Multivariate Conditional Risk Measures

with a view towards systemic risk in financial networks

Hannes Hoffmann

Dissertation an der Fakultät für Mathematik, Informatik und Statistik der Ludwig-Maximilians-Universität München

Eingereicht am 18. April 2017

Multivariate Conditional Risk Measures

with a view towards systemic risk in financial networks

Hannes Hoffmann

Dissertation an der Fakultät für Mathematik, Informatik und Statistik
der Ludwig-Maximilians-Universität München
zur Erlangung des akademischen Doktorgrades der Naturwissenschaften
(Dr. rer. nat.)

Erstgutachter:	Prof. Dr. Thilo Meyer-Brandis
Zweitgutachter:	Prof. Dr. Stefan Weber
Drittgutachter:	Prof. Dr. Ludger Overbeck
Eingereicht am:	18.04.2017
Tag der Disputation:	24.07.2017

Eidesstattliche Versicherung

(Siehe Promotionsordnung vom 12.07.11, §8, Abs. 2 Pkt. 5)

Hiermit erkläre ich, Hannes Hoffmann, an Eides statt, dass die Dissertation von mir selbstständig und ohne unerlaubte Beihilfe angefertigt worden ist.

Ort, Datum	Unterschrift Doktorand

Zusammenfassung

Univariate Risikomaße, welche in erster Linie eingeführt wurden, um die Risiken einzelner Finanzprodukte oder Unternehmen zu erfassen, bilden bereits einen sehr gut erforschten Teilbereich der Finanzmathematik. Jedoch wurde im Zuge der globalen Finanzkrise ab 2007 sichtbar, dass die Risikomessung nicht auf einzelne Einheiten des gesamten Systems beschränkt werden darf. Stattdessen sollte ein ganzheitlicher Ansatz verfolgt werden. Aus diesem Grund befasst sich diese Arbeit mit der Erforschung der multivariaten Risikomaße vor dem Hintergrund der systemischen Risikomessung.
Zunächst beschreiben wir auf axiomatische Art und Weise die Unterklasse der bedingten multivariaten Risikomaße, welche zerlegt werden können in ein univariates bedingtes Risikomaß und eine bedingte Aggregationsfunktion. Diese Klasse ist insbesondere eine Erweiterung der unbedingten Risikomaße auf endlichen Wahrscheinlichkeitsräumen aus Chen et al. (2013). Im Rahmen der systemischen Risikomessung kommt dem bedingten Ansatz besondere Bedeutung zu. Durch das Bedingen auf bestimmte Zustände einzelner Untereinheiten wird es ermöglicht deren Einfluss auf das Aufkommen systemischer Risiken präzise zu erfassen. Im Grunde verlangen wir für das Zerlegungsresultat zunächst nur zwei Arten von Monotonie des multivariaten Risikomaßes. Hierauf aufbauend lassen sich die Auswirkungen zusätzlicher Eigenschaften der axiomatischen Beschreibung des systemischen Risikomaßes auf die einzelnen Funktionen der Zerlegung studieren, sowie umgekehrt.

Auf der anderen Seite wirft das bedingte Modellieren des systemischen Risikos die Frage auf, wie mit Risikomessungen unter verschiedenen Informationen umzugehen ist. Im univariaten Fall wurden hierzu verschiedene Konsistenzbedingungen vorgeschlagen. Hierbei werden wir uns auf das Konzept der starken Konsistenz konzentrieren, welche wir im zweiten Teil dieser Dissertation auf den multivariaten Fall übertragen. Die starke Konsistenz besagt, dass Präferenzen, die durch ein bedingtes Risikomaß ausgedrückt werden, ebenfalls für alle Risikomessungen unter weniger Informationen gelten müssen. Unser Hauptaugenmerk wird hierbei auf der Verbindung zwischen der starken Konsistenz und der Zerlegbarkeit der Risikomaße liegen. Diese Fragestellung ist bisher nicht beachtet worden, da für eindimensionale Risiken eine vorgeschaltete Aggregation unnötig ist. Hierzu betrachten wir unter anderem ein Risikomaß, welches stark konsistent ist bezüglich eines Risikomaßes unter voller Information. Letzteres kann, bis auf einen Vorzeichenwechsel, als eine bedingte Aggregationsfunktion aufgefasst werden. Wir zeigen, dass die starke Konsistenz diese Aggregation auf das Risikomaß unter weniger Informationen überträgt. Umgekehrt zeigen wir, dass wenn ein bedingtes Risikomaß stark konsistent mit einem unbedingten, verteilungsinvarianten Risikomaß ist, dann sind beide von der Form eines verallgemeinerten multivariaten Sicherheitsäquivalentes. Mit Hilfe dieser speziellen Darstellung lässt sich ebenfalls die Zerlegbarkeit etablieren.

Bis zu diesem Punkt haben wir uns nur mit der Risikomessung des Gesamtsystems befasst. Auf der anderen Seite ist es jedoch ebenfalls von Interesse zu bestimmen, welchen Anteil die

einzelnen Untereinheiten dazu beitragen. Damit das Risiko nicht willkürlich aufgeteilt wird, sollten gewisse, auf eine gerechte Verteilung bedachte, Kriterien postuliert werden. In dieser Dissertation betrachten wir dazu Allokationen, die sich im Kern befinden, d.h. dass die aggregierte Allokation für jede Untergruppe des Systems kleiner sein sollte als das entsprechende Risiko des Untersystems. Diese Form der Allokation wird häufig für die Aufteilung des Risikos eines Finanzportfolios herangezogen. Wir zeigen hingegen, dass, wenn Interaktionen zwischen den einzelnen Untereinheiten möglich sind, wie etwa in einem Finanznetzwerk, dann können Kernallokationen durchaus Ungerechtigkeiten hervorbringen. In dem von uns benutzten Finanzsystem zeigen wir zudem eine Allokationsmethode auf, die diese Ungerechtigkeit behebt.

Abstract

Univariate risk measures which were primarily introduced for the risk assessment of single financial products or companies, are already a well-studied research area in the field of financial mathematics. The global financial crisis highlighted that the risk measurement should not be restricted to single entities of the system, instead an integrated approach should be pursued. For this reason, the objective of this thesis is to study multivariate risk measures in the context of systemic risk assessment.

First of all, we give an axiomatic description of multivariate conditional risk measures which allow for a decomposition into a conditional univariate risk measure and a conditional aggregation function. This class extends the unconditional risk measures on finite dimensional spaces proposed in Chen et al. (2013). Within the scope of systemic risk assessment the conditional framework is of considerable importance. Particularly, conditioning on certain states of individual subunits enables us to comprehend the emergence of systemic risk. In order that the decomposition result holds, we basically just have to require two types of monotonicity for the multivariate risk measure. On account of this, we are able to identify the relationship between additional properties of the axiomatic description of the systemic risk measure and of the corresponding functions of the decomposition.

In addition, the conditional framework raises the question if there should be some persistent structure to the conditional risk measurements under different information sets. To this end, many consistency properties have been proposed in the univariate case. Among those, we focus on the strong consistency which we generalize to multivariate conditional risk measures. Strong consistency means that all preferences generated by a conditional risk measure must be preserved under a conditional risk assessment with less information. Our main focus lies on the connection of the strong consistency and the decomposability of the risk measures involved. Note that this question has not been raised so far, since aggregation is superfluous for one-dimensional risks. We consider a risk measure which is strongly consistent with respect to a conditional risk measure under full information. Up to a sign change, the latter allows for an interpretation as a conditional aggregation function. We will show that the strong consistency transmits this aggregation to the risk measure under less information. Conversely, we show that if a conditional risk measure and an unconditional, law-invariant risk measure are strongly consistent, then both are generalized multivariate certainty equivalents. As a result of this particular form, we show that also these risk measures allow for a decomposition as above.

So far, we were just concentrating on the risk measurement of the entire system. In addition, it is also of interest to assess how much the single entities contribute to the total risk. To rule out arbitrariness in the attribution scheme, certain criteria which prevent imbalances need to be postulated. For this purpose we consider core allocations in this thesis. We say that an allocation is in the core, if the aggregated allocation of each subsystem never exceeds the risk of the corresponding subsystem. These allocations are frequently suggested in the context of

portfolio allocation. However, we show that in the presence of interactions between the single entities, core allocations treat the entities in an unfair way. Particular examples of an interacting system are financial networks. Moreover, we will find an alternative allocation method for our proposed financial system which repairs the deficiencies.

Acknowledgments

At this point, I want to thank all the people who supported me elaborating this thesis. Among those, I am particularly indebted to my supervisor Prof. Dr. Thilo Meyer-Brandis for giving me the opportunity to work and research at the workgroup financial mathematics. Throughout the last four years, his excellent guidance and the many inspiring and fruitful discussions helped me to improve this thesis. No less I want to express my profound gratitude to Dr. Gregor Svindland for the scientific advice and his ongoing encouragement. Moreover, I do not want to forget to thank Prof. Dr. Stefan Weber and Prof. Dr. Ludger Overbeck for accepting to review my thesis and for their work in the field of systemic risk which I read with much interest.

To all my friends and colleagues, I am very thankful for the time we had at the workgroup financial mathematics and will keep it in best memories. Special thanks go to Martin Bauer, Rebecca Declara, Nils Detering, Andreas Groll, Jacopo Mancin and Daniel Ritter for the many scientific as well as non-scientific discussions we had at work and during the breaks.

Finally, I could not have reached my goals without the encouragement and support of my family and here in particular my wife, my mother and my sister.

Contents

1 Introduction

For the financial industry, the assessment of risk is an essential task for the operation of its core businesses. Apart from the identification of the return of an investment opportunity, knowing the risk inherent to it is necessary for an investment decision. Since the decision makers, in this case the owners of the financial institutions, are only liable for the firm's losses up to their invested capital, they have an incentive to invest in riskier projects. The resulting excess loss in case of a default is then ultimately passed on to the creditors like other financial institutions, industrial companies or the society. For this reason, there is a demand for the regulation of the financial system. This task is undertaken by a regulatory authority.
Before 2007 it was generally accepted that for the regulation of the financial system it is sufficient to monitor every financial institutions individually. However, in the course of the global financial crisis as well as the subsequent European sovereign debt crisis, this perspective towards risk was questioned. The reason was the observation of various negative feedback effects arising from the financial system itself which are far from negligible. This type of *systemic risk* is typically propagated by financial contagion channels, i.e. a triggering event negatively impacts some market participants which might cause further negative effects on the system and ultimately provokes a downward spiral. Prominent examples for a financial contagion channel are default cascades where losses spread through direct contractual obligations among the financial firms. In this case the triggering event is the bankruptcy of one or more financial firms. Due to the initial failures, the creditors must depreciate their assets with the defaulted counterparties in full or at least partially. This in turn can exceed their own loss absorbing capacity leading to further defaults. As a result, the initial defaults might affect a large portion of the financial system. For the studies in the present thesis, this type of default cascades will be the predominant financial contagion channel. Besides the prior direct interactions via credit relationships, also indirect channels like common market exposures, fire sales and funding liquidity play a fundamental role for systemic risk. For instance, concerns about the well-functioning of the financial system can prompt market participants to hoard liquidity. This hoarding of liquidity constraints the granting of loans or even worse results in a dry out of the credit market. Thus, the financial firms are not able to raise funds or at least only at an increased cost.

So far, these external risks have not been priced in the internal risk management process, but were eventually borne by the real economy and the society. This created a rising interest in integrated risk assessment approaches also known as *systemic risk measures*. The overall goal of this thesis is to contribute to this theory by introducing a feasible framework for the systemic risk measurement and to uncover its underlying structure.

For the rest of this introduction we will proceed as follows: To put our proposed framework into context, we begin in Section 1.1 with a basic review of the most popular methods for the study of systemic risk. Inspecting these examples, we observe that essentially three details play an outstanding roll in the description and assessment of systemic risk: Contagion mechanisms as described earlier, prominent monetary risk measures like the Value at Risk or the expected shortfall and the conditioning on events of the system or parts of it. One aim of the present thesis is to bring together these three features in a unified framework.

In addition to the concrete methods presented in Section 1.1, we supplement our review of currently used approaches towards systemic risk in Section 1.3 by two comprehensive measurement frameworks. Both frameworks are related to the well-studied univariate risk measures which we briefly survey in Section 1.2. An important step in the field of univariate risk measures was the introduction of a set of minimal requirements in order to have a reasonable risk assessment and to define risk measures in terms of these properties. Therefore, we ask in this thesis if it is likewise possible to identify a set of desirable properties for systemic risk measures and thus to describe them axiomatically. For multivariate risks there is more than one possible extension of this framework. We focus in this dissertation on the class of decomposable risk measures, which has been introduced in Chen et al. (2013) and which is one of the frameworks considered in Section 1.3. The underlying idea of this class is to separate systemic risk measurements into a prefixed aggregation function and a univariate risk measure. This decomposition allows us to incorporate two out of the three details we observed in Section 1.1. More precisely, contagion mechanisms can be included in the aggregation function and we can reuse classical univariate risk measures. To embed the third observation, we generalize in the first part of the thesis the class of decomposable risk measures to the conditional framework. To familiarize with conditional risk, we state in Section 1.4 the basic properties of conditional univariate risk measures which have been mainly used for a dynamic risk assessment and briefly comment on the relevant literature.

An important issue for conditional risk measures is their adaptability if new information enters into the risk assessment. To avoid discrepancies in the risk measurement under different sets of information, many notions of consistency have been proposed in the literature. The different types are shortly discussed in the second part of Section 1.4. Among those, we concentrate on the strong consistency which we extended to multivariate conditional risk measures. Our main achievement in the second part of the thesis is the revelation of central connections between the strong consistency and the decomposability of the risk measures. Note that this relation has not been studied in the univariate setting, since aggregation is redundant in that case.

In general, there are two main questions in the area of systemic risk. Besides assessing the overall risk of a system, one is also confronted with the question of how much each unit contributes to the measured risk. Therefore, the last part of the thesis is devoted to the allocation of systemic risk. At first, we review in Section 1.5 the allocation methods which are currently used in the systemic risk literature. Most of these methods have already been used in the context of portfolio allocation, or to put it in other words, for the allocation problem in systems with non-interacting units. In this regard, Denault (2001) postulated a set of desirable properties for an allocation leading to the concept of coherent allocations. Central to his studies is the no-undercut property. This property is also observable in most of the suggested allocation methods for systemic risk and can already be traced back to the game theoretic literature, where it appeared under the name (fuzzy) core. In the last part of Section 1.5 we discuss the problems of the (fuzzy) core for interacting systems. This discussion refers to the set of problems of the final chapter of this thesis where we observe that the (fuzzy) core results in unfair allocations.

1.1 Current approaches towards systemic risk

To contextualize and position our study of systemic risk, we survey the most prominent approaches used in the literature. Throughout we assume a financial system comprising d financial institutions, which we sometimes also call banks for the sake of brevity. Moreover, we assume that the d-dimensional risk factors are random variables on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$.

1.1.1 Market based systemic risk measures

Market based systemic risk measures hypothesize that all systemically relevant information is included in current market prices. The advantage of the market based approach is that each financial firm can continuously monitor the systemic risk on its own, since market prices are accessible to all market participants at any time. In contrast, this is not possible for the models we will consider later as they depend on detailed balance sheet information of the financial firms which are usually not publicly available. Moreover, if the market is efficient, that is all necessary information is indeed contained in the quoted prices, then the market based approaches also take all possible sources of systemic risk into account, whereas the other methods surveyed in this introduction mostly concentrate on a single channel of contagion, like for instance default cascades. However, this comes at the cost that we are also not able to identify the relevant sources of the systemic risk.

The four most prominent examples for market based systemic risk measures with more than a thousand citations are the conditional-Value at Risk (CoVaR) and the co-expected shortfall (CoES) from Adrian and Brunnermeier (2016) and the marginal expected shortfall (MES) and the systemic expected shortfall (SES) from Acharya et al. (2017). Another example is the systemic risk measure SRISK from Acharya et al. (2012) or Brownlees and Engle (2016).

As already indicated by its name, the CoVaR is a conditional generalization of the popular risk measure Value at Risk currently used for financial risk management. Recall that the Value at Risk at level α of a univariate risk factor F, denoted by $\mathrm{VaR}_\alpha(F)$, is up to a sign change the α-(upper-)quantile of the distribution function of F, see further Föllmer and Schied (2011). Akin to this construction, the CoVaR is derived from a conditional distribution function, where the conditioning is on a certain crisis event A of a single financial firm. More precisely, the $\mathrm{CoVaR}_\alpha^{sys|A}$ at level $\alpha \in (0,1)$ is implicitly given by

$$\mathbb{P}\left(X_{sys} \leq \mathrm{CoVaR}_\alpha^{sys|A} \middle| A \right) = \alpha, \tag{1.1.1}$$

where the set A is of the form $\{X_i = \mathrm{VaR}_\alpha(X_i)\}$ for some financial firm i and X_{sys} and X_i are the index return of the financial system, resp. the stock price return of i. That is the conditional-Value at Risk is the threshold such that the probability that the market return X_{sys} drops below $\mathrm{CoVaR}_\alpha^{sys|A}$ in the crisis event A is exactly α. The stress event of the current form has two major drawbacks: firstly, for continuous distributions the probability that the event A occurs is zero and secondly, it does not consider scenarios where the financial firm i is even worse off. As a result extensions to $A = \{X_i \leq \mathrm{VaR}_\alpha(X_i)\}$ were suggested in the literature, see for instance Bernard et al. (2013) or Girardi and Ergün (2013). However, with this extension the quantile regression procedure from Adrian and Brunnermeier (2016) for the estimation of the CoVaR

cannot be applied anymore.
In the example section of Chapter 2, we will see that the CoVaR fits naturally into our proposed framework of decomposable risk measures. Additionally, we will modify the underlying aggregation of the system. That is, we replace the current aggregation $X_{sys} = \sum_{i=1} w_i X_i$ in (1.1.1), where w_i is the index weight of bank i, by a more sophisticated aggregation which introduces an explicit modeling of contagion. With the resulting modified conditional-Value at Risk we perform a small numerical exercise at the end of Section 2.4.

Clearly, the CoVaR as defined in (1.1.1) is motivated by the question of how much the financial system is impacted by a distress of a single institution. In addition to the CoVaR, Adrian and Brunnermeier (2016) also specified a so-called exposure-CoVaR which is given for each financial firm i by

$$\mathbb{P}\left(X_i \leq \mathrm{CoVaR}_\alpha^{i|sys} \,\middle|\, X_{sys} = \mathrm{VaR}_\alpha(X_{sys})\right) = \alpha. \tag{1.1.2}$$

In contrast to (1.1.1) the exposure-CoVaR targets the issue of how much the single institutions participate in a systemic event. This is also the baseline question for the marginal expected shortfall. For the construction of the MES Acharya et al. (2017) use the expected shortfall as risk measure for the overall risk of the system. The expected shortfall ES_α at level α of the market index return is given by

$$\mathrm{ES}_\alpha(X_{sys}) = \mathbb{E}_\mathbb{P}\left[X_{sys} \mid X_{sys} \leq \mathrm{VaR}_\alpha(X_{sys})\right] = \sum_{i=1}^{d} w_i \mathbb{E}_\mathbb{P}\left[X_i \mid X_{sys} \leq \mathrm{VaR}_\alpha(X_{sys})\right].$$

The marginal expected shortfall for institution i is now defined as the marginal contribution of this bank to the overall risk, i.e.

$$\mathrm{MES}_\alpha^i = \frac{\partial \mathrm{ES}_\alpha(X_{sys})}{\partial w_i} = \mathbb{E}_\mathbb{P}\left[X_i \mid X_{sys} \leq \mathrm{VaR}_\alpha(X_{sys})\right]. \tag{1.1.3}$$

Hence the MES can be interpreted as the average loss of a single financial firm given that the whole system is in distress. Note that in contrast to the exposure-CoVaRs in (1.1.2), the marginal expected shortfalls of the single institutions sum up to the overall risk. Thus, the MES is also an allocation procedure for the expected shortfall of the system. We will meet this allocation procedure again in Section 1.5 under the name Euler allocation.
Without going too much into detail we comment that also the SES and the SRISK can be subsumed to be of the type

$$\mathbb{E}_\mathbb{P}\left[X_i \mid \{\text{systemic crisis}\}\right],$$

where X_i is a certain risk factor for the i-th financial firm.

As a final example for a market based systemic risk measurement, we present the more theoretical model of Brunnermeier and Cheridito (2014). This model takes the perspective of a regulating agency which is concerned that a financial system might have negative effects on the society. Here, the state of the society is represented by the real gross domestic product. Brunnermeier and Cheridito assume that the future net-worths X_i of the financial institutions are the main drivers of risk in the system. Clearly, if a net-worth falls below zero the corresponding bank cannot continue its operations which in turn negatively affects the economy. Conversely, if a financial firm performs extremely well this should also have a positive effect on the society. Thus, it is assumed that each financial firm's net-worth enforces an externality on the future gross domestic product Y which is given by

$$E_i = -\alpha_i X_i^- + \beta_i (X_i - \theta_i)^+,$$

where $\alpha_i, \beta_i, \theta_i$ are positive constants and x^- and x^+ denote the negative, resp. positive part of $x \in \mathbb{R}$. Moreover, it is assumed that the preferences of the risk-averse regulator can be described by a utility function U. Then the SystRisk proposed in Brunnermeier and Cheridito (2014) is given by

$$\inf \left\{ m \in \mathbb{R} : U\left(Y + m + \sum_{i=1}^{d} E_i \right) \geq U(\widetilde{Y}) \right\}, \tag{1.1.4}$$

where $\widetilde{Y}$ is some reference economy. Therefore, systemic risk is modeled as the minimal bailout cost of the regulator such that she is indifferent between the resulting future economy and the reference economy.

1.1.2 Mean-field dynamics

Another approach for the study of systemic risk in financial markets is to use a reduced form model of the mean-field type. Although, this dynamic approach does not directly fall into the scope of the approach to systemic risk taken in this thesis, we mention it here, since it is an integral part of the current proceedings in the field of systemic risk and it stresses once again the importance of the inclusion of direct interactions within the system. The general idea is to model bank specific risk factors as continuous-time stochastic processes which are given by a system of interlacing diffusions.

To illustrate this idea, we briefly review the simplified model from Fouque and Sun (2013). In this work the authors consider the log-monetary reserves $X_t = (X_t^1, ..., X_t^d)^\top$ of a system of d banks as risk factors. The log-reserves are assumed to be diffusion processes with corresponding dynamics

$$dX_t^i = \frac{\alpha}{d} \sum_{j=1}^{d} \left(X_t^j - X_t^i \right) dt + \sigma dW_t^i, \quad i = 1, ..., d, \tag{1.1.5}$$

where $\alpha, \sigma > 0$, $W_t = (W_t^1, ..., W_t^d)^\top$ is a d-dimensional standard Brownian motion and $X_0 = x_0 \in \mathbb{R}^d$. Note that (1.1.5) might attain negative values, but since we are considering a logarithmic transformation, the actual monetary reserves are always positive. By (1.1.5) the evolution of the log-reserves can be split into an exchange with the system and a risky bank specific part. Here the interactions between the banks are captured by the drift term in (1.1.5). That is, if the log-reserves X_t^j of bank j exceed the ones of bank i, then the log-reserves of i increase proportional to the difference $X_t^j - X_t^i$ and the factor α, whereas the log-reserves of bank j decrease by the same amount. Hence, we have a simplified model for borrowing and lending relationships within a financial system.

The insights on systemic risk are the following: One observation in this stylized model is that by increasing the parameter α corresponding to a higher borrowing and lending level, the distribution of defaults is more concentrated near zero. Here default means that the log-reserves fall below a fixed threshold $\theta \in \mathbb{R}$ during a specific time period $[0, T]$. Thus, we infer that borrowing and lending commonly increases the stability of the financial system. The downside is that it comes along with the expansion of systemic events. In Fouque and Sun (2013) an event is systemic if the average log-reserve of the financial system falls below a threshold $\theta_{\text{sys}} \in \mathbb{R}$ before time T, i.e.

$$\min_{t\in[0,T]} \frac{1}{d}\sum_{i=1}^{d} X_t^i \leq \theta_{\text{sys}}.$$

Note that by the borrowing and lending activities this event corresponds more likely to many harmless defaults rather than to a huge reserve shortfall of a single institution. The phenomenon of having a small probability of defaults, but at the same time severe losses in case of a systemic event is also known as the robust-yet-fragile property of a financial system, cf. Gai and Kapadia (2010). Consequently, we cannot say a priori if a more integrated system is beneficial or disadvantageous. This will be a major concern in our study of appropriate allocations for systemic risk in Chapter 4. Finally, another important result in this mean-field model is that as the number of banks in the financial system increases, the dynamics of the log-reserves approach independent Ornstein-Uhlenbeck processes. This decoupling is also known as the propagation of chaos.

The mean-field model of Fouque and Sun (2013) has been extended in multiple ways. For instance, Kley et al. (2014) deviated the environment to a more abstract risk factor which they called financial robustness. Moreover, they extended (1.1.5) in the following two ways. Firstly, the idiosyncratic shock is driven by the wider class of Lévy processes instead of Brownian motions and secondly, the interactions are assumed to be inhomogeneous in the sense that the mean reversion level of the drift part is changed from the system average $\frac{1}{d}\sum_{j=1}^{d} X_t^j$ to a weighted average $\sum_{j=1}^{d} w_{ij} X_t^j$ for all $i = 1, ..., d$. Kley et al. used the inhomogeneity to incorporate a core-periphery structure which is frequently observed in real-world financial systems. A core-periphery structure of a financial system postulates the existence of a subsystem which is highly interconnected and the remaining system is only interacting with this dense subsystem. In particular, in Kley et al. (2014) the weights w_{ij} are interpreted as the percentage of the total credits which bank i has extended to bank j. Therefore, the inhomogeneity expresses the idea that a bank's robustness should be more affected by the robustness of its major debtors. A similar approach can also be found in Battiston et al. (2012). For the modeling of monetary reserves an inhomogeneous coupling was considered in Fouque and Ichiba (2013). In this paper the authors also include a new static bank specific drift term, which is interpreted as lending to a central bank. The ability to lend or borrow to a central bank is also the starting point to the subsequent studies of Carmona et al. (2015, 2016). A central feature of their works is the introduction of a stochastic control problem where each financial firm optimizes its lending and borrowing to the central bank under a quadratic cost function. In Carmona et al. (2016) an additional delay in the control is incorporated. In both works, they could find explicit Nash-equilibria for systems with finitely many banks and in Carmona et al. (2015) they further derive the asymptotics for a growing number of banks in the financial system.

All the models above have a major drawback. Once a financial firm receives money from a counterparty, it has no obligation to pay back the money in the future. This deficiency is repaired by the contagion models which explicitly capture the asset and liability structure of the single financial firms and thus improve the description of the lending and borrowing activities within the financial system.

1.1.3 Contagion models

For the market based risk measures from Section 1.1.1 interconnections between financial firms were solely based on a probabilistic dependence. For instance the CoVaR given in (1.1.1) measures the impact of a distress of a single institution on the financial system. However, it cannot detect the source of the systemic risk, that is if the single institution is simply comonotonic to the system or if it is also causal for the worse performance of the system. The contagion models which we review in this section concentrate on the second issue.

Generally speaking, a contagion model is a procedure which determines the losses of each financial institution given that the system or parts of it are negatively affected by some initial (random) shock. In particular, this determination includes possible domino effects arising from the interconnectedness of the system. However, the procedure has to be performed for every possible shock scenario separately. Thus, we see that contagion models form a convenient class of elaborated aggregations, but to assess the systemic risk a further risk evaluation is necessary which corresponds exactly to our proposed framework of decomposable risk measures. Moreover, the inclusion of direct interactions in the system, described by a contagion model, is the essential part which distinguishes the portfolio from the systemic risk allocation problem which we study in Chapter 4.

We begin with a short review of default contagion models in a financial system with interbank lending. Afterwards, we discuss suggested extensions to this model. Throughout it is assumed that the financial system can be depicted as a network of d nodes representing the financial firms which are connected via weighted and directed edges. More precisely, if bank j has extended credits to bank i with a total amount $L_{i,j}$, then this is represented in the network by an edge with weight $L_{i,j}$ pointing from i to j. Knowing the matrix $\mathbf{L} = (L_{i,j})_{i,j=1,\dots,d}$ results in a full description of the interbank assets and liabilities of each institution. We stress that, unless there are no further obligations to disclosure, this information is at most available to supervising agencies which makes these models more difficult to specify than market based risk measures. Moreover, we assume that each institution has sufficiently many external assets such that assets exceed liabilities, i.e. each bank has a positive equity value. For simplicity, we postulate the absence of external liabilities, otherwise we could introduce them as a separate new node in the network. Because the external assets of each financial institution are exposed to the market, their value might deteriorate at a future point in time such that the equity x drops below zero and thus the bank is in default. The former regulatory approach ended at this point, i.e. the task was to minimize the likelihood of a bank's default given that it is decoupled from any systemic effects. This standalone perspective seems to be reasonable under the assumption of an otherwise stable system. However, in a distressed market situation the failure of an institution imposes losses on its already weakened creditors and possibly causing their default. As long as no one steps

into the breach the defaults can spread further to the creditors of the creditors and so on and on. Therefore, especially in the context of systemic risk default contagion should be considered.
If the system is experiencing contagion, then two typical question are, how many institutions are affected and by how much? These questions are answered by contagion models. Generally, we differentiate between two approaches, namely cascade and equilibrium models. Whereas cascade models assume that the creditors of a defaulting bank can recover a fixed percentage of the defaulted loans, equilibrium models focus on the opposite side of the balance sheet, i.e. all assets of the defaulting bank must be liquidated and the residual value is then distributed among the creditors. In short, the recovery is fixed in cascade models, but depends on the severity of the default in equilibrium models. Both procedures can be translated into the following iteration:
Let $(a(x,n))_{n\in\mathbb{N}}$ be a sequence of real-valued vectors representing the shortfalls of capital of the banks. Firstly, we suppose that no institution is in default, or in other words all firms can settle their debt. Thus, the shortfalls of the financial institutions are exclusively given by their initial loss in capital $a(x,0) = x^-$, where $x^- := -\min\{x,0\}$ denotes the negative part of the equity x after the external assets of the financial firms were exposed to an adverse market event. Those banks which have a strictly positive shortfall default and in the next iteration the default propagates to the adjacent banks, that is to the banks to whom they are indebted to. How the losses spread is captured by a function $f : \mathbb{R}^d \to \mathbb{R}^d$. The shortfall after the first round and all subsequent rounds of defaults is then given by

$$a(x,n) = \big(x - f\big(a(x,n-1)\big)\big)^-, \quad n \geq 2. \tag{1.1.6}$$

For the cascade approach, see e.g. Furfine (2003), Upper and Worms (2004), Battiston et al. (2012), Cont et al. (2013), Amini et al. (2016) or Detering et al. (2016), a typical choice for the function $f = (f_1, ..., f_d)$ is given by

$$f_i(a) = \sum_{j:a_j>0} L_{j,i}(1-R_j), \quad \text{for all } i = 1,...,d, \tag{1.1.7}$$

where R_j is the fraction of the liabilities which can be recovered in the liquidation process of bank j. Therefore, f_i specifies the total loss of bank i due to the depreciation of credits extended to defaulting banks. Cont et al. (2013) even argue in favor of the extreme case of zero recovery $R_i = 0$, since liquidation of a financial institution usually extends over several years. Because of the special structure of the function f in (1.1.7) the shortfall of capital in (1.1.6) only varies from one step to the next if there is a new defaulting bank. However, this can only happen d times and thus the iteration terminates in a finite number of steps. The fast convergence makes cascade models particularly appealing for large networks.
Before we proceed with equilibrium models, we present the cascade model suggested in Cont et al. (2013) to measure the riskiness of a single institution. In a first step, Cont et al. introduce the Default Impact which utilizes the procedure (1.1.6) in conjunction with (1.1.7). To be more precise, the Default Impact (DI) is defined as the aggregate loss of capital induced by the contagion in the interbank market after bank j has defaulted, i.e.

$$\mathrm{DI}(x,j) := \sum_{i=1}^{d} \min\{\widetilde{x}_i, a_i(\widetilde{x},d)\},$$

where $\widetilde{x}_j := 0$ and $\widetilde{x}_i := x_i$ for all $i \neq j$. Cont et al. (2013) further extended the Default Impact to incorporate a second source of systemic risk, common asset holdings. For this purpose they introduce a comonotonic shock ε to banks' capital which is given by $\varepsilon_i = x_i g_i(Z)$ for all $i = 1, ..., d$, where $g : \mathbb{R} \to (-1, 0]^d$ is an increasing function and Z is a real-valued random variable representing the state of the economy. Similar to the MES in (1.1.3) they define the Contagion Index (CI) for financial institution j at level α by

$$\mathrm{CI}(j, x) := \mathbb{E}\left[\mathrm{DI}(x + \varepsilon, j) \mid Z \leq \mathrm{VaR}_\alpha(Z)\right].$$

So once more the assessment of systemic risk is based on an aggregation function, here the Default Impact, a univariate risk measure and a conditioning on a systemic event.

Contrarily, a representative example for the function f in (1.1.6) in an equilibrium model is given by

$$f(a) = \mathbf{\Pi}^\top a, \tag{1.1.8}$$

where $\mathbf{\Pi} = (\Pi_{i,j})_{i,j=1,...,d}$ is the matrix of relative liabilities, that is $\Pi_{i,j}$ is the fraction of the total liabilities of bank i which it owes to bank j. Recall, that the relative liability matrix already appeared in the mean-field model proposed in Kley et al. (2014). The function f specified in (1.1.8) is based on the idea that all creditors of a financial firm have equal seniority and thus any losses resulting from a failure of the bank must be distributed uniformly, that is proportional to the nominal value of credits. The propagation of this keynote in the systemic risk literature can be traced back to the seminal paper of Eisenberg and Noe (2001) on the clearing of payment systems. To obtain the original model of Eisenberg and Noe, we do not only have to lower bound the capital shortfalls as it is already done in (1.1.6), but also have to ensure that they do not exceed the corresponding total liabilities of the financial firms.

In the absence of cycles in the interbank network the equilibrium model (1.1.8) converges, like the cascade models, in a finite number of iterations. Conversely, if a number of defaulting banks form a cycle, then each of these financial firms is impacted by its own loss after the loss has circled once. Moreover, this procedure repeats an infinite number of times leading to two possible outcomes: an unbounded increase in losses or a stable fixpoint in (1.1.6). Under rather weak assumptions on the network Eisenberg and Noe (2001) showed that a unique fixpoint exists. They also formulated a so-called fictitious default algorithm which converges in a finite number of steps to the unique fixpoint.

The contagion model of Eisenberg and Noe (2001) has successively been extended to include other channels of contagion or effects which result in an amplification of the losses. A prominent example for the latter is the inclusion of fixed bankruptcy costs like in the cascade models (1.1.7), cf. Rogers and Veraart (2013), Elliott et al. (2014) or Glasserman and Young (2015). Besides the already existing shortfall of capital in case of a default, a further driver for bankruptcy costs is that most of the remaining assets have to be sold in a distressed market environment in the liquidation process. Particularly, this applies to less marketable assets. Furthermore, financial firms might already be obliged to sell illiquid assets before a default occurs. However, if a financial institution is forced to sell a huge number of illiquid assets in a short period of time, it has to accept almost every price. Such an event is called a fire sale. If the accounting standard is based on a mark-to-market valuation, a fire sale not only imposes losses on the selling bank,

but also affects financial firms holding a similar portfolio of illiquid assets due to the deterioration in prices. At the worst this forces a further bank to sell some of its illiquid assets leading possibly to a downward spiral of prices. Because the financial firms do not need to have any business relation with the other banks, this yields an indirect channel of contagion. Cifuentes et al. (2005) showed how the price collapse of an illiquid asset can be incorporated into the equilibrium model of Eisenberg and Noe (2001). Here the price of an illiquid asset is calculated by an inverse demand function which depends on the total amount of illiquid assets sold in the market. For other works which highlight illiquidity as a channel of contagion we also refer to Gai and Kapadia (2010), Amini et al. (2015, 2016) or Feinstein and El-Masri (2015).
So far, the only direct contagion channel we considered was through debt obligations. Elliott et al. (2014) or Elsinger (2011) further studied the effect of mutual equity holdings in the financial system. For a combined treatment of bankruptcy costs, fire sales and mutual equity holdings we refer to Awiszus and Weber (2015).

All of the methods presented so far are more or less concrete models for the study of systemic risk. By contrast, the aim of this thesis is to approach the measurement of systemic risk on a more general level via a characterization in terms of a set of underlying properties. In preparation for this aim, we proceed by first recalling the extensively studied axiomatic description of univariate monetary risk measures in the ensuing section. Thereafter, we have a look at two promising extensions of univariate risk measures to the multivariate setting. Moreover, since one of these extensions is also the keynote to the approach pursued in this dissertation, we further elaborate on the scopes and intersections of the two extensions.

1.2 Univariate risk measures

Denote by $\mathcal{X}$ a linear space of bounded functions $F : \Omega \to \mathbb{R}$ on a given set of scenarios Ω such that $\mathbb{R} \subseteq \mathcal{X}$. The functions $F \in \mathcal{X}$ model risk factors which have a monetary interpretation and particularly $m \in \mathbb{R}$ is a cash amount. Most commonly, $F(\omega)$ models the discounted future profit/loss of a financial position in the scenario $\omega \in \Omega$. We remark that in the literature on risk measures there are two coexisting interpretations of $F(\omega)$, for some authors a positive value means a loss and for others it denotes a profit. In this thesis, we follow the latter convention.
Next we introduce the basic notion of a monetary risk measure. The objective of this concept is to assign to every random position F a cash amount $\eta(F)$ as a measure of its riskiness. In addition, we require that the risk assessment fulfills a certain set of reasonable properties, which we discuss below. This axiomatic description of risk has been triggered by the pioneering work on coherent risk measures in Artzner et al. (1999). Moreover, the coherent risk measures were further generalized to convex risk measures in Frittelli and Rosazza Gianin (2002) and Föllmer and Schied (2002).

Definition 1.2.1. *We say that the mapping $\eta : \mathcal{X} \to \mathbb{R}$ is a* monetary risk measure*, if the following properties hold for all $F, G \in \mathcal{X}$:*

Antitonicity: *$F \leq G$ implies that $\eta(F) \geq \eta(G)$;*

Cash-invariance: *For all $m \in \mathbb{R}$, we have that $\eta(F + m) = \eta(F) - m$.*

A monetary risk measure is called a convex risk measure *if it is additionally*

Convex: *For all $F, G \in \mathcal{X}$ and $\lambda \in [0,1]$, we have that $\eta(\lambda F + (1-\lambda)G) \leq \lambda\eta(F) + (1-\lambda)\eta(G)$.*

Furthermore, a coherent risk measure *is a convex risk measure that additionally fulfills the property of*

Positive homogeneity: *For all $F \in \mathcal{X}$ and $\lambda \geq 0$, we have that $\eta(\lambda F) = \lambda\eta(F)$.*

The antitonicity property is the minimal requirement to ask for in order to have a meaningful risk evaluation. That is a financial position generating a higher profit or a lower loss in each scenario compared to another financial position, should always be considered less risky. The cash-invariance is the main ingredient such that the risk measure allows for a monetary interpretation, since adding a certain amount of cash to a financial position reduces the risk by the same amount. Thus, we have that $\eta(F)$ is the exact cash amount which we must add to a financial position F such that this new position is riskless, i.e. has zero risk. The convexity property is related to diversification effects in the risk measurement. By using the cash-invariance of a monetary risk measure it can be easily shown that convexity is already implied by the weaker property of

Quasi-convexity: For all $F, G \in \mathcal{X}$ and $\lambda \in [0,1]$ it holds that $\eta(\lambda F + (1-\lambda)G) \leq \max\{\eta(F), \eta(G)\}$.

Thus, a convex risk measure is always in line with the diversification idea that combining two financial positions should reduce the risk. For a better understanding of a coherent risk measure, we need to introduce the notion of

Subadditivity: For all $F, G \in \mathcal{X}$ it holds that $\eta(F+G) \leq \eta(F) + \eta(G)$.

The subadditivity property corresponds to the idea that the merger of financial positions is always preferable from a risk management point of view. It can then be easily shown that if the monetary risk measure is normalized in the sense that $\eta(0) = 0$, then two of the properties of convexity, positive homogeneity and subadditivity imply the remaining one. Thus, a coherent risk measure can equivalently be described as a normalized convex risk measure which is subadditive. That is coherent risk measures are monetary risk measures which incorporate diversification and synergy effects.

Apart from the axiomatic description of univariate monetary risk measures, they can alternatively be characterized by their corresponding acceptance set. This characterization is also the starting point for the second generalization to multivariate risk measures which we survey in Section 1.3.

Definition 1.2.2. *We call*

$$\mathcal{A}_\eta := \{F \in \mathcal{X} : \eta(F) \leq 0\}$$

the acceptance set *of the monetary risk measure η.*

As we have already seen above, the cash-invariance implies that the financial position $\eta(F) + F \in \mathcal{A}_\eta$. Moreover, for each cash amount $m < \eta(F)$ we have that $\eta(F+m) = \eta(F) - m > 0$ and thus $m + F \notin \mathcal{A}_\eta$. Hence, it should also be possible to characterize a monetary risk measure as the smallest cash amount which has to be added in order to make a financial position acceptable. This idea is formalized in the following Proposition which is proven in Föllmer and Schied (2011). For this we will need the supremum norm which is given by $\|F\|_\infty := \sup_{\omega\in\Omega} |F(\omega)|$.

Proposition 1.2.3 (See Proposition 4.6 in Föllmer and Schied (2011)). *Let η be a monetary risk measure and $\mathcal{A}_\eta$ the corresponding acceptance set. Then*

(i) $\mathcal{A}_\eta \neq \emptyset$, $\inf\{m \in \mathbb{R} : m \in \mathcal{A}_\eta\} > -\infty$, $\mathcal{A}_\eta$ *is* $\|\cdot\|_\infty$*-closed in* $\mathcal{X}$ *and*

$$F \in \mathcal{A}_\eta, G \in \mathcal{X}, G \geq F \text{ implies that } G \in \mathcal{A}_\eta. \tag{1.2.1}$$

(ii) η *can be represented in terms of its acceptance set, that is for all* $F \in \mathcal{X}$

$$\eta(F) = \inf\{m \in \mathbb{R} : F + m \in \mathcal{A}_\eta\}. \tag{1.2.2}$$

(iii) η *is a convex risk measure if and only if* $\mathcal{A}_\eta$ *is a convex set, i.e. for all* $F, G \in \mathcal{A}_\eta$ *and* $\lambda \in [0,1]$ *we have that* $\lambda F + (1-\lambda)G \in \mathcal{A}_\eta$.

(iv) η *is positively homogeneous if and only if* $\mathcal{A}_\eta$ *is a cone, i.e. for all* $F \in \mathcal{A}_\eta$ *and* $\lambda \geq 0$ *it holds that* $\lambda F \in \mathcal{A}_\eta$. *In particular, this implies that* η *is a coherent risk measure if and only if* $\mathcal{A}_\eta$ *is a convex cone.*

Alternatively, one could start specifying a set $\mathcal{A}$ of acceptable financial positions and define a corresponding minimal capital injection $\eta_\mathcal{A}$ in a similar fashion as in (1.2.2), that is

$$\eta_\mathcal{A}(F) := \inf\{m \in \mathbb{R} : F + m \in \mathcal{A}\} \quad \text{for all } F \in \mathcal{X}. \tag{1.2.3}$$

Proposition 1.2.4 (See Proposition 4.7 in Föllmer and Schied (2011)). *Let* $\emptyset \neq \mathcal{A} \subseteq \mathcal{X}$ *such that* $\inf\{m \in \mathbb{R} : m \in \mathcal{A}\} > -\infty$ *and for all* $F \in \mathcal{A}, G \in \mathcal{X}$ *with* $G \geq F$ *we have that* $G \in \mathcal{A}$. *Then*

(i) $\eta_\mathcal{A}$ *is a monetary risk measure.*

(ii) If $\mathcal{A}$ *is a convex set, then* $\eta_\mathcal{A}$ *is a convex risk measure.*

(iii) If $\mathcal{A}$ *is a cone, then* $\eta_\mathcal{A}$ *is positively homogeneous. Clearly, if* $\mathcal{A}$ *is a convex cone, then* $\eta_\mathcal{A}$ *is a coherent risk measure.*

(iv) It always holds that $\mathcal{A} \subseteq \mathcal{A}_{\eta_\mathcal{A}}$. *Moreover,* $\mathcal{A} = \mathcal{A}_\eta$ *if and only if* $\mathcal{A}$ *is* $\|\cdot\|_\infty$*-closed in* $\mathcal{X}$.

The cash-invariance is the central property such that the monetary risk measure allows for a monetary interpretation. Nevertheless, even if we drop the cash-invariance, i.e. the risk measure is only antitone, we can still represent the risk measure in terms of acceptance sets.

Definition 1.2.5. *We say that* $\eta : \mathcal{X} \to \mathbb{R} \cup \{-\infty, \infty\}$ *is a* risk measure *if it is antitone. Moreover, we denote by*

$$\mathcal{A}_\eta^m := \{F \in \mathcal{X} : \eta(F) \leq m\}, \quad m \in \mathbb{R},$$

the acceptance set at level m *of the risk measure* η.

