X + Y \in \mathcal{V}_+$,

(ii) For all $\alpha \in [0, \infty)$, $X \in \mathcal{V}_+ \Rightarrow \alpha X \in \mathcal{V}_+$,

(iii) $X, -X \in \mathcal{V}_+ \Rightarrow X = 0$.

The positive cone is connected with the order structure in the following way. If $X \leq Y$, then obviously $Y - X \in \mathcal{V}_+$ and vice versa. Conversely, if a vector space has a subset with the above properties, then this set induces a compatible partial order on $\mathcal{V}$ by

$$X \leq Y \quad \Leftrightarrow \quad Y - X \in \mathcal{V}_+.$$

Now, with this partial order $\mathcal{V}_+$ becomes the positive cone of the ordered vector space $\mathcal{V}$. For an arbitrary vector subspace $\mathcal{S} \subset \mathcal{V}$ the cone $\mathcal{V}_+ \cap \mathcal{S}$ induces a partial order on this vector subspace. It is called *canonical ordering*.

From now on, we assume that the ordered vector space $\mathcal{V}$ under consideration is at least a Riesz space. For a given $X \in \mathcal{V}$, we define the *positive part*, *negative part* and *absolute value* by

$$X^+ = \sup\{X, 0\}, \quad X^- = \sup\{-X, 0\}, \quad |X| = \sup\{-X, X\}.$$

Note, that for an arbirtary $X \in \mathcal{V}$ we have $X = X^+ - X^-$. Moreover, we have $\inf\{X^+, X^-\} = 0$. Each two elements X, Y of a Riesz space with $\inf\{|X|, |Y|\} = 0$ are called *disjoint* (or orthogonal). For a given set $\mathcal{U} \subset \mathcal{V}$, the *disjoint complement* $\mathcal{U}^d \subset \mathcal{V}$ consists of all the elements of $\mathcal{V}$ which are disjoint to all the elements in $\mathcal{U}$.

The order structure allows us to define similar concepts to the ones we have already seen in topological spaces. This is possible even if the underlying spaces do not carry any topology. They only rely on the order structure.

Definition A.33. *Let $\mathcal{V}$ be a Riesz space. A net $(X_j)_{j\in\mathcal{J}}$ converges in order (or is order convergent) to some $X \in \mathcal{V}$ if there exists a monotone decreasing net $(Z_j)_{j\in\mathcal{J}}$ with $\inf\{Z_j \mid j \in \mathcal{J}\} = 0$ and for all $j \in \mathcal{J}$ we have*

$$|X - X_j| \leq Z_j.$$

In this situation we simply write $X_j \xrightarrow{o} X$ and call X the order limit (of $(X_j)_{j\in\mathcal{J}}$).

A subset $\mathcal{U} \subseteq \mathcal{V}$ is called *solid* if $|Y| \leq |X|$ and $X \in \mathcal{U}$ implies $Y \in \mathcal{U}$. If the solid subset is also a vector subspace, it is called *ideal*. A subset $\mathcal{U} \subseteq \mathcal{V}$ is called *order closed* if for every order convergent net $(X_j)_{j\in\mathcal{J}} \subset \mathcal{U}$ with limit $X \in \mathcal{V}$ we have $X \in \mathcal{U}$. Finally, an order closed ideal is called *band*.

Let us move on to continuity properties. In the following definition, we will focus our attention on real valued mappings. It is also possible to state a more general version. However, we are mainly interested in real valued mappings. Therefore, we restrict our attention to them.

Definition A.34. *Let $\mathcal{V}$ be a Riesz space and $f\colon \mathcal{V} \to \mathbb{R}$ a real valued mapping.*

1. *If for every order bounded set $\mathcal{U} \subseteq \mathcal{V}$ the set $f(\mathcal{U})$ is bounded in $\mathbb{R}$, f is called order bounded.*

2. *If for every net $(X_j)_{j\in\mathcal{J}}$ with $X_j \xrightarrow{o} X$ we have $f(X_j) \to f(X)$, f is called order continuous.*

3. *If for every net $(X_j)_{j\in\mathcal{J}}$ with $X_j \xrightarrow{o} X$ we have $f(X) \leq \liminf_{j\in\mathcal{J}} f(X_j)$, f is called order l.s.c.*

The space of all order bounded linear functionals, denoted by $\mathcal{V}^\sim$, is called *order dual*. The order strucutre of $\mathcal{V}$ naturally induces an order strucutre on $\mathcal{V}^\sim$. First of all, positive linear functionals are described via $\mathcal{V}_+$. So, $l\colon \mathcal{V} \to \mathbb{R}$ is positive if

$l(X) \geq 0$ for all $X \in \mathcal{V}_+$. Every positive linear functional is monotone and therefore order bounded. Note, that an arbitrary linear functional is order bounded if and only if it is the difference of two positive linear functionals (see [AB06] Corollary 8.25). The set of all positive linear functionals serves as the positive cone and $\mathcal{V}^\sim$ becomes an ordered vector space itself. Moreover, it can be shown that it is in fact a Riesz sapce which is *order complete*, i.e. every nonempty order bounded subset has a supremum (see [AB06] Theorem 8.24). Now, the space of all order continuous linear functionals, denoted by $\mathcal{V}_n^\sim$, is called *order continuous dual*. Since every order continuous linear functional is also order bounded, it is a vector subspace of $\mathcal{V}^\sim$. In particular, it is a band (see [AB03] Theorem 1.73). Any order complete Riesz space is the direct sum of any band in it and its disjoint complement (see [AB03] Theorem 1.46). This means $\mathcal{V}^\sim = \mathcal{V}_n^\sim \oplus (\mathcal{V}_n^\sim)^d$. We call the elements of $\mathcal{V}_s^\sim := (\mathcal{V}_n^\sim)^d$ *singular*.

A.4. Topological Riesz Spaces and Riesz Dual Pairs

The vector space $\mathcal{V}$ under consideration is again a Riesz space. If we equipp $\mathcal{V}$ with a linear topology, there are two questions which immediately arise: Are the lattice operations, i.e. positive part, negative part, absolute value, supremum, and infimum, compatible with the topology in some sense? Is there a connection between the topological and the order dual? *Locally solidness* turns out to be the right concept to give an answer to both questions. As for l.c.t.v.s., a linear topology τ on a Riesz space is called locally solid if it has a neighborhood basis at zero consisting of solid sets. We call $(\mathcal{V}, \tau)$ locally solid Riesz space. The connecting property is *uniform continuity*. A mapping $f : \mathcal{V} \to \mathcal{W}$ with the property that for every neighborhood of zero W in $\mathcal{W}$ there is a neighborhood of zero V in $\mathcal{V}$ such that $X - Y \in V$ implies $f(X) - f(Y) \in W$ is called *uniformly continuous*.

Theorem A.35 ([AB06] Theorem 8.41)**.** *A linear topology τ on a Riesz space $\mathcal{V}$ is locally solid if and only if the lattice operations*

$$\begin{aligned}
\cdot^+ : \mathcal{V} \to \mathcal{V}, &\qquad X \mapsto X^+, \\
\cdot^- : \mathcal{V} \to \mathcal{V}, &\qquad X \mapsto X^-, \\
|\cdot| : \mathcal{V} \to \mathcal{V}, &\qquad X \mapsto |X|, \\
\sup : \mathcal{V} \to \mathcal{V}, &\qquad (X, Y) \mapsto \sup\{X, Y\}, \\
\inf : \mathcal{V} \to \mathcal{V}, &\qquad (X, Y) \mapsto \inf\{X, Y\},
\end{aligned}$$

are uniformly continuous.

For locally solid Riesz spaces, the order dual already contains the topological dual as a subspace.

Theorem A.36 ([AB06] Theorem 8.48). *Let $(\mathcal{V}, \tau)$ be a locally solid Riesz space. Then the topological dual $\mathcal{V}'$ is an ideal in the order dual $\mathcal{V}^\sim$. Moreover, $\mathcal{V}'$ is order complete.*

Consider an arbitrary dual pair $\langle \mathcal{V}, \mathcal{V}' \rangle$. Suppose it is possible to introduce a partial order on these spaces which makes them Riesz spaces. For consistent topologies which are also locally solid, $\mathcal{V}'$ is an ideal in $\mathcal{V}^\sim$. Dual pairs with this property are called *Riesz dual pairs*. Consistency with a Riesz dual pair also requires locally solidness. The next question is: Are there spaces with $\mathcal{V}^\sim = \mathcal{V}'$? *Fréchet lattices* present a wide class of spaces with the demanded property. Recall, a t.v.s. $\mathcal{V}$ is called *metrizable* if there exists a metric d on $\mathcal{V}$ such that the open sets are precisely the open sets in the metric sense. If there is such a metric on $\mathcal{V}$ which makes $\mathcal{V}$ (or more precisely $(\mathcal{V}, d)$) to a complete metric space, we call $\mathcal{V}$ *completely metrizable*. Now, a Fréchet lattice is a locally solid Riesz space which is completely metrizable.