Proposition 1.2.6 (See Theorem 1.7 Drapeau and Kupper (2013))**.** *Let* $\eta : \mathcal{X} \to \mathbb{R} \cup \{-\infty, \infty\}$ *be a quasi-convex risk measure and* $(\mathcal{A}_\eta^m)_{m \in \mathbb{R}}$ *be the corresponding family of acceptance sets at level* m*. Then* $(\mathcal{A}_\eta^m)_{m \in \mathbb{R}}$ *is an increasing family, that is* $\mathcal{A}_\eta^m \subseteq \mathcal{A}_\eta^n$ *for all* $m \leq n$*. Moreover* $\mathcal{A}_\eta^m$ *is a convex set which fulfills the monotonicity property* (1.2.1) *and the right-continuity property* $\mathcal{A}_\eta^m = \bigcap_{n>m} \mathcal{A}_\eta^n$ *for all* $m \in \mathbb{R}$*.*
Conversely, if $(\mathcal{A}^m)_{m \in \mathbb{R}}$ *is a family of sets which fulfills all four properties from above, then*

$$\eta_{\mathcal{A}}(F) := \inf\{m \in \mathbb{R} : F \in \mathcal{A}^m\}, \quad F \in \mathcal{X}, \tag{1.2.4}$$

is a quasi-convex risk measure.
Furthermore $\eta_{\mathcal{A}_\eta} = \eta$ *and* $\mathcal{A}_{\eta_{\mathcal{A}}} = \mathcal{A}$*.*

1.3 Multivariate risk measures

Now let $\mathcal{X}^d$ be a vector space of bounded functions $X : \Omega \to \mathbb{R}^d$, which represent possible risk factors for a financial system. Moreover, assume that $\mathcal{X}^d + \mathbb{R}^d = \mathcal{X}^d$, where $+$ denotes the Minkowski addition. The question arising is if the notion of a univariate (monetary) risk measure can be extended in order to obtain a powerful tool for the systemic risk assessment. The literature on extensions of univariate risk measures to systemic risk measures can generally be subdivided into two branches.
In the first approach the systemic risk measure $\rho : \mathcal{X}^d \to \mathbb{R}$ is a real-valued function which allows for a decomposition

$$\rho(X) = \eta\big(\Lambda(X)\big) \quad \text{for all } X \in \mathcal{X}^d, \tag{1.3.1}$$

where $\Lambda : \mathcal{X}^d \to \mathcal{X}$ is an aggregation function and $\eta : \mathcal{X} \to \mathbb{R}$ a univariate risk measure. Hence, the approach is based on a two-step procedure: first aggregate the single risk factors into a common risk factor for the whole system and secondly evaluate the risk thereof via some well-studied univariate risk measure. Because of this intuitive structure, it is no surprise that this type is the most commonly used in the literature and will be the fundamental concept towards systemic risk in this thesis. Note that the axiomatic study of this approach was initiated by Chen et al. (2013). However, their space of risk factors $\mathcal{X}^d$ was restricted to random vectors on a finite state space. In Kromer et al. (2016) the domain has been extended to more general measurable spaces which comprise the L^p-spaces. We remark that our extension was elaborated independently from Kromer et al. (2016). In comparison to Chen et al. (2013), we extend the study of decomposable risk measures in the following directions: firstly, we allow for a conditional risk assessment which will be discussed further in the next section. Besides conditional risk measures, this also introduces conditional aggregation functions. Secondly, the risk factors are

modeled on the more general space $L_d^\infty(\mathcal{F})$ of essentially bounded random vectors. And thirdly, we reduced the required properties for the decomposition result which in turn forms the basis of a more comprehensive axiomatic characterization.
In the prior section on univariate risk measures, the risk factors represented profits and losses which suggested to measure the risk likewise in terms of a monetary unit. Contrarily, in the multivariate framework the aggregated values $\Lambda(X)$ do not necessarily need to have an interpretation as monetary values, even if the single risk factors X_i have. To illustrate this, suppose you want to express a preference in favor of many firms to have small losses instead of a single firm with a severe damage. With this in mind, a suitable aggregation function is a sum of exponentially transformed losses which obviously lacks any monetary interpretation. Consequently, we do not require that η in the decomposition (1.3.1) is cash-invariant. Nonetheless, we ask for the weaker property of *constancy on constants* which is given by

$$\eta(c) = -c \quad \text{for all } c \in \mathbb{R}. \tag{1.3.2}$$

The constancy on constants is the key to the extraction of the aggregation function from the systemic risk measure ρ. Furthermore, if we have additional properties which apply to the restriction of the systemic risk measure ρ to constant risk factors, like for instance convexity on constants, then these transfer to the aggregation function Λ and vice versa.
More demanding is the question of how we can recover the univariate risk measure η and how it is interlinked with ρ. To this end, we identify the class of risk-consistent properties which brings together the state-wise risk assessment and the overall risk. In the following we consider the risk-antitonicity as exemplary representative of the risk-consistent properties. The risk-antitonicity states that if the scenario-wise risk $X(\omega)$ is less risky than $Y(\omega)$, i.e. $\rho(X(\omega)) \leq \rho(Y(\omega))$, for almost all scenarios $\omega \in \Omega$, then this risk evaluation transfers to the random risks, that is $\rho(X) \leq \rho(Y)$. If we would already know that the decomposition (1.3.1) exists, then the constancy on constants of η yields the following interpretation of the risk-antitonicity: If the aggregated system $\Lambda(X)$ is always better off than $\Lambda(Y)$, then system X must be less risky compared to Y. In particular, for univariate risks there is no aggregation involved and thus risk-antitonicity reduces to the usual antitonicity. Moreover, in the multivariate setting we show that if the risk measure ρ fulfills a risk-consistent property as well as the related property on constants, then the corresponding property also holds in the usual sense. Since risk-consistent properties are an integral part for the decomposition, we call systemic risk measures of type (1.3.1) risk-consistent risk measures. In contrast to Chen et al. (2013) we state the decomposition result for a minimalistic set of axioms on the functions involved. This not only helps us to identify direct relationships between the properties of $\rho_{\mathcal{G}}$ and properties of $\eta_{\mathcal{G}}$ and $\Lambda_{\mathcal{G}}$, but also to cover more examples from the systemic risk literature. For example, the SystRisk given in (1.1.4) could not be covered so far, since it is generally neither positively homogeneous nor convex.

However, before we continue with the risk-consistent systemic risk measures in the conditional framework, we will outline the basic idea of a coexisting perspective towards systemic risk and its connection to our decomposable risk measures.

In (1.2.2) we have seen that every univariate monetary risk measure can equivalently be described by its corresponding acceptance set. By reviewing the risk measure in (1.2.3) we observe that it can equally be split into two consecutive steps: first, identify all cash amounts which make the risk acceptable and second, find the smallest of these cash amounts. Therefore, a univariate monetary risk measure can be represented as a selection of a set-valued function. This is the starting point for the second approach towards systemic risk measurement. Here the systemic risk measure is created from a set-valued function $R : \mathcal{X}^d \to \mathcal{P}(\mathbb{R}^d)$ which is given by

$$R(X) = \left\{ m \in \mathbb{R}^d : X + m \in \mathcal{A} \right\}, \quad \text{for all } X \in \mathcal{X}^d, \tag{1.3.3}$$

where $\mathcal{A} \subseteq \mathcal{X}^d$ and $\mathcal{P}(\mathbb{R}^d)$ denotes the powerset of $\mathbb{R}^d$. The interpretation is analogous to the univariate case, that is $R(X)$ comprises all vectors of cash amounts m such that adding m_i to the i-th entity yields an acceptable system $X + m$. In contrast to the one-dimensional framework where it is always possible to identify the smallest acceptable cash amount, it is a more delicate issue to find an optimal cash injection in the multivariate setting. Therefore, we must specify a procedure which selects an acceptable vector of cash amounts. Finally, aggregating this selection yields a systemic risk measure $\rho : \mathcal{X}^d \to \mathbb{R}$. We remark that Feinstein et al. (2015) already call the set-valued function R a systemic risk measure. A commonly chosen selection criterion in the literature is to minimize the overall injections which corresponds to systemic risk measures of type

$$\rho(X) = \inf \left\{ \sum_{i=1}^{d} m_i : m \in R(X) \right\}, \quad X \in \mathcal{X}^d, \tag{1.3.4}$$

cf. Armenti et al. (2015) or Biagini et al. (2015). In analogy to the univariate case (1.2.4), we can generalize the set-valued functions in (1.3.3) to

$$R(X) = \left\{ m \in \mathbb{R}^d : X \in \mathcal{A}^m \right\}, \quad \text{for all } X \in \mathcal{X}^d, \tag{1.3.5}$$

where $(\mathcal{A}^m)_{m \in \mathbb{R}^d}$ is an appropriate family of acceptance sets. As before, in the process of changing from (1.3.3) to (1.3.5) the monetary interpretation of its elements is lost.

Next, we illustrate the relationship between the systemic risk measures which are based on set-valued risk measures and our class of risk-consistent systemic risk measures. For this purpose, suppose that $\hat{\rho} : \mathcal{X}^d \to \mathbb{R}$ is a systemic risk measure which can be decomposed as in (1.3.1) into a univariate monetary risk measure $\hat{\eta}$ and an aggregation function $\hat{\Lambda}$. Moreover, let $\mathcal{A}_{\hat{\eta}}$ be the acceptance set of $\hat{\eta}$ and define

$$\mathcal{A}^m := \left\{ X \in \mathcal{X}^d : \hat{\Lambda}(X) + \sum_{i=1}^{d} m_i \in \mathcal{A}_{\hat{\eta}} \right\}, \quad \text{for all } m \in \mathbb{R}^d. \tag{1.3.6}$$

Then constructing a systemic risk measure ρ as in (1.3.4)-(1.3.6) yields

$$\begin{aligned} \rho(X) &= \inf \left\{ \sum_{i=1}^{d} m_i : \hat{\Lambda}(X) + \sum_{i=1}^{d} m_i \in \mathcal{A}_{\hat{\eta}} \right\} \\ &= \inf \left\{ m \in \mathbb{R} : \hat{\Lambda}(X) + m \in \mathcal{A}_{\hat{\eta}} \right\} = \hat{\eta}\big(\hat{\Lambda}(X)\big) = \hat{\rho}(X), \quad X \in \mathcal{X}^d. \end{aligned}$$

This example illustrates that the decomposable systemic risk measures can be subsumed under the generalized injection approach. Conversely, in order to use the construction of a systemic risk measure via acceptance sets as specified in (1.3.3), we need to fix a set $\mathcal{A}$ of all acceptable systems. Opposed to (1.3.6), Feinstein et al. (2015) propose a method with an aggregation mechanism which is sensitive to capital levels. To be more precise, they suggest a set-valued risk measure

$$R(X) := \Big\{ m \in \mathbb{R}^d : \Lambda(X + m) \in \mathcal{A}_\eta \Big\}, \quad X \in \mathcal{X}^d,$$

where Λ is an aggregation function and $\mathcal{A}_\eta$ is the acceptance set of a univariate monetary risk measure η. This set-valued function is of type (1.3.3) and by the definition of an acceptance set, we obtain that

$$R(X) := \Big\{ m \in \mathbb{R}^d : \eta\big(\Lambda(X + m)\big) \leq 0 \Big\}, \quad X \in \mathcal{X}^d.$$

Hence, for the proposed method we still rely on a systemic risk measure which is decomposable, since $X \in \mathcal{A}$ if and only if $\rho(X) := \eta(\Lambda(X)) \leq 0$.

Thus, we conclude that the study of one of the two approaches towards systemic risk is not excluding the other, even more they complement each other. If one is just interested in the assessment of the risk of the overall system and less in the contributions of the single entities, the decomposable risk measures are preferable due to their simpler structure. Contrarily, the determination of systemic risk via set-valued functions has the advantage that as a byproduct it provides information on how capital injections into the system should be organized.

For the rest of this thesis we exclusively concentrate on decomposable risk measures. Inspired by the systemic risk approaches from Section 1.1 which condition on systemic events, we particularly aim at analyzing conditional systemic risk measures. This framework is introduced in the following section.

1.4 Conditional risk measures

Besides extending the domain of risk measures to multivariate vector spaces, another plausible generalization is to consider conditional risk measures. In the literature on univariate risk measures they form the basis for the assessment of risk over time as new information on the risk factor is revealed. In addition to the temporal structure, we have seen that in the field of systemic risk it is also of great importance to study the spatial intertwining of the system. That is, in order to detect systemically relevant structures in this context, it is of interest to assess systemic risk conditional on the state of certain subsystems.

In the following we assume that the risk factors are bounded random variables on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and that we have access to the information set $\mathcal{G}$ which is a sub-σ-algebra of $\mathcal{F}$. To include the additional information into the risk assessment, we extent univariate risk measures to functionals of the form

$$\eta_{\mathcal{G}} : L^\infty(\mathcal{F}) \to L^0(\mathcal{G}). \tag{1.4.1}$$

Conditional risk measures of type (1.4.1) have been introduced first in Detlefsen and Scandolo (2005) and Bion-Nadal (2004). As in (1.4.1) most of the literature on conditional risk measures is restricted to bounded risk factors. For the relatively few generalizations to more general L^p

spaces or L^p-type modules, we refer to Filipović et al. (2012) or Acciaio and Goldammer (2013). In analogy to the unconditional case, we define a conditional monetary risk measure as follows:

Definition 1.4.1. *An antitone map* $\eta_{\mathcal{G}} : L^\infty(\mathcal{F}) \to L^\infty(\mathcal{G})$ *is called a* conditional monetary risk measure*, if it is*

Conditional cash-invariant: *For all* $F \in L^\infty(\mathcal{F})$ *and* $\alpha \in L^\infty(\mathcal{G})$*, we have that* $\eta_{\mathcal{G}}(F+\alpha) = \eta_{\mathcal{G}}(F) - \alpha$.

It is further called a conditional convex risk measure *if it is additionally*

Conditional convex: *For all* $F, G \in L^\infty(\mathcal{F})$ *and* $\lambda \in L^\infty(\mathcal{G})$ *with* $0 \leq \lambda \leq 1$*, we have that* $\eta_{\mathcal{G}}(\lambda F + (1-\lambda)G) \leq \lambda \eta_{\mathcal{G}}(F) + (1-\lambda)\eta_{\mathcal{G}}(G)$.

Moreover, a conditional convex risk measure is a coherent risk measure *if it satisfies*

Conditional positive homogeneity: *For all* $F \in L^\infty(\mathcal{F})$ *and* $\lambda \in L^\infty(\mathcal{G})$ *such that* $\lambda \geq 0$*, we have* $\eta_{\mathcal{G}}(\lambda F) = \lambda \eta_{\mathcal{G}}(F)$.

In Chapter 2 we extend the class of systemic risk measures of Chen et al. (2013) to allow for a conditional framework, that is we study risk measures of type

$$\rho_{\mathcal{G}}(X) = \eta_{\mathcal{G}}\big(\Lambda_{\mathcal{G}}(X)\big), \quad X \in L_d^\infty(\mathcal{F}), \tag{1.4.2}$$

where $\eta_{\mathcal{G}} : \operatorname{Im} \Lambda_{\mathcal{G}} \to L^\infty(\mathcal{G})$ is a univariate conditional risk measure and $\Lambda_{\mathcal{G}} : L_d^\infty(\mathcal{F}) \to L^\infty(\mathcal{F})$ a conditional aggregation function. As in the unconditional case, the decomposition relies on the constancy on constants which is given in this setting by

$$\eta_{\mathcal{G}}(\alpha) = -\alpha \quad \text{for all } \alpha \in L^\infty(\mathcal{G}). \tag{1.4.3}$$

However, even if we assume that $\Lambda_{\mathcal{G}}(L_d^\infty(\mathcal{G})) \subseteq L^\infty(\mathcal{G})$, we can still extract the aggregation function via $\Lambda_{\mathcal{G}}(X) := -\rho_{\mathcal{G}}(X)$ for all $X \in L_d^\infty(\mathcal{G})$, but it remains an open question how to extend the aggregation function to the whole of $L_d^\infty(\mathcal{F})$. We will overcome the problem by requiring the existence of a particularly nice realization $\widetilde{\rho}_{\mathcal{G}} : \mathbb{R}^d \times \Omega \to \mathbb{R}$ of the restriction of the multivariate conditional risk measure $\rho_{\mathcal{G}}$ to deterministic risk factors. Then we define the conditional aggregation function for all $X \in L_d^\infty(\mathcal{F})$ by

$$\Lambda_{\mathcal{G}}(X)(\omega) := -\widetilde{\rho}_{\mathcal{G}}(X(\omega), \omega) \quad a.s.$$

Finally, we show that the function $\eta_{\mathcal{G}}$ given by

$$\eta_{\mathcal{G}}(F) := \rho_{\mathcal{G}}(X), \text{ where } \Lambda_{\mathcal{G}}(X) = F,$$

is the desired univariate conditional risk measure. As before a conditional version of the risk-antitonicity is the key property ensuring that $\eta_{\mathcal{G}}$ is well-defined.

Typically we do not have a single information set but a whole stream of information represented by a family of conditional risk measures. For instance, in the dynamic framework more and more details are revealed as time elapses. But also in the spatial setting this is of particular interest. For example, one might think about a manager presiding a financial conglomerate and another which only heads an entity thereof. Usually the manager of the conglomerate has more information which also comprises all the data of the entity. So also the risk assessment should respect this superiority of information, that is the corresponding risk measures are required to be consistent in a certain way. The most commonly used approaches to such consistency in the univariate literature can be pooled by the following general definition proposed in Tutsch (2007).

Definition 1.4.2. *Let $\mathcal{Y} \subseteq L^\infty(\mathcal{F})$ with $\mathcal{Y} + L^\infty(\mathcal{G}) = \mathcal{Y}$ be a reference set. We say two conditional risk measures $\eta_\mathcal{G}$ and $\eta_\mathcal{H}$ with $\mathcal{G} \subseteq \mathcal{H} \subseteq \mathcal{F}$ are* consistent *w.r.t. $\mathcal{Y}$ if*

$$\begin{cases} \eta_\mathcal{H}(F) \geq \eta_\mathcal{H}(\alpha) \Longrightarrow \eta_\mathcal{G}(F) \geq \eta_\mathcal{G}(\alpha), \\ \eta_\mathcal{H}(F) \leq \eta_\mathcal{H}(\alpha) \Longrightarrow \eta_\mathcal{G}(F) \leq \eta_\mathcal{G}(\alpha), \end{cases} \qquad (F \in L^\infty(\mathcal{F}), \alpha \in \mathcal{Y}). \tag{1.4.4}$$

If only the first implication of (1.4.4) *is fulfilled, then we say that $\eta_\mathcal{G}$ and $\eta_\mathcal{H}$ are rejection consistent w.r.t. $\mathcal{Y}$. Contrarily, if only the second implication is valid, then we say that $\eta_\mathcal{G}$ and $\eta_\mathcal{H}$ are acceptance consistent w.r.t. $\mathcal{Y}$.*

Clearly, the smallest benchmark set is $\mathcal{Y} = L^\infty(\mathcal{G})$ and in this case we also speak of *weak consistency*. If the involved conditional risk measures are monetary and normalized in the sense that $\eta_\mathcal{H}(0) = \eta_\mathcal{G}(0)$, then weak acceptance consistency is equivalent to

$$\eta_\mathcal{H}(F) \leq 0 \Longrightarrow \eta_\mathcal{G}(F) \leq 0. \tag{1.4.5}$$

Thus, the interpretation is that if a risky position is always acceptable under the finer information structure $\mathcal{H}$, then it should also be acceptable under less information. Or to put it into the dynamic context, rejecting a position right now is arbitrary, if we already know that at a future point in time we are going to accept it in any case. An analogous argumentation also holds for the weak rejection consistency. The weak consistency is of particular interest in the context of updating a given conditional risk measure to a richer information structure in a consistent way, cf. Tutsch (2008). Furthermore, weak consistency has been considered in Weber (2006), Acciaio and Penner (2011), and Roorda and Schumacher (2013, 2016).
Conversely, the largest reference set $\mathcal{Y}$ such that (1.4.4) is still well defined is $L^\infty(\mathcal{F})$. For this choice of $\mathcal{Y}$ the difference between rejection and acceptance consistency collapses and we simply have

$$\eta_\mathcal{H}(F) \geq \eta_\mathcal{H}(G) \Longrightarrow \eta_\mathcal{G}(F) \geq \eta_\mathcal{G}(G), \quad (F, G \in L^\infty(\mathcal{F})). \tag{1.4.6}$$

If $\eta_\mathcal{H}$ and $\eta_\mathcal{G}$ fulfill (1.4.6) we call them *strongly consistent*. In the pertinent literature strong consistency is the most frequently used requirement to connect conditional risk measures. In contrast to (1.4.5) the strong consistency not only preserves the acceptance or rejection criterion of the decision-maker, but also her complete preferences. Or to reuse the spatial example of the risk managers from above, if the head of a financial conglomerate has a distinct preference, then this should also hold for the entities' manager. Here we also observe that the reverse relation is not meaningful. Even if one risk factor is preferable to some other for a small entity, the group's manager might identify events on her larger information set where this preference fails.

Note that there is a multitude of alternative descriptions of strong consistency. For instance, for conditional convex risk measures strong consistency can be represented in terms of a supermartingale property of the modified risk measures or via an additivity property of the acceptance sets of the risk measures involved. For more details, we refer the interested reader to Acciaio and Penner (2011) and the references therein. Moreover, it can be easily shown that for conditional monetary risk measures strong consistency is further equivalent to the following tower property

$$\eta_{\mathcal{G}}(F) = \eta_{\mathcal{G}}\big(- \eta_{\mathcal{H}}(F)\big) \quad \text{for all } F \in L^{\infty}(\mathcal{F}). \tag{1.4.7}$$

Chapter 3 emphasizes strong consistency in the context of multivariate conditional risk measures. For this purpose we generalize the definition of strong consistency given in (1.4.6). In the first instance, we ask if it is also possible to represent the generalized strong consistency at hand in terms of a tower property similar to (1.4.7). That there is no straightforward extension can be easily seen, since the right hand side of (1.4.7) is not even well-defined in the multivariate setting. In order to repair this deficiency, we need to transfer back the inner systemic risk $\rho_{\mathcal{H}}(X)$ to a d-dimensional risk factor. A possible candidate for this operation is the inverse of the function $L^{\infty}(\mathcal{H}) \ni \alpha \mapsto \rho_{\mathcal{H}}(\alpha(1, ..., 1)^{\top})$ which we denote by $f^{-1}_{\rho_{\mathcal{H}}}$. In the course of Chapter 3 we show that $f^{-1}_{\rho_{\mathcal{H}}}$ is indeed well-defined and that the generalized tower property is given by

$$\rho_{\mathcal{G}}(X) = \rho_{\mathcal{G}}\left(f^{-1}_{\rho_{\mathcal{H}}}\big(\rho_{\mathcal{H}}(X)\big)(1, ..., 1)^{\top}\right), \quad \text{for all } X \in L^{\infty}_{d}(\mathcal{F}). \tag{1.4.8}$$

Note that for a univariate conditional risk measure, where $\eta_{\mathcal{H}}$ is constant on constants, i.e. $f_{\eta_{\mathcal{H}}}(\alpha) = -\alpha$, (1.4.8) reduces to (1.4.7). Making use of (1.4.8), we study the relationship between strong consistency and the risk-consistent conditional systemic risk measures from Chapter 2 in two frameworks. Firstly, we assume consistency of $\rho_{\mathcal{G}}$ with a terminal risk measure $\rho_{\mathcal{F}} : L^{\infty}_{d}(\mathcal{F}) \to L^{\infty}(\mathcal{F})$. The latter risk measurement under $\rho_{\mathcal{F}}$ exhibits essentially no compactification of information and thus can be interpreted as an aggregation of the risk factors. The question is now if this aggregation under full information can be carried over to the risk measurement under $\mathcal{G}$ via the strong consistency. Indeed, we will see in Chapter 3 that under some minor technical assumptions $\rho_{\mathcal{G}}$ is decomposable as in (1.4.2) and the respective conditional aggregation function is strongly consistent with the aggregation under full information. Due to our findings in Chapter 2, the proof thereof basically reduces to showing that $\rho_{\mathcal{G}}$ is risk-antitone. In the second framework, we additionally assume conditional law invariance of the risk measures involved. Moreover, instead of strong consistency with respect to a terminal risk measure, we consider the opposite, that is consistency w.r.t. an unconditional risk measure $\rho : L^{\infty}_{d}(\mathcal{F}) \to \mathbb{R}$. In the univariate setting this has been considered in Föllmer (2014), where he showed that these conditional risk measures are in the class of the conditional certainty equivalents, i.e. they can be represented as

$$\eta_{\mathcal{G}}(F) = -u^{-1}\left(\mathbb{E}_{\mathbb{P}}\left[u(F) \mid \mathcal{G}\right]\right), \quad F \in L^{\infty}(\mathcal{F}), \tag{1.4.9}$$

for some utility function $u : \mathbb{R} \to \mathbb{R}$. Chapter 3 generalizes this result to the multivariate case, that is we show that the members of a family of strongly consistent conditionally law-invariant risk measures are of the form

$$\rho_{\mathcal{G}}(X) = f_{\rho_{\mathcal{G}}}(f^{-1}_{u}(\mathbb{E}_{\mathbb{P}}\left[u(X) \mid \mathcal{G}\right])), \quad X \in L^{\infty}_{d}(\mathcal{F}), \tag{1.4.10}$$

where $u : \mathbb{R}^d \to \mathbb{R}$ is a multivariate utility function and $f_u : \mathbb{R} \to \mathbb{R}$ is the univariate utility given by $f_u(c) := u(c(1, ..., 1)^\top)$. In particular we show that the special structure (1.4.10) implies that the strongly consistent conditional systemic risk measures are again decomposable as in (1.4.2). Moreover, the corresponding univariate conditional risk measure in the decomposition is a univariate conditional certainty equivalent

$$\eta_{\mathcal{G}}(F) = -U_{\mathcal{G}}^{-1}\left(\mathbb{E}_{\mathbb{P}}\left[U_{\mathcal{G}}(F) \mid \mathcal{G}\right]\right), \quad F \in L^\infty(\mathcal{F}),$$

where $U_{\mathcal{G}}$ is a conditional utility function. Note that the stochasticity of $U_{\mathcal{G}}$ solely results from the normalizing function $f_{\rho_{\mathcal{G}}}^{-1}$. Particularly, if $\rho_{\mathcal{G}}$ is normalized on constants, i.e. $f_{\rho_{\mathcal{G}}} \equiv -\operatorname{id}$, then $\eta_{\mathcal{G}}$ is a conditional certainty equivalent as in (1.4.9).

1.5 Allocation of systemic risk

In the context of a multivariate risk measurement it is not only of interest to assess the total risk, but also to determine each risk factor's contribution to the overall risk.

In the framework of set-valued systemic risk measures, one possibility could be to interpret the single components in (1.3.3) as an adequate response to the apportionment of the system's risk, since they represent the monetary amounts which must be added such that the resulting system is deemed acceptable. However, then the contribution problem turns into the task of finding an appropriate selection criterion for the numerous acceptable cash amounts. A possible choice for this selection criterion is to reuse the criterion which was already applied for the assessment of the overall risk, see for instance Armenti et al. (2015). If both problems are solved by the same selection, then the contributions clearly add up to the overall risk. But, apart from this so-called full allocation property, there is no indication why the allocation of the total risk can be considered as fair.

In this thesis we focus on the fairness issue of the allocation problem in the framework of decomposable systemic risk measures. Therefore, the allocation problem can be stated as follows: For a given system $X \in \mathcal{X}^d$, we need to find an allocation $k \in \mathbb{R}^d$ of the total risk $\rho(X) = \eta(\Lambda(X))$ such that the full allocation property

$$\sum_{i=1}^{d} k_i = \rho(X) \tag{1.5.1}$$

holds, where k_i is interpreted as the allocation of the total risk to institution i. Note that there is already a wide-ranging literature in case of the aggregation function Λ being just a simple sum. This situation is also known as portfolio allocation problem, which was studied in Denault (2001), Tasche (2004, 2007), Kalkbrener (2005) or Buch and Dorfleitner (2008). Here the system corresponds to a basket of financial assets and the risk factors are the profits and losses of the assets. Because the investor of the portfolio is just worried about the netted profits and losses, the sum is a suitable aggregation function. Most commonly the portfolio allocation problem is solved via the Shapley-value, the Aumann-Shapley-value or marginal contributions which we further discuss below. The first two originate from the field of game theory and can already be traced back to the works of Shapley (1953) and Aumann and Shapley (1974).

All of the mentioned approaches essentially rely on the formation and risk measurement of subsystems. For a start, we consider subsystems where each entity either participates or is absent from the system. The risk of the corresponding subsystems is given by the subsystem risk measure $\widetilde{\rho} : \mathcal{X}^d \times \mathcal{P} \to \mathbb{R}$, where $\mathcal{P}$ is the set of all possible subsets $J \subseteq \{1, ..., d\}$. Here $\widetilde{\rho}(X, J)$ is the risk of the system X, where only the entities contained in J participate.
Maybe the most intuitive method to determine each entities contribution is to consider its impact on the entire system, i.e.

$$k_i = \rho(X) - \widetilde{\rho}(X, \{1, ..., i-1, i+1, ..., d\}), \quad i = 1, ..., d. \tag{1.5.2}$$

Unfortunately, this procedure lacks the full allocation property (1.5.1) in general. In contrast to this method, the Shapley-value takes all the incremental risks which are generated by a single entity into account and composes a weighted average thereof. To be more precise, the Shapley-value $\mathrm{SV}(X) \in \mathbb{R}^d$ of the system X is formally given by

$$\mathrm{SV}(X)_i = \sum_{J \in \mathcal{P}_i} \frac{(|J|-1)!(d-|J|)!}{d!} \big(\widetilde{\rho}(X, J) - \widetilde{\rho}(X, J \backslash \{i\})\big), \quad i = 1, ..., d,$$

where $\mathcal{P}_i := \{J \in \mathcal{P} : i \in J\}$ are all the subgroups containing entity i. Unlike (1.5.2) the Shapley-value fulfills the full allocation property. We remark that the Shapley-value can also be uniquely described by a set of properties, namely full allocation, symmetry, dummy player and additivity over games, for further details we refer the interested reader to Denault (2001).

One might ask why we always have to remove an entity entirely and not only parts of it. For this purpose, we first need to extent the domain of a subsystem risk measure. In the following we denote by $\rho : \mathcal{X}^d \times \mathbb{R}^d \to \mathbb{R}$ the subsystem risk measure with fractional participation, that is $\rho(X, \lambda)$ is the risk of the system X where λ_i is the level of participation of the i-th entity. Here $\lambda_i = 0$ means absence and $\lambda_i = 1$ full participation. In (1.5.2) we were asking how much risk is generated by the introduction of an entity. Inspired by this question, we are now asking how much risk is introduced by increasing marginally the level of participation of an entity. This leads to the marginal contributions $\mathrm{MC}(X)$ which are given by

$$\mathrm{MC}_i(X) := \left.\frac{\partial \rho(X, \lambda)}{\partial \lambda_i}\right|_{\lambda = (1, ..., 1)^\top}, \quad i = 1, ..., d, \tag{1.5.3}$$

if the partial derivatives exist. It is important to note that if $\lambda \mapsto \rho(X, \lambda)$ is continuously differentiable then Euler's homogeneous function theorem implies that the full allocation property for the marginal contributions holds if and only if the subsystem risk measurement is positively homogeneous in the participation level. This is also the reason why in this case the marginal contributions are commonly referred to as Euler allocation, cf. Tasche (2007).
The marginal contributions have been used in Brunnermeier and Cheridito (2014) for the allocation of systemic risk. However, their model is not positively homogeneous and thus the full allocation property is not satisfied. Nonetheless, they could show that in their setting the sum of the marginal contributions always dominates the systemic risk. Thus, they introduced exogenously given correction terms which reduce the marginal contributions such that the full allocation property holds. Since, they suggested that the correction terms depend on some size

parameter of the financial institutions like corporate taxes, the resulting allocations are called size-shifted marginal contributions.

If we do not want to rely on some artificial correction term, the Aumann-Shapley-value $\mathrm{ASV}(X)$ is a further extension of the marginal contributions to allow for the full allocation property. It is formally given by

$$\mathrm{ASV}(X)_i = \int_0^1 \frac{\partial \rho(X,\lambda)}{\partial \lambda_i}\bigg|_{\lambda=(\gamma,\dots,\gamma)^\top} d\gamma, \quad i = 1, ..., d.$$

The Aumann-Shapley-value is therefore the average marginal contribution as the system evolves uniformly.

In the context of systemic risk, the Aumann-Shapley-value as well as the Shapley-value have been considered extensively, see e.g. Staum (2012), Drehmann and Tarashev (2013), Tarashev et al. (2016) and Staum et al. (2016). Staum et al. (2016) also considered an Eisenberg-Noe type clearing model in a toy financial network and put a special emphasis on the construction of the subsystems. The crucial observation is that, whereas it is clear what absence or full participation means in the context of a portfolio of financial assets, there is a multitude of choices for a complex financial system. That is, with the extraction of an institution it needs to be clarified what should happen with the contractual obligations of this institution with the remaining financial system. For instance, one way to construct the subsystems which we also consider in Chapter 4, is the attribution to external assets scheme. Here an absent bank is still physically present in the network, but it is not exposed to any adverse market events in the future. Moreover, to conclude with the subsystem generation, we remark that Staum et al. (2016) also consider schemes where the underlying network topology changes.

After this brief excursion to the allocation methods currently used in the portfolio as well as the systemic risk literature, we will now focus on our findings of the final Chapter 4. There, we aim at gaining insight into (fuzzy) core allocations for interacting systems. Formally, we say that $k \in \mathbb{R}^d$ is in the fuzzy core $FC_\rho(X)$ if k fulfills the full allocation property (1.5.1) and for all levels of participation $\lambda \in [0,1]^d$

$$\rho(X,\lambda) \geq \sum_{i=1}^{d} \lambda_i k_i. \tag{1.5.4}$$

Similarly, the core concept is the discretized version of (1.5.4), i.e.

$$\widetilde{\rho}(X,J) \geq \sum_{j\in J} k_j, \quad \text{for all } J \in \mathcal{P}.$$

Thus the (fuzzy) core allocations ensure that no subsystem has a lower risk than the amount which is allocated to it. A common justification is that otherwise the disadvantaged subgroup would separate from the system to gain from the decrease in risk. Since no one has an incentive to leave the system, core allocations treat all participants fairly. Both the study of the fuzzy core as well as the core have a long history in the field of game theory. Furthermore, the (fuzzy) core has been studied in Delbaen (2000) for the portfolio allocation problem. In this context, (1.5.4) also appeared under the name of the no-undercut property in Denault (2001). Here the (fuzzy) core is a central property of an axiomatic description of desirable allocations, termed coherent allocations. At the heart of Denault's analysis is the identification of conditions such that the Shapley-value or the Aumann-Shapley-value are coherent allocations and thus also in the core or fuzzy core, resp. In particular, if the Aumann-Shapley-value collapses to the Euler allocation it is contained in the fuzzy core.

In the context of systemic risk allocation, the fuzzy core concept has been considered in Chen et al. (2013) and Kromer et al. (2016). Based on a dual representation of their decomposable positive homogeneous systemic risk measures, Chen et al. and Kromer et al. identify allocations which are closely related to the Euler allocation or the Aumann-Shapley-value. Moreover, they show that the allocations are also in the fuzzy core of the subsystem risk measure

$$\rho(X,\lambda) := \rho\big((\lambda_1 X_1, ..., \lambda_d X_d)^\top\big), \tag{1.5.5}$$

where ρ is the corresponding systemic risk measure.
In this thesis we study the appropriateness of the fuzzy core allocations for systemic risk. Note that in the portfolio framework the merger of two portfolios always results in a decrease of risk due to diversification effects. So the notion of the (fuzzy) core is based on solid grounds, since there is always a benefit which can be distributed among the single financial assets. For systemic risk measurement this fundamental premise is confronted with limits. For instance, a bank might face no uncertainty in isolation, but by integrating it into a financial system it is exposed to risks of its opponents. Conversely, in more integrated systems we have an increased redistribution of losses which might dampen potential contagion effects and thus results in less severe damages. To cut a long story short, for the systemic risk allocation we need to take into account the competing effects of diversification and integration costs. To this end, we will explicitly calculate a (fuzzy) core allocation for the subsystem risk measure in (1.5.5) which is based on a contagion model á la Eisenberg-Noe. Unfortunately, we observe that the (fuzzy) core might lead to obvious discriminations in this systemic risk measurement framework. More precisely, financial institutions which exclusively default due to contagion and afterwards spread further losses into the system are better off under a fuzzy core allocation than banks which do not default at all. One of the reasons, is that (fuzzy) core allocations are by definition upper bounded by the standalone risk, i.e. $k_i \leq \widetilde{\rho}(X,\{i\})$ for all $i = 1,...,d$, and hence no further costs of system integration can be incorporated into the allocation. As a result, we introduce the notion of the reverse (fuzzy) core, where the risk of the subsystem should always be a lower bound of the allocated risk. Moreover, it can be easily seen that by using the same subsystem risk measurement (1.5.5) the corresponding reverse (fuzzy) core is empty. Therefore, we also change the underlying subsystem generation scheme. In the new scheme a zero-participation level means that the bank is extremely well capitalized such that it cannot default at all. Using this scheme, we show that the reverse (fuzzy) core is non-empty and that the corresponding allocations repair the unfairness form before. This result already holds for a more conservative subsystem risk measure which is akin to the attribution to external assets scheme proposed in Staum et al. (2016). Finally, we show that the intersection of the core with the subsystem risk measure (1.5.5) and the reverse core with the new subsystems are basically empty.

2 Risk-Consistent Conditional Systemic Risk Measures

2.0 Contributions of the thesis' author

The current chapter is a joint work with Prof. Dr. Thilo Meyer-Brandis and Dr. Gregor Svindland. It has been published in Stochastic Processes and their Applications, Volume 126 Issue 7, pp. 2014-2037. At the end of the current chapter there is a small study on central counterparties which we excluded in the published version.

The general topic of this section is to extend the framework of Chen et al. (2013). Section 2.2 contains the generalized axiomatic description for conditional multivariate risk measures which allow for a decomposition into a conditional univariate risk measure and a conditional aggregation function. The final framework as well as the presentation of the results has been discussed jointly. H. Hoffmann further introduced explanatory comments like Proposition 2.2.2, Remark 2.2.5 or Remark 2.2.12. In the subsequent Section 2.3 the two preparatory results Lemma 2.3.1 and Lemma 2.3.2 for the main decomposition result have been derived by H. Hoffmann. The proofs of Theorem 2.2.9 as well as Theorem 2.2.11 were established in a close cooperation of the three authors, but with major parts done by H. Hoffmann. Section 2.4 contains extensions of examples appearing in Chen et al. (2013), but also new ones to demonstrate the power of conditional aggregation functions. Here, in particular H. Hoffmann developed the Example 2.4.2, Example 2.4.3 and Example 2.4.5. Moreover, H. Hoffmann performed the concluding numerical study in Example 2.4.9.

2.1 Introduction

The recent financial crisis revealed weaknesses in the financial regulatory framework when it comes to the protection against systemic events. Before, it was generally accepted to measure the risk of financial institutions on a stand alone basis. In the aftermath of the financial crisis risk assessment of financial systems as well as their impact on the real economy has become increasingly important, as is documented by a rapidly growing literature; see e.g. Amini and Minca (2013) or Bisias et al. (2012) for a survey and the references therein. Parts of this literature are concerned with designing appropriate risk measures for financial systems, so-called systemic risk measures. The aim of this paper is to axiomatically characterize the class of systemic risk measures ρ which admit a decomposition of the following form:

$$\rho(X) = \eta\left(\Lambda(X)\right), \tag{2.1.1}$$

where Λ is a state-wise aggregation function over the d-dimensional random risk factors X of the financial system, e.g. profits and losses at a given future time horizon, and η is a univariate risk measure. The aggregation function determines how much a single risk factor contributes to the total risk $\Lambda(X)$ of the financial system in every single state, whereas the so-called base risk measure η quantifies the risk of $\Lambda(X)$. Chen et al. (2013) first introduced axioms for systemic risk measures, and showed that these admit a decomposition of type (2.1.1). Their studies relied on a finite state space and were carried out in an unconditional framework. Kromer et al. (2016) extend this to arbitrary probability spaces, but keep the unconditional setting. The main contributions of this paper are:

1. We axiomatically characterize systemic risk measures of type (2.1.1) in a conditional framework, in particular we consider conditional aggregation functions and conditional base risk measures in (2.1.1).

2. We allow for a very general structure of the aggregation, which is flexible enough to cover examples from the literature which could not be handled in axiomatic approaches to systemic risk so far.

3. We work in a less restrictive axiomatic setting, which gives us the flexibility to study systemic risk measures which for instance need not necessarily be convex or quasi-convex, etc. This again provides enough flexibility to cover a vast amount of systemic risk measures applied in practice or proposed in the literature. It also allows us to identify the relation between properties of ρ and properties of Λ and η, and in particular the mechanisms behind the transfer of properties from ρ to Λ and η, and vice versa. This is related to the following point 4.

4. We identify the underlying structure of the decomposition (2.1.1) by defining systemic risk measures solely in terms of so called risk-consistent properties and properties on constants.

In the following we will elaborate on the points 1.–4. above.

1. A conditional framework for assessing systemic risk

We consider systemic risk in a conditional framework. The standard motivation for considering conditional univariate risk measures (see e.g. Detlefsen and Scandolo (2005) and Acciaio and Penner (2011)) is the conditioning in time, and the argumentation in favor of this also carries over to multivariate risk measures. However, apart from a dynamic assessment of the risk of a financial system, it might be particularly interesting to consider conditioning in space. In that respect Föllmer and Klüppelberg (2014) recently introduced and studied so-called spatial risk measures for univariate risks. Typical examples of spatial conditioning are conditioning on events representing the whole financial system or parts of that system, such as single financial institutions, in distress. This is done to study the impact of such a distress on (parts of) the financial system or the real economy, and thereby to identify systemically relevant structures. For instance the Conditional Value at Risk (CoVaR) introduced in Adrian and Brunnermeier

(2016) considers for $q \in (0,1)$ the q-quantile of the distribution of the netted profits/losses of a financial system $X = (X_1, \dots, X_d)$ conditional on a crisis event $C(X_i)$ of institution i:

$$\mathbb{P}\left(\sum_{i=1}^{d} X_i \leq -\text{CoVaR}_q(X) \,\middle|\, C(X_i) \right) = q; \tag{2.1.2}$$

see Example 2.4.6. More examples can be found in Cont et al. (2013), Engle et al. (2015), Acharya et al. (2017). Such risk measures fit naturally in a conditional framework; cf. Example 2.4.6 and Example 2.4.8.

2. Aggregation of multidimensional risk

A quite common aggregation rule for a multivariate risk $X = (X_1, \dots, X_d)$ is simply the sum

$$\Lambda_{\text{sum}}(X) = \sum_{i=1}^{d} X_i;$$

see the definition of CoVaR in (2.1.2). $\Lambda_{\text{sum}}(X)$ represents the total profit/loss after the netting of all single profits/losses. However, such an aggregation rule might not always be reasonable when measuring systemic risk. The major drawbacks of this aggregation function in the context of financial systems are that profits can be transferred from one institution to another and that losses of a financial institution cannot trigger additional contagion effects. Those deficiencies are overcome by aggregation functions which explicitly account for contagion effects within a financial system. For instance, based on the approach in Eisenberg and Noe (2001), the authors in Chen et al. (2013) introduce such an aggregation rule which however, due to the more restrictive axiomatic setting, exhibits the unrealistic feature that in case of a clearing of the system institutions might decrease their liabilities by more than their total debt level. We will present a more realistic extension of this contagion model together with a small simulation study in Example 2.4.9.

Moreover, we present reasonable aggregation functions which are not comprised by the axiomatic framework of Chen et al. (2013) or Kromer et al. (2016). In particular this includes *conditional aggregation functions* which come naturally into play in our framework; see Example 2.4.5.

3.–4. Axioms for systemic risk measures

Our aim is to identify the relation between properties of ρ and properties of Λ and η in (2.1.1) respectively, and in particular the mechanisms behind the transfer of properties from ρ to Λ and η, and vice versa. We will show that this leads to two different classes of axioms for conditional systemic risk measures. One class concerns the behavior on deterministic risks, so-called properties on constants. The other class of axioms ensures a consistency between state-wise and global - in the sense of over all states - risk assessment. This latter class will be called risk-consistent properties.

The risk-consistent properties ensure a consistency between local - that is ω-wise - risk assessment and the measured global risk. For example, *risk-antitonicity* is expressed by: if for given risk vectors X and Y it holds that $\rho(X(\omega)) \geq \rho(Y(\omega))$ in almost all states ω, then $\rho(X) \geq \rho(Y)$. The naming *risk-antitonicity*, and analogously the naming for the other risk-consistent properties, is motivated by the fact that antitonicity is considered with respect to the order relation $\rho(X(\omega)) \geq \rho(Y(\omega))$ induced by the ω-wise risk comparison of two positions and not with respect to the usual order relation on the space of random vectors.

Note that for a univariate risk measure ρ which is constant on constants, i.e. $\rho(x) = -x$ for all $x \in \mathbb{R}$, risk-antitonicity is equivalent to the 'classical' antitonicity with respect to the usual order relation on the underlying space of random variables. In a general multivariate setting this equivalence does not hold anymore. However, we will show that properties on constants in conjunction with corresponding risk-consistent properties imply the classical properties on the space of risks. This makes our risk model very flexible, since we may identify systemic risk measures where for example the corresponding aggregation function Λ in (2.1.1) is concave, but the base risk measure η is not convex. Moreover, it will turn out that the properties on constants basically determine the underlying aggregation rule in the systemic risk assessment, whereas the risk-consistent properties translate to properties of the base risk measure in the decomposition (2.1.1).

Some of the risk-consistent properties, however partly under different names, also appear in the frameworks of Chen et al. (2013) and Kromer et al. (2016). For instance what we will call risk-antitonicity is called preference consistency in Chen et al. (2013). In our framework we emphasize the link between the risk-consistent properties (and the properties on constants) and the decomposition (2.1.1). This aspect has not been clearly worked out so far. It leads us to introducing a number of new axioms and to classifying all axioms within the mentioned classes of risk-consistent properties and properties on constants.

Structure of the paper

In Section 2.2 we introduce our notation and the main objects of this paper, that is the risk-consistent conditional systemic risk measures, the conditional aggregation functions and the conditional base risk measures as well as their various extensions. At the end of Section 2.2 we state our main decomposition result (Theorem 2.2.9) for risk-consistent conditional systemic risk measures. Moreover, Theorem 2.2.11 reveals the connection between risk-consistent properties and properties on constants on the one hand and the classical properties of risk measures on the other hand. Section 2.3 is devoted to the proofs of Theorem 2.2.9 and Theorem 2.2.11. In Section 2.4 we collect our examples.

2.2 Decomposition of systemic risk measures

Throughout this paper let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and $\mathcal{G}$ be a sub-σ-algebra of $\mathcal{F}$. $\mathscr{L}^\infty(\mathcal{F}) := \mathscr{L}^\infty(\Omega, \mathcal{F}, \mathbb{P})$ refers to the space of $\mathcal{F}$-measurable, $\mathbb{P}$-almost surely (a.s.) bounded random variables and $\mathscr{L}_d^\infty(\mathcal{F})$ to the d-fold cartesian product of $\mathscr{L}^\infty(\mathcal{F})$. As usual, $L^\infty(\mathcal{F})$ and $L_d^\infty(\mathcal{F})$ denote the corresponding spaces of random variables/vectors modulo $\mathbb{P}$-a.s. equality. For $\mathcal{G}$-measurable random variables/vectors analogue notations are used.
In general, upper case letters will represent random variables, where X, Y, Z are multidimensional and F, G, H are one-dimensional, and lower case letters deterministic values.

We will use the usual componentwise orderings on $\mathbb{R}^d$ and $L^\infty_d(\mathcal{F})$, i.e. $x = (x_1, \dots, x_d) \geq y = (y_1, \dots, y_d)$ for $x, y \in \mathbb{R}^d$ if and only if $x_i \geq y_i$ for all $i = 1, \dots, d$, and similarly $X \geq Y$ if and only if $X_i \geq Y_i$ a.s. for all $i = 1, ..., d$. Furthermore $\mathbf{1}_d$ and $\mathbf{0}_d$ denote the d-dimensional vectors whose entries are all equal to 1 or all equal to 0, respectively.

When deriving our main results we will run into similar problems as one faces in the study of stochastic processes: At some point it will not be sufficient to work on equivalence classes, but we will need a specific nice realization or version of the process, for instance a version with continuous paths, etc. In the following, by a realization of a function $\rho_\mathcal{G} : L^\infty_d(\mathcal{F}) \to L^\infty(\mathcal{G})$ we mean a selection of one representative in the equivalence class $\rho_\mathcal{G}(X)$ for each $X \in L^\infty_d(\mathcal{F})$, i.e. a function $\rho_\mathcal{G}(\cdot, \cdot) : L^\infty_d(\mathcal{F}) \times \Omega \to \mathbb{R}$ where $\rho_\mathcal{G}(X, \cdot) \in \mathscr{L}^\infty(\mathcal{G})$ with $\rho_\mathcal{G}(X, \cdot) \in \rho_\mathcal{G}(X)$ for all $X \in L^\infty_d(\mathcal{F})$. We emphasize that in the following we will always denote a realization of a function $\rho_\mathcal{G}$ by its explicit dependence on the two arguments: $\rho_\mathcal{G}(\cdot, \cdot)$. Indeed, our decomposition result in Theorem 2.2.9 will be based on the idea to break down a random variable into every single scenario and evaluating it separately. This implies working with appropriate realizations which will satisfy properties which we will denote *risk-consistent* properties.

Also for risk factors we will work both with equivalence classes of random vectors in $L^\infty_d(\mathcal{F})$ and their corresponding representatives in $\mathscr{L}^\infty_d(\mathcal{F})$. However, in contrast to the realizations of $\rho_\mathcal{G}$ introduced above, here the considerations do not depend on the specific choice of the representative. Hence for risk factors $X \in L^\infty_d(\mathcal{F})$ we will stick to usual abuse of notation of also writing X for an arbitrary representative in $\mathscr{L}^\infty_d(\mathcal{F})$ of the corresponding equivalence class. This will become clear from the context. In particular, $X(\omega)$ denotes an arbitrary representative of the corresponding equivalence class evaluated in the state $\omega \in \Omega$.

Finally, we write $x \in \mathbb{R}^d$ both for real numbers and for (equivalence classes of) constant random variables depending on the context.