Theorem A.37 ([AB06] Theorem 9.11). *Let $(\mathcal{V}, \tau)$ be a Fréchet lattice. Then we have $\mathcal{V}^\sim = \mathcal{V}'$.*

Fréchet lattices already include *Banach lattices*. A Banach lattice is a Riesz space $\mathcal{V}$ equipped with a *lattice norm* $\|\cdot\|$, i.e. $|X| \leq |Y|$ in $\mathcal{V}$ implies $\|X\| \leq \|Y\|$. Additionally, $\mathcal{V}$ (or more precisely $(\mathcal{V}, \|\cdot\|)$) as a normed space is a Banach space. By [AB06] Theorem 8.46, these spaces are locally convex-solid (locally convex and locally solid) when they are equipped with their respective norm topology. Another interesting thing about Fréchet lattices is that every point has a countable neighborhood base. We just have to pick all the metric balls of radius $\frac{1}{n}$ around that point. In general, a topological space with this property is called *first countable*. This property is useful if one tries to analyse a topological property which is described through nets. In fact, for first countable spaces it suffices to consider sequences instad of nets (see [AB06] Theorem 2.40).

Remarkably, every positive linear functional on a Fréchet lattice is continuous (see [AB06] Theorem 9.6). It turns out, that this claim is also true for finite valued convex and increasing mappings. Even for proper convex and increasing mappings, continuity can be preserved on the interior of their domain.

Theorem A.38 ([BF09]). *Let $(\mathcal{V}, \tau)$ be a Fréchet lattice and suppose that $f\colon \mathcal{V} \to \mathbb{R} \cup \{-\infty, \infty\}$ is a proper convex and increasing mapping. Then f is τ-continuous on the interior of its domain.*

Proof. The proof of Theorem 9.6 in [AB06] can be extended to this more general situation. This is done in [BF09] Theorem 1. Let us present this proof here. W.l.o.g., we can assume that $\operatorname{int}(\operatorname{dom}(f))$ is nonempty. We also have that $\operatorname{dom}(f)$ is convex and the interior of a convex set is still convex. Take $Y \in \operatorname{int}(\operatorname{dom}(f))$. Now, the mapping $\tilde{f}\colon \mathcal{V} \to \mathbb{R} \cup \{-\infty, \infty\}$ with $\tilde{f}(X) := f(X+Y) - f(Y)$ is also proper, convex and increasing and we have $0 \in \operatorname{int}\left(\operatorname{dom}\left(\tilde{f}\right)\right)$. So again, there is no loss of generality if we assume $0 \in \operatorname{int}(\operatorname{dom}(f))$. Next, take any metric that generates the topology τ on $\mathcal{V}$ and denote by $\mathcal{B}_r$ the open metric ball of radius $r > 0$ around 0. Fix $N \in \mathbb{N}$ large enough such that $\mathcal{B}_{2/N} \subseteq \operatorname{int}(\operatorname{dom}(f))$. The set $\left\{\mathcal{B}_{1/n}\right\}_{n\in\mathbb{N}}$ forms a countable neighborhood base at 0. Since τ is a locally solid topology, there is a neighborhood base at zero consisting of solid sets. But this means that we can pick a countable neighborhood base at 0 consisting of solid sets. Moreover, there has to be such a solid neighborhood $\mathcal{S}$ with $\mathcal{S} \subseteq \mathcal{B}_{1/2N}$ and hence $\{S + S \mid S \in \mathcal{S}\} =: \mathcal{S}+\mathcal{S} \subseteq \mathcal{B}_{1/N}$. We write $\mathcal{S}_1 := \mathcal{S}$ and consider the solid neighborhood basis at 0 given by $\{\mathcal{S}_n\}_{n\in\mathbb{N}}$ with $\mathcal{S}_{n+1} + \mathcal{S}_{n+1} \subset \mathcal{S}_n$. For each $n \in \mathbb{N}$, we have $\mathcal{S}_{n+1} \subset \mathcal{S}_n \subseteq \mathcal{B}_{1/N}$ and for every sequence $(X_n)_{n\in\mathbb{N}}$ with $X_n \xrightarrow{\tau} 0$ there is a subsequence $(X_{n_k})_{k\in\mathbb{N}}$ with $kX_{n_k} \in \mathcal{S}_k$ for all $k \in \mathbb{N}$. Set $Y_k = \sum_{i=1}^{k} i|X_i|$. We have $Y_k \leq Y_{k+1}$ and $Y_k \in \mathcal{B}_{1/N}$ for all $k \in \mathbb{N}$. Additionally,

$$Y_{k+l} - Y_k = \sum_{i=k+1}^{k+l} i|X_i| \in \mathcal{S}_{k+1} + \mathcal{S}_{k+2} + \ldots + \mathcal{S}_{k+l} \subset \mathcal{S}_k,$$