The following definition introduces our main object of interest in this paper:

Definition 2.2.1 (Risk-consistent Conditional Systemic Risk Measure)**.**
A function $\rho_\mathcal{G} : L^\infty_d(\mathcal{F}) \to L^\infty(\mathcal{G})$ is called a risk-consistent conditional systemic risk measure *(CSRM), if it is*

Antitone on constants: *For all $x, y \in \mathbb{R}^d$ with $x \geq y$ we have $\rho_\mathcal{G}(x) \leq \rho_\mathcal{G}(y)$,*

and if there exists a realization $\rho_\mathcal{G}(\cdot, \cdot)$ such that the restriction

$$\widetilde{\rho}_\mathcal{G} : \mathbb{R}^d \times \Omega \to \mathbb{R};\ x \mapsto \rho_\mathcal{G}(x, \omega) \tag{2.2.1}$$

has continuous paths*, i.e. $\widetilde{\rho}_\mathcal{G}$ is continuous in its first argument a.s., and it satisfies*

Risk-antitonicity: *For all $X, Y \in L^\infty_d(\mathcal{F})$ with $\widetilde{\rho}_\mathcal{G}(X(\omega), \omega) \geq \widetilde{\rho}_\mathcal{G}(Y(\omega), \omega)$ a.s. we have $\rho_\mathcal{G}(X) \geq \rho_\mathcal{G}(Y)$.*

Furthermore, we will consider the following properties of $\rho_\mathcal{G}$ on constants:

Convexity on constants: *$\rho_\mathcal{G}(\lambda x + (1-\lambda)y) \leq \lambda\rho_\mathcal{G}(x) + (1-\lambda)\rho_\mathcal{G}(y)$ for all constants $x, y \in \mathbb{R}^d$ and $\lambda \in [0, 1]$;*

Positive homogeneity on constants: *$\rho_\mathcal{G}(\lambda x) = \lambda\rho_\mathcal{G}(x)$ for all $x \in \mathbb{R}^d$ and $\lambda \geq 0$.*

We will also consider the following risk-consistent properties of $\rho_{\mathcal{G}}$:

Risk-convexity: *If for $X, Y, Z \in L_d^\infty(\mathcal{F})$ there exists an $\alpha \in L^\infty(\mathcal{G})$ with $0 \leq \alpha \leq 1$ such that $\widetilde{\rho}_{\mathcal{G}}(Z(\omega), \omega) = \alpha(\omega)\widetilde{\rho}_{\mathcal{G}}(X(\omega), \omega) + \big(1 - \alpha(\omega)\big)\widetilde{\rho}_{\mathcal{G}}(Y(\omega), \omega)$ a.s., then $\rho_{\mathcal{G}}(Z) \leq \alpha\rho_{\mathcal{G}}(X) + (1-\alpha)\rho_{\mathcal{G}}(Y)$;*

Risk-quasiconvexity: *If for $X, Y, Z \in L_d^\infty(\mathcal{F})$ there exists an $\alpha \in L^\infty(\mathcal{G})$ with $0 \leq \alpha \leq 1$ such that $\widetilde{\rho}_{\mathcal{G}}(Z(\omega), \omega) = \alpha(\omega)\widetilde{\rho}_{\mathcal{G}}(X(\omega), \omega) + \big(1 - \alpha(\omega)\big)\widetilde{\rho}_{\mathcal{G}}(Y(\omega), \omega)$ a.s., then $\rho_{\mathcal{G}}(Z) \leq \rho_{\mathcal{G}}(X) \vee \rho_{\mathcal{G}}(Y)$;*

Risk-positive homogeneity: *If for $X, Y \in L_d^\infty(\mathcal{F})$ there exists an $\alpha \in L^\infty(\mathcal{G})$ with $\alpha \geq 0$ such that $\widetilde{\rho}_{\mathcal{G}}(Y(\omega), \omega) = \alpha(\omega)\widetilde{\rho}_{\mathcal{G}}(X(\omega), \omega)$ a.s., then $\rho_{\mathcal{G}}(Y) = \alpha\rho_{\mathcal{G}}(X)$;*

Risk-regularity: $\rho_{\mathcal{G}}(X, \omega) = \widetilde{\rho}_{\mathcal{G}}(X(\omega), \omega)$ *a.s. for all $X \in L_d^\infty(\mathcal{G})$.*

We will see in Theorem 2.2.9 that risk-antitonicity is the crucial property which guarantees that $\rho_{\mathcal{G}}$ allows a conditional decomposition analogously to (2.1.1). The idea behind all risk-consistent properties is that they ensure a consistency between local - that is ω-wise - risk assessment and the measured global risk. Consider for instance again the risk-antitonicity property and suppose we are given an event $A \in \mathcal{G}$ and random risk factors $Z \in L_d^\infty(\mathcal{F})$ as well as $X, Y \in L_d^\infty(\mathcal{F})$ such that on the level of our realization which satisfies the risk-antitonicity we have $\widetilde{\rho}_{\mathcal{G}}(X(\omega), \omega) \geq \widetilde{\rho}_{\mathcal{G}}(Y(\omega), \omega)$ a.s. on A. In other words for almost all $\omega \in A$, the risk of the constant risk factors $X(\omega)$ evaluated in ω is higher than the corresponding risk of $Y(\omega)$ evaluated in ω. Now consider the modified risk factors $Z_X := X\mathbb{1}_A + Z\mathbb{1}_{A^C}$ and $Z_Y := Y\mathbb{1}_A + Z\mathbb{1}_{A^C}$ where we modify Z on A in such a way that Z_Y is preferred on almost every state in A to Z_X, and otherwise both risk factors are identical. Then risk-antitonicity implies that $\rho_{\mathcal{G}}(Z_Y) \leq \rho_{\mathcal{G}}(Z_X)$.

Our definition of a CSRM is based on properties on constants together with risk-consistent properties. It turns out (see Theorem 2.2.9) that the properties on constants translate into the corresponding properties of the (conditional) aggregation function and the risk-consistent properties translate into the corresponding properties of the (conditional) base risk measure in the decomposition of a CSRM. Moreover, a natural question is to which extend CSRM's also fulfill the established properties of risk measures in the literature. For instance, antitonicity on L_d^∞, i.e. $X \geq Y$ implies $\rho_{\mathcal{G}}(X) \leq \rho_{\mathcal{G}}(Y)$, is commonly accepted as a minimal requirement for risk measures. Further, quasiconvexity or the stronger condition of convexity on L_d^∞ are properties often asked for as they correspond to the requirement that diversification should not be penalized, cf. Cerreia-Vioglio et al. (2011). Also, an important subclass are those CSRM which are positive homogeneous, as for example the CoVaR or the CoES introduced in Adrian and Brunnermeier (2016); see Example 2.4.6 and Example 2.4.7. In general, it will turn out (see Theorem 2.2.11) that properties on constants combined with the corresponding risk-consistent properties will imply properties such as antitonicity, (quasi-) convexity or positive homogeneity of $\rho_{\mathcal{G}}$ on L_d^∞. For example, antitonicity on constants in conjunction with risk-antitonicity implies antitonicity on L_d^∞.

One might ask in which setting it is possible to formulate the risk-consistent properties directly in terms of the function $\rho_{\mathcal{G}}$ without requiring the existence of a particular realization of this function. As we will see in the next Proposition 2.2.2 this is possible if $\rho_{\mathcal{G}}(x)$ has a discrete structure for all $x \in \mathbb{R}^d$. For the sake of brevity we omit the proof.

Proposition 2.2.2. *Let $\rho_{\mathcal{G}} : L_d^\infty(\mathcal{F}) \to L^\infty(\mathcal{G})$ be a function which has a realization with continuous paths. Further suppose that*

$$\rho_{\mathcal{G}}(x) = \sum_{i=1}^{s} a_i(x)\mathbb{1}_{A_i},\ x \in \mathbb{R}^d, \tag{2.2.2}$$

where $a_i(x) \in \mathbb{R}$ and $A_i \in \mathcal{G}$ are pairwise disjoint sets such that $\Omega = \bigcup_{i=1}^{s} A_i$ for $s \in \mathbb{N} \cup \{\infty\}$. Define $k : \Omega \to \mathbb{N};\ \omega \mapsto i$ such that $\omega \in A_i$. Then $\rho_{\mathcal{G}}$ is risk-antitone if and only if

$$\rho_{\mathcal{G}}(X(\omega))\mathbb{1}_{A_{k(\omega)}} \geq \rho_{\mathcal{G}}(Y(\omega))\mathbb{1}_{A_{k(\omega)}} \text{ a.s. implies } \rho_{\mathcal{G}}(X) \geq \rho_{\mathcal{G}}(Y), \tag{2.2.3}$$

where here the point evaluations $X(\omega), Y(\omega) \in \mathbb{R}^d$ have to be understood as equivalence classes of constant random variables. Also the remaining risk-consistent properties can be expressed in a similar way without requiring a particular realization of $\rho_{\mathcal{G}}$.

Remark 2.2.3. Notice that in the setting of Proposition 2.2.2, we had to require that there exists a realization with continuous paths. Sufficient criteria for $\rho_{\mathcal{G}}$ which guarantee that such a continuous realizations exists are well known, e.g. Kolmogorov's criterion (see e.g. Theorem 2.1 in Revuz and Yor (1999)). A sufficient specification of a CSRM solely in terms of $\rho_{\mathcal{G}}$ (without employing any realization) is thus: if $\rho_{\mathcal{G}}$ is antitone on constants, has a discrete structure (2.2.2) and fulfills (2.2.3) and Kolmogorov's criterion, then $\rho_{\mathcal{G}}$ is a CSRM.

In order to state our decomposition result we need to clarify what we mean by a (conditional) aggregation function and a conditional base risk measure. We start with the aggregation function.

Definition 2.2.4 (Aggregation Functions)**.**
We call a function $\widetilde{\Lambda} : \mathbb{R}^d \to \mathbb{R}$ a deterministic aggregation function *(DAF), if it has the following two properties:*

Isotonicity: *If $x, y \in \mathbb{R}^d$ with $x \geq y$, then $\widetilde{\Lambda}(x) \geq \widetilde{\Lambda}(y)$;*

Continuity: *$\widetilde{\Lambda}$ is continuous.*

A DAF is called concave or positive homogeneous, respectively, if it satisfies for all $x, y \in \mathbb{R}^d$

Concavity: *If $\lambda \in [0,1]$, then $\widetilde{\Lambda}\big(\lambda x + (1-\lambda)y\big) \geq \lambda\widetilde{\Lambda}(x) + (1-\lambda)\widetilde{\Lambda}(y)$;*

Positive homogeneity: *$\widetilde{\Lambda}(\lambda x) = \lambda\widetilde{\Lambda}(x)$ for all $\lambda \geq 0$.*

Furthermore, a function $\widetilde{\Lambda}_{\mathcal{G}} : \mathbb{R}^d \times \Omega \to \mathbb{R}$ is a conditional aggregation function *(CAF), if*

(i) $\widetilde{\Lambda}_{\mathcal{G}}(x, \cdot) \in \mathscr{L}^\infty(\mathcal{G})$ for all $x \in \mathbb{R}^d$,

(ii) $\widetilde{\Lambda}_{\mathcal{G}}(\cdot, \omega)$ is a DAF for all $\omega \in \Omega$.

A CAF is called concave (positive homogeneous) if $\widetilde{\Lambda}_{\mathcal{G}}(\cdot, \omega)$ is concave (positive homogeneous) for all $\omega \in \Omega$.

Remark 2.2.5. Note that, functions like CAFs which are continuous in one argument and measurable in the other also appear under the name of Carathéodory functions in the literature on differential equations. For Carathéodory functions it is well known (see e.g. Aubin and Frankowska (2009) Lemma 8.2.6) that they are product measurable, i.e. every CAF $\widetilde{\Lambda}_{\mathcal{G}}$ is $\mathcal{B}(\mathbb{R}) \times \mathcal{G}$-measurable.

Given a CAF $\widetilde{\Lambda}_{\mathcal{G}}$, we extend the aggregation from deterministic to random vectors in the following way (which is well-defined due to Remark 2.2.5 as well as isotonicity and property (i) in the definition of a CAF):

$$\Lambda_{\mathcal{G}} : L_d^{\infty}(\mathcal{F}) \to L^{\infty}(\mathcal{F}), \quad X \mapsto \widetilde{\Lambda}_{\mathcal{G}}(X(\omega), \omega) . \tag{2.2.4}$$

Remark 2.2.6. Notice that the aggregation (2.2.4) of random vectors X is ω-wise in the sense that given a certain state $\omega \in \Omega$, in that state we aggregate the sure payoff $X(\omega)$. Consequently, properties such as isotonicity, concavity or positive homogeneity of the CAF $\widetilde{\Lambda}_{\mathcal{G}}$ translate to the extended CAF $\Lambda_{\mathcal{G}}$. Hence, $\Lambda_{\mathcal{G}}$ always satisfies

$$\Lambda_{\mathcal{G}}(X) \geq \Lambda_{\mathcal{G}}(Y) \text{ for all } X, Y \in L_d^{\infty}(\mathcal{F}) \text{ with } X \geq Y. \tag{2.2.5}$$

If $\widetilde{\Lambda}_{\mathcal{G}}$ is concave, then for all $X, Y \in L_d^{\infty}(\mathcal{F})$ and $\alpha \in L^{\infty}(\mathcal{F})$ with $0 \leq \alpha \leq 1$ we have

$$\Lambda_{\mathcal{G}}(\alpha X + (1-\alpha)Y) \geq \alpha \Lambda_{\mathcal{G}}(X) + (1-\alpha)\Lambda_{\mathcal{G}}(Y), \tag{2.2.6}$$

and if $\widetilde{\Lambda}_{\mathcal{G}}$ is positively homogeneous, then for all $X \in L_d^{\infty}(\mathcal{F})$ and $\alpha \in L^{\infty}(\mathcal{F})$ with $\alpha \geq 0$:

$$\Lambda_{\mathcal{G}}(\alpha X) = \alpha \Lambda_{\mathcal{G}}(X). \tag{2.2.7}$$

The last yet undefined ingredient in our decomposition (2.1.1) is the conditional base risk measure $\eta_{\mathcal{G}}$ which we define next. Notice that the domain $\mathcal{X}$ of $\eta_{\mathcal{G}}$ depends on the underlying aggregation given by $\rho_{\mathcal{G}}$. For example the aggregation function $\widetilde{\Lambda}(x) = \sum_{i=1}^{d} \min\{x_i, 0\}, x \in \mathbb{R}^d$ only considers the losses. Hence, the corresponding base risk measure η a priori only needs to be defined on the negative cone of $L^{\infty}(\mathcal{F})$, even though it in many cases allows for an extension to $L^{\infty}(\mathcal{F})$. We will see in Lemma 2.3.1 that if $\mathcal{X}$ is the image of an extended CAF $\Lambda_{\mathcal{G}}$ then $\mathcal{X}$ is $\mathcal{G}$-conditionally convex, i.e. $F, G \in \mathcal{X}$ and $\alpha \in L^{\infty}(\mathcal{G})$ with $0 \leq \alpha \leq 1$ implies $\alpha F + (1-\alpha)G \in \mathcal{X}$.

Definition 2.2.7 (Conditional Base Risk Measure)**.**
Let $\mathcal{X} \subseteq L^{\infty}(\mathcal{F})$ be a $\mathcal{G}$-conditionally convex set. A function $\eta_{\mathcal{G}} : \mathcal{X} \to L^{\infty}(\mathcal{G})$ is a conditional base risk measure *(CBRM), if it is*

Antitone: *$F \geq G$ implies $\eta_{\mathcal{G}}(F) \leq \eta_{\mathcal{G}}(G)$.*

Moreover, we will also consider CBRM's which fulfill additionally one or more of the following properties:

Constant on constants: *$\eta_{\mathcal{G}}(\alpha) = -\alpha$ for all $\alpha \in \mathcal{X} \cap L^{\infty}(\mathcal{G})$;*

Quasiconvexity: *$\eta_{\mathcal{G}}(\alpha F + (1-\alpha)G) \leq \eta_{\mathcal{G}}(F) \vee \eta_{\mathcal{G}}(G)$ for all $\alpha \in L^{\infty}(\mathcal{G})$ with $0 \leq \alpha \leq 1$;*

Convexity: $\eta_{\mathcal{G}}\left(\alpha F+(1-\alpha) G\right) \leq \alpha \eta_{\mathcal{G}}(F)+(1-\alpha) \eta_{\mathcal{G}}(G)$ *for all* $\alpha \in L^{\infty}(\mathcal{G})$ *with* $0 \leq \alpha \leq 1$*;*

Positive homogeneity: $\eta_{\mathcal{G}}(\alpha F)=\alpha \eta_{\mathcal{G}}(F)$ *for all* $\alpha \in L^{\infty}(\mathcal{G})$ *with* $\alpha \geq 0$ *and* $\alpha F \in \mathcal{X}$.

Constructing a CSRM by composing a CBRM and a CAF as in (2.1.1), we need a property for $\eta_{\mathcal{G}}$ which allows to 'extract' the CAF in order to obtain the properties on constants of $\rho_{\mathcal{G}}$. The constant on constants property serves this purpose, but we will see in Theorem 2.2.9 that the following weaker property is also sufficient.

Definition 2.2.8. *A CBRM* $\eta_{\mathcal{G}}: \mathcal{X} \rightarrow L^{\infty}(\mathcal{G})$ *is called* constant on a CAF $\widetilde{\Lambda}_{\mathcal{G}}$*, if* $\Lambda_{\mathcal{G}}(x) \in \mathcal{X}$ *for all* $x \in \mathbb{R}^{d}$ *and*

$$\eta_{\mathcal{G}}\left(\Lambda_{\mathcal{G}}(x)\right)=-\Lambda_{\mathcal{G}}(x) \text{ for all } x \in \mathbb{R}^{d}. \tag{2.2.8}$$

Clearly, if $\eta_{\mathcal{G}}$ is constant on constants, then it is constant on any CAF with an appropriate image as (2.2.8) is always satisfied.

Conditional risk measures have been widely studied in the literature, see Föllmer and Schied (2011) for an overview. As already explained above the antitonicity is widely accepted as a minimal requirement for risk measures. The constant on constants property is a standard technical assumption, whereas we will only need the weaker property of constancy on an aggregation function for an CBRM. Typically conditional risk measures are also required to be monetary in the sense that they satisfy some translation invariance property which we do not require in our setting, see e.g. Detlefsen and Scandolo (2005). Much of the literature is concerned with the study of quasiconvex or convex conditional risk measures which in our setting implies that the corresponding risk-consistent conditional systemic risk measure will satisfy risk-quasiconvexity resp. risk-convexity, see Theorem 2.2.9.

After introducing all objects and properties of interest we are now able to state our decomposition theorem.

Theorem 2.2.9. *A function* $\rho_{\mathcal{G}}: L_{d}^{\infty}(\mathcal{F}) \rightarrow L^{\infty}(\mathcal{G})$ *is a CSRM if and only if there exists a CAF* $\widetilde{\Lambda}_{\mathcal{G}}: \mathbb{R}^{d} \times \Omega \rightarrow \mathbb{R}$ *and a CBRM* $\eta_{\mathcal{G}}: \operatorname{Im} \Lambda_{\mathcal{G}} \rightarrow L^{\infty}(\mathcal{G})$ *such that* $\eta_{\mathcal{G}}$ *is constant on* $\widetilde{\Lambda}_{\mathcal{G}}$ *(Definition* 2.2.8*) and*

$$\rho_{\mathcal{G}}(X)=\eta_{\mathcal{G}}\left(\Lambda_{\mathcal{G}}(X)\right) \quad \text{for all } X \in L_{d}^{\infty}(\mathcal{F}), \tag{2.2.9}$$

where the extended CAF $\Lambda_{\mathcal{G}}(X):=\widetilde{\Lambda}_{\mathcal{G}}(X(\omega), \omega)$ *was introduced in* (2.2.4)*. The decomposition into* $\eta_{\mathcal{G}}$ *and* $\Lambda_{\mathcal{G}}$ *is unique.*
Furthermore there is a one-to-one correspondence between additional properties of the CBRM $\eta_{\mathcal{G}}$ *and additional risk-consistent properties of the CSRM* $\rho_{\mathcal{G}}$*:*

- $\rho_{\mathcal{G}}$ *is risk-convex iff* $\eta_{\mathcal{G}}$ *is convex;*
- $\rho_{\mathcal{G}}$ *is risk-quasiconvex iff* $\eta_{\mathcal{G}}$ *is quasiconvex;*
- $\rho_{\mathcal{G}}$ *is risk-positive homogeneous iff* $\eta_{\mathcal{G}}$ *is positive homogeneous;*
- $\rho_{\mathcal{G}}$ *is risk-regular iff* $\eta_{\mathcal{G}}$ *is constant on constants.*

Moreover, properties on constants of the CSRM $\rho_{\mathcal{G}}$ are related to properties of the CAF $\widetilde{\Lambda}_{\mathcal{G}}$:

- *$\rho_{\mathcal{G}}$ is convex on constants iff $\widetilde{\Lambda}_{\mathcal{G}}$ is concave;*
- *$\rho_{\mathcal{G}}$ is positive homogeneous on constants iff $\widetilde{\Lambda}_{\mathcal{G}}$ is positive homogeneous.*

The proof of Theorem 2.2.9 is quite lengthy and needs some additional preparation and is thus postponed to Section 2.3. Note that it follows from the proof of Theorem 2.2.9 that the aggregation rule in (2.2.9) is deterministic if and only if $\rho_{\mathcal{G}}(\mathbb{R}^d) \subseteq \mathbb{R}$.

Remark 2.2.10. The decomposition (2.2.9) can also be established without requiring the CSRM to be risk-antitone, but to fulfill the weaker property

$$\widetilde{\rho}_{\mathcal{G}}\left(X(\omega),\omega\right) = \widetilde{\rho}_{\mathcal{G}}\left(Y(\omega),\omega\right) \text{ a.s.} \Longrightarrow \rho_{\mathcal{G}}(X) = \rho_{\mathcal{G}}(Y). \tag{2.2.10}$$

Notice, however, if we only require (2.2.10), then the CBRM $\eta_{\mathcal{G}}$ in (2.2.9) (and also $\rho_{\mathcal{G}}$ itself, see Theorem 2.2.11 below) might not be antitone anymore.

An important question is to which degree CSRM's fulfill the usual (conditional) axioms of risk measures on $L_d^\infty(\mathcal{F})$ (where these axioms on $L_d^\infty(\mathcal{F})$ are defined analogously to the ones on $L^\infty(\mathcal{F})$ in Definition 2.2.7). In the following Theorem 2.2.11 we will investigate the relation between risk-consistent properties and properties on constants on the one side and properties of $\rho_{\mathcal{G}}$ on $L_d^\infty(\mathcal{F})$ on the other.

Theorem 2.2.11. *Let $\rho_{\mathcal{G}}$ be a CSRM. Then*

- *risk-antitonicity together with antitonicity on constants can equivalently be replaced by antitonicity of $\rho_{\mathcal{G}}$ ($X \geq Y$ implies $\rho_{\mathcal{G}}(X) \leq \rho_{\mathcal{G}}(Y)$) together with* (2.2.10).

Moreover:

- *$\rho_{\mathcal{G}}$ is risk-positive homogeneous and positive homogeneous on constants iff $\rho_{\mathcal{G}}$ is positive homogeneous;*
- *If $\rho_{\mathcal{G}}$ is risk-convex and convex on constants, then $\rho_{\mathcal{G}}$ is convex;*
- *If $\rho_{\mathcal{G}}$ is risk-quasiconvex and convex on constants, then $\rho_{\mathcal{G}}$ is quasiconvex.*

As for Theorem 2.2.9 we postpone the proof to Section 2.3.

Remark 2.2.12. We have seen in Theorem 2.2.11 that a property on $L_d^\infty(\mathcal{F})$ of a CSRM is implied by the corresponding risk-consistent property and the property on constants. The reverse is only true for the antitonicity and positive homogeneity. To see this we give a counterexample for the convex case. Suppose that $\widetilde{\Lambda}_{\mathcal{G}}(x) := u^{-1}\left(\sum_{i=1}^d x_i\right)$ and $\eta_{\mathcal{G}}(F) := -u^{-1}\left(\mathbb{E}_{\mathbb{P}}\left[u(F) \mid \mathcal{G}\right]\right)$, where $u : \mathbb{R} \to \mathbb{R}$ is a strictly increasing and convex function. Then it can be easily verified that u^{-1} is strictly increasing and concave. Hence $\widetilde{\Lambda}_{\mathcal{G}}$ is a concave CAF and $\eta_{\mathcal{G}}$ is a CBRM. Nevertheless, there are functions u such that $\eta_{\mathcal{G}}$ is not a convex CBRM, e.g. $u(c) = c\mathbb{1}_{\{c\leq 0\}} + ac\mathbb{1}_{\{c>0\}}, a > 1$. According to Theorem 2.2.9 we get a CSRM $\rho_{\mathcal{G}}$ by composing $\Lambda_{\mathcal{G}}$ and $\eta_{\mathcal{G}}$, which is explicitly given by

$$\rho_{\mathcal{G}}(X) = -u^{-1}\left(\mathbb{E}_{\mathbb{P}}\left[\sum_{i=1}^d X_i \,\middle|\, \mathcal{G}\right]\right).$$

It is obvious that $\rho_{\mathcal{G}}$ is convex. But since $\eta_{\mathcal{G}}$ is not convex, $\rho_{\mathcal{G}}$ cannot be risk-convex by Theorem 2.2.9.

2.3 Proof of Theorem 2.2.9 and 2.2.11

Before we state the proofs of Theorems 2.2.9 and 2.2.11, we provide some auxiliary results.

Lemma 2.3.1. *Let $\widetilde{\Lambda}_{\mathcal{G}} : \mathbb{R}^d \times \Omega \to \mathbb{R}$ be a CAF and let $\mathcal{H}$ be a sub-σ-algebra of $\mathcal{F}$ such that $\mathcal{G} \subseteq \mathcal{H} \subseteq \mathcal{F}$. Then*

$$\Lambda_{\mathcal{G}}\left(L_d^\infty(\mathcal{H})\right) \subseteq L^\infty(\mathcal{H}), \tag{2.3.1}$$

and for every $X, Y \in L^\infty(\mathcal{H})$ and $\alpha \in L^\infty(\mathcal{G})$ with $0 \leq \alpha \leq 1$ there is an $F \in L^\infty(\mathcal{H})$ such that

$$\alpha\Lambda_{\mathcal{G}}(X) + (1-\alpha)\Lambda_{\mathcal{G}}(Y) = \Lambda_{\mathcal{G}}(F\mathbf{1}_d).$$

In particular this implies that the image of $\Lambda_{\mathcal{G}}$ is $\mathcal{G}$-conditionally convex.
Conversely, we have that

$$L^\infty(\mathcal{H}) \cap \operatorname{Im}\Lambda_{\mathcal{G}} \subseteq \Lambda_{\mathcal{G}}\left(L_d^\infty(\mathcal{H})\right).$$

Proof. Let $X \in L_d^\infty(\mathcal{H})$ and set $F(\omega) := \widetilde{\Lambda}_{\mathcal{G}}\left(X(\omega), \omega\right)$, $\omega \in \Omega$. Since $\widetilde{\Lambda}_{\mathcal{G}}$ is a Carathéodory map it follows that F is $\mathcal{H}$-measurable, cf. Lemma 8.2.3 in Aubin and Frankowska (2009). Let $A := \{\omega \in \Omega : \widetilde{\Lambda}_{\mathcal{G}}\left(X(\omega), \omega\right) \leq 0\}$. Then

$$\begin{aligned}\|F\|_\infty = \left\|\widetilde{\Lambda}_{\mathcal{G}}\left(X(\cdot), \cdot\right)\right\|_\infty &\leq \left\|\widetilde{\Lambda}_{\mathcal{G}}\left(\operatorname{essinf} X, \cdot\right)\mathbb{1}_A\right\|_\infty + \left\|\widetilde{\Lambda}_{\mathcal{G}}\left(\operatorname{esssup} X, \cdot\right)\mathbb{1}_{A^C}\right\|_\infty \\ &\leq \left\|\widetilde{\Lambda}_{\mathcal{G}}\left(\operatorname{essinf} X, \cdot\right)\right\|_\infty + \left\|\widetilde{\Lambda}_{\mathcal{G}}\left(\operatorname{esssup} X, \cdot\right)\right\|_\infty < \infty,\end{aligned} \tag{2.3.2}$$

where we used the boundedness condition Definition 2.2.4 (i) in the last step and where $\operatorname{essinf} X := (\operatorname{essinf} X_1, \ldots, \operatorname{essinf} X_d)$, and similarly for esssup. Hence, we conclude that $F \in L^\infty(\mathcal{H})$.
Let $X, Y \in L_d^\infty(\mathcal{H})$ and $\alpha \in L^\infty(\mathcal{G})$ with $0 \leq \alpha \leq 1$. The rest of the proof is based on a measurable selection theorem for which we need that the probability space is complete. However, $L_d^\infty(\Omega, \mathcal{H}, \mathbb{P})$ and $L_d^\infty(\Omega, \widehat{\mathcal{H}}, \widehat{\mathbb{P}})$ are isometric isomorph, where $(\Omega, \widehat{\mathcal{H}}, \widehat{\mathbb{P}})$ denotes the completion of $(\Omega, \mathcal{H}, \mathbb{P})$. Thus for X and Y there exist respective $\widehat{X}, \widehat{Y} \in L_d^\infty(\widehat{\mathcal{H}})$ and it is easily verified that any representatives of the equivalence classes $\widehat{X}$ $(\widehat{Y})$ and X (Y) only differ on a $\widehat{\mathbb{P}}$-nullset. Define

$$\underline{x} := \operatorname{essinf}\left(\min_{i=1,\ldots,d}\left(\min(\widehat{X}_i, \widehat{Y}_i)\right)\right) \quad \text{and} \quad \overline{x} := \operatorname{esssup}\left(\max_{i=1,\ldots,d}\left(\max(\widehat{X}_i, \widehat{Y}_i)\right)\right).$$

Since both $\widehat{X}, \widehat{Y}$ are essentially bounded we have that $\underline{x}, \overline{x} \in \mathbb{R}$. Moreover the random variable G which is given for each $\omega \in \Omega$ by

$$G(\omega) := \alpha(\omega)\widetilde{\Lambda}_{\mathcal{G}}\left(X(\omega), \omega\right) + \big(1 - \alpha(\omega)\big)\widetilde{\Lambda}_{\mathcal{G}}\left(Y(\omega), \omega\right),$$

is contained in an equivalence class in $L^\infty(\mathcal{H})$ by the first part of the proof and thus we can find a corresponding equivalence class $\widehat{G} \in L^\infty(\widehat{\mathcal{H}})$. By isotonicity we have

$$\widetilde{\Lambda}_{\mathcal{G}}\left(\underline{x}\mathbf{1}_d, \omega\right) \leq \widehat{G}(\omega) \leq \widetilde{\Lambda}_{\mathcal{G}}\left(\overline{x}\mathbf{1}_d, \omega\right) \quad \widehat{\mathbb{P}}\text{-}a.s.$$

The continuity of the function $\mathbb{R} \ni x \mapsto \widetilde{\Lambda}_{\mathcal{G}}(x\mathbf{1}_d, \omega)$ for each $\omega \in \Omega$ implies that

$$\widehat{G}(\omega) \in \left\{ \widetilde{\Lambda}_{\mathcal{G}}(x\mathbf{1}_d, \omega) : x \in [\underline{x}, \overline{x}] \right\} \quad \widehat{\mathbb{P}}\text{-}a.s.$$

Finally, we can apply Filippov's theorem (see e.g. Aubin and Frankowska (2009) Theorem 8.2.10), that is there exists a $\widehat{\mathcal{H}}$-measurable selection $\widehat{F}(\omega) \in [\underline{x}, \overline{x}]$ such that

$$\widehat{G}(\omega) = \widetilde{\Lambda}_{\mathcal{G}}\left(\widehat{F}(\omega)\mathbf{1}_d, \omega\right) \quad \widehat{\mathbb{P}}\text{-}a.s.$$

For this measurable selection $\widehat{F}$ we can find an $F \in L^{\infty}(\mathcal{H})$ such that $\widehat{\mathbb{P}}(\widehat{F} \neq F) = 0$. Hence there exists an $F \in L^{\infty}(\mathcal{H})$ such that

$$\alpha\Lambda_{\mathcal{G}}(X) + (1-\alpha)\Lambda_{\mathcal{G}}(Y) = \Lambda_{\mathcal{G}}(F\mathbf{1}_d).$$

For the last part of the proof let $G \in \operatorname{Im}\Lambda_{\mathcal{G}} \cap L^{\infty}(\mathcal{H})$, then by definition there exists an $X \in L_d^{\infty}(\mathcal{F})$ such that $\Lambda_{\mathcal{G}}(X) = G$. Thus by setting $\underline{x} := \operatorname{essinf}(\min_{i=1,\dots,d} X_i)$ and $\overline{x} := \operatorname{esssup}(\max_{i=1,\dots,d} X_i)$ we have that

$$\widetilde{\Lambda}_{\mathcal{G}}(\underline{x}\mathbf{1}_d, \omega) \leq G(\omega) \leq \widetilde{\Lambda}_{\mathcal{G}}(\overline{x}\mathbf{1}_d, \omega) \quad a.s.$$

Moreover, since G is $\mathcal{H}$-measurable, we obtain by a similar argumentation as above that there exists a $\mathcal{H}$-measurable F with $\underline{x} \leq F \leq \overline{x}$ and $\Lambda_{\mathcal{G}}(F\mathbf{1}_d) = G$. □

Lemma 2.3.2. *Let $\widetilde{\Lambda}_{\mathcal{G}}$ be a conditional aggregation function. Then there exists a $\mathbb{P}$-nullset N such that if $x, y \in \mathbb{R}^d$ satisfy $\widetilde{\Lambda}_{\mathcal{G}}(x, \omega) = \widetilde{\Lambda}_{\mathcal{G}}(y, \omega)$ a.s. it holds that $\widetilde{\Lambda}_{\mathcal{G}}(x, \omega) = \widetilde{\Lambda}_{\mathcal{G}}(y, \omega)$ for all $\omega \in N^C$, where N^C denotes the complement of N.*

Proof. Consider the sets $B := \{(x, y) \in \mathbb{Q}^{2d} : \widetilde{\Lambda}_{\mathcal{G}}(x, \omega) \geq \widetilde{\Lambda}_{\mathcal{G}}(y, \omega) \text{ a.s.}\}$ and $N_{(x,y)} := \{\omega \in \Omega : \widetilde{\Lambda}_{\mathcal{G}}(x, \omega) < \widetilde{\Lambda}_{\mathcal{G}}(y, \omega)\}$ for $(x, y) \in B$. By definition $N_{(x,y)}$ is a $\mathbb{P}$-nullset for all $(x, y) \in B$, but since B has only countable many elements, the same holds true for the union $N := \bigcup_{(x,y)\in B} N_{(x,y)}$.
Now consider $x, y \in \mathbb{R}^d$ such that $\widetilde{\Lambda}_{\mathcal{G}}(x, \omega) \geq \widetilde{\Lambda}_{\mathcal{G}}(y, \omega)$ a.s. We can always find sequences $(x_n)_{n\in\mathbb{N}}, (y_n)_{n\in\mathbb{N}} \in \mathbb{Q}^{\mathbb{N}}$ such that $x_n \downarrow x$ and $y_n \uparrow y$ for $n \to \infty$. The isotonicity of $\widetilde{\Lambda}_{\mathcal{G}}$ yields $\widetilde{\Lambda}_{\mathcal{G}}(x_n, \omega) \geq \widetilde{\Lambda}_{\mathcal{G}}(x, \omega) \geq \widetilde{\Lambda}_{\mathcal{G}}(y, \omega) \geq \widetilde{\Lambda}_{\mathcal{G}}(y_n, \omega)$ a.s., thus $(x_n, y_n) \in B$ for all $n \in \mathbb{N}$. Therefore we get for all $\omega \in N^C$ that

$$\widetilde{\Lambda}_{\mathcal{G}}(x, \omega) = \lim_{n\to\infty} \widetilde{\Lambda}_{\mathcal{G}}(x_n, \omega) \geq \lim_{n\to\infty} \widetilde{\Lambda}_{\mathcal{G}}(y_n, \omega) = \widetilde{\Lambda}_{\mathcal{G}}(y, \omega),$$

where we have used that $\widetilde{\Lambda}_{\mathcal{G}}(\cdot, \omega)$ is continuous for every $\omega \in \Omega$. As $\widetilde{\Lambda}_{\mathcal{G}}(x, \omega) = \widetilde{\Lambda}_{\mathcal{G}}(y, \omega)$ a.s. implies $\widetilde{\Lambda}_{\mathcal{G}}(x, \omega) \geq \widetilde{\Lambda}_{\mathcal{G}}(y, \omega)$ a.s. and $\widetilde{\Lambda}_{\mathcal{G}}(x, \omega) \leq \widetilde{\Lambda}_{\mathcal{G}}(y, \omega)$ a.s., the assertion follows. □

Note that the $\mathbb{P}$-nullset N in Lemma 2.3.2 is universal in the sense that it does not depend on the pair $(x, y) \in \mathbb{R}^{2d}$.

Proof of Theorem 2.2.9. For the rest of the proof let $X, Y \in L_d^\infty(\mathcal{F})$.
"$\Leftarrow$":
Suppose that $\widetilde{\Lambda}_\mathcal{G} : \mathbb{R}^d \times \Omega \to \mathbb{R}$ is a CAF with extended CAF $\Lambda_\mathcal{G} : L_d^\infty(\mathcal{F}) \to L^\infty(\mathcal{F})$, and that $\eta_\mathcal{G} : \operatorname{Im} \Lambda_\mathcal{G} \to L^\infty(\mathcal{G})$ is a CBRM which is constant on $\widetilde{\Lambda}_\mathcal{G}$. Moreover, define the function

$$\rho_\mathcal{G} : L_d^\infty(\mathcal{F}) \to L^\infty(\mathcal{G}),\ X \mapsto \eta_\mathcal{G}\left(\Lambda_\mathcal{G}\left(X\right)\right).$$

First we will show that $\rho_\mathcal{G}$ is antitone (and thus in particular antitone on constants): To this end, let $X \geq Y$. As $\widetilde{\Lambda}_\mathcal{G}\left(\cdot, \omega\right)$ is isotone for all $\omega \in \Omega$ we know from (2.2.5) that also the extended CAF is isotone, i.e. $\Lambda_\mathcal{G}\left(X\right) \geq \Lambda_\mathcal{G}\left(Y\right)$. By the antitonicity of $\eta_\mathcal{G}$ we can conclude that

$$\rho_\mathcal{G}\left(X\right) = \eta_\mathcal{G}\left(\Lambda_\mathcal{G}\left(X\right)\right) \leq \eta_\mathcal{G}\left(\Lambda_\mathcal{G}\left(Y\right)\right) = \rho_\mathcal{G}\left(Y\right).$$

Next we will show that there exists a realization of $\rho_\mathcal{G}$ with continuous paths and which fulfills the risk-antitonicity. From (2.2.8) and Lemma 2.3.2 it can be readily seen that we can always find a realization of $\eta_\mathcal{G}$ and a universal $\mathbb{P}$-nullset N such that for all $\omega \in N^C$

$$\eta_\mathcal{G}\left(\Lambda_\mathcal{G}\left(x\right), \omega\right) = -\widetilde{\Lambda}_\mathcal{G}\left(x, \omega\right) \text{ for all } x \in \mathbb{R}^d. \tag{2.3.3}$$

Given this realization of $\eta_\mathcal{G}$ we consider in the following the realization $\rho_\mathcal{G}(\cdot, \cdot)$ of $\rho_\mathcal{G}$ given by

$$\rho_\mathcal{G}\left(X, \omega\right) := \eta_\mathcal{G}\left(\Lambda_\mathcal{G}\left(X\right), \omega\right),\ X \in L_d^\infty(\mathcal{F}), \omega \in \Omega.$$

The function $\widetilde{\rho}_\mathcal{G} : \mathbb{R}^d \times \Omega \to \mathbb{R};\ x \mapsto \rho_\mathcal{G}(x, \omega)$ has continuous paths (a.s.) because $\widetilde{\Lambda}_\mathcal{G}$ has continuous paths. As for the risk-antitonicity, let $\widetilde{\rho}_\mathcal{G}\left(X(\omega), \omega\right) \geq \widetilde{\rho}_\mathcal{G}\left(Y(\omega), \omega\right)$ a.s. By rewriting this in terms of the decomposition, i.e.
$\eta_\mathcal{G}\big(\Lambda_\mathcal{G}\left(X(\omega)\right), \omega\big) \geq \eta_\mathcal{G}\big(\Lambda_\mathcal{G}\left(Y(\omega)\right), \omega\big)$, we realize by (2.3.3) that

$$\widetilde{\Lambda}_\mathcal{G}\left(X(\omega), \omega\right) \leq \widetilde{\Lambda}_\mathcal{G}\left(Y(\omega), \omega\right) \text{ a.s.} \tag{2.3.4}$$

Note that our application of (2.3.3) relies on the fact that the nullset N in (2.3.3) does not depend on $x \in \mathbb{R}^d$. As (2.3.4) is equivalent to $\Lambda_\mathcal{G}\left(X\right) \leq \Lambda_\mathcal{G}\left(Y\right)$, we conclude that

$$\rho_\mathcal{G}(X) = \eta_\mathcal{G}(\Lambda_\mathcal{G}\left(X\right)) \geq \eta_\mathcal{G}(\Lambda_\mathcal{G}\left(Y\right)) = \rho_\mathcal{G}(Y),$$

where we used the antitonicity of $\eta_\mathcal{G}$. Hence, we have proved that $\rho_\mathcal{G}$ is a CSRM.
Next we treat the special cases when $\eta_\mathcal{G}$ and/or $\widetilde{\Lambda}_\mathcal{G}$ satisfy some extra properties.
Risk-regularity: Suppose $\eta_\mathcal{G}$ is constant on constants. Then we have

$$\rho_\mathcal{G}(X) = -\Lambda_\mathcal{G}\left(X\right) \text{ for all } X \in L_d^\infty(\mathcal{G}),$$

and thus we obtain for the realization $\rho_\mathcal{G}\left(\cdot, \cdot\right)$ that for all $X \in L_d^\infty(\mathcal{G})$

$$\rho_\mathcal{G}\left(X, \omega\right) = -\widetilde{\Lambda}_\mathcal{G}\left(X(\omega), \omega\right) \text{ a.s.}$$

As above (2.3.3) implies that for all $\omega \in N^C$

$$-\widetilde{\Lambda}_\mathcal{G}\left(X(\omega), \omega\right) = \eta_\mathcal{G}\left(\Lambda_\mathcal{G}\left(X(\omega)\right), \omega\right) = \widetilde{\rho}_\mathcal{G}\left(X(\omega), \omega\right).$$

Risk-quasiconvexity/convexity: Suppose that $\eta_{\mathcal{G}}$ is quasiconvex. We show that $\rho_{\mathcal{G}}$ is risk-quasiconvex. To this end, suppose there exist $X, Y, Z \in L_d^\infty(\mathcal{F})$ and an $\alpha \in L^\infty(\mathcal{G})$ with $0 \leq \alpha \leq 1$ such that

$$\widetilde{\rho}_{\mathcal{G}}\left(Z(\omega), \omega\right) = \alpha(\omega)\widetilde{\rho}_{\mathcal{G}}\left(X(\omega), \omega\right) + \big(1-\alpha(\omega)\big)\widetilde{\rho}_{\mathcal{G}}\left(Y(\omega), \omega\right) \text{ a.s.}$$

Then, as above, by using (2.3.3), it follows that

$$\Lambda_{\mathcal{G}}\left(Z\right) = \alpha\Lambda_{\mathcal{G}}\left(X\right) + (1-\alpha)\Lambda_{\mathcal{G}}\left(Y\right).$$

Hence the quasiconvexity of $\eta_{\mathcal{G}}$ yields

$$\begin{aligned}\rho_{\mathcal{G}}(Z) = \eta_{\mathcal{G}}\left(\Lambda_{\mathcal{G}}\left(Z\right)\right) &= \eta_{\mathcal{G}}\left(\alpha\Lambda_{\mathcal{G}}\left(X\right) + (1-\alpha)\Lambda_{\mathcal{G}}\left(Y\right)\right)\\ &\leq \eta_{\mathcal{G}}\left(\Lambda_{\mathcal{G}}\left(X\right)\right) \vee \eta_{\mathcal{G}}\left(\Lambda_{\mathcal{G}}\left(Y\right)\right)\\ &= \rho_{\mathcal{G}}(X) \vee \rho_{\mathcal{G}}(Y).\end{aligned}$$

Similarly it follows that $\rho_{\mathcal{G}}$ is risk-convex whenever $\eta_{\mathcal{G}}$ is convex.
Risk-positive homogeneity: Finally, if $\eta_{\mathcal{G}}$ is positively homogeneous, then it is straightforward to see that also $\rho_{\mathcal{G}}$ is risk-positively homogeneous.
Properties on constants: Suppose that $\widetilde{\Lambda}_{\mathcal{G}}$ is concave or positive homogeneous, then it is an immediate consequence of (2.3.3) that $\rho_{\mathcal{G}}$ is convex on constants or positive homogeneous on constants, resp.
"$\Rightarrow$":
Let $\rho_{\mathcal{G}}\left(\cdot,\cdot\right)$ denote a realization of the CSRM $\rho_{\mathcal{G}}$ such that $\widetilde{\rho}_{\mathcal{G}}$ has continuous paths and the risk-antitonicity holds. We define the function $\widehat{\Lambda}_{\mathcal{G}} : \mathbb{R}^d \times \Omega \to \mathbb{R}$ by

$$\widehat{\Lambda}_{\mathcal{G}}\left(x, \omega\right) := -\widetilde{\rho}_{\mathcal{G}}\left(x, \omega\right). \tag{2.3.5}$$

We show that $\widehat{\Lambda}_{\mathcal{G}}(\cdot, \omega)$ is a DAF for almost all $\omega \in \Omega$, i.e. that it is isotone and continuous. The continuity is obvious by (2.3.5). For the isotonicity consider the sets $B := \big\{(x, y) \in \mathbb{Q}^{2d} : \ x \geq y\big\}$ and $A^{(1)}_{(x,y)} := \{\omega \in \Omega : \ \widehat{\Lambda}_{\mathcal{G}}\left(x, \omega\right) < \widehat{\Lambda}_{\mathcal{G}}\left(y, \omega\right)\}$ for $(x, y) \in B$. Since $\rho_{\mathcal{G}}$ is antitone on constants we obtain that $A^{(1)} := \bigcup_{(x,y)\in B} A^{(1)}_{(x,y)}$ is a $\mathbb{P}$-nullset. Moreover, let $A^{(2)}$ denote the $\mathbb{P}$-nullset on which $\widehat{\Lambda}_{\mathcal{G}}$ has discontinuous sample paths. Consider $x, y \in \mathbb{R}^d$ such that $x \geq y$, and let $(x_n, y_n) \in B^{\mathbb{N}}$ be a sequence which converges to (x, y) for $n \to \infty$. Then we get for all $\omega \in \left(A^{(1)} \cup A^{(2)}\right)^C$ that

$$\widehat{\Lambda}_{\mathcal{G}}\left(x, \omega\right) = \lim_{n\to\infty} \widehat{\Lambda}_{\mathcal{G}}\left(x_n, \omega\right) \geq \lim_{n\to\infty} \widehat{\Lambda}_{\mathcal{G}}\left(y_n, \omega\right) = \widehat{\Lambda}_{\mathcal{G}}\left(y, \omega\right),$$

and thus the paths $\widehat{\Lambda}_{\mathcal{G}}\left(\cdot, \omega\right)$ are isotone a.s.

The fact that the paths $\widehat{\Lambda}_{\mathcal{G}}\left(\cdot, \omega\right)$ are concave (positively homogeneous) a.s. whenever $\rho_{\mathcal{G}}$ is convex on constants (positively homogeneous on constants) follows by a similar approximation argument on the continuous paths which are concave (positively homogeneous) on $\mathbb{Q}^d$.