which means that $(Y_k)_{k\in\mathbb{N}}$ is Cauchy sequence and there exists some $Y \in \mathcal{V}$ with $Y_n \xrightarrow{\tau} Y$ and $Y = \sup_{k\in\mathbb{N}}\{Y_k\}$ (see [AB06] Theorem 8.43). We also have that $Y \in \operatorname{cl}\left(\mathcal{B}_{1/N}\right) \subset \mathcal{B}_{2/N} \subseteq \operatorname{int}(\operatorname{dom}(f))$ and hence $f(Y)$ is finite. Since f is increasing, we have $f(X) \leq f(|X|)$ and $f(-|X|) \leq f(X)$. Using (A.5) (with $X = X$, $Y = 0$ and $\theta = -1$) we obtain $-f(X) \leq f(|X|)$ and finally $|f(X)| \leq f(|X|)$. Now, we have

$$|f(X_{n_k})| \leq \frac{1}{k} f(k|X_{n_k}|) \leq \frac{1}{k} f(Y_k) \leq \frac{1}{k} f(Y) \to 0.$$

So, for every sequence $(X_n)_{n\in\mathbb{N}}$ with $X_n \xrightarrow{\tau} 0$ we can find a subsequence $(X_{n_k})_{k\in\mathbb{N}}$ with $|f(X_{n_k})| \to f(0) = 0$. But this means we also have $f(X_n) \to f(0) = 0$ for every sequence $(X_n)_{n\in\mathbb{N}}$ with $X_n \xrightarrow{\tau} 0$. Indeed, suppose there is some sequence $(X_n)_{n\in\mathbb{N}}$ with $X_n \xrightarrow{\tau} 0$ such that $f(X_n)$ does not converge to $f(0) = 0$. Then, we can find a ball of radius $\epsilon > 0$ around 0 in $\mathbb{R}$ such that $f(X_l) > \epsilon$ for infinitely many $l \in \mathbb{N}$. The subsequence $(X_{n_l})_{l\in\mathbb{N}}$ would still satisfy $X_{n_l} \xrightarrow{\tau} 0$, but it is impossible to extract an other subsequence from $(f(X_{n_l}))_{l\in\mathbb{N}}$ - and hence from $X_{n_l} \xrightarrow{\tau} 0$ - which converges to $f(0) = 0$. But we just showed that this is possible for all sequences in $\mathcal{V}$ with limit $0 \in \mathcal{V}$. Since every Fréchet lattice is first countable, we showed that f is continuous at 0. Finally, Theorem 5.43 in [AB06] implies that f is continuous on the whole $\operatorname{int}(\operatorname{dom}(f))$. □

Since τ-continuity implies τ-l.s.c., a real valued proper convex and increasing mapping $f\colon \mathcal{V} \to \mathbb{R}$ on a locally convex Fréchet lattice $(\mathcal{V}, \tau)$ admits the representation

$$f(X) = \sup_{X'_+ \in \mathcal{V}'} \{\langle X, X' \rangle - f^*(X')\},$$

thanks to the Fenchel-Moreau Theorem A.32. In some situations, however, it might be useful to refine this dual representation. With refinement we mean to reduce the set of dual objects over which the supremum is taken. In particular, a representation without singular elements might be of special interest. For mappings on a locally convex Fréchet lattice this can be derived by a representation only through the order continuous dual $\mathcal{V}_n^\sim$. The following Theorem collects properties linking the order structure and topological structure. This yields the desired representation. It uses the following convexity property for topologies:

Definition A.39 ([BF09]). *A linear topology τ on a Riesz space $\mathcal{V}$ has the (CC) property if for every τ-converging net $(X_j)_{j\in\mathcal{J}}$ with limit $X \in \mathcal{V}$, there exists a subsequence $(X_{j_n})_{n\in\mathbb{N}}$ of $(X_j)_{j\in\mathcal{J}}$ and a sequence $(Z_n)_{n\in\mathbb{N}}$ consisting of convex combinations of elements of $(X_{j_n})_{n\in\mathbb{N}}$ with $Z_n \xrightarrow{o} X$.*

Theorem A.40. *Let $(\mathcal{V}, \tau)$ be a locally convex Fréchet lattice and suppose that $f\colon \mathcal{V} \to \mathbb{R} \cup \{\infty\}$ is a proper convex and increasing mapping which is τ-l.s.c..*

(i) Suppose that τ is order continuous, i.e. for every net $(X_j)_{j\in J}$ $X_j \xrightarrow{o} 0$ implies $X_j \xrightarrow{\tau} 0$, we have $\mathcal{V}' = \mathcal{V}_n^\sim$.