Given the above considerations, we choose a modification $\widetilde{\Lambda}_{\mathcal{G}}$ of $\widehat{\Lambda}_{\mathcal{G}}$ such that $\widetilde{\Lambda}_{\mathcal{G}}(\cdot, \omega)$, is a (concave/positively homogeneous) DAF for all $\omega \in \Omega$. Note that for $\widetilde{\Lambda}_{\mathcal{G}}$ relation (2.3.5) is only valid a.s., that is there is a $\mathbb{P}$-nullset N such that for all $x \in \mathbb{R}^d$ and $\omega \in N^C$

$$\widetilde{\Lambda}_{\mathcal{G}}\left(x, \omega\right) = -\widetilde{\rho}_{\mathcal{G}}\left(x, \omega\right). \tag{2.3.6}$$

As $-\widetilde{\rho}_{\mathcal{G}}(x,\cdot) \in \mathscr{L}^\infty(\mathcal{G})$ and thus also $\widetilde{\Lambda}_{\mathcal{G}}(x,\cdot) \in \mathscr{L}^\infty(\mathcal{G})$ for all $x \in \mathbb{R}^d$ (note that $N \in \mathcal{G}$), we have shown that $\widetilde{\Lambda}_{\mathcal{G}}$ is indeed a CAF.

Next, we will construct a CBRM $\eta_{\mathcal{G}} : \operatorname{Im}\Lambda_{\mathcal{G}} =: \mathcal{X} \to L^\infty(\mathcal{G})$ such that $\rho_{\mathcal{G}} = \eta_{\mathcal{G}} \circ \Lambda_{\mathcal{G}}$ where $\Lambda_{\mathcal{G}}$ is the extended CAF of $\widetilde{\Lambda}_{\mathcal{G}}$. For $F \in \mathcal{X}$ we define

$$\eta_{\mathcal{G}}(F) := \rho_{\mathcal{G}}(X), \tag{2.3.7}$$

where $X \in L^\infty_d(\mathcal{F})$ is given by

$$\Lambda_{\mathcal{G}}(X) = F. \tag{2.3.8}$$

Since $F \in \mathcal{X}$ the existence of such X is always ensured. By (2.3.8) and (2.3.7) we obtain the desired decomposition

$$\eta_{\mathcal{G}}(\Lambda_{\mathcal{G}}(X)) = \rho_{\mathcal{G}}(X),$$

if $\eta_{\mathcal{G}}$ is well-defined. In order to show the latter, let $X^{(1)}, X^{(2)} \in L^\infty_d(\mathcal{F})$ such that

$$\Lambda_{\mathcal{G}}\left(X^{(1)}\right) = \Lambda_{\mathcal{G}}\left(X^{(2)}\right) = F,$$

which by definition of $\Lambda_{\mathcal{G}}$ in (2.2.4) can be rewritten as

$$\widetilde{\Lambda}_{\mathcal{G}}\left(X^{(1)}(\omega),\omega\right) = F(\omega) = \widetilde{\Lambda}_{\mathcal{G}}\left(X^{(2)}(\omega),\omega\right) \text{ a.s.}$$

By (2.3.6) this can be restated in terms of $\widetilde{\rho}_{\mathcal{G}}(\cdot,\cdot)$ as

$$\widetilde{\rho}_{\mathcal{G}}\left(X^{(1)}(\omega),\omega\right) = \widetilde{\rho}_{\mathcal{G}}\left(X^{(2)}(\omega),\omega\right) \text{ a.s.}$$

Now the risk-antitonicity of $\rho_{\mathcal{G}}$ yields $\rho_{\mathcal{G}}\left(X^{(1)}\right) = \rho_{\mathcal{G}}\left(X^{(2)}\right)$, so $\eta_{\mathcal{G}}$ in (2.3.7) is indeed well-defined.

Next we will show that $\eta_{\mathcal{G}}$ is a CBRM. For this purpose, let in the following $F, G \in \mathcal{X}$ and $X, Y \in L^\infty_d(\mathcal{F})$ be such that $\Lambda_{\mathcal{G}}(X) = F$, $\Lambda_{\mathcal{G}}(Y) = G$.

Antitonicity: Assume $F \geq G$. Then, by (2.3.6) for almost every $\omega \in \Omega$

$$-\widetilde{\rho}_{\mathcal{G}}(X(\omega),\omega) = \widetilde{\Lambda}_{\mathcal{G}}(X(\omega),\omega) = F(\omega) \geq G(\omega) = \widetilde{\Lambda}_{\mathcal{G}}(Y(\omega),\omega) = -\widetilde{\rho}_{\mathcal{G}}(Y(\omega),\omega).$$

Hence, risk-antitonicity ensures that $\rho_{\mathcal{G}}(X) \leq \rho_{\mathcal{G}}(Y)$. But by (2.3.7) this is equivalent to $\eta_{\mathcal{G}}(F) \leq \eta_{\mathcal{G}}(G)$.

Constancy on $\widetilde{\Lambda}_{\mathcal{G}}$: Constancy on $\widetilde{\Lambda}_{\mathcal{G}}$ is an immediate consequence of (2.3.6)-(2.3.8), since for $x \in \mathbb{R}^d$

$$\eta_{\mathcal{G}}(\Lambda_{\mathcal{G}}(x)) = \rho_{\mathcal{G}}(x) = -\Lambda_{\mathcal{G}}(x).$$

Hence, the decomposition (2.2.9) is proved.

Uniqueness: Let $\eta^{(1)}_{\mathcal{G}}, \eta^{(2)}_{\mathcal{G}}$ be CBRM's and $\widetilde{\Lambda}^{(1)}_{\mathcal{G}}, \widetilde{\Lambda}^{(2)}_{\mathcal{G}}$ be CAF's such that $\eta^{(1)}_{\mathcal{G}}$ and $\eta^{(2)}_{\mathcal{G}}$ are constant on $\widetilde{\Lambda}^{(1)}_{\mathcal{G}}$ and $\widetilde{\Lambda}^{(2)}_{\mathcal{G}}$ resp. and it holds that

$$\eta^{(1)}_{\mathcal{G}}\left(\Lambda^{(1)}_{\mathcal{G}}(X)\right) = \rho_{\mathcal{G}}(X) = \eta^{(2)}_{\mathcal{G}}\left(\Lambda^{(2)}_{\mathcal{G}}(X)\right) \quad \text{for all } X \in L^\infty_d(\mathcal{F}).$$

Then it follows from the constancy on the respective CAF's that for all $x \in \mathbb{R}^d$ $\Lambda_{\mathcal{G}}^{(1)}(x) = \Lambda_{\mathcal{G}}^{(2)}(x)$, i.e.

$$\widetilde{\Lambda}_{\mathcal{G}}^{(1)}(x,\omega) = \widetilde{\Lambda}_{\mathcal{G}}^{(2)}(x,\omega) \quad a.s. \tag{2.3.9}$$

Note that by a similar argumentation as in the proof of Lemma 2.3.2 (2.3.9) holds true on a universal $\mathbb{P}$-nullset N for all $x \in \mathbb{R}^d$. In order to show that $\Lambda_{\mathcal{G}}^{(1)}$ and $\Lambda_{\mathcal{G}}^{(2)}$ are not only equal on constants let $X \in L_d^\infty(\mathcal{F})$. Then X can be approximated by simple $\mathcal{F}$-measurable random vectors, i.e. there exists a sequence $(X_n)_{n\in\mathbb{N}}$ with $X_n \to X$ $\mathbb{P}$-a.s. and $X_n = \sum_{i=1}^{k_n} x_i^n \mathbb{1}_{A_i^n}$ for all $n \in \mathbb{N}$, where $x_i^n \in \mathbb{R}$ and $A_i^n \in \mathcal{F}, i = 1, ..., k_n$ are disjoint sets such that $\mathbb{P}(A_i^n) > 0$ and $\mathbb{P}\big(\bigcup_{i=1}^{k_n} A_i^n\big) = 1$. Denote by M the $\mathbb{P}$-nullset on which $(X_n)_{n\in\mathbb{N}}$ does not converge. Then by the continuity property of a CAF and (2.3.9) we have for all $\omega \in (N \cup M)^C$ that

$$\begin{aligned}
\widetilde{\Lambda}_{\mathcal{G}}^{(1)}(X(\omega),\omega) = \widetilde{\Lambda}_{\mathcal{G}}^{(1)}\left(\lim_{n\to\infty} X_n(\omega),\omega\right) \quad &= \quad \lim_{n\to\infty} \widetilde{\Lambda}_{\mathcal{G}}^{(1)}(X_n(\omega),\omega) \\
= \lim_{n\to\infty} \sum_{i=1}^{k_n} \widetilde{\Lambda}_{\mathcal{G}}^{(1)}(x_i^n,\omega)\, \mathbb{1}_{A_i^n}(\omega) \quad &= \quad \lim_{n\to\infty} \sum_{i=1}^{k_n} \widetilde{\Lambda}_{\mathcal{G}}^{(2)}(x_i^n,\omega)\, \mathbb{1}_{A_i^n}(\omega) \\
= \widetilde{\Lambda}_{\mathcal{G}}^{(2)}(X(\omega),\omega)\,,
\end{aligned}$$

and thus $\Lambda_{\mathcal{G}}^{(1)}(X) = \Lambda_{\mathcal{G}}^{(2)}(X)$ for all $X \in L_d^\infty(\mathcal{F})$. Finally for all $F \in \operatorname{Im}\Lambda_{\mathcal{G}}^{(1)} = \operatorname{Im}\Lambda_{\mathcal{G}}^{(2)}$ there is an $X \in L_d^\infty(\mathcal{F})$ such that $\Lambda_{\mathcal{G}}^{(1)}(X) = \Lambda_{\mathcal{G}}^{(2)}(X) = F$ and hence

$$\eta_{\mathcal{G}}^{(1)}(F) = \rho_{\mathcal{G}}(X) = \eta_{\mathcal{G}}^{(2)}(F).$$

Next we consider the cases when $\rho_{\mathcal{G}}$ fulfills some additional properties.
Constant on constants: Let $\rho_{\mathcal{G}}$ be risk-regular. Then (2.3.6) implies that for all $X \in L_d^\infty(\mathcal{G})$

$$\rho_{\mathcal{G}}(X,\omega) = \widetilde{\rho}_{\mathcal{G}}(X(\omega),\omega) = -\widetilde{\Lambda}_{\mathcal{G}}(X(\omega),\omega) \text{ a.s.}$$

and hence $\rho_{\mathcal{G}}(X) = -\Lambda_{\mathcal{G}}(X)$. Let now $F \in \mathcal{X} \cap L^\infty(\mathcal{G})$. By the definition of $\mathcal{X}$ and Lemma 2.3.1 we know that there exists a $X \in L_d^\infty(\mathcal{G})$ such that $\Lambda_{\mathcal{G}}(X) = F$. We thus obtain by (2.3.7) that

$$\eta_{\mathcal{G}}(F) = \rho_{\mathcal{G}}(X) = -\Lambda_{\mathcal{G}}(X) = -F.$$

Quasiconvexity/convexity: Let $\rho_{\mathcal{G}}$ be risk-quasiconvex. Let $\alpha \in L^\infty(\mathcal{G})$ with $0 \le \alpha \le 1$ and set $H := \alpha F + (1-\alpha)G$, where $F, G \in \mathcal{X}$, and $X, Y \in L_d^\infty(\mathcal{F})$ are such that $\Lambda_{\mathcal{G}}(X) = F$, $\Lambda_{\mathcal{G}}(Y) = G$. Note that since $\mathcal{X}$ is $\mathcal{G}$-conditionally convex, $H \in \mathcal{X}$ and thus there exists a $Z \in L_d^\infty(\mathcal{F})$ with $\Lambda_{\mathcal{G}}(Z) = H$. Then

$$\begin{aligned}
\widetilde{\Lambda}_{\mathcal{G}}(Z(\omega),\omega) = H(\omega) &= \alpha(\omega)F(\omega) + (1-\alpha(\omega))G(\omega) \\
&= \alpha(\omega)\widetilde{\Lambda}_{\mathcal{G}}(X(\omega),\omega) + (1-\alpha(\omega))\widetilde{\Lambda}_{\mathcal{G}}(Y(\omega),\omega) \text{ a.s.}
\end{aligned}$$

Thus it follows by (2.3.6)

$$\widetilde{\rho}_{\mathcal{G}}(Z(\omega),\omega) = \alpha(\omega)\widetilde{\rho}_{\mathcal{G}}(X(\omega),\omega) + (1-\alpha(\omega))\widetilde{\rho}_{\mathcal{G}}(Y(\omega),\omega) \text{ a.s.,}$$

which in conjunction with the risk-quasiconvexity of $\rho_{\mathcal{G}}$ results in

$$\eta_{\mathcal{G}}(H) = \rho_{\mathcal{G}}(Z) \leq \rho_{\mathcal{G}}(X) \vee \rho_{\mathcal{G}}(Y) = \eta_{\mathcal{G}}(F) \vee \eta_{\mathcal{G}}(G).$$

Similarly one shows that $\eta_{\mathcal{G}}$ is convex if $\rho_{\mathcal{G}}$ is risk-convex.
Positive homogeneity: Let $\rho_{\mathcal{G}}$ be risk-positively homogeneous. Further let $F \in \mathcal{X}$, $X \in L^\infty_d(\mathcal{F})$ with $\Lambda_{\mathcal{G}}(X) = F$, and let $\alpha \in L^\infty(\mathcal{G})$ with $\alpha \geq 0$ and $\alpha F =: G \in \mathcal{X}$. Then there is also a $Y \in L^\infty_d(\mathcal{F})$ with $\Lambda_{\mathcal{G}}(Y) = G$. Moreover, $\widetilde{\Lambda}_{\mathcal{G}}(Y(\omega), \omega) = \alpha(\omega)\widetilde{\Lambda}_{\mathcal{G}}(X(\omega), \omega)$ a.s. Hence, by (2.3.6) in conjunction with the risk-positive homogeneity we obtain that $\rho_{\mathcal{G}}(Y) = \alpha\rho_{\mathcal{G}}(X)$. Consequently,

$$\eta_{\mathcal{G}}(\alpha F) = \rho_{\mathcal{G}}(Y) = \alpha\rho_{\mathcal{G}}(X) = \alpha\eta_{\mathcal{G}}(F).$$

□

Proof of Theorem 2.2.11. As $\rho_{\mathcal{G}}$ is risk-antitone and antitone on constants, it is obvious that $\rho_{\mathcal{G}}$ also fulfills (2.2.10). Furthermore, we already showed, based on the antitonicity on constants and continuous paths requirements, in the proof of Theorem 2.2.9 that $\omega \mapsto \widetilde{\rho}_{\mathcal{G}}(x, \omega)$ $(= -\widehat{\Lambda}_{\mathcal{G}}(x, \omega))$ has almost surely antitone paths. Hence, we have for all $X, Y \in L^\infty_d(\mathcal{F})$ with $X \geq Y$, that $\widetilde{\rho}_{\mathcal{G}}(X(\omega), \omega) \leq \widetilde{\rho}_{\mathcal{G}}(Y(\omega), \omega)$ a.s. and thus the risk-antitonicity yields $\rho_{\mathcal{G}}(X) \leq \rho_{\mathcal{G}}(Y)$. Hence we conclude that $\rho_{\mathcal{G}}$ is antitone.
For the converse implication let $\rho_{\mathcal{G}}$ be antitone and let $\rho_{\mathcal{G}}(\cdot, \cdot)$ be a realization with corresponding restriction $\widetilde{\rho}_{\mathcal{G}}(\cdot, \cdot)$ which fulfills (2.2.10). The antitonicity on constants is an immediate consequence of the much stronger antitonicity on L^∞ of $\rho_{\mathcal{G}}$. By reconsidering the proof of Theorem 2.2.9, we observe that we may replace the risk-antitonicity by (2.2.10) when extracting the aggregation function. Hence, (2.2.10) is sufficient to construct a modification of $\rho_{\mathcal{G}}(\cdot, \cdot)$ and thus of $\widetilde{\rho}_{\mathcal{G}}(\cdot, \cdot)$ such that $\widetilde{\rho}_{\mathcal{G}}(\cdot, \cdot)$ has surely continuous and antitone paths. Therefore, suppose that $\rho_{\mathcal{G}}(\cdot, \cdot)$ is already this realization. Now let $X, Y \in L^\infty_d(\mathcal{F})$ with $\widetilde{\rho}_{\mathcal{G}}(X(\omega), \omega) \leq \widetilde{\rho}_{\mathcal{G}}(Y(\omega), \omega)$ a.s. According to Lemma 2.3.1 with $\widetilde{\Lambda}_{\mathcal{G}}$ as in (2.3.6) there are $F, G \in L^\infty(\mathcal{F})$ such that

$$\widetilde{\rho}_{\mathcal{G}}(F(\omega)\mathbf{1}_d, \omega) = \widetilde{\rho}_{\mathcal{G}}(X(\omega), \omega) \leq \widetilde{\rho}_{\mathcal{G}}(Y(\omega), \omega) = \widetilde{\rho}_{\mathcal{G}}(G(\omega)\mathbf{1}_d, \omega) \text{ a.s.}$$

As the paths of $\widetilde{\rho}_{\mathcal{G}}(\cdot, \cdot)$ are antitone, it can be readily seen that $F \geq G$ on $A := \{\omega \in \Omega : \widetilde{\rho}_{\mathcal{G}}(X(\omega), \omega) < \widetilde{\rho}_{\mathcal{G}}(Y(\omega), \omega)\}$. Now set $H := G\mathbb{1}_A + F\mathbb{1}_{A^C} \in L^\infty(\mathcal{F})$. Then $F \geq H$ and $\widetilde{\rho}_{\mathcal{G}}(Y(\omega), \omega) = \widetilde{\rho}_{\mathcal{G}}(H(\omega)\mathbf{1}_d, \omega)$ a.s. Hence it follows from (2.2.10) and the antitonicity of $\rho_{\mathcal{G}}$ that

$$\rho_{\mathcal{G}}(X) = \rho_{\mathcal{G}}(F\mathbf{1}_d) \leq \rho_{\mathcal{G}}(H\mathbf{1}_d) = \rho_{\mathcal{G}}(Y).$$

This completes the proof of the first equivalence in Theorem 2.2.11.

Let $\rho_{\mathcal{G}}$ be risk-positive homogeneous and positive homogeneous on constants. Since all requirements of Theorem 2.2.9 are met, we also have that $\mathbb{R}^d \ni x \mapsto \widetilde{\rho}_{\mathcal{G}}(x, \omega)$ is almost surely positive homogeneous. Therefore we obtain for all $X \in L^\infty_d(\mathcal{F})$ and $\alpha \in L^\infty(\mathcal{G})$ with $\alpha \geq 0$ that

$$\widetilde{\rho}_{\mathcal{G}}(\alpha(\omega)X(\omega), \omega) = \alpha(\omega)\widetilde{\rho}_{\mathcal{G}}(X(\omega), \omega) \text{ a.s.,}$$

and hence the risk-positive homogeneity implies $\rho_{\mathcal{G}}(\alpha X) = \alpha\rho_{\mathcal{G}}(X)$ which is positive homogeneity of $\rho_{\mathcal{G}}$.

Conversely, if $\rho_{\mathcal{G}}$ is positive homogeneous, then it is also positive homogeneous on constants as well as for almost all paths of the realization. Hence, if there exists $X, Z \in L_d^\infty(\mathcal{F})$ and $\alpha \in L^\infty(\mathcal{G})$ with $\alpha \geq 0$ such that

$$\widetilde{\rho}_{\mathcal{G}}\left(Z(\omega), \omega\right) = \alpha(\omega)\widetilde{\rho}_{\mathcal{G}}\left(X(\omega), \omega\right) \text{ a.s.,}$$

then the right-hand-side equals $\widetilde{\rho}_{\mathcal{G}}\left(\alpha(\omega)X(\omega), \omega\right)$ a.s. Using (2.2.10) and the positive homogeneity of $\rho_{\mathcal{G}}$ we conclude that

$$\rho_{\mathcal{G}}(Z) = \rho_{\mathcal{G}}(\alpha X) = \alpha\rho_{\mathcal{G}}(X).$$

Let $\rho_{\mathcal{G}}$ be risk-convex and convex on constants. First we will show that risk-convexity is equivalent to the following property: If for $X, Y, Z \in L_d^\infty(\mathcal{F})$ there exists a $\alpha \in L^\infty(\mathcal{G})$ with $0 \leq \alpha \leq 1$ such that

$$\widetilde{\rho}_{\mathcal{G}}\left(Z(\omega), \omega\right) \leq \alpha(\omega)\widetilde{\rho}_{\mathcal{G}}\left(X(\omega), \omega\right) + \big(1-\alpha(\omega)\big)\widetilde{\rho}_{\mathcal{G}}\left(Y(\omega), \omega\right) \text{ a.s.,} \qquad (2.3.10)$$
$$\text{then } \rho_{\mathcal{G}}(Z) \leq \alpha\rho_{\mathcal{G}}(X) + \big(1-\alpha\big)\rho_{\mathcal{G}}(Y).$$

On the one hand, it is obvious that (2.3.10) implies risk-convexity. On the other hand, let $Z^{(1)} \in L_d^\infty(\mathcal{F})$ such that

$$\widetilde{\rho}_{\mathcal{G}}\left(Z^{(1)}(\omega), \omega\right) \leq \alpha(\omega)\widetilde{\rho}_{\mathcal{G}}\left(X(\omega), \omega\right) + \big(1-\alpha(\omega)\big)\widetilde{\rho}_{\mathcal{G}}\left(Y(\omega), \omega\right) \text{ a.s.}$$

We know by Lemma 2.3.1 that there is a $Z^{(2)} \in L_d^\infty(\mathcal{F})$ such that

$$\alpha(\omega)\widetilde{\rho}_{\mathcal{G}}\left(X(\omega), \omega\right) + \big(1-\alpha(\omega)\big)\widetilde{\rho}_{\mathcal{G}}\left(Y(\omega), \omega\right) = \widetilde{\rho}_{\mathcal{G}}\left(Z^{(2)}(\omega), \omega\right) \text{ a.s.}$$

By the risk-convexity we obtain that

$$\rho_{\mathcal{G}}(Z^{(2)}) \leq \alpha\rho_{\mathcal{G}}(X) + (1-\alpha)\rho_{\mathcal{G}}(Y).$$

As risk-antitonicity implies $\rho_{\mathcal{G}}(Z^{(1)}) \leq \rho_{\mathcal{G}}(Z^{(2)})$, we conclude that risk-convexity and (2.3.10) are equivalent. Next we show the convexity of $\rho_{\mathcal{G}}$. To this end let $X, Y \in L_d^\infty(\mathcal{F})$ and $\alpha \in L^\infty(\mathcal{G})$ with $0 \leq \alpha \leq 1$. Once again we can reason as in the proof of Theorem 2.2.9 that $\mathbb{R}^d \ni x \mapsto \widetilde{\rho}_{\mathcal{G}}\left(x, \omega\right)$ is almost surely convex, because $\rho_{\mathcal{G}}$ has continuous paths and is convex on constants. Thus we have that

$$\widetilde{\rho}_{\mathcal{G}}\left((\alpha X + (1-\alpha)Y)(\omega), \omega\right) \leq \alpha(\omega)\widetilde{\rho}_{\mathcal{G}}\left(X(\omega), \omega\right) + \big(1-\alpha(\omega)\big)\widetilde{\rho}_{\mathcal{G}}\left(Y(\omega), \omega\right) \text{ a.s.}$$

Now (2.3.10) implies that

$$\rho_{\mathcal{G}}(\alpha X + (1-\alpha)Y) \leq \alpha\rho_{\mathcal{G}}(X) + (1-\alpha)\rho_{\mathcal{G}}(Y),$$

which is the desired convexity of $\rho_{\mathcal{G}}$.

The other assertion concerning risk-quasiconvexity follows in a similar way. □

2.4 Examples

Example 2.4.1. As already mentioned in the introduction a typical aggregation function when dealing with multidimensional risks is

$$\widetilde{\Lambda}_{\text{sum}}(x) = \sum_{i=1}^{d} x_i,\ x \in \mathbb{R}^d.$$

However, such an aggregation rule might not always be reasonable when measuring systemic risk. The main reason for this is the limited transferability of profits and losses between institutions of a financial system. An alternative popular aggregation function which does not allow for a subsidization of losses by other profitable institutions is given by

$$\widetilde{\Lambda}_{\text{loss}}(x) = \sum_{i=1}^{d} -x_i^-,\ x \in \mathbb{R}^d,$$

where $x_i^- = -\min\{x_i, 0\}$; see Example 2.4.8. Obviously, both $\widetilde{\Lambda}_{\text{sum}}$ and $\widetilde{\Lambda}_{\text{loss}}$ are DAF's which are additionally concave and positive homogeneous.

Example 2.4.2 (Countercyclical regulation)**.** Risk charges based on systemic risk measures typically will increase drastically in a distressed market situation which might even worsen the crisis further. Therefore one might argue that, for instance in a recession where also the real economy is affected, the financial regulation should be relaxed in order to stabilize the real economy, cf. Brunnermeier and Cheridito (2014). In our setup we can incorporate such a dynamic countercyclical regulation as follows:

Let $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in\{0,\dots,T\}}, \mathbb{P})$ be a filtered probability space, where $\mathcal{F}_T = \mathcal{F}$. Let $(x, y) \in \mathbb{R}^{2d}$ be the profits/losses of the financial system, where the first d components x are the profits/losses from contractual obligations with the real economy and y are the profits/losses from other obligations. Moreover let $Y(t)$, $t = -1, 0, ..., T-1$ be the gross domestic product (GDP) process with $Y(t) \in \mathscr{L}^\infty(\mathcal{F}_t)$, $t = 0, ..., T-1$, and $Y(-1) \in \mathbb{R}^+\backslash\{0\}$. Suppose that the regulator sees the economy in distress at time t, if the GDP process $Y(t)$ is less than $(1+\theta)Y(t-1)$ for some $\theta \in \mathbb{R}$. We assume that in those scenarios the regulator is interested to lower the regulation in order to give incentives to the financial system for the supply of additional credit to the real economy. This policy might lead to the following dynamic conditional aggregation function from the perspective of the regulator

$$\widetilde{\Lambda}\big((x,y), t, \omega\big) := -\sum_{i=1}^{d} \Big(\alpha \mathbb{1}_{A(t)}(\omega) + \mathbb{1}_{A(t)^C}(\omega)\Big)\, x_i^- + y_i^-, t = 0, ..., T-1,$$

where $\alpha \in [0,1)$ and $A(t) = \{Y(t) \leq (1+\theta)Y(t-1)\}$ for $t = 0, ..., T-1$. Obviously, $\widetilde{\Lambda}\big((x,y), t, \omega\big)$ is a CAF with respect to $\mathcal{F}_t$ which is positive homogeneous and concave.

Example 2.4.3 (Too big to fail)**.** In this example we will consider a dynamic conditional aggregation function which depends on the relative size of the interbank liabilities. For instance, Cont et al. (2013) find that for the Brasilian banking network there is a strong connection between the size of the interbank liabilities of a financial institution and its systemic importance. This fact is often quoted as 'too big to fail'.

Let $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in\{0,\dots,T\}}, \mathbb{P})$ be a filtered probability space, where $\mathcal{F}_T = \mathcal{F}$. Moreover, let $L_i(t) \in \mathscr{L}^\infty(\mathcal{F}_t)$ denote the sum of all liabilities at time t of institution $i \in \{1, \dots, d\}$ to any other banks. Then

$$\alpha_i(t) := \frac{L_i(t)}{\sum_{j=1}^d L_j(t)}, \quad t = 0, ..., T-1,$$

is the relative size of its interbank liabilities. Now consider the following conditional extension of an aggregation function which was proposed in Brunnermeier and Cheridito (2014):

$$\widetilde{\Lambda}_{\text{BC}}(x, t, \omega) = \sum_{i=1}^d -\alpha_i(t, \omega) x_i^- + \beta_i(\theta_i - x_i)^-, \quad t = 0, ..., T-1,$$

where $\beta, \theta \geq 0$. Firstly, this conditional aggregation function always takes losses into consideration, whereas profits of a financial institution i are only accounted for if they are above a firm specific threshold θ_i. Secondly, profits are weighted by the deterministic factor β and the losses are weighted proportional to the liability size of the corresponding financial institution at time t. Therefore losses from large institutions, which are more likely to be systemically relevant, contribute more to the total risk.
$\widetilde{\Lambda}_{\text{BC}}(\cdot, t, \cdot)$ is a CAF which, however, in general is neither quasiconvex nor positively homogeneous as it may be partly flat depending on θ.

Example 2.4.4. Suppose that the regulator of the financial system has certain preferences on the distribution of the total loss amongst the financial institutions. For instance he might prefer a situation when a number of financial institutions face a relatively small loss each in front of a situation in which one financial institution experiences a relatively large loss. Such a preference can be incorporated by the following aggregation function

$$\widetilde{\Lambda}_{\text{exp}}(x) = \sum_{i=1}^d -x_i^- \mathbb{1}_{\{x_i > \theta_i\}} + \left(\frac{1}{\gamma_i} \left(1 - e^{\gamma_i (x_i^- + \theta_i)} \right) + \theta_i \right) \mathbb{1}_{\{x_i \leq \theta_i\}},$$

where $\theta_i \leq 0$ and $\gamma_i > 0$ for $i = 1, ..., d$. That is, if the losses of firm i exceed a certain threshold θ_i, e.g. a certain percentage of the equity value, then the losses are accounted for exponentially.

Example 2.4.5 (Stochastic discount)**.** Suppose that $D \in \mathscr{L}^\infty(\mathcal{F})$ is some $\mathcal{G}$-measurable stochastic discount factor. A typical approach to define monetary risk measurement of some future risk is to consider the discounted risks. Consider any (conditional) aggregation function $\widetilde{\Lambda}$, which does not discount in aggregation, such as $\widetilde{\Lambda}_{\text{sum}}$, $\widetilde{\Lambda}_{\text{loss}}$, or $\widetilde{\Lambda}_{\text{BC}}$, etc. Then the discounted monetary aggregated risk is $D\widetilde{\Lambda}(X)$. If $\widetilde{\Lambda}$ is positively homogeneous, then $D\widetilde{\Lambda}(X) = \widetilde{\Lambda}(DX)$ which is the aggregated risk of the discounted system DX. However, if $\widetilde{\Lambda}$ is not positively homogeneous - such as $\widetilde{\Lambda}_{\text{BC}}$ or $\widetilde{\Lambda}_{\text{exp}}$ - then the discounted aggregated risk can only be formulated in terms of the conditional aggregation function

$$\widetilde{\Lambda}_{\mathcal{G}}(x, \omega) := \widetilde{\Lambda}(x) D(\omega).$$

Example 2.4.6 (CoVaR)**.** In this example we will consider the CoVaR proposed in Adrian and Brunnermeier (2016); see (2.1.2). To this end, we first recall the (conditional) Value at Risk: We denote the Value at Risk at level $q \in (0,1)$ by

$$\mathrm{VaR}_q(F) = -\inf_{x\in\mathbb{R}}\{\mathbb{P}(F \le x) > q\}.$$

Furthermore, the conditional VaR at level $q \in (0,1)$ is defined as

$$\mathrm{VaR}_q(F|\mathcal{G}) := -\operatorname*{essinf}_{\alpha\in L^\infty(\mathcal{G})}\left\{\mathbb{P}\big(F \le \alpha \mid \mathcal{G}\big) > q\right\},$$

c.f. Föllmer and Schied (2011). The conditional VaR is positive homogeneous, antitone, and constant on constants. Thus it is a CBRM which is constant on every possible CAF. Note that, as is well-known for the unconditional case, the conditional VaR is not quasiconvex. By composing $\mathrm{VaR}_q(\cdot|\mathcal{G})$ with a CAF $\widetilde{\Lambda}_\mathcal{G}$ we obtain a CSRM

$$\rho_\mathcal{G}(X) = \mathrm{VaR}_q\left(\Lambda_\mathcal{G}(X)|\,\mathcal{G}\right), \quad X \in L_d^\infty(\mathcal{F}), \tag{2.4.1}$$

which is risk-positive homogeneous and risk-regular.
Now we consider the case where X represents a financial system and the CAF in (2.4.1) is $\widetilde{\Lambda}_{\text{sum}}$. Moreover consider the sub-σ-algebra $\mathcal{G} := \sigma(A)$ of $\mathcal{F}$, where $A := \{X_j \le -\mathrm{VaR}_q(X_j)\}$ for a fixed $j \in \{1,...,d\}$. Then the CSRM $\rho_\mathcal{G}(X)$ from (2.4.1) evaluated in the event A equals

$$\mathrm{VaR}_q\left(\sum_{i=1}^d X_i\,\middle|\, \{X_j \le -\mathrm{VaR}_q(X_j)\}\right) \tag{2.4.2}$$

which is the CoVaR proposed in Adrian and Brunnermeier (2016).
As we have already pointed out in the introduction, it is more reasonable to use an aggregation function which incorporates an explicit contagion structure. We will modify the CoVaR in this direction in Example 2.4.9.

Example 2.4.7 (CoES and SES)**.** The conditional Average Value at Risk at level $q \in (0,1)$ is given by

$$\mathrm{AVaR}_q(F|\mathcal{G}) := \operatorname*{esssup}_{\mathbb{Q}\in\mathcal{P}^q}\mathbb{E}_\mathbb{Q}\left[-F \mid \mathcal{G}\right], \quad F \in L^\infty(\mathcal{F}),$$

where $\mathcal{P}^q$ is the set of probability measures $\mathbb{Q}$ on $(\Omega,\mathcal{F})$ which are absolutely continuous w.r.t. $\mathbb{P}$ such that $\mathbb{Q}|_\mathcal{G} = \mathbb{P}$ and $\frac{d\mathbb{Q}}{d\mathbb{P}} \le 1/q$ a.s. $\mathrm{AVaR}_q(\cdot|\mathcal{G})$ is a convex and positive homogeneous CBRM. Notice that the conditional Average Value at Risk can also be written as

$$\mathrm{AVaR}_q(F|\mathcal{G}) = \frac{1}{q}\mathbb{E}_\mathbb{P}\left[(F + \mathrm{VaR}_q(F|\mathcal{G}))^- \;\middle|\; \mathcal{G}\right] + \mathrm{VaR}_q(F|\mathcal{G}), \tag{2.4.3}$$

cf. Föllmer and Schied (2011), where $\mathrm{VaR}_q(\cdot|\mathcal{G})$ is discussed in Example 2.4.6.
As in Example 2.4.6 let $\mathcal{G} = \sigma(A)$ with $A = \{X_j \le -\mathrm{VaR}_q(X_j)\}$ for a fixed $j \in \{1,...,d\}$ and $q \in (0,1)$. Using (2.4.3), if

$$\mathbb{P}(F \le -\mathrm{VaR}_q(F|\mathcal{G})|\,\mathcal{G}) = q,$$

then

$$\begin{aligned}\mathrm{AVaR}_q(F|\mathcal{G}) = \mathbb{E}_{\mathbb{P}}\left[-F \mid \{F \leq -\mathrm{VaR}_q(F|A)\} \cap A\right] \mathbb{1}_A \\ + \mathbb{E}_{\mathbb{P}}\left[-F \mid \{F \leq -\mathrm{VaR}_q(F|A^C)\} \cap A^C\right] \mathbb{1}_{A^C}.\end{aligned} \tag{2.4.4}$$

Therefore, $\rho_{\mathcal{G}}(X) = \mathrm{AVaR}_q(\widetilde{\Lambda}_{\text{sum}}(X)|\mathcal{G})$ evaluated in the event A equals

$$\mathbb{E}_{\mathbb{P}}\left[-\sum_{i=1}^d X_i \,\middle|\, \left\{\sum_{i=1}^d X_i \leq -\mathrm{VaR}_q\left(\sum_{i=1}^d X_i \,\middle|\, A\right)\right\} \cap A\right].$$

In other words, $\rho_{\mathcal{G}}(X)|_A$ is the expected loss of the financial system X given that the loss X_j of institution j is below $\mathrm{VaR}_q(X_j)$ and simultaneously the loss of the system is below its CoVaR $\mathrm{VaR}_q(\sum_{i=1}^d X_i|A)$. (2.4.4) corresponds to the conditional expected shortfall (CoES) proposed in Adrian and Brunnermeier (2016).

Now we change the point of view and consider the losses of a financial institution X_j given that the financial system is in distress, that is if

$$\sum_{i=1}^d X_i \leq -\mathrm{VaR}_q\left(\sum_{i=1}^d X_i\right).$$

Let $\mathcal{G} := \sigma\big(\{\sum_{i=1}^d X_i \leq -\mathrm{VaR}_q(\sum_{i=1}^d X_i)\}\big)$. By composing the DAF $\widetilde{\Lambda}(x) := x_j$ and the CBRM $\eta_{\mathcal{G}}(F) = \mathbb{E}_{\mathbb{P}}\left[-F \mid \mathcal{G}\right]$ we obtain a convex and positive homogeneous CSRM

$$\rho_{\mathcal{G}}(Y) = \mathbb{E}_{\mathbb{P}}\left[Y_j \mid \mathcal{G}\right], \quad Y \in L_d^\infty(\mathcal{F}).$$

$\rho_{\mathcal{G}}(X)$ evaluated on the event $\{\sum_{i=1}^d X_i \leq -\mathrm{VaR}_q(\sum_{i=1}^d X_i)\}$ which is the so-called systemic expected shortfall (SES^j) introduced in Acharya et al. (2017).

Example 2.4.8 (DIP)**.** In this example we recall the distress insurance premium (DIP) proposed by Huang et al. (2012). It is closely related to CoES and SES discussed in Example 2.4.7. However, instead of $\widetilde{\Lambda}_{\text{sum}}$, the aggregation function is $\widetilde{\Lambda}_{\text{loss}}$, that is losses cannot be subsidized by profits from the other institutions. The event representing the financial system in distress is $\{\Lambda_{\text{loss}}(X) \leq \theta\}$ for a fixed $\theta \in \mathbb{R}$, i.e. the financial system is in distress if the total losses fall below a certain threshold θ. Let $\mathcal{G} := \sigma\left(\{\Lambda_{\text{loss}}(X) \leq \theta\}\right)$. As a CBRM choose $\eta_{\mathcal{G}}(F) = \mathbb{E}_{\mathbb{Q}}\left[-F \mid \mathcal{G}\right]$, where $\mathbb{Q}$ is a risk neutral measure which is equivalent to $\mathbb{P}$. The resulting positive homogeneous and convex CSRM evaluated in $\{\Lambda_{\text{loss}}(X) \leq \theta\}$ is given by

$$\mathbb{E}_{\mathbb{Q}}\left[\sum_{i=1}^d Y_i^- \,\middle|\, \Lambda_{\text{loss}}(X) \leq \theta\right], \quad Y \in L_d^\infty(\mathcal{F}),$$

which corresponds to the DIP for $Y = X$. Since the expectation is under a risk neutral measure it can be interpreted as the premium of an aggregate excess loss reinsurance contract.

Example 2.4.9 (Contagion model)**.** In this example we want to specify an aggregation function that explicitly models the default mechanisms in a financial system and perform a small simulation study. For this purpose we will assume the simplified balance sheet structure given in Table 2.4.1 for each of the d financial institutions. Let $X \in L_d^\infty(\mathcal{F})$ be the vector of equity values of the financial institutions after some market shock on the external assets/liabilities. Moreover let $\mathbf{\Pi}$ be the relative liability matrix of size $d \times d$, i.e. the i,jth entry represents the proportion of the total interbank liabilities of institution i which it owes to institution j. We denote the d-dimensional vector of the total interbank liabilities by L.

<table>
<tr><th>Assets</th><th>Liabilities</th></tr>
<tr><td rowspan="2">External Assets</td><td>Equity</td></tr>
<tr><td rowspan="2">External Liabilities</td></tr>
<tr><td rowspan="2">Interbank Assets</td></tr>
<tr><td>Interbank Liabilities</td></tr>
</table>

Table 2.4.1: Stylized balance sheet.

We now consider an extension of the aggregation function proposed by Chen et al. (2013) which is based on the default model in Eisenberg and Noe (2001):
For a deterministic vector of equity values $x \in \mathbb{R}^d$ we define the DAF $\widetilde{\Lambda}_{\text{CM1}}$ by the optimization problem:

$$\widetilde{\Lambda}_{\text{CM1}}(x) := \max_{y,b \in \mathbb{R}_+^d} \quad \sum_{i=1}^{d} -\Big(x_i + b_i - (\mathbf{\Pi}^\top y)_i\Big)^- - \gamma b_i \tag{2.4.5}$$

$$\text{subject to} \quad y = \max\Big(\min\Big(\mathbf{\Pi}^\top y - x - b, L\Big), 0\Big), \tag{2.4.6}$$

where y_i is the amount by which financial institution i decreases its total liabilities to the remaining institutions and $b \in \mathbb{R}^d$ represents the option of an external participant, e.g. a lender of last resort, to inject a capital amount b_i into institution i. The cost of the injected capital of the lender of last resort is modeled by the parameter $\gamma > 1$.
There are two possible ways a financial institution can default: First it might default due to the market shock right at the beginning ($x_i < 0$). Secondly, if it still has sufficient capital endowment after the market shock, the losses from other institutions might force it into default by contagion effects ($x_i - (\mathbf{\Pi}^\top y)_i < 0$). The constraint (2.4.6) expresses that if a financial institution defaults, it can either reduce its payments to other institutions or the lender of last resort has to inject capital to cover the default losses. As opposed to the framework in Chen et al. (2013) we are able to incorporate the limited liability assumption ($y \leq L$) proposed in Eisenberg and Noe (2001). Furthermore the lender of last resort will only inject capital into a financial institution as long as the benefit from preventing further contagion exceeds the costs of the injection of the lender of last resort.
It can be readily seen that $\widetilde{\Lambda}_{\text{CM1}}$ is isotone and continuous. The aggregation function $\widetilde{\Lambda}_{\text{CM1}}$

given in (2.4.5) is deterministic. One possible extension within our framework is now to consider conditional modifications of $\widetilde{\Lambda}_{\text{CM1}}$. For example, if there exists only partial information or uncertainty about the future of the interbank liability structure then the relative liability matrix $\mathbf{\Pi}(\omega)$ and/or the total interbank liabilities $L(\omega)$ might be modeled stochastically. In this case it can be easily seen that the corresponding aggregation function is a CAF.

We will complete this example by employing the aggregation function $\widetilde{\Lambda}_{\text{CM1}}$ in a small simulation study. The simulation serves illustration purposes only and does not have the objective to represent a real world financial system. We begin with the construction of network with 10 institutions as a realization of an Erdös-Rényi graph with success probability $p = 0.35$, that is there exists a directed edge between institution i and j with a probability of 35% independent of the other connections. Furthermore we assume that the exposures between financial institutions follow a half-normal distribution. So far we have only knowledge about the size of the interbank assets/liabilities. For the remaining parts of the balance sheet (see Table 2.4.1) we assume that the value of equity is a fixed proportion of the total assets and that the external assets/liabilities are chosen such that the balance sheet balances out. The resulting financial system can be found in Figure 2.4.1.

In the following we want to investigate the impact on the financial system if the institutions are exposed to a shock on their external, i.e. non-interbank, assets and liabilities. For this purpose we add a shock to the initial equity which is normally distributed with mean zero and a standard deviation which is proportional to the financial institutions external assets/liabilities. The single shocks are positively correlated with $\rho = 0.1$.

In Table 2.4.2 we list some comparative statistics of the financial system for 30'000 shock scenarios and for different costs of the regulator. The first two rows consider the CSRM's obtained by composing the aggregation function $\widetilde{\Lambda}_{\text{CM1}}$ with the negative expectation and the VaR at level 5%, resp. Note that we also included the asymptotic case of $\gamma \to \infty$, which corresponds to the situation in which the regulator does not intervene.

γ	1.6	2.6	'∞'
$-\mathbb{E}\left[\Lambda_{\text{CM1}}(X)\right]$	70.62	88.00	109.30
$\text{VaR}_{0.05}(\Lambda_{\text{CM1}}(X))$	213.34	291.59	442.45
$\sum b_i$	23.01	10.67	0.00
$\sum x_i^-$	52.33		
Initially defaulted banks	2.57		
Defaulted banks after contagion	2.87	3.25	3.58

Table 2.4.2: Statistics of the financial system for 30'000 shock scenarios.

We observe that with an increasing γ the regulator is less willing to inject capital and thus the contagion effects increase which results in a higher risk in terms of the expectation and the Value at Risk. Moreover without a regulator on average round about one financial institution defaults due to contagion effects.

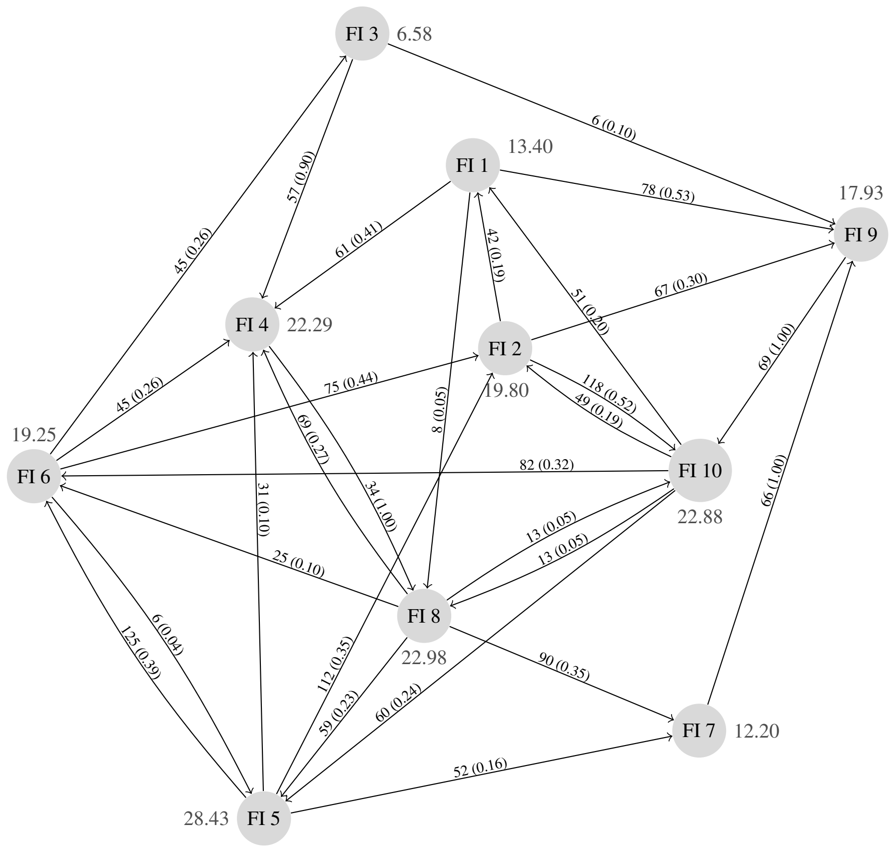


Figure 2.4.1: Exemplary financial system.

In the next step we want to investigate the systemic importance of the single institutions. For this purpose we modify the CoVaR in Example 2.4.6, that is, instead of the summing the losses we use the more realistic CAF $\widetilde{\Lambda}_{\text{CM1}}$. Thus we define for a $q \in (0,1)$:

$$\text{CoVaR}_q^j := \text{VaR}_q\left(\Lambda_{\text{CM2}}(X) \middle| \, X_j \leq -\text{VaR}_q(X_j)\right), \quad j = 1, ..., d,$$

where

$$\begin{aligned} \widetilde{\Lambda}_{\text{CM2}}(x) := \max_{y,b \in \mathbb{R}_+^d} \quad & \sum_{i=1}^{d} -y_i - \gamma b_i \\ \text{subject to} \quad & y = \max\left(\min\left(\mathbf{\Pi}^\top y - x - b, L\right), 0\right). \end{aligned}$$

The difference between $\widetilde{\Lambda}_{\text{CM2}}$ and $\widetilde{\Lambda}_{\text{CM1}}$ is that losses in case of a default are only taken into consideration up to the total interbank liabilities of this institution, i.e. only the losses which spread into the system are taken into account. For example consider an isolated institution in the system which has a huge exposure to the outside of the system, then in order to identify systemically relevant institution it is not meaningful to aggregate the losses from those exposures, nevertheless from the perspective of the total risk of the system those losses should also contribute as it was done in our prior study. As for $\widetilde{\Lambda}_{\text{CM1}}$ it can be easily seen that $\widetilde{\Lambda}_{\text{CM2}}$ is a CAF. The results for this risk-consistent systemic risk measures $\text{CoVaR}_q^j, j = 1, ..., d$ can be found in Table 2.4.3. We observe that the systemic importance is always a trade-off between the

$\gamma = 2.6$	FI j	2	3	6	4	7	1	10	9	5	8
	$\text{CoVaR}_{0.1}^j$	266.94	297.28	298.49	308.61	320.58	322.56	332.94	355.23	362.27	367.68
$\gamma = \infty$	FI j	2	4	3	7	9	6	1	10	8	5
	$\text{CoVaR}_{0.1}^j$	397.73	419.11	423.18	459.33	471.81	473.61	481.40	548.21	563.60	601.09
	FI j	2	6	10	3	1	7	5	9	8	4
	$-\text{VaR}_{0.1}(X_j)$	13.30	-7.67	-15.05	-17.01	-20.69	-22.98	-26.89	-30.48	-32.11	-33.41
	FI j	4	3	7	9	1	6	2	10	8	5
	L_j	34	63	66	69	147	171	227	255	256	320

Table 2.4.3: Systemic importance ranking based on $\text{CoVaR}_{0.1}^j$.

possibility of high downward shocks and the ability to transmit them. For instance institution 2 can transfer losses up to 227, but it is also the institution which is the least exposed to the market, which makes it also the least systemic important institution. Contrarily institution 4 is the most exposed institution, but does not have the ability to transmit those losses which also results in a low position in the systemic importance ranking. Finally institution 5 or 8 are very vulnerable to the market and have the largest total interbank liabilities and are thus identified as the most systemic institutions.

After measuring the total risk of this financial system and identifying the most systemic institutions, we want to investigate if the introduction of a central counterparty (CCP) will lower the overall risk and how the CCP should be capitalized. The CCP is a new institution in the system which clears centrally all exposures in system, i.e. it is the only counterparty for the financial institutions. Furthermore all interbank assets and liabilities with a CCP are netted. The resulting financial system can be found in Figure 2.4.2.

By construction the CCP has no initial capital endowment and thus if a debtor bank of the CCP defaults this loss is immediately transferred to the creditor banks of the CCP. Therefore the creditor banks should make an upfront payment to the CCP in terms of a percentage α_A of their assets with the central counterparty. This payment can also be interpreted as the premium for credit insurance up to the total deposits. Now the question arises, if there is an optimal contribution α_A in terms of the aggregation function $\widetilde{\Lambda}_{\text{CM1}}$. Since the creditor banks of the central counterparty have no interbank liabilities, their losses would not be considered by the CAF $\widetilde{\Lambda}_{\text{CM2}}$. This in turn implies that the optimal contribution scheme is to transfer all of the creditors' equity to the central counterparty. To study the effect of the central counterparty, we assume that the regulator is only rarely intervening. Note that the absence of the regulator in this model is already achieved by choosing $\gamma > 3$. The results for $\text{VaR}_{0.05}(\Lambda_{\text{CM1}}(X))$ and $-\mathbb{E}_{\mathbb{P}}\left[\Lambda_{\text{CM1}}(X)\right]$ for 30'000 scenarios and for different values of α_A can be found in Figure 2.4.3. We observe

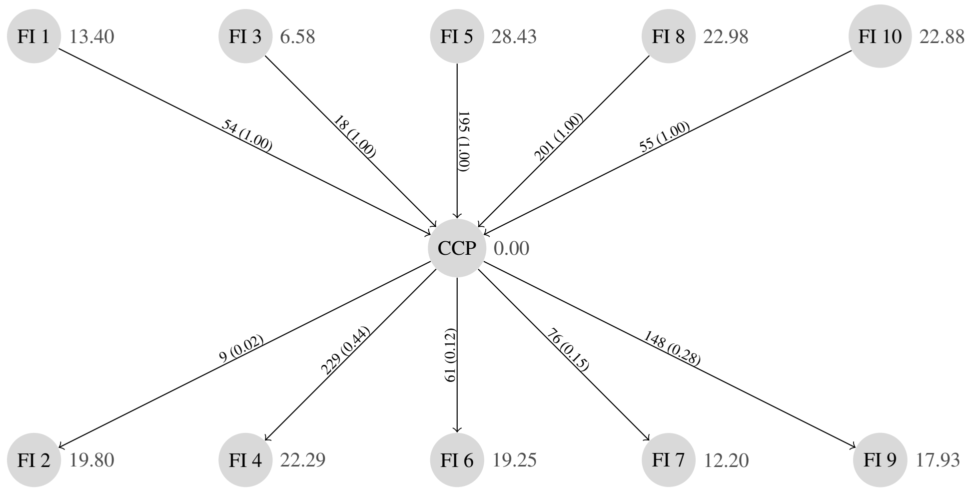
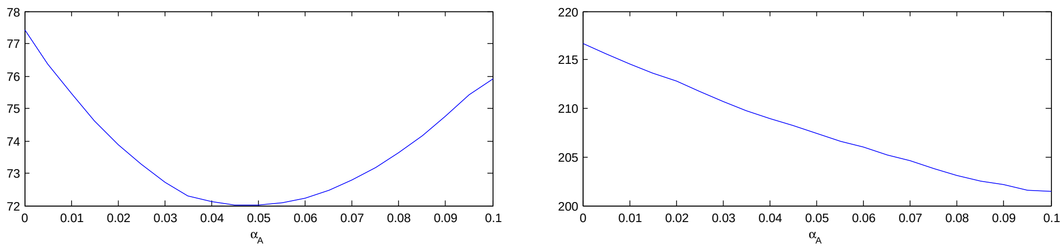


Figure 2.4.2: Financial system from Figure 2.4.1 with central counterparty



(a) $-\mathbb{E}_{\mathbb{P}}\left[\Lambda_{\text{CM1}}(X)\right]$ (b) $\text{VaR}_{0.05}(\Lambda_{\text{CM1}}(X))$

Figure 2.4.3: Statistics of Λ_{CM1} for different creditor bank contribution percentages and $\gamma = 2.9$

that there is an optimal contribution to the CCP which reduces the average risk significantly, whereas in extreme scenarios like the 5% quantile it is always optimal to transfer as much capital as possible to the CCP. The last observation is due to the fact that in those extreme scenarios the CCP is hit by a huge shock from a debtor bank, thus by transferring money to the CCP this loss will be reduced by the same amount whereas the losses on the creditor banks resulting from the transfer are comparably small.

3 Strongly Consistent Multivariate Conditional Risk Measures

3.0 Contributions of the thesis' author

This chapter is a joint work with Prof. Dr. Thilo Meyer-Brandis and Dr. Gregor Svindland. It has been submitted to the journal Mathematics and Financial Economics. A preprint is also available at *http://www.fm.mathematik.uni-muenchen.de/download/publications/consist_syst_rm.pdf*

This chapter studies strong consistency and its implication for multivariate conditional risk measures. Section 3.2 contains the definitions of the objects of interest and preliminary results on the function $f_{\rho_{\mathcal{G}}}$ which have solely been stated by H. Hoffmann. In Section 3.3 the author of the thesis showed some minor results on strong consistency and translated the main decomposition theorem of the preceding Chapter 2 in Proposition 3.3.11 to the new setup. In joint discussions Theorem 3.3.12 which connects decomposability and strongly consistency w.r.t. some terminal risk measure, was established. For Section 3.4 H. Hoffmann, T. Meyer-Brandis and G. Svindland discussed how the representation result in Föllmer (2014) can be extended to multivariate conditional risk measures, which yielded Theorem 3.4.5. Moreover, the three authors talked a lot about the presentation of the implications of this result on the decomposition of the risk measures involved. The corresponding results have mainly been proved by H. Hoffmann. Finally, in Section 3.5 the theory of strong consistency of multivariate conditional risk measures was extended to two-dimensional information structures. Here in particular H. Hoffmann proved the Proposition 3.5.9 and the ensuing findings.

3.1 Introduction

Over the recent years the study of *multivariate risk measures*

$$\rho : L_d^\infty(\mathcal{F}) \to \mathbb{R}, \tag{3.1.1}$$

that associate a risk level $\rho(X)$ to a d-dimensional vector $X = (X_1, ..., X_d)$ of random risk factors at a given future time horizon T has increasingly gained importance. Here, $L_d^\infty(\mathcal{F})$ denotes the space of d-dimensional bounded random vectors on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, i.e. we restrict the analysis to bounded risk factors X for technical simplicity.

A natural extension of the static viewpoint of deterministic risk measurement in (3.1.1) is to consider *conditional risk measures* which allow for risk measurement under varying information. A conditional multivariate risk measure is a map

$$\rho_{\mathcal{G}} : L_d^\infty(\mathcal{F}) \to L^\infty(\mathcal{G}), \tag{3.1.2}$$

that associates to a d-dimensional risk factor a $\mathcal{G}$-measurable bounded random variable, where $\mathcal{G} \subseteq \mathcal{F}$ is a sub-σ-algebra. We interpret $\rho_{\mathcal{G}}(X)$ as the risk of X given the information $\mathcal{G}$. In the present literature, conditional risk measures have mostly been studied within the framework of univariate *dynamic* risk measures, where one adjusts the risk measurement in response to the flow of information that is revealed when time elapses. For a good overview on univariate dynamic risk measures we refer the reader to Acciaio and Penner (2011) or Tutsch (2007). One possible motivation to study conditional multivariate risk measures is thus the extension from univariate to multivariate dynamic risk measures, and to study the question of what happens to the risk of a system as new information arises in the course of time. In the context of multivariate risk measures, however, also a second interesting and important dimension of conditioning arises, besides dynamic conditioning: Risk measurement conditional on information in space in order to identify systemic relevant structures. In that case $\mathcal{G}$ represents for example information on the state of a subsystem, and one is interested in questions of the type: How is the overall risk of the system affected, given that a subsystem is in distress? Or how is the risk of a single institution affected, given the entire system is in distress? In Föllmer (2014) and Föllmer and Klüppelberg (2014) the authors analyze such spatial conditioning in the context of univariate conditional risk measures, so-called spatial risk measures. Another field of application where these questions are important are the systemic risk measures, which measure the risk of a financial network. In particular the systemic risk measures CoVaR of Adrian and Brunnermeier (2016) or the systemic expected shortfall of Acharya et al. (2017) can be considered to be examples of conditional multivariate risk measures.

When dealing with families of conditional risk measures, a frequently imposed requirement is that the conditional risk measurement behaves consistent in a certain way with respect to the flow of information. In particular, in the literature on univariate dynamic risk measures most often the so-called *strong consistency* is studied; c.f. Detlefsen and Scandolo (2005), Cheridito et al. (2006), Cheridito and Kupper (2011), Kupper and Schachermayer (2009), Penner (2007). Two univariate conditional risk measures $\rho_{\mathcal{G}}$ and $\rho_{\mathcal{H}}$ with corresponding σ-algebras $\mathcal{G} \subseteq \mathcal{H} \subseteq \mathcal{F}$ are called strongly consistent if for all $X, Y \in L^\infty(\mathcal{F})$

$$\rho_{\mathcal{H}}(X) \leq \rho_{\mathcal{H}}(Y) \Longrightarrow \rho_{\mathcal{G}}(X) \leq \rho_{\mathcal{G}}(Y), \tag{3.1.3}$$

i.e. strong consistency states that if Y is riskier than X given the information $\mathcal{H}$, then this risk preference also holds under less information.

The purpose of this paper is to study the concept of strong consistency for multivariate conditional risk measures. Note that the motivation and interpretation of strong consistency in (3.1.3) remains perfectly meaningful when extending to the multivariate case. In analogy to the univariate case we thus define strong consistency of two multivariate conditional risk measures $\rho_{\mathcal{G}}$ and $\rho_{\mathcal{H}}$ with $\mathcal{G} \subseteq \mathcal{H} \subseteq \mathcal{F}$ as in (3.1.3) for any d-dimensional risk vectors X and Y in $L_d^\infty(\mathcal{F})$. As a first main result we then prove that the members of any family of strongly consistent multivariate conditional risk measures are necessarily of the following form:

$$\rho_{\mathcal{G}}(X) = \eta_{\mathcal{G}}\left(\Lambda_{\mathcal{G}}(X)\right), \tag{3.1.4}$$

where $\eta_{\mathcal{G}} : L^\infty(\mathcal{F}) \to L^\infty(\mathcal{G})$ is a univariate conditional risk measure, and $\Lambda_{\mathcal{G}} : L_d^\infty(\mathcal{F}) \to L^\infty(\mathcal{F})$ is a (conditional) aggregation function. This subclass of multivariate conditional risk measures corresponds to the idea that we first aggregate the risk factors X and then evaluate the risk of the aggregated values. In fact many prominent examples of multivariate conditional risk measures are of type (3.1.4), for instance the Contagion Index of Cont et al. (2013) or the SystRisk of Brunnermeier and Cheridito (2014) from the systemic risk literature. Chen et al. (2013) were the first to axiomatically describe this intuitive type of multivariate risk measures on a finite state space, and in Kromer et al. (2016) this has been extended to general L^p-spaces, whereas the conditional framework was studied in Hoffmann et al. (2016). We also remark that in Kromer et al. (2014) the authors study consistency of risk measures over time which can be decomposed as in (3.1.4). However, their definition of consistency differs from ours in (3.1.3) as they require consistency of the underlying univariate risk measure and the aggregation function in (3.1.4) simultaneously.

A requirement on the strongly consistent family of multivariate conditional risk measures we ask for here—which is automatically satisfied in the univariate case—is that it contains a terminal risk measure $\rho_{\mathcal{F}} : L^\infty_d(\mathcal{F}) \to L^\infty(\mathcal{F})$ under full information $\mathcal{F}$. Such a terminal risk measure is nothing but a statewise aggregation rule for the components of a risk $X \in L^\infty_d(\mathcal{F})$. In the univariate case, if $X \in L^\infty(\mathcal{F})$, there is of course no aggregation necessary. Indeed letting the terminal risk measure correspond to the identity mapping, i.e. $\rho_{\mathcal{F}} = -\operatorname{id}$, we have that any univariate risk measure $\rho_{\mathcal{G}}$ with $\mathcal{G} \subseteq \mathcal{F}$ is strongly consistent with $\rho_{\mathcal{F}}$ by monotonicity, so the existence of such a terminal risk measure which is strongly consistent with the other risk measures of the family is no further restriction. In the truly multivariate case, however, it is very natural that also under full information there is a rule for aggregating risk over the dimensions, and the risk measures in the family should be consistent with this terminal aggregation rule. If this is the case, we show, as already mentioned, that the members of the family are necessarily of type (3.1.4). Indeed we show that by strong consistency the risk measures inherit a property called *risk-antitonicity* in Hoffmann et al. (2016) from the terminal risk measure. This property is the essential axiom behind allowing for a decomposition of type (3.1.4); see Theorem 3.3.12.

Along the path to this result we characterize strong consistency in terms of a tower property. It is well-known, see e.g. Tutsch (2007), that for univariate conditional risk measures which are normalized on constants ($\eta_{\mathcal{G}}(a) = -a$ for all $a \in L^\infty(\mathcal{G})$), strong consistency (3.1.3) is equivalent to the following tower property:

$$\rho_{\mathcal{G}}(X) = \rho_{\mathcal{G}}\big(-\rho_{\mathcal{H}}(X)\big) \text{ for all } X \in L^\infty(\mathcal{F}). \tag{3.1.5}$$

The recursive formulation (3.1.5) is often more useful than (3.1.3) when analyzing strong consistency. The formulation (3.1.5), however, cannot be extended in a straight forward manner to the multivariate case. Firstly, note that (3.1.5) is not even well-defined in the multivariate case since $\rho_{\mathcal{H}}(X)$ is not a d-dimensional random vector but a random number. Secondly, also in the univariate case the equivalence (3.1.3) $\Leftrightarrow$ (3.1.5) only holds for risk measures that are normalized on constants, which in the *monetary* univariate case is implied up to a normalization by requiring that this class of risk measures satisfy cash-additivity ($\eta_{\mathcal{G}}(X + a) = \eta_{\mathcal{G}}(X) - a$). For multivariate risk measures there is neither a canonical extension of the concept of cash-additivity nor is it clear that such a property is desirable at all. In a first step we therefore derive a generalization of the recursive formulation (3.1.5) of strong consistency for not necessarily cash-additive multivariate risk measures. Indeed, under some typical regularity assumptions, one of our first results is that two multivariate conditional risk measures $\rho_{\mathcal{G}}$ and $\rho_{\mathcal{H}}$ with $\mathcal{G} \subseteq \mathcal{H} \subseteq \mathcal{F}$ are strongly consistent if and only if for all $X \in L_d^\infty(\mathcal{F})$

$$\rho_{\mathcal{G}}(X) = \rho_{\mathcal{G}}\big(f_{\rho_{\mathcal{H}}}^{-1}(\rho_{\mathcal{H}}(X))\mathbf{1}_d\big), \tag{3.1.6}$$

where $\mathbf{1}_d$ is a d-dimensional vector with all entries equal to 1, and $f_{\rho_{\mathcal{H}}}^{-1}$ is the (well-defined) inverse of the function $f_{\rho_{\mathcal{H}}}$ associated to $\rho_{\mathcal{H}}$ given by

$$f_{\rho_{\mathcal{H}}} : L^\infty(\mathcal{H}) \to L^\infty(\mathcal{H}); \alpha \mapsto \rho_{\mathcal{H}}(\alpha \mathbf{1}_d). \tag{3.1.7}$$

The map $f_{\rho_{\mathcal{H}}}$ describes the risk of a system where each component is equipped with the same amount of ($\mathcal{H}$-constant) cash α. Note that if $\rho_{\mathcal{H}}$ is a univariate risk measure that is normalized on constants then $f_{\rho_{\mathcal{H}}} = -\,\mathrm{id}$ is minus the identity map and (3.1.6) reduces to (3.1.5). In this sense, for a multivariate risk measure $\rho_{\mathcal{H}}$ the generalization of the normalization on constants property that is suited for our purposes is the requirement $f_{\rho_{\mathcal{H}}} = -\,\mathrm{id}$. Further, we remark that one can always "normalize" a given conditional risk measure $\rho_{\mathcal{H}}$ by putting

$$\bar{\rho}_{\mathcal{H}}(X) := -f_{\rho_{\mathcal{H}}}^{-1} \circ \rho_{\mathcal{H}}(X). \tag{3.1.8}$$

Then $\bar{\rho}_{\mathcal{H}}$ is a multivariate conditional risk measure with $f_{\bar{\rho}_{\mathcal{H}}} = -\,\mathrm{id}$.

After studying strong consistency for general families of multivariate conditional risk measures, we move on to give a characterization of strongly consistent multivariate conditional risk measures which are also conditionally law-invariant. In contrast to before we do not require consistency with respect to a risk measure under full information, but with respect to the initial risk measure given the trivial information $\{\emptyset, \Omega\}$. These studies were triggered by the results obtained in Föllmer (2014) for univariate risk measures, where it is shown that the only family of univariate, strongly consistent, conditional, cash-additive, convex risk measures is the family of conditional entropic risk measures, i.e. the conditional risk measures are conditional certainty equivalents of the form

$$\rho_{\mathcal{H}}(X) = -u^{-1}\left(\mathbb{E}_{\mathbb{P}}\left[u(X) \mid \mathcal{H}\right]\right), \quad X \in L^\infty(\mathcal{F}),$$

with deterministic utility function $u(x) = a + be^{\beta x}$ or $u(x) = a + bx$, where $a \in \mathbb{R}$ and $b, \beta > 0$ are constants. We also remark that Kupper and Schachermayer (2009) showed this characterization for the case of dynamic risk measures by an alternative proof. In the multivariate case we will see that every strongly consistent family of multivariate conditionally law-invariant conditional risk measures consists of risk measures of type

$$\rho_{\mathcal{H}}(X) = f_{\rho_{\mathcal{H}}}\left(f_u^{-1}\big(\mathbb{E}_{\mathbb{P}}\left[u(X) \mid \mathcal{H}\right]\big)\right), \quad X \in L_d^\infty(\mathcal{F}), \tag{3.1.9}$$

where $u : \mathbb{R}^d \to \mathbb{R}$ is a multivariate utility function and $f_u(x) := u(x\mathbf{1}_d)$, $x \in \mathbb{R}$. In other words they can be decomposed into the function $f_{\rho_{\mathcal{H}}}$ in (3.1.7) applied to a multivariate conditional certainty equivalent $\big(f_u^{-1}\big(\mathbb{E}_{\mathbb{P}}\left[u(X) \mid \mathcal{H}\right]\big)\big)$. For the study of univariate conditional certainty equivalents and their dynamic behavior we refer the interested reader to Frittelli and Maggis (2011). Moreover, we will derive the decomposition (3.1.4) from (3.1.9), i.e. in terms of u and f_u.

Structure of the paper

In Section 3.2 we introduce our notation and multivariate conditional risk measures. Moreover, we give the definition and some auxiliary results for the function $f_{\rho_{\mathcal{H}}}$ mentioned in (3.1.7). In Sections 3.3 and 3.4 we prove our main results outlined above for two strongly consistent conditional risk measures, where the law-invariant case is studied in Section 3.4. Throughout Section 3.5 we extend these results to families of multivariate conditional risk measures.

3.2 Definitions and basic results

Throughout this paper $(\Omega, \mathcal{F}, \mathbb{P})$ is a probability space. For $d \in \mathbb{N}$ we denote by $L_d^\infty(\mathcal{F}) := \{X = (X_1, ..., X_d) : X_i \in L^\infty(\Omega, \mathcal{F}, \mathbb{P})\ \forall i\}$ the space of equivalence classes of $\mathcal{F}$-measurable, $\mathbb{P}$-almost surely (a.s.) bounded random vectors. It is a Banach space when equipped with the norm $\|X\|_{d,\infty} := \max_{i=1,...,d} \|X_i\|_\infty$ where $\|F\|_\infty := \operatorname{esssup} |F|$ is the supremum norm for $F \in L^\infty(\Omega, \mathcal{F}, \mathbb{P})$. We will use the usual componentwise orderings on $\mathbb{R}^d$ and $L_d^\infty(\mathcal{F})$, i.e. $x = (x_1, \ldots, x_d) \geq y = (y_1, \ldots, y_d)$ for $x, y \in \mathbb{R}^d$ if and only if $x_i \geq y_i$ for all $i = 1, \ldots, d$, and similarly $X \geq Y$ if and only if $X_i \geq Y_i$ $\mathbb{P}$-a.s. for all $i = 1, ..., d$. Furthermore $\mathbf{1}_d$ and $\mathbf{0}_d$ denote the d-dimensional vectors whose entries are all equal to 1 or all equal to 0, respectively.

Definition 3.2.1. *Let $\mathcal{G} \subseteq \mathcal{F}$. A conditional risk measure (CRM) is a function*

$$\rho_{\mathcal{G}} : L_d^\infty(\mathcal{F}) \to L^\infty(\mathcal{G}),$$

possessing the following properties:

i) There exists a position with zero risk, i.e. $0 \in \operatorname{Im} \rho_{\mathcal{G}}$.

ii) ***Strict Antitonicity:*** $X \geq Y$ *and* $\mathbb{P}(X > Y) > 0$ *implies* $\rho_{\mathcal{G}}(X) \leq \rho_{\mathcal{G}}(Y)$ *and* $\mathbb{P}\big(\rho_{\mathcal{G}}(X) < \rho_{\mathcal{G}}(Y)\big) > 0$.

iii) ***$\mathcal{G}$-Locality:*** *For all* $A \in \mathcal{G}$ *we have* $\rho_{\mathcal{G}}(X\mathbb{1}_A + Y\mathbb{1}_{A^C}) = \rho_{\mathcal{G}}(X)\mathbb{1}_A + \rho_{\mathcal{G}}(Y)\mathbb{1}_{A^C}$.

iv) ***Lebesgue property:*** *If* $(X_n)_{n\in\mathbb{N}} \subset L_d^\infty(\mathcal{F})$ *is a* $\|\cdot\|_{d,\infty}$*-bounded sequence such that* $X_n \to X$ $\mathbb{P}$*-a.s., then*

$$\rho_{\mathcal{G}}(X) = \lim_{n\to\infty} \rho_{\mathcal{G}}(X_n) \quad \mathbb{P}\text{-a.s.}$$

We remark that the properties in Definition 3.2.1 are standard in the literature on conditional risk measures. Note that strict antitonicity is sometimes also referred to as strong sensitivity in the literature. In order to stress the dimension we often use the term univariate conditional risk measure for a conditional risk measure as defined in Definition 3.2.1 with $d = 1$ and we typically denote it by $\eta_{\mathcal{G}}$. For $d > 1$ the risk measure $\rho_{\mathcal{G}}$ of Definition 3.2.1 is called multivariate conditional risk measure.

A standard assumption on univariate CRMs is cash-additivity, i.e. $\eta_{\mathcal{G}}(X + \alpha) = \eta_{\mathcal{G}}(X) - \alpha$ for all $\alpha \in L^\infty(\mathcal{G})$, which in particular implies that we postulate a certain behavior of the risk measure $\eta_{\mathcal{G}}$ on $(\mathcal{G})$-constants $\alpha \in L^\infty(\mathcal{G})$ which turns out to be helpful in the study of consistency. Since we do not require this property - given that a multivariate analogue is tricky to define and probably not reasonable to ask for - we will have to extract the behavior of a CRM on constants in the following way.

Definition 3.2.2. *For every CRM $\rho_{\mathcal{G}}$ we introduce the function*

$$f_{\rho_{\mathcal{G}}} : L^\infty(\mathcal{G}) \to L^\infty(\mathcal{G}); \alpha \mapsto \rho_{\mathcal{G}}(\alpha \mathbf{1}_d)$$

and the corresponding inverse function

$$f_{\rho_{\mathcal{G}}}^{-1} : \operatorname{Im} f_{\rho_{\mathcal{G}}} \to L^\infty(\mathcal{G}); \beta \mapsto \alpha \textit{ such that } f_{\rho_{\mathcal{G}}}(\alpha) = \beta.$$

Remark 3.2.3. Note that the strict antitonicity of $\rho_{\mathcal{G}}$ implies that the inverse function $f_{\rho_{\mathcal{G}}}^{-1}$ in Definition 3.2.2 is well-defined. Indeed let $\beta \in \operatorname{Im} f_{\rho_{\mathcal{G}}}$ and $\alpha_1, \alpha_2 \in L^\infty(\mathcal{G})$ such that $f_{\rho_{\mathcal{G}}}(\alpha_1) = \beta = f_{\rho_{\mathcal{G}}}(\alpha_2)$. Suppose that $\mathbb{P}(A) > 0$ where $A := \{\alpha_1 > \alpha_2\} \in \mathcal{G}$. Then by strict antitonicity and $\mathcal{G}$-locality we obtain that

$$\begin{aligned}
\beta \mathbb{1}_A + \rho_{\mathcal{G}}(\mathbf{0}_d)\mathbb{1}_{A^C} &= \rho_{\mathcal{G}}(\alpha_1 \mathbf{1}_d)\mathbb{1}_A + \rho_{\mathcal{G}}(\mathbf{0}_d)\mathbb{1}_{A^C} = \rho_{\mathcal{G}}(\alpha_1 \mathbf{1}_d \mathbb{1}_A) \\
&\le \rho_{\mathcal{G}}(\alpha_2 \mathbf{1}_d \mathbb{1}_A) = \rho_{\mathcal{G}}(\alpha_2 \mathbf{1}_d)\mathbb{1}_A + \rho_{\mathcal{G}}(\mathbf{0}_d)\mathbb{1}_{A^C} \\
&= \beta \mathbb{1}_A + \rho_{\mathcal{G}}(\mathbf{0}_d)\mathbb{1}_{A^C},
\end{aligned}$$

and the inequality is strict with positive probability which is a contradiction. Thus we have that $\mathbb{P}(\alpha_1 > \alpha_2) = 0$. The same argument for $\{\alpha_1 < \alpha_2\}$ yields $\alpha_1 = \alpha_2$ $\mathbb{P}$-a.s.

Next we will show that properties of $\rho_{\mathcal{G}}$ transfer to $f_{\rho_{\mathcal{G}}}$ and $f_{\rho_{\mathcal{G}}}^{-1}$. Since the domain of $f_{\rho_{\mathcal{G}}}^{-1}$ might be only a subset of $L^\infty(\mathcal{G})$, we need to adapt the definition of the Lebesgue property for $f_{\rho_{\mathcal{G}}}^{-1}$ in the following way: If $(\beta_n)_{n\in\mathbb{N}} \subset \operatorname{Im} f_{\rho_{\mathcal{G}}}$ is a sequence which is lower- and upper-bounded by some $\underline{\beta}, \overline{\beta} \in \operatorname{Im} f_{\rho_{\mathcal{G}}}$, i.e. $\underline{\beta} \le \beta_n \le \overline{\beta}$ for all $n \in \mathbb{N}$, and such that $\beta_n \to \beta$ $\mathbb{P}$-a.s., then $f_{\rho_{\mathcal{G}}}^{-1}(\beta_n) \to f_{\rho_{\mathcal{G}}}^{-1}(\beta)$ $\mathbb{P}$-a.s. Note that this alternative definition of the Lebesgue property is equivalent to Definition 3.2.1 (iv) if the domain is $L^\infty(\mathcal{G})$. The properties 'strict antitonicity' and 'locality' of $f_{\rho_{\mathcal{G}}}$ or $f_{\rho_{\mathcal{G}}}^{-1}$ are defined analogous to Definition 3.2.1 (ii) and (iii).

Lemma 3.2.4. *Let $f_{\rho_{\mathcal{G}}}$ and $f_{\rho_{\mathcal{G}}}^{-1}$ be as in Definition* 3.2.2. *Then $f_{\rho_{\mathcal{G}}}$ and $f_{\rho_{\mathcal{G}}}^{-1}$ are strictly antitone, $\mathcal{G}$-local and fulfill the Lebesgue property.*

Proof. For $f_{\rho_{\mathcal{G}}}$ the statement follows immediately from the definition and the corresponding properties of $\rho_{\mathcal{G}}$. Concerning the properties of $f_{\rho_{\mathcal{G}}}^{-1}$, we start by proving strict antitonicity. Let $\beta_1, \beta_2 \in \operatorname{Im} f_{\rho_{\mathcal{G}}}$ such that $\beta_1 \ge \beta_2$ and $\mathbb{P}(\beta_1 > \beta_2) > 0$. Suppose that $\mathbb{P}(A) > 0$ where $A := \left\{ f_{\rho_{\mathcal{G}}}^{-1}(\beta_1) > f_{\rho_{\mathcal{G}}}^{-1}(\beta_2) \right\} \in \mathcal{G}$. Then

$$\begin{aligned}
\beta_1 \mathbb{1}_A + f_{\rho_{\mathcal{G}}}(0)\mathbb{1}_{A^C} &= f_{\rho_{\mathcal{G}}}\left(f_{\rho_{\mathcal{G}}}^{-1}(\beta_1)\right)\mathbb{1}_A + f_{\rho_{\mathcal{G}}}(0)\mathbb{1}_{A^C} = f_{\rho_{\mathcal{G}}}\left(f_{\rho_{\mathcal{G}}}^{-1}(\beta_1)\mathbb{1}_A\right) \\
&\le f_{\rho_{\mathcal{G}}}\left(f_{\rho_{\mathcal{G}}}^{-1}(\beta_2)\mathbb{1}_A\right) = \beta_2 \mathbb{1}_A + f_{\rho_{\mathcal{G}}}(0)\mathbb{1}_{A^C},
\end{aligned}$$

and the inequality is strict on a set with positive probability since $f_{\rho_{\mathcal{G}}}$ is strictly antitone. This of course contradicts $\beta_1 \ge \beta_2$. Hence $f_{\rho_{\mathcal{G}}}^{-1}(\beta_1) \le f_{\rho_{\mathcal{G}}}^{-1}(\beta_2)$. Moreover, as

$$\mathbb{P}(\beta_1 > \beta_2) = \mathbb{P}\left(f_{\rho_{\mathcal{G}}}\left(f_{\rho_{\mathcal{G}}}^{-1}(\beta_1)\right) > f_{\rho_{\mathcal{G}}}\left(f_{\rho_{\mathcal{G}}}^{-1}(\beta_2)\right)\right) > 0$$

we must have $f_{\rho_{\mathcal{G}}}^{-1}(\beta_1) \neq f_{\rho_{\mathcal{G}}}^{-1}(\beta_2)$ with positive probability, i.e.

$$\mathbb{P}\left(f_{\rho_{\mathcal{G}}}^{-1}(\beta_1) < f_{\rho_{\mathcal{G}}}^{-1}(\beta_2)\right) > 0.$$

Now we show that $f_{\rho_{\mathcal{G}}}^{-1}$ is $\mathcal{G}$-local. Let $\beta_1, \beta_2 \in \operatorname{Im} f_{\rho_{\mathcal{G}}}$ as well as $A \in \mathcal{G}$ be arbitrary. Further let $\alpha_i = f_{\rho_{\mathcal{G}}}^{-1}(\beta_i), i = 1, 2$, i.e. $f_{\rho_{\mathcal{G}}}(\alpha_i) = \beta_i$. Then we have that

$$f_{\rho_{\mathcal{G}}}(\alpha_1 \mathbb{1}_A + \alpha_2 \mathbb{1}_{A^C}) = f_{\rho_{\mathcal{G}}}(\alpha_1)\mathbb{1}_A + f_{\rho_{\mathcal{G}}}(\alpha_2)\mathbb{1}_{A^C} = \beta_1 \mathbb{1}_A + \beta_2 \mathbb{1}_{A^C}.$$

Thus $f_{\rho_{\mathcal{G}}}^{-1}(\beta_1 \mathbb{1}_A + \beta_2 \mathbb{1}_{A^C}) = \alpha_1 \mathbb{1}_A + \alpha_2 \mathbb{1}_{A^C}$.

Finally for the Lebesgue property let $\underline{\beta}, \overline{\beta} \in \operatorname{Im} f_{\rho_{\mathcal{G}}}$ and let $(\beta_n)_{n\in\mathbb{N}} \subset \operatorname{Im} f_{\rho_{\mathcal{G}}}$ be a sequence with $\underline{\beta} \leq \beta_n \leq \overline{\beta}$ for all $n \in \mathbb{N}$ and $\beta_n \to \beta$ $\mathbb{P}$-a.s. Consider the bounded sequences $\beta_n^u := \sup_{k\geq n} \beta_k$ and $\beta_n^d := \inf_{k\geq n} \beta_k$, $n \in \mathbb{N}$ which converge monotonically almost surely to β, i.e. $\beta_n^u \downarrow \beta$ $\mathbb{P}$-a.s. and $\beta_n^d \uparrow \beta$ $\mathbb{P}$-a.s. Since $\underline{\beta} \leq \beta_n^u \leq \overline{\beta}$ for all $n \in \mathbb{N}$ which by antitonicity of $f_{\rho_{\mathcal{G}}}^{-1}$ yields $f_{\rho_{\mathcal{G}}}^{-1}(\overline{\beta}) \leq f_{\rho_{\mathcal{G}}}^{-1}(\beta_n^u) \leq f_{\rho_{\mathcal{G}}}^{-1}(\underline{\beta})$, we observe that the sequence $\left(f_{\rho_{\mathcal{G}}}^{-1}(\beta_n^u)\right)_{n\in\mathbb{N}}$ is uniformly bounded in $L^\infty(\mathcal{G})$. Note that by the same argumentation also the sequences $\left(f_{\rho_{\mathcal{G}}}^{-1}(\beta_n^d)\right)_{n\in\mathbb{N}}$ and $\left(f_{\rho_{\mathcal{G}}}^{-1}(\beta_n)\right)_{n\in\mathbb{N}}$ are uniformly bounded in $L^\infty(\mathcal{G})$. Next we will show that $\beta_n^u \in \operatorname{Im} f_{\rho_{\mathcal{G}}}$ for all $n \in \mathbb{N}$. Fix $n \in \mathbb{N}$ and set recursively

$$A_{n-1}^n := \{\beta_n^u = \beta\} \quad \text{and} \quad A_k^n := \{\beta_n^u = \beta_k\} \setminus \bigcup_{i=n-1}^{k-1} A_i^n,\ k \geq n,$$

then it follows from induction that $A_k^n \in \mathcal{G}, k \geq n-1$. Since $\sup\{\beta, \beta_k : k \geq n\} = \max\{\beta, \beta_k : k \geq n\}$, we have that $\left(\bigcup_{k\geq n-1} A_k^n\right)^C$ is a $\mathbb{P}$-nullset. It follows from $\mathcal{G}$-locality and the Lebesgue property of $f_{\rho_{\mathcal{G}}}$ that

$$\begin{aligned}
f_{\rho_{\mathcal{G}}} & \left(f_{\rho_{\mathcal{G}}}^{-1}(\beta)\mathbb{1}_{A_{n-1}^n} + \sum_{k\geq n} f_{\rho_{\mathcal{G}}}^{-1}(\beta_k)\mathbb{1}_{A_k^n}\right) \\
&= \beta \mathbb{1}_{A_{n-1}^n} + f_{\rho_{\mathcal{G}}}\left(\lim_{m\to\infty} \sum_{k=n}^m f_{\rho_{\mathcal{G}}}^{-1}(\beta_k)\mathbb{1}_{A_k^n}\right) \mathbb{1}_{\bigcup_{k\geq n} A_k^n} \\
&= \beta \mathbb{1}_{A_{n-1}^n} + \lim_{m\to\infty}\left(\sum_{k=n}^m \beta_k \mathbb{1}_{A_k^n} + f_{\rho_{\mathcal{G}}}(0)\, \mathbb{1}_{\bigcup_{k\geq m} A_k^n}\right) \\
&= \beta \mathbb{1}_{A_{n-1}^n} + \sum_{k\geq n} \beta_k \mathbb{1}_{A_k^n} = \beta_n^u,
\end{aligned}$$

which implies $\beta_n^u \in \operatorname{Im} f_{\rho_{\mathcal{G}}}$. By a similar argumentation we obtain $\beta_n^d \in \operatorname{Im} f_{\rho_{\mathcal{G}}}$. Recall that $\beta_n^u \downarrow \beta$ $\mathbb{P}$-a.s. which by antitonicity of $f_{\rho_{\mathcal{G}}}^{-1}$ implies that the sequence $\left(f_{\rho_{\mathcal{G}}}^{-1}(\beta_n^u)\right)_{n\in\mathbb{N}}$ is isotone and thus $\alpha = \lim_{n\to\infty} f_{\rho_{\mathcal{G}}}^{-1}(\beta_n^u)$ exists in $L^\infty(\mathcal{G})$. It follows from antitonicity and the Lebesgue property of $f_{\rho_{\mathcal{G}}}$ that

$$\beta = \lim_{n\to\infty} \beta_n^u = \lim_{n\to\infty} f_{\rho_{\mathcal{G}}}\left(f_{\rho_{\mathcal{G}}}^{-1}(\beta_n^u)\right) = f_{\rho_{\mathcal{G}}}(\alpha),$$

and hence that indeed $\alpha = f_{\rho_{\mathcal{G}}}^{-1}(\beta)$. Analogously, we obtain that $f_{\rho_{\mathcal{G}}}(\hat{\alpha}) = \beta$ for $\hat{\alpha} = \lim_{n\to\infty} f_{\rho_{\mathcal{G}}}^{-1}(\beta_n^d)$, and thus $\hat{\alpha} = \alpha = f_{\rho_{\mathcal{G}}}^{-1}(\beta)$. Hence, by antitonicity of $f_{\rho_{\mathcal{G}}}^{-1}$

$$\begin{aligned} f_{\rho_{\mathcal{G}}}^{-1}(\beta) &= \lim_{n\to\infty} f_{\rho_{\mathcal{G}}}^{-1}(\beta_n^u) \leq \liminf_{n\to\infty} f_{\rho_{\mathcal{G}}}^{-1}(\beta_n) \\ &\leq \limsup_{n\to\infty} f_{\rho_{\mathcal{G}}}^{-1}(\beta_n) \leq \lim_{n\to\infty} f_{\rho_{\mathcal{G}}}^{-1}(\beta_n^d) = f_{\rho_{\mathcal{G}}}^{-1}(\beta), \end{aligned}$$

so $\lim_{n\to\infty} f_{\rho_{\mathcal{G}}}^{-1}(\beta_n) = f_{\rho_{\mathcal{G}}}^{-1}(\beta)$, i.e. $f_{\rho_{\mathcal{G}}}^{-1}$ has the Lebesgue property. □

An important observation that will be needed later on is that the domain of $f_{\rho_{\mathcal{G}}}^{-1}$ is equal to the image of $\rho_{\mathcal{G}}$, i.e. $f_{\rho_{\mathcal{G}}}^{-1}(\rho_{\mathcal{G}}(X))$ is well-defined for all $X \in L_d^\infty(\mathcal{F})$.

Lemma 3.2.5. *For a CRM $\rho_{\mathcal{G}} : L_d^\infty(\mathcal{F}) \to L^\infty(\mathcal{G})$ it holds that*

$$\rho_{\mathcal{G}}(L_d^\infty(\mathcal{F})) = f_{\rho_{\mathcal{G}}}(L^\infty(\mathcal{G})).$$

Proof. Clearly, $\rho_{\mathcal{G}}(L_d^\infty(\mathcal{F})) \supseteq f_{\rho_{\mathcal{G}}}(L^\infty(\mathcal{G}))$.
For the reverse inclusion let $X \in L_d^\infty(\mathcal{F})$. Our aim is to show that there exists an $\alpha^* \in L^\infty(\mathcal{G})$ such that

$$\rho_{\mathcal{G}}(X) = f_{\rho_{\mathcal{G}}}(\alpha^*). \tag{3.2.1}$$

Define

$$P := \left\{\alpha \in L^\infty(\mathcal{G}) \,:\, f_{\rho_{\mathcal{G}}}(\alpha) \geq \rho_{\mathcal{G}}(X)\right\}.$$

As $-\|X\|_{d,\infty}\mathbf{1}_d \leq X \leq \|X\|_{d,\infty}\mathbf{1}_d$ we have that $-\|X\|_{d,\infty} \in P$, so $P \neq \emptyset$. Moreover, P is bounded from above by $\|X\|_{d,\infty}$ since if $A := \{\alpha > \|X\|_{d,\infty}\}$ for $\alpha \in L^\infty(\mathcal{G})$ has positive probability, then by $\mathcal{G}$-locality and strict antitonicity

$$f_{\rho_{\mathcal{G}}}(\alpha)\mathbb{1}_A = f_{\rho_{\mathcal{G}}}(\alpha\mathbb{1}_A)\mathbb{1}_A \leq f_{\rho_{\mathcal{G}}}(\|X\|_{d,\infty}\mathbb{1}_A)\mathbb{1}_A = f_{\rho_{\mathcal{G}}}(\|X\|_{d,\infty})\mathbb{1}_A \leq \rho_{\mathcal{G}}(X)\mathbb{1}_A$$

where the first inequality is strict with positive probability, so $\alpha \notin P$. By $\mathcal{G}$-locality it also follows that P is upwards directed. Hence, for

$$\alpha^* := \operatorname{esssup} P$$

there is a uniformly bounded sequence $(\alpha_n)_{n\in\mathbb{N}} \subset P$ such that $\alpha^* = \lim_{n\to\infty} \alpha_n$ $\mathbb{P}$-a.s.; see Föllmer and Schied (2011) Theorem A.33. Thus it follows that $\alpha^* \in L^\infty(\mathcal{G})$ and

$$f_{\rho_{\mathcal{G}}}(\alpha^*) = \lim_{n\to\infty} f_{\rho_{\mathcal{G}}}(\alpha_n) \geq \rho_{\mathcal{G}}(X),$$

i.e. $\alpha^* \in P$. Let

$$B := \{f_{\rho_{\mathcal{G}}}(\alpha^*) > \rho_{\mathcal{G}}(X)\}$$

and note that by the Lebesgue property

$$B = \bigcup_{n\in\mathbb{N}} \{f_{\rho_{\mathcal{G}}}(\alpha^* + 1/n) > \rho_{\mathcal{G}}(X)\} \quad \mathbb{P}\text{-a.s.}$$

Hence, if $\mathbb{P}(B) > 0$ it follows that $\mathbb{P}(B_n) > 0$ for some $B_n := \{f_{\rho_{\mathcal{G}}}(\alpha^* + 1/n) > \rho_{\mathcal{G}}(X)\}$. Note that $B_n \in \mathcal{G}$ and that

$$\alpha^* \mathbb{1}_{B_n^C} + (\alpha^* + 1/n)\mathbb{1}_{B_n} \in P$$

by $\mathcal{G}$-locality of $f_{\rho_{\mathcal{G}}}$. But this contradicts the definition of α^*. Hence, $\mathbb{P}(B) = 0$. □

Sometimes it will be useful to normalize the CRM in the following sense:

Definition 3.2.6. *We call a CRM $\rho_{\mathcal{G}} : L_d^\infty(\mathcal{F}) \to L^\infty(\mathcal{G})$* normalized on constants *if*

$$f_{\rho_{\mathcal{G}}}(\alpha) = -\alpha \quad \textit{for all } \alpha \in L^\infty(\mathcal{G}).$$

Indeed let $\rho_{\mathcal{G}} : L_d^\infty(\mathcal{F}) \to L^\infty(\mathcal{G})$ be a CRM and define $\bar{\rho}_{\mathcal{G}} := -f_{\rho_{\mathcal{G}}}^{-1} \circ \rho_{\mathcal{G}}$. Then $\bar{\rho}_{\mathcal{G}}$ is a CRM (Lemma 3.2.4 and Lemma 3.2.5) which is normalized in the sense of being normalized on constants as defined above. We call $\bar{\rho}_{\mathcal{G}}$ the *normalized CRM* of $\rho_{\mathcal{G}}$.

3.3 Strong consistency

In this section we study consistency of CRMs. We consider the most frequently used consistency condition for univariate risk measures in the literature which is known as strong consistency and extend it to the multivariate case. We refer to Detlefsen and Scandolo (2005), Cheridito et al. (2006), Cheridito and Kupper (2011), Kupper and Schachermayer (2009), and Penner (2007) for more information on strong consistency of univariate risk measures. Kromer et al. (2014) also study a kind of consistency for multivariate risk measures, however, as we will point out in Remark 3.4.11 below, their definition of consistency differs from our approach. For the remainder of this section we let $\mathcal{G}$ and $\mathcal{H}$ be two sub-σ-algebras of $\mathcal{F}$ such that $\mathcal{G} \subseteq \mathcal{H}$, and let $\rho_{\mathcal{G}} : L_d^\infty(\mathcal{F}) \to L^\infty(\mathcal{G})$ and $\rho_{\mathcal{H}} : L_d^\infty(\mathcal{F}) \to L^\infty(\mathcal{H})$ be the corresponding CRMs.

Definition 3.3.1 (Strong consistency)**.** *The pair $\{\rho_{\mathcal{G}}, \rho_{\mathcal{H}}\}$ is called* strongly consistent *if*

$$\rho_{\mathcal{H}}(X) \leq \rho_{\mathcal{H}}(Y) \;\Rightarrow\; \rho_{\mathcal{G}}(X) \leq \rho_{\mathcal{G}}(Y) \quad (X, Y \in L_d^\infty(\mathcal{F})). \tag{3.3.1}$$

Strong consistency states that if one risk is preferred to another risk in almost surely all states under more information, then this preference already holds under less information. Our first result shows that strong consistency can be equivalently defined by a recursive relation.