(ii) Suppose that f is order l.s.c.. If $\sigma(\mathcal{V}, \mathcal{V}_n^\sim)$ has the (CC) property, then f is also $\sigma(\mathcal{V}, \mathcal{V}_n^\sim)$-l.s.c..

Proof. The proofs can be found in [BF09]. Let us state the proofs here. To see (i), we first notice that by Theorem A.37 we always have $\mathcal{V}_n^\sim \subseteq \mathcal{V}^\sim = \mathcal{V}'$. Now, the order continuity of the topology gives us that every τ-continuous mapping f is also order continuous. Hence $\mathcal{V}' \subseteq \mathcal{V}_n^\sim$. For part (ii), let $c \in \mathbb{R}$ be arbitrary and consider the set $A = \{X \in \mathcal{V} \mid f(X) \leq c\}$. Moreover, let $(X_j)_{j\in\mathcal{J}} \subset A$ be a $\sigma(\mathcal{V}, \mathcal{V}_n^\sim)$-convergent net with limit $X \in \mathcal{V}$. Now the (CC) property of the topology $\sigma(\mathcal{V}, \mathcal{V}_n^\sim)$ yields us a subsequence $(X_{j_n})_{n\in\mathbb{N}}$ of $(X_j)_{j\in\mathcal{J}}$ and a sequence $(Z_n)_{n\in\mathbb{N}}$ consisting of convex combinations of elements of $(X_{j_n})_{n\in\mathbb{N}}$ with $Z_n \xrightarrow{o} X$. Now, the convexity and order l.s.c. of f imply

$$f(X) \leq \liminf_{n\in\mathbb{N}} f(Z_n) \leq c,$$

and therefore $X \in A$. □

Note, that part (ii) of Theorem A.40 fails to reduce the dual representation to $\mathcal{V}_n^\sim$ if $\mathcal{V}_n^\sim$ does not separate the points of $\mathcal{V}$. This condition is necessary and sufficient for $\sigma(\mathcal{V}, \mathcal{V}_n^\sim)$ to be a Hausdorff topology (see [AB06] Section 2.13). In this case, $\sigma(\mathcal{V}, \mathcal{V}_n^\sim)$ is indeed a locally convex Hausdorff topology, $(\mathcal{V}, \sigma(\mathcal{V}, \mathcal{V}_n^\sim))' = \mathcal{V}_n^\sim$, and we can apply the Fenchel-Moreau Theorem A.32.

A.5. Differentiability and Subgradients

We will need a generalization of the concept of directional derivatives for operators and (extended) real valued mappings. In the following $\mathcal{U}, \mathcal{V}$ and $\mathcal{W}$ are arbitrary Hausdorff l.c.t.v.s.. Each space has a corresponding (topological) dual space and we view them together as dual pairs (see A.2).

Definition A.41. *Let $F : \mathcal{V} \to \mathcal{W}$, $\mathcal{O} \subset \mathcal{V}$, $X \in int(\mathcal{O})$ and $V \in \mathcal{V}$. Define*

$$\delta_+ F(X, V) = \lim_{h \to 0^+} \frac{F(X + hV) - F(X)}{h}.$$

If the limit exists, we call $\delta_+ F(X, V)$ the right directional derivative of F at X in the direction V.

In the same setup, we are able to define

$$\delta_- F(X, V) = \lim_{h \to 0^-} \frac{F(X + hV) - F(X)}{h}.$$

If the limit exists we call $\delta_- F(X, V)$ the *left directional derivative* of F at X in the direction V. However, the left directional derivative can be expressed in terms of the right directional derivative, i.e.

$$\delta_- F(X, V) = -\delta_+ F(X, -V).$$

Therefore we call $\delta_+ F(X, V)$ the *directional derivative* and omit the term right. In situations where $\delta_+ F(X, V)$ and $\delta_- F(X, V)$ exist and coincide for all $V \in \mathcal{V}$ we have the following:

Definition A.42. *Let $F : \mathcal{V} \to \mathcal{W}$, $\mathcal{O} \subset \mathcal{V}$, $X \in int(\mathcal{O})$ and $V \in \mathcal{V}$. Define*

$$\delta F(X, V) = \lim_{h \to 0} \frac{F(X + hV) - F(X)}{h}.$$

If the limit exists for all $V \in \mathcal{V}$, we call the mapping $\delta F(X, \cdot) : \mathcal{V} \to \mathcal{W}$ Gâteaux differential[1] of F at X and say that F is Gâteaux differentiable at X. If $f : \mathcal{V} \to \mathbb{R} \cup \{-\infty, \infty\}$ is Gâteaux differentiable at X and there exists a $V' \in \mathcal{V}'$ such that $\delta f(X, \cdot) = \langle \cdot, V' \rangle$, we call f Gâteaux differentiable with (Gâteaux) derivative $\nabla f(X) \coloneqq V'$ at X.