Lemma 3.3.2. *Equivalent are:*

(i) $\{\rho_{\mathcal{G}}, \rho_{\mathcal{H}}\}$ *is strongly consistent;*

(ii) For all $X \in L_d^\infty(\mathcal{F})$ it holds that

$$\rho_{\mathcal{G}}(X) = \rho_{\mathcal{G}}\Big(f_{\rho_{\mathcal{H}}}^{-1}\big(\rho_{\mathcal{H}}(X)\big)\mathbf{1}_d\Big),$$

where $f_{\rho_{\mathcal{H}}}^{-1}$ was defined in Definition 3.2.2.

Proof. (i)⇒(ii): As for all $X \in L_d^\infty(\mathcal{F})$

$$\rho_{\mathcal{H}}(X) = \rho_{\mathcal{H}}\left(f_{\rho_{\mathcal{H}}}^{-1}\left(\rho_{\mathcal{H}}(X)\right)\mathbf{1}_d\right),$$

it follows from strong consistency that

$$\rho_{\mathcal{G}}(X) = \rho_{\mathcal{G}}\left(f_{\rho_{\mathcal{H}}}^{-1}\left(\rho_{\mathcal{H}}(X)\right)\mathbf{1}_d\right).$$

(ii)⇒(i): Let $X, Y \in L_d^\infty(\mathcal{F})$ be such that $\rho_{\mathcal{H}}(X) \le \rho_{\mathcal{H}}(Y)$. Then by antitonicity of $f_{\rho_{\mathcal{H}}}^{-1}$ and $\rho_{\mathcal{G}}$ it follows that

$$\rho_{\mathcal{G}}(X) = \rho_{\mathcal{G}}\Big(f_{\rho_{\mathcal{H}}}^{-1}\big(\rho_{\mathcal{H}}(X)\big)\mathbf{1}_d\Big) \le \rho_{\mathcal{G}}\Big(f_{\rho_{\mathcal{H}}}^{-1}\big(\rho_{\mathcal{H}}(Y)\big)\mathbf{1}_d\Big) = \rho_{\mathcal{G}}(Y).$$

□

Remark 3.3.3. Let $\eta_{\mathcal{G}}$ and $\eta_{\mathcal{H}}$ be two univariate CRMs, where $\eta_{\mathcal{H}}$ is normalized on constants, i.e. $\eta_{\mathcal{H}}(\alpha) = -\alpha$ for all $\alpha \in L^\infty(\mathcal{H})$. Then $f_{\eta_{\mathcal{H}}}(\alpha) = f_{\eta_{\mathcal{H}}}^{-1}(\alpha) = -\alpha$ and thus strong consistency is equivalent to

$$\eta_{\mathcal{G}}(F) = \eta_{\mathcal{G}}\big(-\eta_{\mathcal{H}}(F)\big), \quad F \in L^\infty(\mathcal{F}).$$

Remark 3.3.4. If $\{\rho_{\mathcal{G}}, \rho_{\mathcal{H}}\}$ is strongly consistent so is the pair of normalized CRMs $\{\bar\rho_{\mathcal{G}}, \bar\rho_{\mathcal{H}}\}$ as defined in Definition 3.2.6 and vice versa. Since $f_{\bar\rho_{\mathcal{G}}} = f_{\bar\rho_{\mathcal{H}}} = -\,\mathrm{id}$ strong consistency of the normalized CRMs is equivalent to

$$\bar\rho_{\mathcal{G}}(F) = \bar\rho_{\mathcal{G}}\big(-\bar\rho_{\mathcal{H}}(F)\mathbf{1}_d\big), \quad F \in L^\infty(\mathcal{F}),$$

in analogy to Remark 3.3.3.

In the following lemma we will show that strong consistency of $\{\rho_{\mathcal{G}}, \rho_{\mathcal{H}}\}$ uniquely determines the normalized CRM $\bar\rho_{\mathcal{H}}$.

Lemma 3.3.5. *If $\{\rho_{\mathcal{G}}, \rho_{\mathcal{H}}\}$ is strongly consistent, then $\rho_{\mathcal{G}}$ uniquely determines the normalized CRM $\bar\rho_{\mathcal{H}} = -f_{\rho_{\mathcal{H}}}^{-1} \circ \rho_{\mathcal{H}}$.*

Proof. Suppose that there are two CRMs $\rho_{\mathcal{H}}^1$ and $\rho_{\mathcal{H}}^2$ which are strongly consistent with respect to $\rho_{\mathcal{G}}$, i.e.

$$\rho_{\mathcal{G}}\Big(f_{\rho_{\mathcal{H}}^1}^{-1}\big(\rho_{\mathcal{H}}^1(X)\big)\mathbf{1}_d\Big) = \rho_{\mathcal{G}}(X) = \rho_{\mathcal{G}}\Big(f_{\rho_{\mathcal{H}}^2}^{-1}\big(\rho_{\mathcal{H}}^2(X)\big)\mathbf{1}_d\Big), \quad X \in L_d^\infty(\mathcal{F}).$$

We will show that $f_{\rho_{\mathcal{H}}^1}^{-1}\left(\rho_{\mathcal{H}}^1(X)\right) = f_{\rho_{\mathcal{H}}^2}^{-1}\left(\rho_{\mathcal{H}}^2(X)\right)$. Suppose that there exists an $X \in L_d^\infty(\mathcal{F})$ such that $A := \left\{f_{\rho_{\mathcal{H}}^1}^{-1}\left(\rho_{\mathcal{H}}^1(X)\right) > f_{\rho_{\mathcal{H}}^2}^{-1}\left(\rho_{\mathcal{H}}^2(X)\right)\right\} \in \mathcal{H}$ has positive probability. Then, by the $\mathcal{H}$-locality of $\rho_{\mathcal{H}}^1$ and $\rho_{\mathcal{H}}^2$, we obtain

$$\begin{aligned}
\rho_{\mathcal{G}}(X\mathbb{1}_A) &= \rho_{\mathcal{G}}\Big(f_{\rho_{\mathcal{H}}^1}^{-1}\big(\rho_{\mathcal{H}}^1(X\mathbb{1}_A)\big)\mathbf{1}_d\Big) = \rho_{\mathcal{G}}\Big(f_{\rho_{\mathcal{H}}^1}^{-1}\big(\rho_{\mathcal{H}}^1(X)\big)\mathbb{1}_A\mathbf{1}_d\Big) \\
&\le \rho_{\mathcal{G}}\Big(f_{\rho_{\mathcal{H}}^2}^{-1}\big(\rho_{\mathcal{H}}^2(X)\big)\mathbb{1}_A\mathbf{1}_d\Big) = \rho_{\mathcal{G}}\Big(f_{\rho_{\mathcal{H}}^2}^{-1}\big(\rho_{\mathcal{H}}^2(X\mathbb{1}_A)\big)\mathbf{1}_d\Big) \\
&= \rho_{\mathcal{G}}(X\mathbb{1}_A). \qquad (3.3.2)
\end{aligned}$$

where the inequality (3.3.2) is strict with positive probability as $\rho_{\mathcal{G}}$ is strictly antitone, and hence we have a contradiction. Reverting the role of $\rho_{\mathcal{H}}^1$ and $\rho_{\mathcal{H}}^2$ in the definition of A proves the lemma. □

In Hoffmann et al. (2016) we studied under which conditions a (multivariate) conditional risk measure can be decomposed as in (3.1.4), i.e. into a conditional aggregation function and a univariate conditional risk measure. We will pursue showing that strong consistency of $\{\rho_{\mathcal{G}}, \rho_{\mathcal{F}}\}$ is already sufficient to guarantee a decomposition (3.1.4) for both $\rho_{\mathcal{G}}$ and $\rho_{\mathcal{F}}$. To this end we need to clarify what we mean by a conditional aggregation function:

Definition 3.3.6. *We call a function* $\Lambda : L_d^\infty(\mathcal{F}) \to L^\infty(\mathcal{F})$ *a* conditional aggregation function *if it fulfills the following properties:*

Strict isotonicity: $X \geq Y$ *and* $\mathbb{P}(X > Y) > 0$ *implies* $\Lambda(X) \geq \Lambda(Y)$ *and* $\mathbb{P}\big(\Lambda(X) > \Lambda(Y)\big) > 0$.

$\mathcal{F}$-***Locality:*** $\Lambda(X\mathbb{1}_A + Y\mathbb{1}_{A^C}) = \Lambda(X)\mathbb{1}_A + \Lambda(Y)\mathbb{1}_{A^C}$ *for all* $A \in \mathcal{F}$*;*

Lebesgue property: *For any uniformly bounded sequence* $(X_n)_{n\in\mathbb{N}}$ *in* $L_d^\infty(\mathcal{F})$ *such that* $X_n \to X$ $\mathbb{P}$*-a.s., we have that*

$$\Lambda(X) = \lim_{n\to\infty} \Lambda(X_n) \quad \mathbb{P}\text{-a.s.}$$

Moreover for $\mathcal{H} \subset \mathcal{F}$*, we call* Λ *a* $\mathcal{H}$-conditional aggregation function *if in addition*

$$\Lambda(L_d^\infty(\mathcal{J})) \subseteq L^\infty(\mathcal{J}) \text{ for all } \mathcal{H} \subseteq \mathcal{J} \subseteq \mathcal{F}.$$

Remark 3.3.7. The name $\mathcal{H}$-conditional aggregation function refers to the fact that $\Lambda(x) \in L^\infty(\mathcal{H})$ for all $x \in \mathbb{R}^d$. Thus every conditional aggregation function is at least a $\mathcal{F}$-conditional aggregation function.

As for conditional risk measures we define:

Definition 3.3.8. *For a conditional aggregation function* $\Lambda : L_d^\infty(\mathcal{F}) \to L^\infty(\mathcal{F})$ *let*

$$f_\Lambda : L^\infty(\mathcal{F}) \to L^\infty(\mathcal{F}); F \mapsto \Lambda(F\mathbf{1}_d)$$

and

$$f_\Lambda^{-1} : \operatorname{Im} f_\Lambda \to L^\infty(\mathcal{F}); G \mapsto F \text{ such that } f_\Lambda(F) = G.$$

Lemma 3.3.9. *Let* $\Lambda : L_d^\infty(\mathcal{F}) \to L^\infty(\mathcal{F})$ *be a conditional aggregation function. Then* f_Λ *and* f_Λ^{-1} *are strictly isotone,* $\mathcal{F}$*-local, and fulfill the Lebesgue property. Moreover,* $\Lambda(L_d^\infty(\mathcal{F})) = f_\Lambda(L^\infty(\mathcal{F}))$ *and* $\Lambda(X) = \Lambda\big(f_\Lambda^{-1}(\Lambda(X))\mathbf{1}_d\big)$ *for all* $X \in L_d^\infty(\mathcal{F})$*.*

The well-definedness of f_Λ^{-1} follows as in Remark 3.2.3. Further the proof of Lemma 3.3.9 is analogous to the proofs of Lemma 3.2.4 and Lemma 3.2.5 and therefore omitted here.

In order to state the decomposition result for strongly consistent CRMs, we first recall the main result from Hoffmann et al. (2016) adapted to the framework of this paper in Proposition 3.3.11 for which we need the following definition.

Definition 3.3.10. *We say that a function* $\rho_{\mathcal{G}} : L_d^\infty(\mathcal{F}) \to L^\infty(\mathcal{G})$ *has a continuous realization* $\rho_{\mathcal{G}}(\cdot,\cdot)$*, if for all* $X \in L_d^\infty(\mathcal{F})$ *there exists a representative* $\rho_{\mathcal{G}}(X,\cdot)$ *of the equivalence class* $\rho_{\mathcal{G}}(X)$ *such that* $\widetilde{\rho}_{\mathcal{G}} : \mathbb{R}^d \times \Omega \to \mathbb{R}; (x,\omega) \mapsto \rho_{\mathcal{G}}(x,\omega)$ *is continuous in its first argument* $\mathbb{P}$*-a.s.*

Proposition 3.3.11. *Let $\rho_{\mathcal{G}} : L_d^\infty(\mathcal{F}) \to L^\infty(\mathcal{G})$ be a CRM and suppose that there exists a continuous realization $\rho_{\mathcal{G}}(\cdot,\cdot)$ which satisfies* risk-antitonicity*:*

$$\widetilde{\rho}_{\mathcal{G}}(X(\omega),\omega) \geq \widetilde{\rho}_{\mathcal{G}}(Y(\omega),\omega)\ \mathbb{P}\text{-a.s., implies } \rho_{\mathcal{G}}(X) \geq \rho_{\mathcal{G}}(Y).$$

Then there exists a $\mathcal{G}$-conditional aggregation function $\Lambda_{\mathcal{G}} : L_d^\infty(\mathcal{F}) \to L^\infty(\mathcal{F})$ and a univariate CRM $\eta_{\mathcal{G}} : \operatorname{Im}\Lambda_{\mathcal{G}} \to L^\infty(\mathcal{G})$ such that

$$\rho_{\mathcal{G}}(X) = \eta_{\mathcal{G}}(\Lambda_{\mathcal{G}}(X)) \quad \text{for all } X \in L_d^\infty(\mathcal{F})$$

and

$$\eta_{\mathcal{G}}(\Lambda_{\mathcal{G}}(X)) = -\Lambda_{\mathcal{G}}(X) \quad \text{for all } X \in L_d^\infty(\mathcal{G}). \tag{3.3.3}$$

This decomposition is unique.

Proof. Since $\rho_{\mathcal{G}}$ is antitone, $\mathbb{R}^d \ni x \mapsto \rho_{\mathcal{G}}(x)$ is antitone. It has been shown in Hoffmann et al. (2016) Theorem 2.10 that this property in conjunction with the fact that $\rho_{\mathcal{G}}$ has a continuous realization which fulfills risk-antitonicity is sufficient for the existence and uniqueness of a function $\Lambda_{\mathcal{G}} : L_d^\infty(\mathcal{F}) \to L^\infty(\mathcal{F})$ which is isotone, $\mathcal{F}$-local and fulfills the Lebesgue property and a function $\eta_{\mathcal{G}} : \operatorname{Im}\Lambda_{\mathcal{G}} \to L^\infty(\mathcal{G})$ which is antitone such that

$$\rho_{\mathcal{G}} = \eta_{\mathcal{G}} \circ \Lambda_{\mathcal{G}} \quad \text{and} \quad \eta_{\mathcal{G}}\big(\Lambda_{\mathcal{G}}(x)\big) = -\Lambda_{\mathcal{G}}(x) \text{ for all } x \in \mathbb{R}^d. \tag{3.3.4}$$

Note that in the proof of Theorem 2.10 in Hoffmann et al. (2016) $\Lambda_{\mathcal{G}}$ is basically constructed by setting $\Lambda_{\mathcal{G}}(X)(\omega) = -\widetilde{\rho}_{\mathcal{G}}(X(\omega),\omega)$, which implies that $\Lambda_{\mathcal{G}}$ is necessarily $\mathcal{F}$-local even though this is not directly mentioned in the paper. Indeed in Hoffmann et al. (2016) we do not require or mention locality at all.

It remains to be shown that $\Lambda_{\mathcal{G}}$ is a $\mathcal{G}$-conditional aggregation function, $\eta_{\mathcal{G}}$ is a univariate CRM on $\operatorname{Im}\Lambda_{\mathcal{G}}$, and that (3.3.3) holds. First of all, we show that $\mathcal{F}$-locality and (3.3.4) imply (3.3.3). To this end denote by $\mathcal{S}$ the set of $\mathcal{F}$-measurable simple random vectors, i.e. $X \in \mathcal{S}$ if X is of the form $X = \sum_{i=1}^k x_i \mathbb{1}_{A_i}$, where $k \in \mathbb{N}$, $x_i \in \mathbb{R}^d$ and $A_i \in \mathcal{F}$, $i = 1,...,k$, are disjoint sets such that $\mathbb{P}(A_i) > 0$ and $\mathbb{P}(\bigcup_{i=1}^k A_i) = 1$. Now let $X \in L_d^\infty(\mathcal{G})$. Pick a uniformly bounded sequence $(X_n)_{n\in\mathbb{N}} = \left(\sum_{i=1}^{k_n} x_i^n \mathbb{1}_{A_i^n}\right)_{n\in\mathbb{N}} \subset \mathcal{S}$ such that $A_i^n \in \mathcal{G}$ for all $i = 1,\ldots,k_n$, $n \in \mathbb{N}$, and $X_n \to X$ $\mathbb{P}$-a.s. Then by (3.3.4), $\mathcal{F}$-locality and the Lebesgue property of $\Lambda_{\mathcal{G}}$ and $\rho_{\mathcal{G}}$ we infer that

$$\begin{aligned}
-\Lambda_{\mathcal{G}}(X) &= -\lim_{n\to\infty} \Lambda_{\mathcal{G}}(X_n) = \lim_{n\to\infty} \sum_{i=1}^{k_n} -\Lambda_{\mathcal{G}}(x_i^n)\mathbb{1}_{A_i^n} \\
&= \lim_{n\to\infty} \sum_{i=1}^{k_n} \rho_{\mathcal{G}}(x_i^n)\mathbb{1}_{A_i^n} = \lim_{n\to\infty} \rho_{\mathcal{G}}(X_n) = \rho_{\mathcal{G}}(X),
\end{aligned}$$

which proves (3.3.3). Next we show that $\Lambda_{\mathcal{G}}$ is a $\mathcal{G}$-conditional aggregation function. The yet missing properties which need to be verified are strict antitonicity and that $\Lambda_{\mathcal{G}}$ is $\mathcal{G}$-conditional. The latter follows from Hoffmann et al. (2016) Lemma 3.1. As for strict antitonicity let $X, Y \in$

$L^\infty_d(\mathcal{F})$ with $X \geq Y$ such that $\mathbb{P}(X > Y) > 0$. Then by isotonicity of $\Lambda_\mathcal{G}$ we have that $\Lambda_\mathcal{G}(X) \geq \Lambda_\mathcal{G}(Y)$. Suppose that $\Lambda_\mathcal{G}(X) = \Lambda_\mathcal{G}(Y)$ $\mathbb{P}$-a.s., then

$$\rho_\mathcal{G}(X) = \eta_\mathcal{G}(\Lambda_\mathcal{G}(X)) = \eta_\mathcal{G}(\Lambda_\mathcal{G}(Y)) = \rho_\mathcal{G}(Y)$$

which contradicts strict antitonicity of $\rho_\mathcal{G}$. Thus $\Lambda_\mathcal{G}$ fulfills all properties of a $\mathcal{G}$-conditional aggregation function.
As for $\eta_\mathcal{G}$, note that by Lemma 3.3.9 for all $F \in \operatorname{Im}\Lambda_\mathcal{G}$ we have that

$$\eta_\mathcal{G}(F) = \eta_\mathcal{G}\big(\Lambda_\mathcal{G}\big(f^{-1}_{\Lambda_\mathcal{G}}(F)\mathbf{1}_d\big)\big) = \rho_\mathcal{G}\big(f^{-1}_{\Lambda_\mathcal{G}}(F)\mathbf{1}_d\big), \tag{3.3.5}$$

where $f^{-1}_{\Lambda_\mathcal{G}}$ was defined in Definition 3.3.8. Since $\rho_\mathcal{G}$ and $f^{-1}_{\Lambda_\mathcal{G}}$ are strictly monotone, $\mathcal{G}$-local, and fulfill the Lebesgue property, so does $\eta_\mathcal{G}$, i.e. $\eta_\mathcal{G}$ is a univariate CRM on $\operatorname{Im}\Lambda_\mathcal{G}$. □

Theorem 3.3.12. *Let $\rho_\mathcal{G} : L^\infty_d(\mathcal{F}) \to L^\infty(\mathcal{G})$ and $\rho_\mathcal{F} : L^\infty_d(\mathcal{F}) \to L^\infty(\mathcal{F})$ be CRMs such that $\{\rho_\mathcal{G}, \rho_\mathcal{F}\}$ is strongly consistent. Moreover, suppose that*

$$f^{-1}_{\rho_\mathcal{F}} \circ \rho_\mathcal{F}(x) \in \mathbb{R} \quad \text{for all } x \in \mathbb{R}^d. \tag{3.3.6}$$

If $\rho_\mathcal{G}$ has a continuous realization $\rho_\mathcal{G}(\cdot,\cdot)$, then there exists a $\mathcal{G}$-conditional aggregation function $\Lambda_\mathcal{G} : L^\infty_d(\mathcal{F}) \to L^\infty(\mathcal{F})$ and a univariate CRM $\eta_\mathcal{G} : \operatorname{Im}\Lambda_\mathcal{G} \to L^\infty(\mathcal{G})$ such that

$$\rho_\mathcal{G}(X) = \eta_\mathcal{G}\big(\Lambda_\mathcal{G}(X)\big) \quad \text{for all } X \in L^\infty_d(\mathcal{F}) \tag{3.3.7}$$

and

$$\eta_\mathcal{G}\big(\Lambda_\mathcal{G}(X)\big) = -\Lambda_\mathcal{G}(X) \quad \text{for all } X \in L^\infty_d(\mathcal{G}).$$

Let $\Lambda_\mathcal{F} := -\rho_\mathcal{F}$ and $\eta_\mathcal{F} := -\operatorname{id}$ so that $\rho_\mathcal{F} = \eta_\mathcal{F} \circ \Lambda_\mathcal{F}$ for the $\mathcal{F}$-conditional aggregation function $\Lambda_\mathcal{F}$ and the univariate CRM $\eta_\mathcal{F}$. Then

$$\Lambda_\mathcal{F}(X) \leq \Lambda_\mathcal{F}(Y) \quad \Longrightarrow \quad \Lambda_\mathcal{G}(X) \leq \Lambda_\mathcal{G}(Y) \quad (X, Y \in L^\infty_d(\mathcal{F})), \tag{3.3.8}$$

i.e. $\Lambda_\mathcal{G}$ and $\Lambda_\mathcal{F}$ are strongly consistent.

Conversely, suppose that the CRM $\rho_\mathcal{G} : L^\infty_d(\mathcal{F}) \to L^\infty(\mathcal{G})$ satisfies (3.3.7)*, then $\{\rho_\mathcal{G}, \rho_\mathcal{F}\}$ is strongly consistent where $\rho_\mathcal{F} := -\Lambda_\mathcal{G}$ is a CRM.*

We remark that in Theorem 3.3.12 we require consistency of the pair $\{\rho_\mathcal{G}, \rho_\mathcal{F}\}$ where $\rho_\mathcal{F}$ is a CRM given the full information $\mathcal{F}$. Note that $\rho_\mathcal{F}$ is (apart from the sign) simply a conditional aggregation function as defined in Definition 3.3.6, so $\rho_\mathcal{G}$ is required to be consistent with some aggregation function under full information. This also explains $\Lambda_\mathcal{F}$. For $d = 1$ this consistency is automatically satisfied by monotonicity (and the aggregation is simply the identity function), and clearly the assertion is trivial anyway. For higher dimensions, Theorem 3.3.12 states that if there exists an aggregation function which is consistent with $\rho_\mathcal{G}$, then $\rho_\mathcal{G}$ is automatically of type (3.3.7). Clearly, if we already know that (3.3.7) holds true, then $\rho_\mathcal{G}$ is consistent with $\rho_\mathcal{F} = -\Lambda_\mathcal{G}$. Consistency with an aggregation under full information is a very natural requirement, because even under full information, so without risk, typically the losses still need to be aggregated in some way, and therefore any CRM under less information $\mathcal{G}$ should respect this aggregation.

Note also that the condition (3.3.6) is a slight strengthening of being normalized on constants, the latter being automatically satisfied by the very definition of the normalization $f_{\rho_{\mathcal{F}}}^{-1} \circ \rho_{\mathcal{F}}$; see above.

The following proof of Theorem 3.3.12 is based on two observations: $\rho_{\mathcal{F}}$ is necessarily risk-antitone as defined in Proposition 3.3.11. Strong consistency in turn implies that risk-antitonicity of $\rho_{\mathcal{F}}$ is passed on (backwards) to $\rho_{\mathcal{G}}$, and hence Proposition 3.3.11 applies.

Proof of Theorem 3.3.12*:* In case we already know that (3.3.7) holds, then by antitonicity of $\eta_{\mathcal{G}}$ it follows that $\{\rho_{\mathcal{G}}, -\Lambda_{\mathcal{G}}\}$ is strongly consistent, and clearly $-\Lambda_{\mathcal{G}} : L_d^{\infty}(\mathcal{F}) \to L^{\infty}(\mathcal{F})$ is also a CRM. Thus the last assertion of Theorem 3.3.12 is proved.

In order to show the first part of Theorem 3.3.12, we recall that the only property which remains to be shown in order to apply Proposition 3.3.11 is risk-antitonicity of $\rho_{\mathcal{G}}$: For this purpose we first consider simple random vectors $X, Y \in \mathcal{S}$ where $\mathcal{S}$ was defined in the proof of Proposition 3.3.11. Note that there is no loss of generality by assuming that $X = \sum_{i=1}^{n} x_i \mathbb{1}_{A_i} \in \mathcal{S}$ and $Y = \sum_{i=1}^{n} y_i \mathbb{1}_{A_i} \in \mathcal{S}$, i.e. the partition $(A_i)_{i=1,\dots,n}$ of Ω is the same for X and Y. Suppose that $\widetilde{\rho}_{\mathcal{G}}(X(\omega),\omega) \geq \widetilde{\rho}_{\mathcal{G}}(Y(\omega),\omega)$ $\mathbb{P}$-a.s. It follows that

$$\widetilde{\rho}_{\mathcal{G}}(x_i,\omega) \geq \widetilde{\rho}_{\mathcal{G}}(y_i,\omega) \quad \text{for all } \omega \in A_i \backslash N, i = 1, ..., n,$$

where N is a $\mathbb{P}$-nullset. We claim that this implies

$$f_{\rho_{\mathcal{G}}}^{-1}\big(\rho_{\mathcal{G}}(x_i)\big) \leq f_{\rho_{\mathcal{G}}}^{-1}\big(\rho_{\mathcal{G}}(y_i)\big) \quad \text{for all } i = 1, ..., n. \tag{3.3.9}$$

In order to verify this, we first notice that as $\rho_{\mathcal{G}}$ and $\rho_{\mathcal{F}}$ are strongly consistent and by (3.3.6) we have for all $x \in \mathbb{R}^d$ that

$$f_{\rho_{\mathcal{G}}}^{-1}\big(\rho_{\mathcal{G}}(x)\big) = f_{\rho_{\mathcal{G}}}^{-1}\big(\rho_{\mathcal{G}}\big(f_{\rho_{\mathcal{F}}}^{-1}\big(\rho_{\mathcal{F}}(x)\big)\mathbf{1}_d\big)\big) = f_{\rho_{\mathcal{F}}}^{-1}\big(\rho_{\mathcal{F}}(x)\big) \in \mathbb{R}. \tag{3.3.10}$$

Here we also used that the normalization $-f_{\rho_{\mathcal{G}}}^{-1} \circ \rho_{\mathcal{G}}$ is normalized on constants. In other words $f_{\rho_{\mathcal{G}}}^{-1}\big(\rho_{\mathcal{G}}(x_i)\big)$ and $f_{\rho_{\mathcal{G}}}^{-1}\big(\rho_{\mathcal{G}}(y_i)\big)$ are real numbers. Next we define $B_i := \{\omega \in \Omega \mid \widetilde{\rho}_{\mathcal{G}}(x_i,\omega) \geq \widetilde{\rho}_{\mathcal{G}}(y_i,\omega)\} \in \mathcal{G}$. Then $(A_i \setminus N) \subseteq B_i$ and hence $\mathbb{P}(B_i) > 0$ for all $i = 1, ..., n$. Using antitonicity and $\mathcal{G}$-locality of $f_{\rho_{\mathcal{G}}}^{-1}$ we obtain

$$f_{\rho_{\mathcal{G}}}^{-1}\big(\rho_{\mathcal{G}}(x_i)\big)\mathbb{1}_{B_i} = f_{\rho_{\mathcal{G}}}^{-1}\big(\rho_{\mathcal{G}}(x_i)\mathbb{1}_{B_i}\big)\mathbb{1}_{B_i} \leq f_{\rho_{\mathcal{G}}}^{-1}\big(\rho_{\mathcal{G}}(y_i)\mathbb{1}_{B_i}\big)\mathbb{1}_{B_i} = f_{\rho_{\mathcal{G}}}^{-1}\big(\rho_{\mathcal{G}}(y_i)\big)\mathbb{1}_{B_i}.$$

As $f_{\rho_{\mathcal{G}}}^{-1}\big(\rho_{\mathcal{G}}(y_i)\big)$ are indeed real numbers, (3.3.9) follows.

Now by strong consistency of $\{\rho_{\mathcal{G}}, \rho_{\mathcal{F}}\}$, $\mathcal{F}$-locality of $\rho_{\mathcal{F}}$ and $f_{\rho_{\mathcal{F}}}^{-1}$, and by (3.3.10) as well as antitonicity of $\rho_{\mathcal{G}}$ we obtain

$$\begin{aligned}
\rho_{\mathcal{G}}(X) &= \rho_{\mathcal{G}}\left(f_{\rho_{\mathcal{F}}}^{-1}\big(\rho_{\mathcal{F}}(X)\big)\mathbf{1}_d\right) = \rho_{\mathcal{G}}\left(\sum_{i=1}^{n} f_{\rho_{\mathcal{F}}}^{-1}\big(\rho_{\mathcal{F}}(x_i)\big)\mathbb{1}_{A_i}\mathbf{1}_d\right) \\
&= \rho_{\mathcal{G}}\left(\sum_{i=1}^{n} f_{\rho_{\mathcal{G}}}^{-1}\big(\rho_{\mathcal{G}}(x_i)\big)\mathbb{1}_{A_i}\mathbf{1}_d\right) \geq \rho_{\mathcal{G}}\left(\sum_{i=1}^{n} f_{\rho_{\mathcal{G}}}^{-1}\big(\rho_{\mathcal{G}}(y_i)\big)\mathbb{1}_{A_i}\mathbf{1}_d\right) \\
&= \rho_{\mathcal{G}}(Y),
\end{aligned}$$

which proves risk-antitonicity for simple random vectors $X, Y \in \mathcal{S}$. For general $X, Y \in L_d^\infty(\mathcal{F})$ with $\widetilde{\rho}_\mathcal{G}(X(\omega),\omega) \geq \widetilde{\rho}_\mathcal{G}(Y(\omega),\omega)$ for $\mathbb{P}$-a.e. $\omega \in \Omega$ we can find uniformly bounded sequences $(X_n)_{n\in\mathbb{N}}, (Y_n)_{n\in\mathbb{N}} \subset \mathcal{S}$ such that $X_n \nearrow X$ and $Y_n \searrow Y$ $\mathbb{P}$-a.s. for $n \to \infty$. Then by antitonicity

$$\widetilde{\rho}_\mathcal{G}(X_n(\omega),\omega) \geq \widetilde{\rho}_\mathcal{G}(X(\omega),\omega) \geq \widetilde{\rho}_\mathcal{G}(Y(\omega),\omega) \geq \widetilde{\rho}_\mathcal{G}(Y_n(\omega),\omega) \text{ for } \mathbb{P}\text{-a.s.}$$

Therefore, $\rho_\mathcal{G}(X_n) \geq \rho_\mathcal{G}(Y_n)$ and the Lebegue property of $\rho_\mathcal{G}$ yield

$$\rho_\mathcal{G}(X) = \lim_{n\to\infty} \rho_\mathcal{G}(X_n) \geq \lim_{n\to\infty} \rho_\mathcal{G}(Y_n) = \rho_\mathcal{G}(Y).$$

Thus $\rho_\mathcal{G}$ is risk-antitone and we apply Proposition 3.3.11. Hence, there is a $\mathcal{G}$-conditional aggregation function $\Lambda_\mathcal{G} : L_d^\infty(\mathcal{F}) \to L^\infty(\mathcal{F})$ and a univariate CRM $\eta_\mathcal{G} : \operatorname{Im}\Lambda_\mathcal{G} \to L^\infty(\mathcal{G})$ such that $\rho_\mathcal{G} = \eta_\mathcal{G} \circ \Lambda_\mathcal{G}$ and $\eta_\mathcal{G}\big(\Lambda_\mathcal{G}(X)\big) = -\Lambda_\mathcal{G}(X)$ for all $X \in L_d^\infty(\mathcal{G})$.
Let $X, Y \in L_d^\infty(\mathcal{F})$ such that

$$\Lambda_\mathcal{F}(X) = -\rho_\mathcal{F}(X) \leq -\rho_\mathcal{F}(Y) = \Lambda_\mathcal{F}(Y)$$

and let $(X_n)_{n\in\mathbb{N}} \subset \mathcal{S}$ and $(Y_n)_{n\in\mathbb{N}} \subset \mathcal{S}$ be uniformly bounded sequences such that $X_n \nearrow X$ and $Y_n \searrow Y$ $\mathbb{P}$-a.s. for $n \to \infty$. Again there is no loss in assuming that both X_n and Y_n for given $n \in \mathbb{N}$ are defined over the same partition, i.e. $X_n = \sum_{i=1}^{k_n} x_i^n \mathbb{1}_{A_i^n}$ and $Y_n = \sum_{i=1}^{k_n} y_i^n \mathbb{1}_{A_i^n}$. By the $\mathcal{F}$-locality and antitonicity of $\rho_\mathcal{F}$ it follows that for all $n \in \mathbb{N}$

$$\sum_{i=1}^{k_n} \rho_\mathcal{F}(x_i^n)\mathbb{1}_{A_i^n} \geq \rho_\mathcal{F}(X) \geq \rho_\mathcal{F}(Y) \geq \sum_{i=1}^{k_n} \rho_\mathcal{F}(y_i^n)\mathbb{1}_{A_i^n}.$$

As $f_{\rho_\mathcal{F}}^{-1} \circ \rho_\mathcal{F}(x_i^n)$ and $f_{\rho_\mathcal{F}}^{-1} \circ \rho_\mathcal{F}(y_i^n)$ are real numbers according to assumption (3.3.6) and as the above computation shows that $f_{\rho_\mathcal{F}}^{-1} \circ \rho_\mathcal{F}(x_i^n) \leq f_{\rho_\mathcal{F}}^{-1} \circ \rho_\mathcal{F}(y_i^n)$ on A_i^n, we obtain, as above that indeed $f_{\rho_\mathcal{F}}^{-1} \circ \rho_\mathcal{F}(x_i^n) \leq f_{\rho_\mathcal{F}}^{-1} \circ \rho_\mathcal{F}(y_i^n)$, $i = 1, \ldots, k_n$. Now strong consistency and (3.3.3) imply that

$$\begin{aligned}
-\Lambda_\mathcal{G}(x_i^n) &= \rho_\mathcal{G}(x_i^n) = \rho_\mathcal{G}(f_{\rho_\mathcal{F}}^{-1} \circ \rho_\mathcal{F}(x_i^n)) \\
&\geq \rho_\mathcal{G}(f_{\rho_\mathcal{F}}^{-1} \circ \rho_\mathcal{F}(y_i^n)) = \rho_\mathcal{G}(y_i^n) \\
&= -\Lambda_\mathcal{G}(y_i^n)
\end{aligned}$$

and hence by $\mathcal{G}$-locality of $\Lambda_\mathcal{G}$

$$\Lambda_\mathcal{G}(X_n) = \sum_{i=1}^{k_n} \Lambda_\mathcal{G}(x_i^n)\mathbb{1}_{A_i^n} \leq \sum_{i=1}^{k_n} \Lambda_\mathcal{G}(y_i^n)\mathbb{1}_{A_i^n} = \Lambda_\mathcal{G}(Y_n).$$

Finally we conclude with the Lebesgue property that

$$\Lambda_\mathcal{G}(X) = \lim_{n\to\infty} \Lambda_\mathcal{G}(X_n) \leq \lim_{n\to\infty} \Lambda_\mathcal{G}(Y_n) = \Lambda_\mathcal{G}(Y).$$

□

Remark 3.3.13. We know from Lemma 3.3.9 that the inverse function $f_{\Lambda_\mathcal{G}}^{-1}$ of $f_{\Lambda_\mathcal{G}}$ is isotone and that $\Lambda_\mathcal{G}(X) = \Lambda_\mathcal{G}\big(f_{\Lambda_\mathcal{G}}^{-1}(\Lambda_\mathcal{G}(X))\mathbf{1}_d\big)$ for all $X \in L_d^\infty(\mathcal{F})$. Therefore it can be shown as in Lemma 3.3.2, that (3.3.8) is equivalent to

$$f_{\Lambda_\mathcal{G}}^{-1}\big(\Lambda_\mathcal{G}(X)\big) = f_{\Lambda_\mathcal{F}}^{-1}\big(\Lambda_\mathcal{F}(X)\big), \text{ for all } X \in L_d^\infty(\mathcal{F}).$$

Note that we cannot write the recursive form of the strong consistency of two CRMs $\rho_\mathcal{G}$ and $\rho_\mathcal{F}$ as above, since $f_{\rho_\mathcal{G}}$ is only defined on $L^\infty(\mathcal{G})$ and not on $L^\infty(\mathcal{F})$ in constrast to $f_{\Lambda_\mathcal{G}}$.

In the following Theorem we summarize our findings from Proposition 3.3.11 and Theorem 3.3.12 on CRMs which extend the results in Hoffmann et al. (2016) for strong consistency:

Theorem 3.3.14. *If $\rho_\mathcal{G} : L_d^\infty(\mathcal{F}) \to L^\infty(\mathcal{G})$ is a CRM with a continuous realization $\rho_\mathcal{G}(\cdot,\cdot)$ and satisfies $f_{\rho_\mathcal{G}}^{-1} \circ \rho_\mathcal{G}(x) \in \mathbb{R}$ for all $x \in \mathbb{R}^d$, then the following three statements are equivalent*

(i) $\rho_\mathcal{G}(\cdot,\cdot)$ *is risk-antitone;*

(ii) $\rho_\mathcal{G}$ *is decomposable as in* (3.3.7)*;*

(iii) $\rho_\mathcal{G}$ *is strongly consistent with some aggregation function* $\Lambda : L_d^\infty(\mathcal{F}) \to L^\infty(\mathcal{F})$*, i.e.* $\{\rho_\mathcal{G}, -\Lambda\}$ *is strongly consistent.*

Proof. The equivalence of (ii) and (iii) has been shown in Theorem 3.3.12 and that (i) implies (ii) follows from Proposition 3.3.11. Finally, the proof of Theorem 3.3.12 shows that (iii) implies (i).

□

3.4 Conditional law-invariance and strong consistency

As in the previous section, if not otherwise stated, throughout this section we let $\mathcal{G}$ and $\mathcal{H}$ be two sub-σ-algebras of $\mathcal{F}$ such that $\mathcal{G} \subseteq \mathcal{H}$, and let $\rho_\mathcal{G} : L_d^\infty(\mathcal{F}) \to L^\infty(\mathcal{G})$ and $\rho_\mathcal{H} : L_d^\infty(\mathcal{F}) \to L^\infty(\mathcal{H})$ be the corresponding CRMs.

Definition 3.4.1. *A CRM $\rho_\mathcal{G}$ is conditional law-invariant if $\rho_\mathcal{G}(X) = \rho_\mathcal{G}(Y)$ whenever the $\mathcal{G}$-conditional distributions $\mu_X(\cdot|\mathcal{G})$ and $\mu_Y(\cdot|\mathcal{G})$ of $X, Y \in L_d^\infty(\mathcal{F})$ are equal, i.e. if $\mathbb{P}(X \in A \mid \mathcal{G}) = \mathbb{P}(Y \in A \mid \mathcal{G})$ for all Borel sets $A \in \mathcal{B}(\mathbb{R}^d)$. In case $\mathcal{G} = \{\emptyset, \Omega\}$ is trivial, conditional law-invariance of $\rho_\mathcal{G}$ is also referred to as law-invariance.*

In the law-invariant case we will often have to require a little more regularity of the underlying probability space $(\Omega, \mathcal{F}, \mathbb{P})$:

Definition 3.4.2. *We say that $(\Omega, \mathcal{F}, \mathbb{P})$ is*

atomless, *if $(\Omega, \mathcal{F}, \mathbb{P})$ supports a random variable with continuous distribution;*

conditionally atomless given $\mathcal{H} \subset \mathcal{F}$, *if $(\Omega, \mathcal{F}, \mathbb{P})$ supports a random variable with continuous distribution which is independent of $\mathcal{H}$.*

The next lemma shows that conditional law-invariance is passed from $\rho_{\mathcal{G}}$ (forward) to $\rho_{\mathcal{H}}$ by strong consistency. The proof is based on Föllmer (2014).

Lemma 3.4.3. *If $\{\rho_{\mathcal{G}}, \rho_{\mathcal{H}}\}$ is strongly consistent and $\rho_{\mathcal{G}}$ is conditionally law-invariant, then $\rho_{\mathcal{H}}$ is also conditionally law-invariant.*

Proof. Let $X, Y \in L^\infty(\mathcal{F})$ such that $\mu_X(\cdot|\mathcal{H}) = \mu_Y(\cdot|\mathcal{H})$ and let $A := \{\rho_{\mathcal{H}}(X) > \rho_{\mathcal{H}}(Y)\} \in \mathcal{H}$. Then the random variables $X\mathbb{1}_A$ and $Y\mathbb{1}_A$ have the same conditional distribution given $\mathcal{G}$. As $\rho_{\mathcal{G}}$ is conditionally law-invariant and strongly consistent with $\rho_{\mathcal{H}}$ we obtain

$$\begin{aligned}\rho_{\mathcal{G}}\Big(f_{\rho_{\mathcal{H}}}^{-1}\big(\rho_{\mathcal{H}}(X)\mathbb{1}_A + \rho_{\mathcal{H}}(\mathbf{0}_d)\mathbb{1}_{A^C}\big)\mathbf{1}_d\Big) &= \rho_{\mathcal{G}}(X\mathbb{1}_A) = \rho_{\mathcal{G}}(Y\mathbb{1}_A)\\ &= \rho_{\mathcal{G}}\Big(f_{\rho_{\mathcal{H}}}^{-1}\big(\rho_{\mathcal{H}}(Y)\mathbb{1}_A + \rho_{\mathcal{H}}(\mathbf{0}_d)\mathbb{1}_{A^C}\big)\mathbf{1}_d\Big).\end{aligned}$$

On the other hand, by strict antitonicity of $\rho_{\mathcal{G}}$ and $f_{\rho_{\mathcal{H}}}^{-1}$

$$\rho_{\mathcal{G}}\Big(f_{\rho_{\mathcal{H}}}^{-1}\big(\rho_{\mathcal{H}}(X)\mathbb{1}_A + \rho_{\mathcal{H}}(\mathbf{0}_d)\mathbb{1}_{A^C}\big)\mathbf{1}_d\Big) \geq \rho_{\mathcal{G}}\Big(f_{\rho_{\mathcal{H}}}^{-1}\big(\rho_{\mathcal{H}}(Y)\mathbb{1}_A + \rho_{\mathcal{H}}(\mathbf{0}_d)\mathbb{1}_{A^C}\big)\mathbf{1}_d\Big),$$

and the inequality is strict with positive probability if $\mathbb{P}(A) > 0$. Thus A must be a $\mathbb{P}$-nullset and interchanging X and Y in the definition of A shows that indeed $\rho_{\mathcal{H}}(X) = \rho_{\mathcal{H}}(Y)$. □

While in Theorem 3.3.12 we had to require that the strongly consistent pair $\{\rho_{\mathcal{G}}, \rho_{\mathcal{H}}\}$ satisfies $\mathcal{H} = \mathcal{F}$, in this section we in some sense require the opposite extreme, namely that $\mathcal{G} = \{\emptyset, \Omega\}$ is trivial while $\mathcal{H} \subseteq \mathcal{F}$.

Assumption 1. For the rest of the section we assume that $\mathcal{G} = \{\emptyset, \Omega\}$. For simplicity we will write $\rho := \rho_{\mathcal{G}} = \rho_{\{\emptyset,\Omega\}}$.

Lemma 3.4.4. *Let $\{\rho, \rho_{\mathcal{H}}\}$ be strongly consistent and suppose that ρ is law-invariant (and thus $\rho_{\mathcal{H}}$ is conditionally law-invariant by Lemma 3.4.3). If $(\Omega, \mathcal{H}, \mathbb{P})$ is an atomless probability space and $X \in L_d^\infty(\mathcal{F})$ is independent of $\mathcal{H}$, then*

$$f_{\rho_{\mathcal{H}}}^{-1}\big(\rho_{\mathcal{H}}(X)\big) = f_\rho^{-1}\big(\rho(X)\big).$$

The proof of Lemma 3.4.4 is adapted from Kupper and Schachermayer (2009).

Proof. We distinguish three cases:

- Suppose that $f_{\rho_{\mathcal{H}}}^{-1}\big(\rho_{\mathcal{H}}(X)\big) \leq f_\rho^{-1}\big(\rho(X)\big)$ and strictly smaller with positive probability. Then by strong consistency

  $$\begin{aligned}f_\rho^{-1}\big(\rho(X)\big) &= f_\rho^{-1}\left(\rho\left(f_{\rho_{\mathcal{H}}}^{-1}\big(\rho_{\mathcal{H}}(X)\big)\mathbf{1}_d\right)\right)\\ &< f_\rho^{-1}\left(\rho\left(f_\rho^{-1}\big(\rho(X)\big)\mathbf{1}_d\right)\right) = f_\rho^{-1}\big(\rho(X)\big),\end{aligned}$$

  by strict antitonicity of ρ which is a contradiction.

- Analogously it follows that it is not possible that $f_{\rho_{\mathcal{H}}}^{-1}\big(\rho_{\mathcal{H}}(X)\big) \geq f_\rho^{-1}\big(\rho(X)\big)$ and $\mathbb{P}(f_{\rho_{\mathcal{H}}}^{-1}\big(\rho_{\mathcal{H}}(X)\big) > f_\rho^{-1}\big(\rho(X)\big)) > 0$.

- There exist $A, B \in \mathcal{H}$ such that $\mathbb{P}(A) = \mathbb{P}(B) > 0$ and
$$f_{\rho_\mathcal{H}}^{-1}\big(\rho_\mathcal{H}(X)\big) > f_\rho^{-1}\big(\rho(X)\big) \text{ on } A \text{ and } f_{\rho_\mathcal{H}}^{-1}\big(\rho_\mathcal{H}(X)\big) < f_\rho^{-1}\big(\rho(X)\big) \text{ on } B.$$
Then we have for an arbitrary $m = a\mathbf{1}_d$ where $a \in \mathbb{R}$ that
$$\begin{aligned}\rho(X\mathbb{1}_A + m\mathbb{1}_{A^C}) &= \rho\left(f_{\rho_\mathcal{H}}^{-1}\big(\rho_\mathcal{H}(X\mathbb{1}_A + m\mathbb{1}_{A^C})\big)\mathbf{1}_d\right)\\ &= \rho\left(f_{\rho_\mathcal{H}}^{-1}\big(\rho_\mathcal{H}(X)\big)\mathbb{1}_A\mathbf{1}_d + m\mathbb{1}_{A^C}\right)\\ &< \rho\left(f_\rho^{-1}\big(\rho(X)\big)\mathbb{1}_A\mathbf{1}_d + m\mathbb{1}_{A^C}\right) \end{aligned} \tag{3.4.1}$$
and similarly
$$\rho(X\mathbb{1}_B + m\mathbb{1}_{B^C}) > \rho\left(f_\rho^{-1}\big(\rho(X)\big)\mathbb{1}_B\mathbf{1}_d + m\mathbb{1}_{B^C}\right). \tag{3.4.2}$$
However, as X is independent of $\mathcal{H}$ the random vector $X\mathbb{1}_A + m\mathbb{1}_{A^C}$ has the same distribution under $\mathbb{P}$ as $X\mathbb{1}_B + m\mathbb{1}_{B^C}$. Note that also $f_\rho^{-1}\big(\rho(X)\big)\mathbb{1}_A + a\mathbb{1}_{A^C}$ and $f_\rho^{-1}\big(\rho(X)\big)\mathbb{1}_B + a\mathbb{1}_{B^C}$ share the same distribution under $\mathbb{P}$. Hence, as ρ is law-invariant, (3.4.1) and (3.4.2) yield a contradiction.

□

Now we are able to extend the representation result of Föllmer (2014) to multivariate CRMs.

Theorem 3.4.5. *Let $(\Omega, \mathcal{H}, \mathbb{P})$ be atomless and let $(\Omega, \mathcal{F}, \mathbb{P})$ be conditionally atomless given $\mathcal{H}$. Suppose that ρ is law-invariant. Then, $\{\rho, \rho_\mathcal{H}\}$ is strongly consistent if and only if ρ and $\rho_\mathcal{H}$ are of the form*
$$\rho(X) = g\left(f_u^{-1}\big(\mathbb{E}_\mathbb{P}[u(X)]\big)\right) \quad \textit{for all } X \in L_d^\infty(\mathcal{F}) \tag{3.4.3}$$
and
$$\rho_\mathcal{H}(X) = g_\mathcal{H}\left(f_u^{-1}\big(\mathbb{E}_\mathbb{P}[u(X) \mid \mathcal{H}]\big)\right) \quad \textit{for all } X \in L_d^\infty(\mathcal{F}) \tag{3.4.4}$$
where $u : \mathbb{R}^d \to \mathbb{R}$ is strictly increasing and continuous, $f_u^{-1} : \operatorname{Im} f_u \to \mathbb{R}$ is the inverse function of
$$f_u : \mathbb{R} \to \mathbb{R}; \quad x \mapsto u(x\mathbf{1}_d)$$
and $g : \mathbb{R} \to \mathbb{R}$ and $g_\mathcal{H} : L^\infty(\mathcal{H}) \to L^\infty(\mathcal{H})$ are strictly antitone, fulfill the Lebesgue property, $0 \in \operatorname{Im} g \cap \operatorname{Im} g_\mathcal{H}$, and $g_\mathcal{H}$ is $\mathcal{H}$-local.

In particular, for any CRM of type (3.4.3) *(or* (3.4.4)*) we have that $g = f_\rho$ ($g_\mathcal{H} = f_{\rho_\mathcal{H}}$), where f_ρ and $f_{\rho_\mathcal{H}}$ are defined in Definition* 3.2.2.

The common function $u : \mathbb{R}^d \to \mathbb{R}$ appearing in (3.4.3) and (3.4.4) can be seen as a multivariate utility where u being strictly increasing means that $x, y \in \mathbb{R}^d$ with $x \geq y$ and $x \neq y$ implies $u(x) > u(y)$. So $f_u^{-1}\big(\mathbb{E}_\mathbb{P}[u(\cdot)]\big)$ and $f_u^{-1}\big(\mathbb{E}_\mathbb{P}[u(\cdot) \mid \mathcal{H}]\big)$ are (conditional) certainty equivalents – in the univariate case ($d = 1$) we clearly have $f_u^{-1} = u^{-1}$. Thus if ρ and/or $\rho_\mathcal{H}$ in Theorem 3.4.5 are normalized on constants (and hence $f_\rho \equiv -\operatorname{id}$ or $f_{\rho_\mathcal{H}} \equiv -\operatorname{id}$), then ρ and/or $\rho_\mathcal{H}$ equal (minus) certainty equivalents. But (3.4.3) and (3.4.4) also comprise other prominent classes of risk measures. For instance if $f_\rho = -f_u$ or $f_{\rho_\mathcal{H}} = -f_u$, then $\rho_\mathcal{H}(X) = -\mathbb{E}_\mathbb{P}[u(X)]$ is an multivariate expected utility whereas $\rho_\mathcal{H}(X) = -\mathbb{E}_\mathbb{P}[u(X) \mid \mathcal{H}]$ is a multivariate conditional expected utility.

Proof. For the last assertion of the theorem note that since u is a deterministic function, we have for $\alpha \in L^\infty(\mathcal{H})$ that

$$\begin{aligned} f_{\rho_{\mathcal{H}}}(\alpha) = \rho_{\mathcal{H}}(\alpha \mathbf{1}_d) &= g_{\mathcal{H}}\left(f_u^{-1}\big(\mathbb{E}_{\mathbb{P}}\left[u(\alpha \mathbf{1}_d) \mid \mathcal{H}\right]\big)\right) \\ &= g_{\mathcal{H}}\left(f_u^{-1}\big(f_u(\alpha)\big)\right) = g_{\mathcal{H}}(\alpha) \end{aligned}$$

and analogously we obtain $f_\rho \equiv g$.

Next we prove sufficiency in the first statement of the theorem: Let $\rho_{\mathcal{H}}$ and ρ be as in (3.4.4) and (3.4.3). It is easily verified that $\rho_{\mathcal{H}}$ and ρ are (conditionally) law-invariant CRMs. Furthermore, since f_u^{-1} is strictly increasing and $g_{\mathcal{H}}$ is strictly antitone and $\mathcal{H}$-local, we have for each $X, Y \in L_d^\infty(\mathcal{F})$ with $\rho_{\mathcal{H}}(X) \geq \rho_{\mathcal{H}}(Y)$ that

$$\mathbb{E}_{\mathbb{P}}\left[u(X) \mid \mathcal{H}\right] \leq \mathbb{E}_{\mathbb{P}}\left[u(Y) \mid \mathcal{H}\right].$$

But this implies that also $\mathbb{E}_{\mathbb{P}}\left[u(X)\right] \leq \mathbb{E}_{\mathbb{P}}\left[u(Y)\right]$ and thus that $\rho(X) \geq \rho(Y)$, i.e. $\{\rho, \rho_{\mathcal{H}}\}$ is strongly consistent.

Now we prove necessity in the first statement of the theorem: We assume in the following that ρ and $\rho_{\mathcal{H}}$ are normalized on constants and follow the approach of Föllmer (2014) Theorem 3.4. The idea is to introduce a preference order $\prec$ on multivariate distributions μ, ν on $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$ with bounded support given by

$$\mu \prec \nu \quad \Longleftrightarrow \quad \rho(X) > \rho(Y), \quad \text{with } X \sim \mu \text{ and } Y \sim \nu.$$

Here $\mathcal{B}(\mathbb{R}^d)$ denotes the Borel-σ-algebra on $\mathbb{R}^d$ and $X \sim \mu$ means that the distribution of $X \in L_d^\infty(\mathcal{F})$ under $\mathbb{P}$ is μ. It is well-known that if this preference order fulfills a set of conditions, then there exists a von Neumann-Morgenstern representation, that is

$$\mu \prec \nu \quad \Longleftrightarrow \quad \int u(x)\,\mu(dx) < \int u(x)\,\nu(dx), \tag{3.4.5}$$

where $u : \mathbb{R}^d \to \mathbb{R}$ is a continuous function. Sufficient conditions to guarantee (3.4.5) are that $\prec$ is continuous and fulfills the independence axiom; cf. Föllmer and Schied (2011) Corollary 2.28. We refer to Föllmer and Schied (2011) for a definition and comprehensive discussion of preference orders and the mentioned properties. Suppose for the moment that we have already proved (3.4.5). Note that strict antitonicity of ρ implies that $\delta_x \succ \delta_y$ whenever $x, y \in \mathbb{R}^d$ satisfy $x \geq y$ and $x \neq y$. Hence $u(x) = \int u(s)\,\delta_x(ds) > \int u(s)\,\delta_y(ds) = u(y)$, and we conclude that u is necessarily strictly increasing as claimed.

Now we prove (3.4.5): The proof of continuity of $\prec$ is completely analogous to the corresponding proof in Föllmer (2014) Theorem 3.4, so we omit it here. The crucial property is the independence axiom, which states that for any three distributions μ, ν, ϑ such that $\mu \preceq \nu$ and for all $\lambda \in (0, 1]$, we have

$$\lambda\mu + (1-\lambda)\vartheta \preceq \lambda\nu + (1-\lambda)\vartheta.$$

Since $(\Omega, \mathcal{F}, \mathbb{P})$ is conditionally atomless given $\mathcal{H}$, we can find $X, Y, Z \in L_d^\infty(\mathcal{F})$ which are independent of $\mathcal{H}$ such that $X \sim \mu, Y \sim \nu$ and $Z \sim \vartheta$. Furthermore, since $(\Omega, \mathcal{H}, \mathbb{P})$ is atomless, we can find an $A \in \mathcal{H}$ with $\mathbb{P}(A) = \lambda$. It can be easily seen that $X\mathbb{1}_A + Z\mathbb{1}_{A^C} \sim$

$\lambda\mu+(1-\lambda)\vartheta$ and $Y\mathbb{1}_A+Z\mathbb{1}_{A^C}\sim\lambda\nu+(1-\lambda)\vartheta$. Moreover, since $\mu\preceq\nu$, we have that $\rho(X)\geq\rho(Y)$. As $\{\rho,\rho_{\mathcal{H}}\}$ is strongly consistent and as ρ is law-invariant, we know from Lemma 3.4.3 that $\rho_{\mathcal{H}}$ is conditionally law-invariant. This ensures that we can apply Lemma 3.4.4 to the random vectors X and Y which are independent of $\mathcal{H}$. Therefore, by $\mathcal{H}$-locality of $\rho_{\mathcal{H}}$ and recalling Remark 3.3.4

$$\begin{aligned}\rho\left(X\mathbb{1}_A+Z\mathbb{1}_{A^C}\right)&=\rho\left(-\rho_{\mathcal{H}}\left(X\mathbb{1}_A+Z\mathbb{1}_{A^C}\right)\mathbf{1}_d\right)\\&=\rho\left(-\rho_{\mathcal{H}}(X)\mathbb{1}_A\mathbf{1}_d-\rho_{\mathcal{H}}(Z)\mathbb{1}_{A^C}\mathbf{1}_d\right)\\&=\rho\left(-\rho(X)\mathbb{1}_A\mathbf{1}_d-\rho_{\mathcal{H}}(Z)\mathbb{1}_{A^C}\mathbf{1}_d\right)\\&\geq\rho\left(-\rho(Y)\mathbb{1}_A\mathbf{1}_d-\rho_{\mathcal{H}}(Z)\mathbb{1}_{A^C}\mathbf{1}_d\right)\;=\;\rho\left(Y\mathbb{1}_A+Z\mathbb{1}_{A^C}\right),\end{aligned}$$

which is equivalent to $\lambda\mu+(1-\lambda)\vartheta\preceq\lambda\nu+(1-\lambda)\vartheta$. Thus there exists a von Neumann-Morgenstern representation (3.4.5) with a continuous and strictly increasing utility function $u:\mathbb{R}^d\to\mathbb{R}$.
In the next step we define $f_u:\mathbb{R}\to\mathbb{R};x\mapsto u(x\mathbf{1}_d)$. Then f_u is strictly increasing and continuous and thus f_u^{-1} exists. Let μ be an arbitrary distribution on $(\mathbb{R}^d,\mathcal{B}(\mathbb{R}^d))$ with bounded support and $X\sim\mu$. Then

$$\rho\big(\|X\|_{d,\infty}\mathbf{1}_d\big)\leq\rho(X)\leq\rho\big(-\|X\|_{d,\infty}\mathbf{1}_d\big)$$

and hence

$$\begin{aligned}f_u(-\|X\|_{d,\infty})=\int u(x)\;\delta_{-\|X\|_{d,\infty}\mathbf{1}_d}(dx)&\leq\int u(x)\;\mu(dx)\\&\leq\int u(x)\;\delta_{\|X\|_{d,\infty}\mathbf{1}_d}(dx)=f_u(\|X\|_{d,\infty}).\end{aligned}$$

The intermediate value theorem now implies the existence of a constant $c(\mu)\in\mathbb{R}$ such that

$$f_u\big(c(\mu)\big)=\int u(x)\;\mu(dx)\quad\iff\quad c(\mu)=f_u^{-1}\left(\int u(x)\;\mu(dx)\right).$$

Finally, since $\delta_{c(\mu)\mathbf{1}_d}\approx\mu$, we have

$$\rho(X)=\rho\big(c(\mu)\mathbf{1}_d\big)=-c(\mu)=-f_u^{-1}\left(\int u(x)\;\mu(dx)\right)=-f_u^{-1}\big(\mathbb{E}_{\mathbb{P}}\left[u(X)\right]\big).$$

Hence, we have proved (3.4.3) (with $g\equiv-\,\mathrm{id}$). Define

$$\psi_{\mathcal{H}}(X):=-f_u^{-1}\big(\mathbb{E}_{\mathbb{P}}\left[u(X)\mid\mathcal{H}\right]\big),\quad X\in L_d^\infty(\mathcal{F}),$$

then we have seen in the first part of the proof that $\psi_{\mathcal{H}}$ is a CRM which is strongly consistent with ρ. Moreover, $\psi_{\mathcal{H}}$ is normalized on constants. Thus it follows by Lemma 3.3.5 that $\rho_{\mathcal{H}}=\psi_{\mathcal{H}}$. If ρ and/or $\rho_{\mathcal{H}}$ are not normalized on constants, then considering the normalized CRMs $-f_\rho^{-1}\circ\rho$ and $-f_{\rho_{\mathcal{H}}}^{-1}\circ\rho_{\mathcal{H}}$ as introduced after Definition 3.2.6, the result follows from $\rho=f_\rho\circ\big(-(-f_\rho^{-1}\circ\rho)\big)$ and $\rho_{\mathcal{H}}=f_{\rho_{\mathcal{H}}}\circ\big(-(-f_{\rho_{\mathcal{H}}}^{-1}\circ\rho_{\mathcal{H}})\big)$, i.e. $g=f_\rho$ and $g_{\mathcal{H}}=f_{\rho_{\mathcal{H}}}$.

□

Recall Theorem 3.3.12 where we proved that if a multivariate CRM $\rho_{\mathcal{H}}$ is strongly consistent in a forward looking way with an aggregation $\rho_{\mathcal{F}}$ under full information $\mathcal{F}$ (and $\rho_{\mathcal{F}}$ fulfills (3.3.6)), then the multivariate CRM can be decomposed as in (3.3.7). The following Theorem 3.4.6 shows that we also obtain such a decomposition (3.3.7) under law-invariance by requiring strong consistency of $\rho_{\mathcal{H}}$ in a backward looking way with ρ given trivial information $\{\emptyset, \Omega\}$.

When stating Theorem 3.4.6 we will need an extension of $f_{\rho_{\mathcal{H}}}$ to $L^\infty(\mathcal{F})$: Suppose that the process $\mathbb{R} \ni a \mapsto f_{\rho_{\mathcal{H}}}(a)$ allows for a continuous realization. Due to the fact that $\rho_{\mathcal{H}}$ is strictly antitone and $\mathcal{H}$-local, we can find a possibly different realization $f_{\rho_{\mathcal{H}}}(\cdot,\cdot)$ such that $\widetilde{f}_{\rho_{\mathcal{H}}} : \mathbb{R} \times \Omega \to \mathbb{R} : x \mapsto f_{\rho_{\mathcal{H}}}(x, \omega)$ is continuous and strictly decreasing in the first argument for all $\omega \in \Omega$. Note that there exists a well-defined inverse $\widetilde{f}^{-1}_{\rho_{\mathcal{H}}}(\cdot, \omega)$ of $\widetilde{f}_{\rho_{\mathcal{H}}}(\cdot, \omega)$ for all $\omega \in \Omega$. Now define the functions

$$\bar{f}_{\rho_{\mathcal{H}}} : L^\infty(\mathcal{F}) \to L^\infty(\mathcal{F}); \quad F \mapsto \widetilde{f}_{\rho_{\mathcal{H}}}(F(\omega), \omega) \tag{3.4.6}$$

and

$$\bar{f}^{-1}_{\rho_{\mathcal{H}}} : \operatorname{Im} \bar{f}_{\rho_{\mathcal{H}}} \to L^\infty(\mathcal{F}); \quad F \mapsto \widetilde{f}^{-1}_{\rho_{\mathcal{H}}}(F(\omega), \omega),$$

where we with the standard abuse of notation identify the random variable $\widetilde{f}_{\rho_{\mathcal{H}}}(F(\omega), \omega)$ or $\widetilde{f}^{-1}_{\rho_{\mathcal{H}}}(F(\omega), \omega)$ with the equivalence classes they generate in $L^\infty(\mathcal{F})$.
By construction of $\bar{f}_{\rho_{\mathcal{H}}}$ we have that

$$\bar{f}_{\rho_{\mathcal{H}}}(L^\infty(\mathcal{J})) \subseteq L^\infty(\mathcal{J})$$

for all σ-algebras $\mathcal{J}$ such that $\sigma\left(f_{\rho_{\mathcal{H}}}(a, \cdot), a \in \mathbb{R}\right) \subseteq \mathcal{J} \subseteq \mathcal{F}$, c.f. Hoffmann et al. (2016) Lemma 3.1. By definition $\bar{f}_{\rho_{\mathcal{H}}}$ is also $\mathcal{F}$-local and has the Lebesgue property due to continuity of $\mathbb{R} \ni a \mapsto \widetilde{f}_{\rho_{\mathcal{H}}}(a, \omega)$. Moreover, $\mathcal{H}$-locality and continuity also imply that indeed $\bar{f}_{\rho_{\mathcal{H}}}(X) = f_{\rho_{\mathcal{H}}}(X)$ for all $X \in \mathcal{H}$ (approximation by simple random variables), so $\bar{f}_{\rho_{\mathcal{H}}}$ is indeed an extension of $f_{\rho_{\mathcal{H}}}$ to $L^\infty(\mathcal{F})$.

Theorem 3.4.6. *Under the same conditions as in Theorem* 3.4.5 *let $\{\rho, \rho_{\mathcal{H}}\}$ be strongly consistent. Then ρ can be decomposed as*

$$\rho = \eta \circ \Lambda,$$

where

$$\Lambda : L^\infty_d(\mathcal{F}) \to L^\infty(\mathcal{F}); \quad X \mapsto -f_\rho\left(f_u^{-1}\left(u(X)\right)\right)$$

is a $\{\emptyset, \Omega\}$-conditional aggregation function,

$$\eta : \operatorname{Im} \Lambda \to \mathbb{R}; \quad F \mapsto -U^{-1}\left(\mathbb{E}_{\mathbb{P}}\left[U(F)\right]\right)$$

is a law-invariant univariate certainty equivalent given by the (deterministic) utility

$$U : \operatorname{Im} \rho \to \mathbb{R}; \quad a \mapsto f_u\left(f_\rho^{-1}(-a)\right)$$

which is strictly increasing and continuous. Here $u : \mathbb{R}^d \to \mathbb{R}$ is the multivariate utility function from Theorem 3.4.5.
If the function $\mathbb{R} \ni a \mapsto f_{\rho_{\mathcal{H}}}(a)$ has a continuous realization, then $\rho_{\mathcal{H}}$ can be decomposed as

$$\rho_{\mathcal{H}} = \eta_{\mathcal{H}} \circ \Lambda_{\mathcal{H}},$$

with

$$\eta_{\mathcal{H}}\left(\Lambda_{\mathcal{H}}(X)\right) = -\Lambda_{\mathcal{H}}(X), \quad \textit{for all } X \in L_d^\infty(\mathcal{H}),$$

where

- $\Lambda_{\mathcal{H}} : L_d^\infty(\mathcal{F}) \to L^\infty(\mathcal{F}); X \mapsto -\bar{f}_{\rho_{\mathcal{H}}}\left(f_u^{-1}\left(u(X)\right)\right)$ *is a* $\sigma\left(f_{\rho_{\mathcal{H}}}(a,\cdot) : a \in \mathbb{R}\right)$*-conditional aggregation function (*$f_{\rho_{\mathcal{H}}}(a,\cdot)$ *denotes a continuous realization with strictly increasing paths);*
- $\eta_{\mathcal{H}} : \operatorname{Im}\Lambda_{\mathcal{H}} \to L^\infty(\mathcal{H}); \quad F \mapsto -U_{\mathcal{H}}^{-1}\left(\mathbb{E}_{\mathbb{P}}\left[U_{\mathcal{H}}(F) \mid \mathcal{H}\right]\right)$ *is a univariate conditional certainty equivalent;*
- *the stochastic utility* $U_{\mathcal{H}} : \operatorname{Im}\Lambda_{\mathcal{H}} \to L^\infty(\mathcal{F}); \quad F \mapsto f_u\left(\bar{f}_{\rho_{\mathcal{H}}}^{-1}(-F)\right)$ *is strictly isotone,* $\mathcal{F}$*-local, fulfills the Lebesgue property and* $U_{\mathcal{H}}^{-1}(\operatorname{Im}U_{\mathcal{H}} \cap L^\infty(\mathcal{H})) \subseteq L^\infty(\mathcal{H})$*;*
- $\bar{f}_{\rho_{\mathcal{H}}}$ *is given in* (3.4.6).

Moreover, it holds that

$$U_{\mathcal{H}} \circ \Lambda_{\mathcal{H}} = u = U \circ \Lambda \tag{3.4.7}$$

are deterministic and independent of the chosen information $\mathcal{H}$ *or* $\{\Omega, \emptyset\}$*.*
Finally we also have that $f_{\Lambda_{\mathcal{H}}}^{-1} \circ \Lambda_{\mathcal{H}} = f_u^{-1} \circ u = f_\Lambda^{-1} \circ \Lambda$*, i.e.* $\{\Lambda, \Lambda_{\mathcal{H}}\}$ *is strongly consistent as defined in* (3.3.8).

Proof. By Theorem 3.4.5 we have that

$$\begin{aligned}\rho_{\mathcal{H}}(X) &= f_{\rho_{\mathcal{H}}}\left(f_u^{-1}\left(\mathbb{E}_{\mathbb{P}}\left[u(X) \mid \mathcal{H}\right]\right)\right)\\ &= \bar{f}_{\rho_{\mathcal{H}}}\left(f_u^{-1}\left(\mathbb{E}_{\mathbb{P}}\left[f_u\left(\bar{f}_{\rho_{\mathcal{H}}}^{-1}\left(\bar{f}_{\rho_{\mathcal{H}}}\left(f_u^{-1}\left(u(X)\right)\right)\right)\right) \mid \mathcal{H}\right]\right)\right),\end{aligned}$$

where u and f_u are given in Theorem 3.4.5. Hence, recalling the definitions of $U_{\mathcal{H}}$, $\eta_{\mathcal{H}}$, and $\Lambda_{\mathcal{H}}$, we have $\rho_{\mathcal{H}} = \eta_{\mathcal{H}} \circ \Lambda_{\mathcal{H}}$. It can be readily seen that $U_{\mathcal{H}}$ as well as $U_{\mathcal{H}}^{-1}$, and thus also $\Lambda_{\mathcal{H}}$, are $\mathcal{F}$-local, strictly isotone, and fulfill the Lebesgue property. As $\bar{f}_{\rho_{\mathcal{H}}}(L^\infty(\mathcal{J})) \subseteq L^\infty(\mathcal{J})$ for all σ-algebras $\mathcal{J}$ such that $\sigma\left(f_{\rho_{\mathcal{H}}}(a,\cdot) : a \in \mathbb{R}\right) \subseteq \mathcal{J} \subseteq \mathcal{F}$, the same also applies to $\Lambda_{\mathcal{H}} = -\bar{f}_{\rho_{\mathcal{H}}} \circ f_u^{-1} \circ u$ and we conclude that $\Lambda_{\mathcal{H}}$ is a $\sigma\left(f_{\rho_{\mathcal{H}}}(a,\cdot) : a \in \mathbb{R}\right)$-conditional aggregation function. Moreover, for $X \in L_d^\infty(\mathcal{H})$

$$\eta_{\mathcal{H}}\left(\Lambda_{\mathcal{H}}(X)\right) = \bar{f}_{\rho_{\mathcal{H}}}\left(f_u^{-1}(u(X))\right) = -U_{\mathcal{H}}^{-1}\big(u(X)\big) = -\Lambda_{\mathcal{H}}(X).$$

The result for ρ follows similarly to the proof above without requiring a continuous realization and by using the canonical extension of f_ρ from $\mathbb{R}$ to $L_d^\infty(\mathcal{F})$, i.e. $\bar{f}_\rho(F)(\omega) = f_\rho(F(\omega))$ for all $\omega \in \Omega$ and $F \in L^\infty(\mathcal{F})$. □

We remark that (3.4.7) is the crucial fact which ensures that ρ and $\rho_{\mathcal{H}}$ are strongly consistent and (conditionally) law-invariant.

In Theorem 3.4.6 we have seen that basically every CRM which is strongly consistent with a law-invariant CRM under trivial information can be decomposed into a conditional aggregation function and a univariate conditional certainty equivalent. For the rest of this section we study the effect of additional properties of the CRMs on this decomposition. For instance, we want to identify conditions under which the univariate conditional certainty equivalent is generated by a deterministic (instead of a stochastic) utility function; see Corollary 3.4.7. Also we study what happens if the univariate CRMs η and $\eta_{\mathcal{H}}$ from Theorem 3.4.6 are required to be strongly consistent; see Corollary 3.4.9.

Corollary 3.4.7. *In the situation of Theorem* 3.4.6*, if ρ is normalized on constants, then*

$$\Lambda(X) = f_u^{-1}(u(X)), \quad X \in L_d^\infty(\mathcal{F}),$$

and

$$\eta(F) = \rho(F\mathbf{1}_d) = -f_u^{-1}(\mathbb{E}_\mathbb{P}[f_u(F)]), \quad F \in L^\infty(\mathcal{F}).$$

If $\rho_{\mathcal{H}}$ is normalized on constants, then similarly

$$\Lambda_{\mathcal{H}}(X) = f_u^{-1}(u(X)), \quad X \in L_d^\infty(\mathcal{F}),$$

and

$$\eta_{\mathcal{H}}(F) = \rho_{\mathcal{H}}(F\mathbf{1}_d) = -f_u^{-1}(\mathbb{E}_\mathbb{P}[f_u(F) \mid \mathcal{H}]), \quad F \in L^\infty(\mathcal{F}).$$

In particular the univariate conditional certainty equivalent $\eta_{\mathcal{H}}$ is now given by the deterministic univariate utility function f_u, and thus $\eta_{\mathcal{H}}$ is conditionally law-invariant.
If both ρ and $\rho_{\mathcal{H}}$ are normalized on constants, then $\Lambda = \Lambda_{\mathcal{H}}$.

Remark 3.4.8. Suppose that ρ and $\rho_{\mathcal{H}}$ from Theorem 3.4.6 are normalized on constants and that for all $F, G \in L^\infty(\mathcal{F})$, $m, \lambda \in \mathbb{R}$ with $\lambda \in (0,1)$

$$\rho(F\mathbf{1}_d + m\mathbf{1}_d) = \rho(F\mathbf{1}_d) - m \tag{3.4.8}$$

as well as

$$\rho\big(\lambda F\mathbf{1}_d + (1-\lambda)G\mathbf{1}_d\big) \leq \lambda\rho(F\mathbf{1}_d) + (1-\lambda)\rho(G\mathbf{1}_d). \tag{3.4.9}$$

Recalling Corollary 3.4.7 it follows that $\eta(F) = \rho(F\mathbf{1}_d)$ is cash-additive (3.4.8) and convex (3.4.9). Since f_u is a deterministic function it can be easily checked that η and $\eta_{\mathcal{H}}$ are strongly consistent (conditionally) law-invariant univariate CRMs. Therefore we are in the framework of Föllmer (2014). There it is shown that the univariate CRMs must be either linear or of entropic type, i.e.

$$f_u(x) = ax + b \quad \text{or} \quad f_u(x) = -ae^{-\beta x} + b, \quad x \in \mathbb{R},$$

for constants $a, b, \beta \in \mathbb{R}$ with $a, \beta > 0$, which implies that

$$\eta_{\mathcal{H}}(F) = \mathbb{E}_\mathbb{P}[-F \mid \mathcal{H}] \quad \text{or} \quad \eta_{\mathcal{H}}(F) = \frac{1}{\beta}\log\left(\mathbb{E}_\mathbb{P}\left[e^{-\beta F} \,\middle|\, \mathcal{H}\right]\right)$$

and similarly for η. Clearly, this also has consequences for the aggregation function $\Lambda = \Lambda_{\mathcal{H}} = f_u^{-1} \circ u$ since $x \mapsto u(x\mathbf{1}_d) = f_u(x)$ is either of linear or exponential form. For instance, a possible aggregation would be given by $u(x_1, \dots, x_d) = a\sum_{i=1}^d w_i x_i + b$, where $w_i \in (0,1)$ for $i = 1, ..., d$ such that $\sum_{i=1}^d w_i = 1$, because $f_u(x) = ax + b$. In this case the aggregation function is simply $\Lambda(x) = \sum_{i=1}^d w_i x_i$.

Corollary 3.4.9. *In the situation of Theorem* 3.4.6*, suppose that* η *and* $\eta_{\mathcal{H}}$ *are defined on all of* $L^\infty(\mathcal{F})$*. Then* $\{\eta, \eta_{\mathcal{H}}\}$ *are strongly consistent if and only if*

$$\eta = -\widetilde{u}^{-1}\left(\mathbb{E}_{\mathbb{P}}\left[\widetilde{u}(F)\right]\right) \quad \textit{and} \quad \eta_{\mathcal{H}} = -\widetilde{u}^{-1}\left(\mathbb{E}_{\mathbb{P}}\left[\widetilde{u}(F) \mid \mathcal{H}\right]\right)$$

for a continuous and strictly increasing utility function $\widetilde{u} : \mathbb{R} \to \mathbb{R}$*. Moreover, the corresponding (conditional) aggregation functions are given by*

$$\Lambda = -f_\rho \circ f_u^{-1} \circ u \quad \textit{and} \quad \Lambda_{\mathcal{H}} = -f_\rho \circ f_u^{-1} \circ a_{\mathcal{H}} \circ u,$$

where $a_{\mathcal{H}}(F) = \alpha F + \beta$*,* $F \in L^\infty(\mathcal{F})$*, is a positive affine transformation given by* $\alpha, \beta \in L^\infty(\mathcal{H})$ *with* $\mathbb{P}(\alpha > 0) = 1$.

Proof. As η is law-invariant, it follows from Lemma 3.4.3 that $\eta_{\mathcal{H}}$ is conditionally law-invariant. Moreover, $f_\eta \equiv f_{\eta_{\mathcal{H}}} \equiv -\operatorname{id}$, i.e. η and $\eta_{\mathcal{H}}$ are normalized on constants. Thus by Theorem 3.4.5 we obtain that

$$\eta = -\widetilde{u}^{-1}\left(\mathbb{E}_{\mathbb{P}}\left[\widetilde{u}(F)\right]\right) \quad \text{and} \quad \eta_{\mathcal{H}} = -\widetilde{u}^{-1}\left(\mathbb{E}_{\mathbb{P}}\left[\widetilde{u}(F) \mid \mathcal{H}\right]\right)$$

for a continuous and strictly increasing function $\widetilde{u} : \mathbb{R} \to \mathbb{R}$. It follows from Lemma 3.4.10 below that U as well as $U_{\mathcal{H}}$ are affine transformations of $\widetilde{u}$. This in turn implies that $U_{\mathcal{H}} = \tilde{a}_{\mathcal{H}} \circ U$, where $\tilde{a}_{\mathcal{H}}(F) = \tilde{\alpha} F + \tilde{\beta}$ for $\tilde{\alpha}, \tilde{\beta} \in L^\infty(\mathcal{H})$ with $\mathbb{P}(\tilde{\alpha} > 0) = 1$. Finally we obtain that the $\sigma\left(f_{\rho_{\mathcal{H}}}(a,\omega), a \in \mathbb{R}\right)$-conditional aggregation function $\Lambda_{\mathcal{H}}$ is given by

$$\Lambda_{\mathcal{H}} = U_{\mathcal{H}}^{-1} \circ u = U^{-1} \circ \tilde{a}_{\mathcal{H}}^{-1} \circ u = -f_\rho \circ f_u^{-1} \circ \tilde{a}_{\mathcal{H}}^{-1} \circ u.$$

Since the inverse $a_{\mathcal{H}} := \tilde{a}_{\mathcal{H}}^{-1}$ of an affine function is affine the result follows. □

Lemma 3.4.10. *Let* $U_{\mathcal{H}}$ *be the stochastic utility from Theorem* 3.4.6 *and let* $\widetilde{U}_{\mathcal{H}} : \operatorname{Im}\Lambda_{\mathcal{H}} \to L^\infty(\mathcal{F})$ *be another function which is strictly isotone,* $\mathcal{F}$*-local, fulfills the Lebesgue property and* $\widetilde{U}_{\mathcal{H}}(\operatorname{Im}\Lambda_{\mathcal{H}} \cap L^\infty(\mathcal{H})) \subseteq L^\infty(\mathcal{H})$*, such that*

$$\widetilde{U}_{\mathcal{H}}^{-1}\left(\mathbb{E}_{\mathbb{P}}\left[\widetilde{U}_{\mathcal{H}}(F) \,\middle|\, \mathcal{H}\right]\right) = U_{\mathcal{H}}^{-1}\left(\mathbb{E}_{\mathbb{P}}\left[U_{\mathcal{H}}(F) \mid \mathcal{H}\right]\right), \quad \textit{for all } F \in \operatorname{Im}\Lambda_{\mathcal{H}}. \tag{3.4.10}$$

Then $\widetilde{U}_{\mathcal{H}}$ *is an* $\mathcal{H}$*-measurable positive affine transformation of* $U_{\mathcal{H}}$*, i.e. there exist* $\alpha, \beta \in L^\infty(\mathcal{H})$ *with* $\mathbb{P}(\alpha > 0) = 1$ *such that* $\widetilde{U}_{\mathcal{H}}(F) = \alpha U_{\mathcal{H}}(F) + \beta$ *for all* $F \in \operatorname{Im}\Lambda_{\mathcal{H}}$.

Proof. We have seen in Theorem 3.4.6 that $U_{\mathcal{H}} \circ \Lambda_{\mathcal{H}} = u$, where u is strictly increasing and continuous. Thus

$$\mathcal{X} := \operatorname{Im} U_{\mathcal{H}} = u(L_d^\infty(\mathcal{F})) \subseteq L^\infty(\mathcal{F})$$

and it follows that for all $F \in \mathcal{X}$ there exists a sequence of $\mathcal{F}$-simple random variables $(F_n)_{n\in\mathbb{N}} \subseteq \mathcal{X}$ such that $F_n \to F$ $\mathbb{P}$-a.s. Moreover, by the intermediate value theorem we can find for each $X, Y \in L_d^\infty(\mathcal{F})$ and $\lambda \in L^\infty(\mathcal{F})$ with $0 \leq \lambda \leq 1$ a random variable Z such that $\min\{-\|X\|_{d,\infty}, -\|Y\|_{d,\infty}\} \leq Z \leq \max\{\|X\|_{d,\infty}, \|Y\|_{d,\infty}\}$ and for all $\mathbb{P}$-almost all $\omega \in \Omega$

$$\lambda(\omega) u\big(X(\omega)\big) + (1-\lambda) u\big(Y(\omega)\big) = u\big(Z(\omega)\mathbf{1}_d\big)$$

where $X(\cdot), Y(\cdot)$ and $\lambda(\cdot)$ are arbitrary representatives of X, Y and λ. Indeed, it can be shown by a measurable selection argument that Z can be chosen to be $\mathcal{F}$-measurable and hence $\mathcal{X}$ is $\mathcal{F}$-conditionally convex in the sense that $\lambda F + (1-\lambda)G \in \mathcal{X}$ for all $F, G \in \mathcal{X}$ and $\lambda \in L^\infty(\mathcal{F})$ with $0 \leq \lambda \leq 1$.

Next define the strictly isotone and $\mathcal{F}$-local function

$$V_\mathcal{H} : \mathcal{X} \to L^\infty(\mathcal{F});\ X \mapsto \widetilde{U}_\mathcal{H}\left(U_\mathcal{H}^{-1}(F)\right),$$

that is $\widetilde{U}_\mathcal{H} = V_\mathcal{H} \circ U_\mathcal{H}$. Moreover, it easily follows that $V_\mathcal{H}$ fulfills the Lebesgue property and $V_\mathcal{H}(\mathcal{X} \cap L^\infty(\mathcal{H})) \subseteq L^\infty(\mathcal{H})$. We show that $V_\mathcal{H}$ is an affine function, that is $V_\mathcal{H}(F) = \alpha F + \beta$ for all $F \in \mathcal{X}$, where $\alpha, \beta \in L^\infty(\mathcal{F})$. Note that affinity can be equivalently expressed via $V_\mathcal{H}(\lambda F + (1-\lambda)G) = \lambda V_\mathcal{H}(F) + (1-\lambda)V_\mathcal{H}(G)$ for all $F, G \in \mathcal{X}$ and $\lambda \in L^\infty(\mathcal{F})$ with $0 \leq \lambda \leq 1$.
We suppose that $V_\mathcal{H}$ is not affine, i.e. there are $F, G \in \mathcal{X}$ and $\lambda \in L^\infty(\mathcal{F})$ with $0 \leq \lambda \leq 1$ such that

$$\mathbb{P}\left(V_\mathcal{H}(\lambda F + (1-\lambda)G) \neq \lambda V_\mathcal{H}(F) + (1-\lambda)V_\mathcal{H}(G)\right) > 0. \tag{3.4.11}$$

First note that it suffices to assume that (3.4.11) holds for deterministic F, G and λ. To see this suppose that $V_\mathcal{H}$ is affine on deterministic values, but not on the whole of $\mathcal{X}$, i.e. (3.4.11) holds for some $F, G \in \mathcal{X}$ and $\lambda \in L^\infty(\mathcal{F})$ with $0 \leq \lambda \leq 1$. We know that there exist sequences of $\mathcal{F}$-simple functions $(F_n)_{n\in\mathbb{N}}, (G_n)_{n\in\mathbb{N}} \subset \mathcal{X} \cap \mathcal{S}$ and $(\lambda_n)_{n\in\mathbb{N}} \subset L^\infty(\mathcal{F}) \cap \mathcal{S}$ with $0 \leq \lambda_n \leq 1$ for all $n \in \mathbb{N}$ such that $F_n \to F, G_n \to G, \lambda_n \to \lambda$ $\mathbb{P}$-a.s., where $\mathcal{S}$ was defined in the proof of Proposition 3.3.11. Without loss of generality we might assume that $F_n = \sum_{i=1}^{k_n} F_i^n \mathbb{1}_{A_i^n}, G_n = \sum_{i=1}^{k_n} G_i^n \mathbb{1}_{A_i^n}$ and $\lambda_n = \sum_{i=1}^{k_n} \lambda_i^n \mathbb{1}_{A_i^n}$ have the same disjoint $\mathcal{F}$-partition $(A_i^n)_{i=1,\dots,k_n}$. By the $\mathcal{F}$-locality and Lebesgue property and since $F_i^n, G_i^n, \lambda_i^n \in \mathbb{R}$ for all $i = 1, \dots, k_n$ and $n \in \mathbb{N}$ we have

$$\begin{aligned}
V_\mathcal{H}(\lambda F + (1-\lambda)G) &= \lim_{n\to\infty} V_\mathcal{H}(\lambda_n F_n + (1-\lambda_n)G_n) \\
&= \lim_{n\to\infty} V_\mathcal{H}\left(\sum_{i=1}^{k_n} (\lambda_i^n F_i^n + (1-\lambda_i^n)G_i^n)\mathbb{1}_{A_i^n}\right) \\
&= \lim_{n\to\infty} \sum_{i=1}^{k_n} V_\mathcal{H}\big(\lambda_i^n F_i^n + (1-\lambda_i^n)G_i^n\big)\mathbb{1}_{A_i^n} \\
&= \lim_{n\to\infty} \sum_{i=1}^{k_n} \Big(\lambda_i^n V_\mathcal{H}(F_i^n) + (1-\lambda_i^n)V_\mathcal{H}(G_i^n)\Big)\mathbb{1}_{A_i^n} \\
&= \lim_{n\to\infty} \lambda_n V_\mathcal{H}(F_n) + (1-\lambda_n)V_\mathcal{H}(G_n) \\
&= \lambda V_\mathcal{H}(F) + (1-\lambda)V_\mathcal{H}(G),
\end{aligned}$$

which contradicts (3.4.11). Moreover we assume that $0 < \lambda < 1$ since otherwise this would also contradict (3.4.11). Finally, we assume w.l.o.g. that

$$A := \{V_\mathcal{H}(\lambda F + (1-\lambda)G) < \lambda V_\mathcal{H}(F) + (1-\lambda)V_\mathcal{H}(G)\} \in \mathcal{H}$$

has positive probability. Next define $H_1 := F\mathbb{1}_A + G\mathbb{1}_{A^C}$ and $H_2 := G$, then $H_i \in \mathcal{X} \cap L^\infty(\mathcal{H}), i = 1, 2$ and by $\mathcal{F}$-locality of $V_\mathcal{H}$

$$V_\mathcal{H}(\lambda H_1 + (1-\lambda)H_2) \leq \lambda V_\mathcal{H}(H_1) + (1-\lambda)V_\mathcal{H}(H_2)$$

and the inequality is strict with positive probability.

Since $(\Omega, \mathbb{P}, \mathcal{F})$ is conditionally atomless given $\mathcal{H}$ there exists a $B \in \mathcal{F}$ with $\mathbb{P}(B) = \lambda$ and which is independent of $\mathcal{H}$. Since $H_1, H_2 \in \mathcal{X}$ and $\mathcal{X}$ is $\mathcal{F}$-conditionally convex

$$H := H_1\mathbb{1}_B + H_2\mathbb{1}_{B^C} \in \mathcal{X}.$$

Now by $\mathcal{F}$-locality of $V_\mathcal{H}$, $V_\mathcal{H}(\mathcal{X} \cap L^\infty(\mathcal{H})) \subseteq L^\infty(\mathcal{H})$ and $B \perp\!\!\!\perp \mathcal{H}$ we get

$$\begin{aligned}
\mathbb{E}_\mathbb{P}\left[V_\mathcal{H}(H) \mid \mathcal{H}\right] &= \mathbb{E}_\mathbb{P}\left[V_\mathcal{H}\left(H_1\mathbb{1}_B + H_2\mathbb{1}_{B^C}\right) \mid \mathcal{H}\right] \\
&= V_\mathcal{H}(H_1)\mathbb{E}_\mathbb{P}\left[\mathbb{1}_B \mid \mathcal{H}\right] + V_\mathcal{H}(H_2)\mathbb{E}_\mathbb{P}\left[\mathbb{1}_{B^C} \mid \mathcal{H}\right] \\
&= V_\mathcal{H}(H_1)\mathbb{E}_\mathbb{P}\left[\mathbb{1}_B\right] + V_\mathcal{H}(H_2)\mathbb{E}_\mathbb{P}\left[\mathbb{1}_{B^C}\right] \\
&= \lambda V_\mathcal{H}(H_1) + (1-\lambda)V_\mathcal{H}(H_2) \\
&\geq V_\mathcal{H}(\lambda H_1 + (1-\lambda)H_2) \\
&= V_\mathcal{H}\left(\mathbb{E}_\mathbb{P}\left[H_1\mathbb{1}_B + H_2\mathbb{1}_{B^C} \mid \mathcal{H}\right]\right) \\
&= V_\mathcal{H}\left(\mathbb{E}_\mathbb{P}\left[H \mid \mathcal{H}\right]\right),
\end{aligned}$$

and the inequality is strict with positive probability. Moreover $\mathcal{X} = \operatorname{Im} U_\mathcal{H}$ implies the existence of a $\widetilde{H} \in \operatorname{Im} \Lambda_\mathcal{H}$ such that $H = U_\mathcal{H}(\widetilde{H})$. Finally we get

$$\begin{aligned}
\widetilde{U}_\mathcal{H}^{-1}\left(\mathbb{E}_\mathbb{P}\left[\widetilde{U}_\mathcal{H}(\widetilde{H}) \,\middle|\, \mathcal{H}\right]\right) &= U_\mathcal{H}^{-1}\left(V_\mathcal{H}^{-1}\left(\mathbb{E}_\mathbb{P}\left[V_\mathcal{H}\left(U_\mathcal{H}(\widetilde{H})\right) \,\middle|\, \mathcal{H}\right]\right)\right) \\
&= U_\mathcal{H}^{-1}\left(V_\mathcal{H}^{-1}\left(\mathbb{E}_\mathbb{P}\left[V_\mathcal{H}(H) \mid \mathcal{H}\right]\right)\right) \\
&\geq U_\mathcal{H}^{-1}\left(V_\mathcal{H}^{-1}\left(V_\mathcal{H}\left(\mathbb{E}_\mathbb{P}\left[H \mid \mathcal{H}\right]\right)\right)\right) \\
&= U_\mathcal{H}^{-1}\left(\mathbb{E}_\mathbb{P}\left[H \mid \mathcal{H}\right]\right) \\
&= U_\mathcal{H}^{-1}\left(\mathbb{E}_\mathbb{P}\left[U_\mathcal{H}(\widetilde{H}) \,\middle|\, \mathcal{H}\right]\right),
\end{aligned}$$

and the inequality is strict with positive probability, since $\widetilde{U}_\mathcal{H}^{-1}$ and $U_\mathcal{H}^{-1}$ are strictly isotone (c.f. Lemma 3.2.4). Thus we have the desired contradiction of (3.4.10) and hence $V_\mathcal{H}$ is affine, i.e. $V_\mathcal{H}(F) = \alpha F + \beta$ for all $F \in \mathcal{X}$, where $\alpha, \beta \in L^\infty(\mathcal{F})$. Moreover, since we know that $V_\mathcal{H}(x) \in L^\infty(\mathcal{H})$ for all $x \in \mathbb{R} \cap \mathcal{X}$, we obtain that α, β are actually $\mathcal{H}$-measurable. That $\alpha > 0$ follows immediately from the fact that $\widetilde{U}_\mathcal{H}, U_\mathcal{H}^{-1}$ are strictly isotone. □

Remark 3.4.11. Our notion of consistency is defined in terms of the multivariate CRMs. In contrast in Kromer et al. (2014) it is a priori assumed that the multivariate CRMs are of the decomposable form $\rho = \eta \circ \Lambda$ as in (3.3.7) and they define "consistency" of $\{\rho_\mathcal{G}, \rho_\mathcal{H}\}$ by requiring strong consistency of both pairs $\{\eta_\mathcal{G}, \eta_\mathcal{H}\}$ and $\{\Lambda_\mathcal{G}, \Lambda_\mathcal{H}\}$. Note that these definitions of consistency are not equivalent, in particular strong consistency of both $\{\eta_\mathcal{G}, \eta_\mathcal{H}\}$ and $\{\Lambda_\mathcal{G}, \Lambda_\mathcal{H}\}$ does not imply strong consistency of $\{\rho_\mathcal{G}, \rho_\mathcal{H}\}$. Kromer et al. (2014) also study the interplay of the strong consistency of $\{\rho_\mathcal{G}, \rho_\mathcal{H}\}$ and of strong consistency of both $\{\eta_\mathcal{G}, \eta_\mathcal{H}\}$ and $\{\Lambda_\mathcal{G}, \Lambda_\mathcal{H}\}$. As Corollary 3.4.9 shows in the law-invariant case this requirement is quite restrictive.

3.5 Consistency of a family of conditional risk measures

So far we only considered consistency for two multivariate CRMs. In this section we extend our results on strong consistency to families of multivariate CRMs. We begin with some motivating examples.

Example 3.5.1 (Dynamic risk measures)**.** If one is interested in a dynamic risk measurement under growing information in time up to a terminal time $T > 0$, this can be modeled by a family of CRMs $(\rho_t)_{t\in[0,T]}$ and a filtration $(\mathcal{F}_t)_{t\in[0,T]}$ such that $\rho_t : L_d^\infty(\mathcal{F}_T) \to L^\infty(\mathcal{F}_t)$.

In systemic risk measurement conditioning on varying information in space rather than in time is of interest. In that situation, as opposed to Example 3.5.1, the family of multivariate CRMs is not necessarily indexed by a filtration. To exemplify this we recall a multivariate version of the spatial risk measures which have been introduced by Föllmer (2014) in a univariate framework.