The following Proposition collects some rules for Gâteaux differentials.

Proposition A.43. *Let $F, G \colon \mathcal{U} \to \mathcal{V}$, $H \colon \mathcal{W} \to \mathcal{U}$, $U \in \mathcal{U}$ and $W \in \mathcal{W}$.*

[1] Some authors call it *weak differential*.

(i) Suppose that F and G are Gâteaux differentiable at $X \in \mathcal{U}$. Then

$$\delta(F \pm G)(X, \cdot) = \delta F(X, \cdot) \pm \delta G(X, \cdot).$$

(ii) Suppose that F and G are Gâteaux differentiable at $X \in \mathcal{U}$. Then

$$\delta(FG)(X, \cdot) = \delta F(X, \cdot) G(X) + F(X) \delta G(X, \cdot),$$

where FG describes the element-wise product of F and G.

(iii) Suppose that H is Gâteaux differentiable at $Y \in \mathcal{W}$ and G is Gâteaux differentiable at $H(Y)$. Then

$$\delta(G \circ H)(Y, \cdot) = \delta G(H(Y), \delta H(Y, \cdot))$$

Proof. Part (i) follows directly from Definition A.42. To prove part (ii) and (iii), first notice that

$$F(X + hU) = F(X) + h\delta F(X, U) + o(h),$$

where $o(h)$ describes some q with

$$\lim_{h \to 0} \frac{q}{h} = 0.$$

Now, for all $U \in \mathcal{U}$ we obtain

$$\begin{aligned}
&\delta(FG)(X, U) \\
&= \lim_{h \to 0} \frac{F(X + hU) G(X + hU) - F(X) G(X)}{h} \\
&= \lim_{h \to 0} \frac{(F(X) + h\delta F(X, U) + o(h))(G(X) + h\delta G(X, U) + o(h)) - F(X) G(X)}{h} \\
&= \delta F(X, U) G(X) + F(X) \delta G(X, U) \\
&\quad + \lim_{h \to 0} \frac{F(X) o(h) + G(X) o(h) + o(h)^2}{h} \\
&\quad + \lim_{h \to 0} (h\delta F(X, U) \delta G(X, U) + \delta F(X, U) o(h) + \delta G(X, U) o(h)) \\
&= \delta F(X, U) G(X) + F(X) \delta G(X, U),
\end{aligned}$$

which proves part (ii). For part (iii) we have for all $W \in \mathcal{W}$

$$\begin{aligned}
&\delta (G \circ H) (Y, W) \\
&= \lim_{h \to 0} \frac{G (H (Y + hW)) - G (H (Y))}{h} \\
&= \lim_{h \to 0} \frac{G (H (Y) + h\delta H (Y, W) + o (h)) - G (H (Y))}{h} \\
&= \lim_{h \to 0} \frac{G (H (Y) + h (\delta H (Y, W) + h^{-1} o (h))) - G (H (Y))}{h} \\
&= \lim_{h \to 0} \frac{G (H (Y)) + h\delta G (H (Y), \delta H (Y, W) + h^{-1} o (h)) + o (h) - G (H (Y))}{h} \\
&= \delta G (H (Y), \delta H (Y, W)) .
\end{aligned}$$

□

For a proper convex mapping $f : \mathcal{V} \to \mathbb{R} \cup \{\infty\}$ and some $X \in \operatorname{dom}(f)$, the directional derivative $\delta_+ f (X, V)$ exists (in $\mathbb{R} \cup \{-\infty, \infty\}$, for points $X \in \operatorname{int}(\operatorname{dom}(f))$ it is finite) for all $V \in \mathcal{V}$. Moreover, it satisfies

$$\delta_+ f (X, V) \leq f (X + V) - f (X) , \tag{A.44}$$

for all $V \in \mathcal{V}$ (see [Zal02] Theorem 2.1.13). Now, if f is actually Gâteaux differentiable with derivative $\nabla f (X) \in \mathcal{V}'$ at X, the above inequality becomes

$$\langle X, \nabla f (X) \rangle \leq f (X + V) - f (X) . \tag{A.45}$$

Even if f is not Gâteaux differentiable with derivative, for all $X \in \mathcal{V}$ we can consider the set

$$\begin{aligned}
\partial f (X) &= \{V' \in \mathcal{V}' \mid \langle V - X, V' \rangle \leq f (V) - f (X) \quad \forall V \in \mathcal{V}\} \\
&= \{V' \in \mathcal{V}' \mid \langle V, V' \rangle + f (V) \leq f (X + V) \quad \forall V \in \mathcal{V}\} ,
\end{aligned}$$

where we set $\partial f (X) = \emptyset$ for $X \notin \operatorname{dom}(f)$. It is called *subdifferential* (at X) and we call f sub-differentiable (at X) if $\partial f (X)$ is nonempty. There is a connection between directional derivatives and subdifferentials.

Theorem A.46 ([Zal02] Theorem 2.4.4, Theorem 2.4.9)**.** *Let $f : \mathcal{V} \to \mathbb{R} \cup \{\infty\}$ be a proper convex mapping and $X \in dom(f)$. Then*

$$\partial f (X) = \partial \delta_+ f (X, \cdot) (0) .$$

Moreover, the following holds true:

(i) If X is in the interior of the domain and f is Gâteaux differentiable with derivative $\nabla f(X)$, then

$$\partial f(X) = \{\nabla f(X)\}.$$

(ii) If f is continuous at X, then $\partial f(X)$ is nonempty, $\delta_+ f(X, \cdot)$ is continuous, and

$$\delta_+ f(X, \cdot) = \max_{V' \in \partial f(X)} \{\langle \cdot, V' \rangle\}.$$

The previous Theorem readily implies that a proper convex mapping $f : \mathcal{V} \to \mathbb{R} \cup \{\infty\}$ which is continuous at $X \in \operatorname{dom}(f)$ is Gâteaux differentiable with derivative at X if and only if $\partial f(X)$ consists of exactly one element. We close this section with the following Theorem.

Theorem A.47 ([BF09])**.** *Let $\mathcal{V}$ be a locally convex Fréchet lattice and suppose that $f : \mathcal{V} \to \mathbb{R} \cup \{\infty\}$ is a proper convex and increasing mapping. Then f is sub-differentiable on the interior of its domain.*

Proof. In Theorem A.38, we have already seen that such an f has to be continuous on $\operatorname{int}(\operatorname{dom}(f))$. Part (ii) of the previous Theorem directly yields the claim. □

References

[AB03] Charalambos D. Aliprantis and Owen Burkinshaw. *Locally solid Riesz spaces with applications to economics.* American Mathematical Soc., 2003.

[AB06] Charalambos D. Aliprantis and Kim C. Border. Infinite dimensional analysis: A hitchhikers guide, 2006.

[ADEH99] Philippe Artzner, Freddy Delbaen, Jean-Marc Eber, and David Heath. Coherent measures of risk. *Mathematical finance*, 9(3):203–228, 1999.

[AKMM21] Maria Arduca, Pablo Koch-Medina, and Cosimo Munari. Dual representations for systemic risk measures based on acceptance sets. *Mathematics and Financial Economics*, 15(1):155–184, 2021.

[AR20] Çağın Ararat and Birgit Rudloff. Dual representations for systemic risk measures. *Mathematics and Financial Economics*, 14(1):139–174, 2020.

[BF09] Sara Biagini and Marco Frittelli. 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*, pages 1–28. Springer, 2009.

[BFFMB19] Francesca Biagini, Jean-Pierre Fouque, Marco Frittelli, and Thilo Meyer-Brandis. A unified approach to systemic risk measures via acceptance sets. *Mathematical Finance*, 29(1):329–367, 2019.

[BFFMB20] Francesca Biagini, Jean-Pierre Fouque, Marco Frittelli, and Thilo Meyer-Brandis. On fairness of systemic risk measures. *Finance and Stochastics*, pages 1–52, 2020.

[CG18] Francesca Centrone and Emanuela Rosazza Gianin. Capital allocation à la aumann–shapley for non-differentiable risk measures. *European Journal of Operational Research*, 267(2):667–675, 2018.

[CIM13] Chen Chen, Garud Iyengar, and Ciamac C Moallemi. An axiomatic approach to systemic risk. *Management Science*, 59(6):1373–1388, 2013.

[CK11] Patrick Cheridito and Eduard Kromer. Ordered contribution allocations: theoretical properties and applications. *Journal of Risk*, 14(1):123, 2011.

[Del02] Freddy Delbaen. Coherent risk measures on general probability spaces. In *Advances in finance and stochastics*, pages 1–37. Springer, 2002.

[Den01] Michel Denault. Coherent allocation of risk capital. *Journal of risk*, 4:1–34, 2001.

[DF21] Alessandro Doldi and Marco Frittelli. Real-valued systemic risk measures. *Mathematics*, 9(9):1016, 2021.

[DFG24] Alessandro Doldi, Marco Frittelli, and Emanuela Rosazza Gianin. Are shortfall systemic risk measures one dimensional? *SIAM Journal on Financial Mathematics*, 15(1):SC1–SC14, 2024.

[EN01] Larry Eisenberg and Thomas H Noe. Systemic risk in financial systems. *Management Science*, 47(2):236–249, 2001.

[FG02] Marco Frittelli and Emanuela Rosazza Gianin. Putting order in risk measures. *Journal of Banking & Finance*, 26(7):1473–1486, 2002.

[FRW17] Zachary Feinstein, Birgit Rudloff, and Stefan Weber. Measures of systemic risk. *SIAM Journal on Financial Mathematics*, 8(1):672–708, 2017.

[FS08] Hans Föllmer and Alexander Schied. *Stochastic Finance: An Introduction in Discrete Time*. Walter de Gruyter, 2008.

[FW15] Hans Föllmer and Stefan Weber. The axiomatic approach to risk measures for capital determination. *Annual Review of Financial Economics*, 7:301–337, 2015.

[Kal05] Michael Kalkbrener. An axiomatic approach to capital allocation. *Mathematical Finance: An International Journal of Mathematics, Statistics and Financial Economics*, 15(3):425–437, 2005.

[KO14] Eduard Kromer and Ludger Overbeck. Representation of bsde-based dynamic risk measures and dynamic capital allocations. *International Journal of Theoretical and Applied Finance*, 17(05):1450032, 2014.

[KOZ16] Eduard Kromer, Ludger Overbeck, and Katrin A. Zilch. Systemic risk measures on general measurable spaces. *Mathematical Methods of Operations Research*, 84(2):323–357, 2016.

[KR09] Mareike Kaina and Ludger Rüschendorf. On convex risk measures on l p-spaces. *Mathematical methods of operations research*, 69(3):475–495, 2009.

[RGMP16] Ben-Zion A Rubshtein, Genady Ya Grabarnik, Mustafa A Muratov, and Yulia S Pashkova. Foundations of symmetric spaces of measurable functions. *Development in Mathematics*, 45, 2016.

[Roc97] R Tyrrell Rockafellar. *Convex analysis*, volume 11. Princeton university press, 1997.

[RS06] Andrzej Ruszczyński and Alexander Shapiro. Optimization of convex risk functions. *Mathematics of operations research*, 31(3):433–452, 2006.

[RV13] Leonard CG Rogers and Luitgard AM Veraart. Failure and rescue in an interbank network. *Management Science*, 59(4):882–898, 2013.

[Tas07] Dirk Tasche. Capital allocation to business units and sub-portfolios: the euler principle. *arXiv preprint arXiv:0708.2542*, 2007.

[Tsa09] Andreas Tsanakas. To split or not to split: Capital allocation with convex risk measures. *Insurance: Mathematics and Economics*, 44(2):268–277, 2009.

[Zal02] Constantin Zalinescu. *Convex analysis in general vector spaces*. World scientific, 2002.

Index of Notations

Declaration of Authorship

I declare that I have completed this dissertation single-handedly without the unauthorized help of a second party and only with the assistance acknowledged therein. I have appropriately acknowledged and cited all text passages that are derived verbatim from or are based on the content of published work of others, and all information relating to verbal communications. I consent to the use of an anti-plagiarism software to check my thesis. I have abided by the principles of good scientific conduct laid down in the charter of the Justus Liebig University Giessen „Satzung der Justus-Liebig-Universität Gießen zur Sicherung guter wissenschaftlicher Praxis“ in carrying out the investigations described in the dissertation.

Selbstständigkeitserklärung

Ich erkläre: Ich habe die vorgelegte Dissertation selbstständig und ohne unerlaubte fremde Hilfe und nur mit den Hilfen angefertigt, die ich in der Dissertation angegeben habe. Alle Textstellen, die wörtlich oder sinngemäß aus veröffentlichten Schriften entnommen sind, und alle Angaben, die auf mündlichen Auskünften beruhen, sind als solche kenntlich gemacht. Ich stimme einer evtl. Überprüfung meiner Dissertation durch eine Antiplagiat-Software zu. Bei den von mir durchgeführten und in der Dissertation erwähnten Untersuchungen habe ich die Grundsätze guter wissenschaftlicher Praxis, wie sie in der „Satzung der Justus-Liebig-Universität Gießen zur Sicherung guter wissenschaftlicher Praxis“ niedergelegt sind, eingehalten.

Pohlheim, April 2024