Example 3.5.2 (Multivariate spatial risk measures)**.** Let $I = \{1, ..., d\}$ denote a set of financial institutions and let $(S, \mathcal{S})$ be a measurable space. Each financial institution $i \in I$ can be in some state $s \in S$, and $\Omega = S^I = \{\omega = (\omega)_{i\in I} : \omega_i \in S\}$ denotes all possible states of the system. Then the σ-algebra $\mathcal{F}_J$ on Ω which is generated by the canonical projections on the j-th coordinate for $j \in J$ describes the observable information within the subsystem of financial institutions $J \subseteq I$. Finally let $\mathbb{P}$ be a probability measure on $(\Omega, \mathcal{F})$, where $\mathcal{F} := \mathcal{F}_I$. Then the risk evolution under varying spatial information can be modeled by the family of CRMs $(\rho_J)_{J\subseteq I}$, where each $\rho_J : L_d^\infty(\mathcal{F}) \to L^\infty(\mathcal{F}_J)$, i.e. ρ_J is the risk of the system given the information on the state of the financial institutions within the subsystem J.

From the viewpoint of a regulator, systemic risk measurement contingent on information in space is helpful in identifying systemic relevant structures, i.e. in analyzing questions like: "How much is the system affected given that a specific institution or subgroup of institutions is in distress?", or "How resilient is a specific institution or subgroup of institutions given that the system is in distress?". In Example 3.5.2 the spatial conditioning is based on a σ-algebra which is generated by all possible states of the institutions within a given subsystem. To treat questions of the type mentioned before one might alternatively consider conditioning with respect to more granular information in space. For instance, in the spirit of the systemic risk measures CoVaR in Adrian and Brunnermeier (2016) or Systemic Expected Shortfall in Acharya et al. (2017) one could condition on a single crisis event with respect to a given subsystem, e.g. that all financial institutions within the subsystem are below their individual value-at-risk levels.

In Example 3.5.1 as well as Example 3.5.2 the families of CRMs are indexed by one-dimensional information structure. However, in Frittelli and Maggis (2011), they propose conditional certainty equivalents based on a two-dimensional information structure caused by the fact that utilities of agents may vary over time:

Example 3.5.3 (Conditional certainty equivalents)**.** Let $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in\mathbb{R}^+}, \mathbb{P})$ be an atomless filtered probability space and let $u_t : \mathbb{R} \times \Omega \to \mathbb{R}$ be a function which is strictly increasing and continuous in the first argument and $\mathcal{F}_t$-measurable in the second argument for all $t \in \mathbb{R}^+$. Suppose that the range $\mathcal{R}_t := \{u_t(x, \omega) : x \in \mathbb{R}\}$ is independent of $\omega \in \Omega$, that $\mathcal{R}_t \subseteq \mathcal{R}_s$ for all $s \leq t$, and denote the pathwise inverse function of u by $u_t^{-1}(y) \in L^\infty(\mathcal{F}_t)$ for all $y \in \mathcal{R}_t$, where $u_t(x)$ and $u_t^{-1}(y)$ is the shorthand for $u_t(x, \cdot)$ and $u_t^{-1}(y, \cdot)$, resp. Then the backward conditional certainty equivalent is given by

$$C_{s,t} : L^\infty(\mathcal{F}_t) \to L^\infty(\mathcal{F}_s); F \mapsto C_{s,t}(F) = -u_s^{-1}\big(\mathbb{E}_\mathbb{P}\left[u_t(F) \mid \mathcal{F}_s\right]\big).$$

It has been shown in Frittelli and Maggis (2011) Proposition 1.1 that for a fixed $T \in \mathbb{R}^+$, we have that the family $(C_{t,T})_{t \leq T}$ is consistent, i.e. for all $s \leq t \leq T$

$$C_{t,T}(F) \geq C_{t,T}(G) \Longrightarrow C_{s,T}(F) \geq C_{s,T}(G) \quad (F, G \in L^\infty(\mathcal{F}_T)).$$

Also in the context of conditioning on spatial information a two-dimensional information structure could be of interest, for example to represent risk measurement policies that differ locally in the financial system.

Example 3.5.4 (Local regulatory policies)**.** In the context of Example 3.5.2, let $I = \{1, ..., d\}$ be a network of financial institutions that is of interest for supervisory authorities associated to different levels with possibly different regulatory policies. For example, think of I as the European financial system. Then regulatory policies of authorities on the European level might differ from policies on the national levels which again might differ from regional policies. To include these different regulatory viewpoints into the framework of spatial risk measures one could consider a family of CRMs $(\rho_{J,K})_{J \subseteq K \subseteq I}$, where each $\rho_{J,K} : L_d^\infty(\mathcal{F}_K) \to L^\infty(\mathcal{F}_J)$. Here the first index J has the same meaning as in Example 3.5.2, i.e. the risk measurement is performed conditioned on the state of the institutions in subsystem J. The second index K identifies the type of regulatory policy on the risk management prevailing in subsystem K, for example expected shortfall measures at different significance levels according to European ($K = I$), national, or regional standards. Even though regulatory policies may differ depending on the level of authority, it might still be desirable that these policies behave consistently in some way, i.e. the family $(\rho_{J,K})_{J \subseteq K \subseteq I}$ should be consistent not only with respect to the contingent information implied by the index J but also with respect to the different policies implied by the index K. In the following, this question will be considered.

Motivated by the examples above, we will consider the following types of families of CRMs in this section: Let $\mathcal{I}_1$ and $\mathcal{I}_2$ be sets of sub-σ-algebras of $\mathcal{F}$ such that $\mathcal{I}_1$ contains the trivial σ-algebra and denote by $\mathcal{E} := \{(\mathcal{H}, \mathcal{T}) \in \mathcal{I}_1 \times \mathcal{I}_2 : \mathcal{H} \subseteq \mathcal{T}\}$. In the following we denote by $\rho_{\mathcal{H},\mathcal{T}}$ a multivariate CRM which maps $L_d^\infty(\mathcal{T})$ to $L^\infty(\mathcal{H})$ and we consider families of CRMs of type $(\rho_{\mathcal{H},\mathcal{T}})_{(\mathcal{H},\mathcal{T}) \in \mathcal{E}}$. In order to allow for a comparison of the risks of two random risk factors under different information, we assume for the rest of this section that $\rho_{\mathcal{H},\mathcal{T}_1}(L_d^\infty(\mathcal{T}_1)) = \rho_{\mathcal{H},\mathcal{T}_2}(L_d^\infty(\mathcal{T}_2))$ for all $(\mathcal{H}, \mathcal{T}_1), (\mathcal{H}, \mathcal{T}_2) \in \mathcal{E}$. Sometimes it will also be convenient to consider only a subfamily of $\mathcal{E}$ where the second σ-algebra is fixed. In that case we denote the corresponding index set by $\mathcal{E}(\mathcal{T}) := \{\mathcal{H} \in \mathcal{I}_1 : \mathcal{H} \subseteq \mathcal{T}\}$ for $\mathcal{T} \in \mathcal{I}_2$. Note that the structure of the families of CRMs discussed in Example 3.5.1 and Example 3.5.2 is covered by this framework by letting $\mathcal{I}_2 := \{\mathcal{F}\}$.

Definition 3.5.5. *A family of CRMs $(\rho_{\mathcal{H},\mathcal{T}})_{(\mathcal{H},\mathcal{T}) \in \mathcal{E}}$ is strongly consistent if for all $\mathcal{G} \subseteq \mathcal{H} \subseteq \mathcal{T}_1 \cap \mathcal{T}_2$*

$$\rho_{\mathcal{H},\mathcal{T}_1}(X) \geq \rho_{\mathcal{H},\mathcal{T}_2}(Y) \Longrightarrow \rho_{\mathcal{G},\mathcal{T}_1}(X) \geq \rho_{\mathcal{G},\mathcal{T}_2}(Y), \quad (X \in L^\infty(\mathcal{T}_1), Y \in L_d^\infty(\mathcal{T}_2)).$$

It can be easily checked that the conditional certainty equivalents of Frittelli and Maggis (2011) (see Example 3.5.3) are strongly consistent. Analogously to Lemma 3.3.2 strong consistency is equivalent to the following recursive relation between the CRMs.

Lemma 3.5.6. *Let $(\rho_{\mathcal{H},\mathcal{T}})_{(\mathcal{H},\mathcal{T})\in\mathcal{E}}$ be family of CRMs, then the following statements are equivalent:*

(i) $(\rho_{\mathcal{H},\mathcal{T}})_{(\mathcal{H},\mathcal{T})\in\mathcal{E}}$ *is strongly consistent;*

(ii) For all $\mathcal{G} \subseteq \mathcal{H} \subseteq \mathcal{T}_1 \cap \mathcal{T}_2$ and $X \in L_d^\infty(\mathcal{T}_1)$

$$\rho_{\mathcal{G},\mathcal{T}_1}(X) = \rho_{\mathcal{G},\mathcal{T}_2}\left(f^{-1}_{\rho_{\mathcal{H},\mathcal{T}_2}}\big(\rho_{\mathcal{H},\mathcal{T}_1}(X)\big)\mathbf{1}_d\right).$$

Clearly, our results from the previous sections carry over to families of CRM. We illustrate this in the following by giving the straightforward extensions of Theorem 3.3.12 and Theorem 3.4.5 to a family of CRMs.

Theorem 3.5.7. *Let $(\rho_{\mathcal{H},\mathcal{T}})_{(\mathcal{H},\mathcal{T})\in\mathcal{E}}$ be a family of strongly consistent CRMs. Moreover, if there exists a $\mathcal{T} \in \mathcal{I}_2$ such that*

$$f^{-1}_{\rho_{\mathcal{T},\mathcal{T}}} \circ \rho_{\mathcal{T},\mathcal{T}}(x) \in \mathbb{R}, \quad \forall x \in \mathbb{R}^d, \tag{3.5.1}$$

then each multivariate CRM $\rho_{\mathcal{H},\mathcal{T}}$ of the subfamily $(\rho_{\mathcal{H},\mathcal{T}})_{\mathcal{H}\in\mathcal{E}(\mathcal{T})}$ which has a continuous realization $\rho_{\mathcal{H},\mathcal{T}}(\cdot,\cdot)$ can be decomposed into a $\mathcal{H}$-conditional aggregation function $\Lambda_{\mathcal{H},\mathcal{T}} : L_d^\infty(\mathcal{T}) \to L^\infty(\mathcal{T})$ and a univariate CRM $\eta_{\mathcal{H},\mathcal{T}} : \operatorname{Im}\Lambda_{\mathcal{H},\mathcal{T}} \to L^\infty(\mathcal{H})$ such that

$$\rho_{\mathcal{H},\mathcal{T}} = \eta_{\mathcal{H},\mathcal{T}} \circ \Lambda_{\mathcal{H},\mathcal{T}}$$

and $\rho_{\mathcal{H},\mathcal{T}}(X) = \eta_{\mathcal{H},\mathcal{T}}\big(\Lambda_{\mathcal{H},\mathcal{T}}(X)\big) = -\Lambda_{\mathcal{H},\mathcal{T}}(X)$ for all $X \in L_d^\infty(\mathcal{H})$. Moreover, for those $\rho_{\mathcal{H},\mathcal{T}}, \mathcal{H} \in \mathcal{E}(\mathcal{T})$, for which a decomposition exists the corresponding conditional aggregation functions are strongly consistent.

Theorem 3.5.8. *Let $(\rho_{\mathcal{H},\mathcal{T}})_{(\mathcal{H},\mathcal{T})\in\mathcal{E}}$ be a family of CRMs. Furthermore, suppose that there exists an $(\mathcal{G},\mathcal{T}) \in \mathcal{E}$ such that $(\Omega,\mathcal{T},\mathbb{P})$ is a conditionally atomless probability space given $\mathcal{G}$, $(\Omega,\mathcal{G},\mathbb{P})$ is atomless and $\rho_\mathcal{T} := \rho_{\{\emptyset,\Omega\},\mathcal{T}}$ is law-invariant. Then the subfamily $(\rho_{\mathcal{H},\mathcal{T}})_{\mathcal{H}\in\mathcal{E}(\mathcal{T})}$ is strongly consistent if and only if for each $\mathcal{H} \in \mathcal{E}(\mathcal{T})$ the CRM $\rho_{\mathcal{H},\mathcal{T}}$ is of the form*

$$\rho_{\mathcal{H},\mathcal{T}}(X) = g_{\mathcal{H},\mathcal{T}}\left(f^{-1}_{u_\mathcal{T}}\big(\mathbb{E}_\mathbb{P}\left[u_\mathcal{T}(X) \mid \mathcal{H}\right]\big)\right), \quad \text{for all } X \in L_d^\infty(\mathcal{T}), \tag{3.5.2}$$

where $u_\mathcal{T} : \mathbb{R}^d \to \mathbb{R}$ is strictly increasing and continuous, $f^{-1}_{u_\mathcal{T}} : \operatorname{Im} f_{u_\mathcal{T}} \to \mathbb{R}$ is the unique inverse function of $f_{u_\mathcal{T}} : \mathbb{R} \to \mathbb{R}; x \mapsto u_\mathcal{T}(x\mathbf{1}_d)$ and $g_{\mathcal{H},\mathcal{T}} : L^\infty(\mathcal{H}) \to L^\infty(\mathcal{H})$ is strictly antitone, $\mathcal{H}$-local, fulfills the Lebesgue property and $0 \in \operatorname{Im} g_{\mathcal{H},\mathcal{T}}$.
In particular, for any CRM of type (3.5.2) *we have that $g_{\mathcal{H},\mathcal{T}} = f_{\rho_{\mathcal{H},\mathcal{T}}}$, where $f_{\rho_{\mathcal{H},\mathcal{T}}}$ is defined in Definition* 3.2.2.

Note that the latter results, being extensions from the two-CRM-case of the previous sections, only used the strong consistency as a pairwise strong consistency of the elements in subfamilies $(\rho_{\mathcal{H},\mathcal{T}})_{(\mathcal{H},\mathcal{T})\in\mathcal{E}(\mathcal{T})}$ of $(\rho_{\mathcal{H},\mathcal{T}})_{(\mathcal{H},\mathcal{T})\in\mathcal{E}}$. But if $\mathcal{I}_2$ contains more than just one σ-algebra, then the definition of strong consistency given in Definition 3.5.5 also has implications on the relations between these subfamilies corresponding to different sets $\mathcal{E}(\mathcal{T})$ for $\mathcal{T} \in \mathcal{I}_2$.

Assumption 2. In order to have sufficiently many subfamilies we suppose for the remainder of this section that $\mathcal{I}_1 = \mathcal{I}_2 =: \mathcal{I}$.

Proposition 3.5.9. *Let $(\rho_{\mathcal{H},\mathcal{T}})_{(\mathcal{H},\mathcal{T})\in\mathcal{E}}$ be a strongly consistent family such that* (3.5.2) *holds for all $(\mathcal{H},\mathcal{T}) \in \mathcal{E}$. Then for all $\mathcal{T}_1, \mathcal{T}_2 \in \mathcal{I}$ and $\mathcal{H} \in \mathcal{T}_1 \cap \mathcal{T}_2$, $\mathcal{H} \in \mathcal{I}$,*

$$\rho_{\mathcal{H},\mathcal{T}_1}(X) = f_{\rho_{\mathcal{H},\mathcal{T}_2}}\left(f^{-1}_{u_{\mathcal{T}_2}}\left(a_{\mathcal{T}_1,\mathcal{T}_2}\mathbb{E}_{\mathbb{P}}\left[u_{\mathcal{T}_1}(X) \mid \mathcal{H}\right] + b_{\mathcal{H},\mathcal{T}_1,\mathcal{T}_2}\right)\right),$$

where $a_{\mathcal{T}_1,\mathcal{T}_2} \in \mathbb{R}^+\backslash\{0\}$, $b_{\mathcal{H},\mathcal{T}_1,\mathcal{T}_2} \in L^\infty(\mathcal{H})$ and $\mathbb{E}_{\mathbb{P}}\left[b_{\mathcal{H},\mathcal{T}_1,\mathcal{T}_2} \mid \mathcal{G}\right] = b_{\mathcal{G},\mathcal{T}_1,\mathcal{T}_2}$ for all $\mathcal{G} \in \mathcal{I}$ with $\mathcal{G} \subseteq \mathcal{H}$.

In order to prove Proposition 3.5.9 we need some auxiliary lemmas and therefore the proof is deferred to the end of this section. From Proposition 3.5.9 it follows that any strongly consistent family $(\rho_{\mathcal{H},\mathcal{T}})_{(\mathcal{H},\mathcal{T})\in\mathcal{E}}$ (under Assumption 2) is basically a family of conditional certainty equivalents as in Frittelli and Maggis (2011):

Corollary 3.5.10. *In the situation of Proposition* 3.5.9, *if $a_{\mathcal{T}_1,\mathcal{T}_2} = 1$, $b_{\mathcal{H},\mathcal{T}_1,\mathcal{T}_2} = 0$ for all $\mathcal{H} \subseteq \mathcal{T}_1 \cap \mathcal{T}_2$ where $\mathcal{H} \in \mathcal{I}$ and $\mathcal{T}_1, \mathcal{T}_2 \in \mathcal{I}$, and if $\rho_{\mathcal{T},\mathcal{T}}$ are normalized on constants for all $\mathcal{T} \in \mathcal{I}$, then $(\rho_{\mathcal{H},\mathcal{T}})_{(\mathcal{H},\mathcal{T})\in\mathcal{E}}$ satisfies*

$$\rho_{\mathcal{H},\mathcal{T}}(X) = -f^{-1}_{u_{\mathcal{H}}}\left(\mathbb{E}_{\mathbb{P}}\left[u_{\mathcal{T}}(X) \mid \mathcal{H}\right]\right), \quad X \in L^\infty_d(\mathcal{T}). \tag{3.5.3}$$

Proof. If $a_{\mathcal{T}_1,\mathcal{T}_2} = 1$ and $b_{\mathcal{H},\mathcal{T}_1,\mathcal{T}_2} = 0$ for all $\mathcal{H} \subseteq \mathcal{T}_1 \cap \mathcal{T}_2$, then

$$\rho_{\mathcal{H},\mathcal{T}_1}(X) = f_{\rho_{\mathcal{H},\mathcal{T}_2}}\left(f^{-1}_{u_{\mathcal{T}_2}}\left(\mathbb{E}_{\mathbb{P}}\left[u_{\mathcal{T}_1}(X) \mid \mathcal{H}\right]\right)\right),$$

and thus by choosing $\mathcal{T}_2 = \mathcal{H}$ and since $\rho_{\mathcal{H},\mathcal{H}}$ is normalized on constants we get (3.5.3). □

Next we prepare the proof of Proposition 3.5.9:

Lemma 3.5.11. *Let $u : \mathbb{R}^d \to \mathbb{R}$ be a deterministic utility, i.e. u is strictly increasing and continuous, and let $\mathcal{G}$ and $\mathcal{H}$ be sub-σ-algebras of $\mathcal{F}$ such that $\mathcal{G} \subseteq \mathcal{H}$. Then*

$$\mathbb{E}_{\mathbb{P}}\left[u(L^\infty_d(\mathcal{H})) \mid \mathcal{G}\right] = u(L^\infty_d(\mathcal{G})).$$

Proof. "$\supseteq$": Obvious. "$\subseteq$": Define the CRM $\rho_{\mathcal{G}} : L^\infty_d(\mathcal{H}) \to L^\infty(\mathcal{G}); X \mapsto -\mathbb{E}_{\mathbb{P}}\left[u(X) \mid \mathcal{G}\right]$. By Lemma 3.2.5 it follows that

$$\begin{aligned}\mathbb{E}_{\mathbb{P}}\left[u(L^\infty_d(\mathcal{H})) \mid \mathcal{G}\right] &= -\rho_{\mathcal{G}}(L^\infty_d(\mathcal{H})) = -f_{\rho_{\mathcal{G}}}(L^\infty(\mathcal{G})) = \mathbb{E}_{\mathbb{P}}\left[u(L^\infty(\mathcal{G})\mathbf{1}_d) \mid \mathcal{G}\right] \\ &\subseteq \mathbb{E}_{\mathbb{P}}\left[u(L^\infty_d(\mathcal{G})) \mid \mathcal{G}\right] = u(L^\infty_d(\mathcal{G})).\end{aligned}$$

□

Lemma 3.5.12. *For an arbitrary $\mathcal{T} \in \mathcal{I}$ let $u_{\mathcal{T}} : \mathbb{R}^d \to \mathbb{R}$ be a deterministic utility and define $\mathcal{X}_{\mathcal{H}} := u_{\mathcal{T}}(L^\infty_d(\mathcal{H}))$ for all $\mathcal{H} \in \mathcal{E}(\mathcal{T})$. Moreover, let $p_{\mathcal{H}} : \mathcal{X}_{\mathcal{H}} \to L^\infty(\mathcal{H})$ be functions such that $p_{\mathcal{H}}$ is $\mathcal{H}$-local, strictly isotone and fulfills the Lebesgue-property. If for all $\mathcal{G}, \mathcal{H} \in \mathcal{E}(\mathcal{T})$ with $\mathcal{G} \subseteq \mathcal{H}$ and $\mathcal{H}$ atomless it holds that*

$$p_{\mathcal{G}}\left(\mathbb{E}_{\mathbb{P}}\left[F \mid \mathcal{G}\right]\right) = \mathbb{E}_{\mathbb{P}}\left[p_{\mathcal{H}}(F) \mid \mathcal{G}\right] \text{ for all } F \in \mathcal{X}_{\mathcal{H}}, \tag{3.5.4}$$

then

$$p_{\mathcal{H}}(F) = aF + \beta_{\mathcal{H}},$$

where $a \in \mathbb{R}^{+}\backslash\{0\}$ *and* $\beta_{\mathcal{H}} \in L^{\infty}(\mathcal{H})$ *such that* $\mathbb{E}_{\mathbb{P}}[\beta_{\mathcal{H}} \mid \mathcal{G}] = \beta_{\mathcal{G}}$.
Note that (3.5.4) *is well-defined by Lemma* 3.5.11.

Proof. Firstly, we consider the case where $\mathcal{G}$ is the trivial σ-algebra. We write $p := p_{\{\Omega,\emptyset\}}$. Note that, since p is a deterministic function, $p(\mathbb{E}_{\mathbb{P}}[F])$ is law-invariant and thus by (3.5.4) also $\mathbb{E}_{\mathbb{P}}[p_{\mathcal{H}}(F)]$.
Now suppose that there exist $x, y \in \mathcal{X} := \mathcal{X}_{\{\Omega,\emptyset\}}$ with $p_{\mathcal{H}}(x) - p_{\mathcal{H}}(y) \notin \mathbb{R}$, i.e. there exists a $c \in \mathbb{R}$ such that $\mathbb{P}(p_{\mathcal{H}}(x) \leq p_{\mathcal{H}}(y) + c) \in (0,1)$. Since $\mathcal{H}$ is an atomless space we can choose $A_1, A_2, A_3 \in \mathcal{H}$ with

$$\mathbb{P}(A_1) = \mathbb{P}(A_2) := q > 0$$

such that

$$A_1 \subseteq \{p_{\mathcal{H}}(x) \leq p_{\mathcal{H}}(y) + c\}, A_2 \subseteq \{p_{\mathcal{H}}(x) > p_{\mathcal{H}}(y) + c\}, A_3 := (A_1 \cup A_2)^C.$$

Moreover, we define

$$F_1 := x\mathbb{1}_{A_1} + y\mathbb{1}_{A_2} + x\mathbb{1}_{A_3} \quad \text{and} \quad F_2 := y\mathbb{1}_{A_1} + x\mathbb{1}_{A_2} + x\mathbb{1}_{A_3}.$$

Obviously $F_1, F_2 \sim q\delta_y + (1-q)\delta_x$, that is $F_1 \overset{\mathrm{d}}{=} F_2$. However, since $p_{\mathcal{H}}$ is $\mathcal{H}$-local, we have

$$\begin{aligned}
\mathbb{E}_{\mathbb{P}}[p_{\mathcal{H}}(F_1)] + cq &= \mathbb{E}_{\mathbb{P}}[p_{\mathcal{H}}(x)\mathbb{1}_{A_1}] + \mathbb{E}_{\mathbb{P}}[(p_{\mathcal{H}}(y) + c)\mathbb{1}_{A_2}] + \mathbb{E}_{\mathbb{P}}[p_{\mathcal{H}}(x)\mathbb{1}_{A_3}] \\
&< \mathbb{E}_{\mathbb{P}}[(p_{\mathcal{H}}(y) + c)\mathbb{1}_{A_1}] + \mathbb{E}_{\mathbb{P}}[p_{\mathcal{H}}(x)\mathbb{1}_{A_2}] + \mathbb{E}_{\mathbb{P}}[p_{\mathcal{H}}(x)\mathbb{1}_{A_3}] \\
&= \mathbb{E}_{\mathbb{P}}[p_{\mathcal{H}}(F_2)] + cq,
\end{aligned}$$

which contradicts the law-invariance of $F \mapsto \mathbb{E}_{\mathbb{P}}[p_{\mathcal{H}}(F)]$.
Hence we have that $p_{\mathcal{H}}(x) - p_{\mathcal{H}}(y) \in \mathbb{R}$ for all $x, y \in \mathcal{X}$. Choose an arbitrary $\widetilde{x} \in \mathcal{X}$, and let

$$a(x) := p_{\mathcal{H}}(x) - p_{\mathcal{H}}(\widetilde{x}), \quad x \in \mathcal{X},$$

so $a : \mathcal{X} \to \mathbb{R}$. Define $\widetilde{\beta}_{\mathcal{H}} := p_{\mathcal{H}}(\widetilde{x}) \in L^{\infty}(\mathcal{H})$, then $p_{\mathcal{H}}(x) = a(x) + \widetilde{\beta}_{\mathcal{H}}$. The function a is continuous, since otherwise there would exist a sequence $(x_n)_{n\in\mathbb{N}} \subset \mathcal{X}$ with $x_n \to x \in \mathcal{X}$, but $a(x_n) \not\to a(x)$ and the Lebesgue-property would imply the contradiction

$$p_{\mathcal{H}}(x) = \lim_{n\to\infty} p_{\mathcal{H}}(x_n) = \lim_{n\to\infty} a(x_n) + \widetilde{\beta}_{\mathcal{H}} \neq a(x) + \widetilde{\beta}_{\mathcal{H}} = p_{\mathcal{H}}(x).$$

Let $F \in \mathcal{X}_{\mathcal{H}}$. Since the $\mathcal{H}$-measurable simple random vectors are dense in $L^{\infty}_{d}(\mathcal{H})$ and by the definition of $\mathcal{X}_{\mathcal{H}}$ there exists a sequence of $\mathcal{H}$-measurable simple random variables $(F_n)_{n\in\mathbb{N}} \subset \mathcal{X}_{\mathcal{H}} \cap \mathcal{S}$ with $F_n = \sum_{i=1}^{k_n} x_i^n 1_{A_i^n} \to F$ $\mathbb{P}$-a.s. Thus

$$p_{\mathcal{H}}(F) = \lim_{n\to\infty} p_{\mathcal{H}}(F_n) = \lim_{n\to\infty} \sum_{i=1}^{k_n} p_{\mathcal{H}}(x_i^n)\mathbb{1}_{A_i^n} = \lim_{n\to\infty} \sum_{i=1}^{k_n} a(x_i^n)\mathbb{1}_{A_i^n} + \widetilde{\beta}_{\mathcal{H}}$$

$$= \lim_{n\to\infty} a\left(\sum_{i=1}^{k_n} x_i^n \mathbb{1}_{A_i^n}\right) + \widetilde{\beta}_{\mathcal{H}} = \lim_{n\to\infty} a(F_n) + \widetilde{\beta}_{\mathcal{H}} = a(F) + \widetilde{\beta}_{\mathcal{H}}.$$

The function $\mathcal{X}_{\mathcal{H}} \ni F \mapsto \mathbb{E}_{\mathbb{P}}[F]$ induces a preference relation on $\mathcal{M} := \{\mu : \exists F \in \mathcal{X}_{\mathcal{H}}$ such that $F \sim \mu\}$ via

$$\mu \succcurlyeq \nu \quad \Longleftrightarrow \quad \mathbb{E}_{\mathbb{P}}[F] \geq \mathbb{E}_{\mathbb{P}}[G], F \sim \mu, G \sim \nu.$$

Moreover the function $x \mapsto p^{-1}(x + \mathbb{E}[\widetilde{\beta}_{\mathcal{H}}])$ is strictly increasing and by (3.5.4)

$$\mathbb{E}_{\mathbb{P}}[F] = p^{-1}\left(\mathbb{E}_{\mathbb{P}}[p_{\mathcal{H}}(F)]\right) = p^{-1}\left(\mathbb{E}_{\mathbb{P}}[a(F)] + \mathbb{E}[\widetilde{\beta}_{\mathcal{H}}]\right).$$

Thus $\mathbb{E}_{\mathbb{P}}[a(F)]$ is another affine numerical representation of $\succcurlyeq$. It is well-known that the affine numerical representation of $\succcurlyeq$ is unique up to a positive affine transformation (see e.g. Föllmer and Schied (2011) Theorem 2.21), i.e. there exist $\tilde{a}, b \in \mathbb{R}, \tilde{a} > 0$ such that $\mathbb{E}_{\mathbb{P}}[a(F)] = \tilde{a}\mathbb{E}_{\mathbb{P}}[F] + b$ for all $F \in \mathcal{X}_{\mathcal{H}}$. In particular this implies that for all $x \in \mathcal{X}$

$$a(x) = \mathbb{E}_{\mathbb{P}}[a(x)] = a\mathbb{E}_{\mathbb{P}}[x] + b = \tilde{a}x + b.$$

By setting $b + \widetilde{\beta}_{\mathcal{H}} =: \beta_{\mathcal{H}} \in L^{\infty}(\mathcal{H})$ we get for all $F \in \mathcal{X}_{\mathcal{H}}$ that

$$p_{\mathcal{H}}(F) = a(F) + \widetilde{\beta}_{\mathcal{H}} = \tilde{a}F + b + \widetilde{\beta}_{\mathcal{H}} = \tilde{a}F + \beta_{\mathcal{H}}.$$

Finally we obtain by (3.5.4) that for every $\mathcal{G} \subseteq \mathcal{H}$ and for all $F \in \mathcal{X}_{\mathcal{G}}$

$$p_{\mathcal{G}}(F) = p_{\mathcal{G}}\left(\mathbb{E}_{\mathbb{P}}[F \mid \mathcal{G}]\right) = \mathbb{E}_{\mathbb{P}}[p_{\mathcal{H}}(F) \mid \mathcal{G}] = aF + \mathbb{E}_{\mathbb{P}}[\beta_{\mathcal{H}} \mid \mathcal{G}],$$

which proves the martingale property of $(\beta_{\mathcal{G}})_{\mathcal{G}\subseteq\mathcal{H}}$. □

Proof of Proposition 3.5.9*:* Let $(\rho_{\mathcal{H},\mathcal{T}})_{(\mathcal{H},\mathcal{T})\in\mathcal{E}}$ be a strongly consistent family such that (3.5.2) holds for all $(\mathcal{H},\mathcal{T}) \in \mathcal{E}$, i.e.

$$\rho_{\mathcal{H},\mathcal{T}}(X) = f_{\rho_{\mathcal{H},\mathcal{T}}}\left(f_{u_{\mathcal{T}}}^{-1}\big(\mathbb{E}_{\mathbb{P}}[u_{\mathcal{T}}(X) \mid \mathcal{H}]\big)\right), \quad \text{for all } X \in L_d^{\infty}(\mathcal{T}),$$

We define the functions

$$h_{\mathcal{H},\mathcal{T}} : u_{\mathcal{T}}(L_d^{\infty}(\mathcal{H})) \to L^{\infty}(\mathcal{H}); F \mapsto f_{\rho_{\mathcal{H},\mathcal{T}}} \circ f_{u_{\mathcal{T}}}^{-1}(F)$$

and

$$p_{\mathcal{H},\mathcal{T}_1,\mathcal{T}_2} : u_{\mathcal{T}_1}(L_d^{\infty}(\mathcal{H})) \to L^{\infty}(\mathcal{H}); F \mapsto h_{\mathcal{H},\mathcal{T}_2}^{-1} \circ h_{\mathcal{H},\mathcal{T}_1}(F).$$

By strong consistency, we obtain for $\mathcal{G} \subseteq \mathcal{H} \subseteq \mathcal{T}_1 \cap \mathcal{T}_2$, $X \in L_d^{\infty}(\mathcal{T}_1)$ and $F := \mathbb{E}_{\mathbb{P}}[u_{\mathcal{T}_1}(X) \mid \mathcal{H}]$ that

$$\begin{aligned}
p_{\mathcal{G},\mathcal{T}_1,\mathcal{T}_2}\left(\mathbb{E}_{\mathbb{P}}[F \mid \mathcal{G}]\right) &= h_{\mathcal{G},\mathcal{T}_2}^{-1}\left(h_{\mathcal{G},\mathcal{T}_1}\big(\mathbb{E}_{\mathbb{P}}[\mathbb{E}_{\mathbb{P}}[u_{\mathcal{T}_1}(X) \mid \mathcal{H}] \mid \mathcal{G}]\big)\right) \\
&= h_{\mathcal{G},\mathcal{T}_2}^{-1}\left(\rho_{\mathcal{G},\mathcal{T}_1}(X)\right) \\
&= h_{\mathcal{G},\mathcal{T}_2}^{-1}\left(\rho_{\mathcal{G},\mathcal{T}_2}\left(f_{\rho_{\mathcal{H},\mathcal{T}_2}}^{-1}(\rho_{\mathcal{H},\mathcal{T}_1}(X))\mathbf{1}_d\right)\right)
\end{aligned}$$

$$= \mathbb{E}_{\mathbb{P}}\left[h^{-1}_{\mathcal{H},\mathcal{T}_2}\left(h_{\mathcal{H},\mathcal{T}_1}\left(\mathbb{E}_{\mathbb{P}}\left[u_{\mathcal{T}_1}(X) \mid \mathcal{H}\right]\right)\right) \,\middle|\, \mathcal{G}\right]$$
$$= \mathbb{E}_{\mathbb{P}}\left[p_{\mathcal{H},\mathcal{T}_1,\mathcal{T}_2}(F) \mid \mathcal{G}\right]. \tag{3.5.5}$$

By Lemma 3.5.12 (3.5.5) is fulfilled, if and only if

$$p_{\mathcal{H},\mathcal{T}_1,\mathcal{T}_2}(F) = a_{\mathcal{T}_1,\mathcal{T}_2}F + b_{\mathcal{H},\mathcal{T}_1,\mathcal{T}_2}, \quad \text{for all } F \in u_{\mathcal{T}_1}(L^\infty_d(\mathcal{H})),$$

where $a_{\mathcal{T}_1,\mathcal{T}_2} \in \mathbb{R}^+\backslash\{0\}$, $b_{\mathcal{H},\mathcal{T}_1,\mathcal{T}_2} \in L^\infty(\mathcal{H})$ and $\mathbb{E}_{\mathbb{P}}\left[b_{\mathcal{H},\mathcal{T}_1,\mathcal{T}_2} \mid \mathcal{G}\right] = b_{\mathcal{G},\mathcal{T}_1,\mathcal{T}_2}$ for all $\mathcal{G} \in \mathcal{I}$ with $\mathcal{G} \subseteq \mathcal{H}$. Thus

$$h_{\mathcal{H},\mathcal{T}_1}(F) = h_{\mathcal{H},\mathcal{T}_2}(a_{\mathcal{T}_1,\mathcal{T}_2}F + b_{\mathcal{H},\mathcal{T}_1,\mathcal{T}_2}), \quad F \in u_{\mathcal{T}_1}(L^\infty_d(\mathcal{H})),$$

which implies that

$$\rho_{\mathcal{H},\mathcal{T}_1}(X) = f_{\rho_{\mathcal{H},\mathcal{T}_2}}\left(f^{-1}_{u_{\mathcal{T}_2}}\left(a_{\mathcal{T}_1,\mathcal{T}_2}\mathbb{E}_{\mathbb{P}}\left[u_{\mathcal{T}_1}(X) \mid \mathcal{H}\right] + b_{\mathcal{H},\mathcal{T}_1,\mathcal{T}_2}\right)\right).$$

□

4 Allocation of Systemic Risk

4.0 Contributions of the thesis' author

This chapter is a joint work between the author of the thesis H. Hoffmann, the supervisor T. Meyer-Brandis and G. Svindland. A preprint is also available at *http://www.fm.mathematik.uni-muenchen.de/download/publications/systallo.pdf*.

This final chapter investigates the appropriateness of the (fuzzy) core in the context of systemic risk allocation. Section 4.2 just contains definitions and a small review of the classical (fuzzy) core. The authors developed jointly the contagion model presented at the beginning of Section 4.3 which is related to the framework of Eisenberg and Noe (2001). The remaining calculation in this section have been performed by H. Hoffmann. He also suggested to consider the reverse (fuzzy) core in conjunction with the new subsystem generation scheme in (4.3.4). Moreover, the major part for the statement of the reverse fuzzy core in the contagion model in Section 4.5 has been done by the thesis' author. The final example and the subsequent appendix has been elaborated by H. Hoffmann.

4.1 Introduction

In this work our aim is to study the appropriateness of the transfer of a classical game theoretic allocation concept to the allocation problem for financial systems with interacting institutions. This is due to the fact that in the recent financial crisis it became apparent that a risk evaluation of a financial network on the basis of the single institutions is not sufficient in order to capture the systemic risk inherent from the various feedback mechanisms between the institutions.

For this purpose we position ourself in a stylized market clearing framework for interbank loans. This framework traces back to the seminal work of Eisenberg and Noe (2001). Briefly speaking, we have a system of financial institutions which are connected via bilateral credit agreements. If now a financial institution defaults due to some adverse market event, then it has to be liquidated immediately and the remaining assets are distributed among the creditors of the institution proportionately to their liabilities. Since the proceeds of liquidation are less than the total liabilities owed to the other banks in the system, these banks face additional losses which might result in a default of one or more creditor banks. These potential defaults can trigger further failures of banks and thus the initial default spreads into the financial system.

In order to allow for a comprehensive risk assessment of financial systems like the above, systemic risk measures have been introduced. As we have already described how losses propagate into the financial system, we can easily calculate the total losses of the system by summing up the losses of the single institutions after all possible contagion has taken place. Now the risk of the system can be easily obtained by using a well-known univariate risk measure. An axiomatic description of this particular type of systemic risk measures which allow for a decomposition into an aggregation function and a univariate risk measure has been studied in Chen et al. (2013), Kromer et al. (2016) and Hoffmann et al. (2016).

Besides the assessment of the total risk of a financial system, it is also of interest to identify the individual contributions of the single institutions to the total risk. In particular, the contributions should add up to the system's risk, that is we look for appropriate allocation schemes. Chen et al. (2013) and Kromer et al. (2016) propose an allocation procedure which is based on a dual representation of the systemic risk measures. These allocations are essentially equal to the Aumann-Shapley value which is known from the game theoretic literature, cf. Aumann and Shapley (1974). The Aumann-Shapley value is also an example for a coherent allocation as defined in Denault (2001) which gained much attention for the portfolio allocation problem. Among the properties of a coherent allocation the no-undercut property is the most crucial. Moreover, it is also the main building block of the core from the game theory literature, cf. Aubin (1979). The no-undercut property says that for all subgroups the amount of the total risk which is allocated to it should be smaller than the measured risk of this subgroup. This property is commonly justified by the following consideration, if a subgroup would get a share of the total risk which is higher than its own risk, then this subsystem would split from the system and consequently obtains a lower risk. Whereas for a portfolio of financial assets it can be easily answered how a separation of a subportfolio should be implemented, this task is much more complex for a financial network. In this work we concentrate on two possible ways to measure the risk of subsystems after they separated from the financial system. In both approaches we assume that the underlying financial network topology remains intact. For examples where also the interbank liabilities in the financial network are modified, we refer to the works of Drehmann and Tarashev (2013) or Staum et al. (2016). However, in their work they do not study the implications on the core.

For obvious reasons, the core is only meaningful if the risk measurement of the subsystems is subadditive, that is merging two disjoint subgroups should reduce the risk. This diversification effect is usually assumed for the classical risk management of a firm or a portfolio of financial assets. Similarly, for decomposable systemic risk measures, one can argue in favor of a diversification benefit for the risk measurement of the aggregated values of the financial system. However, for the aggregation itself it depends on the applied model for the construction of the subsystems if the merger of two subsystems decreases or increases the risk. This is because the introduction of a new financial firm to the financial system might serve as both a transmitter and a buffer of losses. Thus, the corresponding aggregation function can exhibit diversification benefits as well as integration costs.

In Chen et al. (2013) and Kromer et al. (2016) the authors overcome this problem by considering an aggregation of the subsystems which corresponds to a worst-case view, when it comes to the spreading of risk within the system. As a result they also have a diversification benefit on this level. The considered subsystem generation is a generalization from the classical portfolio approach, where the risk factors correspond to profits and losses of certain financial instruments. Thus, considering the risk of the accumulated profits and losses of a subsystem suggests itself as the subsystem risk. That is simply summing up all risk factors which are in the subsystem and equate the remainders to zero.

Unfortunately, we will see that in financial networks where contagion might take place this allocation procedure creates wrong incentives to the financial institutions. The reason is that, whereas it was sufficient in the classical approach to measure how much each subsystem spreads into the system, we have now also a second origin of risk, namely the ability of a subsystem to transfer the losses. For example consider two financial systems which are connected exclusively via one intermediate institution having no other operations. Obviously, the intermediate institution can be considered systemic, since it is the only possible way that losses of one financial system can be carried over to the other. However, the systemic relevance of the intermediary cannot be expressed by a core allocation, since each core allocation must be bounded from above by its standalone risk and the intermediary has no other sources of risk apart from the losses from the financial systems.

In order to tackle this problem we invert the definition of the core, i.e. the allocated risk for each subsystem should be at least as much as the risk of this subsystem. We call this allocation principle the reverse core. Clearly, by reversing the core definition there is now also a need for changing the underlying subsystem risk management in such a way that instead of a diversification benefit we have that there is a consolidation cost. In our analysis this is provided by supposing that all institutions outside of a subsystem are equipped with such a high amount of capital that a default is excluded. Contrarily to the classical subsystem generation discussed earlier this supports a best-case view. For our interaction model we will see that this new definition resolves the unfairnesses from before. Moreover, we identify under which assumptions there exist allocations in the intersection of both approaches.

Structure of the paper

In Section 4.2 we state our notation and review the (fuzzy) core concept from the game theoretic literature adapted to more general aggregation functions. In Section 4.3 we apply the core concept to a financial system with contagion. Based on the deficiencies of this allocation we alter the underlying risk measurement for the subsystems from a worst-case to a best-case perspective. Due to this change we introduce in Section 4.4 the notion of the reverse core and study its relation to the core concept from before. Finally, in Section 4.5 we determine a reverse core element for our financial system and show that in most cases it does not coincide with the elements from Section 4.3. In the appendix 4.A we discuss how the non-emptiness of the cores for random risk factors can be inferred from deterministic risk factors.

4.2 Standard game theoretic approach to systemic risk

Throughout this work we consider a financial system $\mathcal{I} := \{1, ..., d\}$ which consists of $d \in \mathbb{N}$ different financial institutions. In the analysis of the financial system $\mathcal{I}$ subsystems will play a decisive role. Here $\mathcal{P} := \mathcal{P}(\mathcal{I})$ is the powerset of $\mathcal{I}$ and represents the set of all subsystems. Since $\mathcal{I}$ is the largest system we consider, J^C denotes the complementary set of $J \in \mathcal{P}$ with respect to $\mathcal{I}$, i.e. $J^C := \mathcal{I} \backslash J$. We will denote the i-th unit vector of $\mathbb{R}^d$ by e_i, i.e. all components are equal to zero except the i-th component which is equal to one. Moreover, $\mathbf{0}_d$ and $\mathbf{1}_d$ is the notation for the d-dimensional vectors where all components are equal to zero or one, resp. As usual $\mathbb{R}^d_+$ is the space of d-dimensional non-negative real valued vectors. By $\mathbf{I}_d$ we denote the

$d \times d$ dimensional identity matrix and by $\mathbf{A}^\top$ the transpose of the matrix $\mathbf{A}$. Apart from the usual matrix multiplication, we will sometimes also need the Hadamard product (component-wise multiplication) which we denote by $*$, i.e. $x * y = (x_1 y_1, ..., x_d y_d)^\top$ for $x, y \in \mathbb{R}^d$.

Let $\mathcal{X}^d$ be a space of $\mathbb{R}^d$-valued functions on some measurable space $(\Omega, \mathcal{F})$ representing d-dimensional risk factors of the financial system $\mathcal{I}$. We evaluate the systemic risk of these risk factors via a systemic risk measure.

Definition 4.2.1. *A function $\rho : \mathcal{X}^d \to \mathbb{R}$ is called systemic risk measure if it is antitone, that is for $X, Y \in \mathcal{X}^d$ with $X \leq Y$ we have that $\rho(X) \geq \rho(Y)$.*

In addition to the measurement of the risk of the whole financial system, we also suppose that we have excess to information on the risk of each subgroup of financial institutions. For this measurement we introduce the notion of a subsystem risk measure.

Definition 4.2.2. *We say that $\widetilde{\rho} : \mathcal{X}^d \times \mathcal{P} \to \mathbb{R}$ is subsystem risk measure for the systemic risk measure $\rho : \mathcal{X}^d \to \mathbb{R}$ if the function $\mathcal{X}^d \ni X \mapsto \widetilde{\rho}(X, J)$ is a systemic risk measure for all $J \in \mathcal{P}$ and $\widetilde{\rho}(X, \mathcal{I}) = \rho(X)$.*
Moreover, if we just consider deterministic risk factors we will call it a subsystem construction scheme and denote it by $\widetilde{\Lambda} : \mathbb{R}^d \times \mathcal{P} \to \mathbb{R}$.

A fairness criterion which is known as the core in the game theoretic literature are individually and coalitionally stable allocations, i.e. allocations where no entity or group of entities has an incentive to deviate from the allocation by splitting from the system, cf. Aubin (1979). More precisely, the core is formally defined as follows:

Definition 4.2.3 (Allocation and core)**.** *For a given $X \in \mathcal{X}^d$, we say that $k \in \mathbb{R}^d$ is an allocation of the systemic risk $\rho(X)$ if*

$$\sum_{i=1}^{d} k_i = \rho(X).$$

Moreover, let $\widetilde{\rho} : \mathcal{X}^d \times \mathcal{P} \to \mathbb{R}$ be a subsystem risk measure of ρ. We say that k is in the core $C^-_{\widetilde{\rho}}(X)$ if k is an allocation which additionally fulfills that for all subsystems $J \in \mathcal{P}$

$$\widetilde{\rho}(X, J) \geq \sum_{j \in J} k_j. \tag{4.2.1}$$

The core has been prominently used for the allocation of the risk of a portfolio consisting of financial assets. In the following example we will review this framework and the motivation for the core.

Example 4.2.4. The core C^- is a superset of the coherent allocations as postulated by Denault (2001). Here (4.2.1) also appeared under the name of the no undercut property. Denault (2001) exclusively concentrates on the risk measurement of a portfolio of financial assets. This corresponds in our framework to a financial system with no interactions between the single institutions, i.e. the well-being of a financial firm is unaffected by the state of the other banks.

However, note that in the absence of feedback mechanisms, the single risk factors, here the profits and losses, might still be dependent in a probabilistic sense. In the portfolio framework of Denault (2001) the risk of a multivariate risk factor $X \in \mathcal{X}^d$ is measured by

$$\eta\left(\sum_{j=1}^{d} X_j\right),$$

where $\eta : \mathcal{X} \to \mathbb{R}$ is some coherent risk measure. Recall that a coherent risk measure is a functional which is antitone, cash-additive, convex and positive homogeneous. For more details on coherent risk measures we refer to Föllmer and Schied (2011). Since the removal or the adding of a financial asset to a portfolio does not affect the performance of the remaining assets, the sum is an appropriate aggregation function in this setup. By an analogous argumentation, the risk of a subsystem $J \in \mathcal{P}$ should be measured via

$$\widetilde{\rho}(X, J) := \eta\left(\sum_{j\in J} X_j\right). \tag{4.2.2}$$

Recall that every coherent risk measure η is subadditive and thus we have that for all disjoint $J_1, J_2 \in \mathcal{P}$

$$\widetilde{\rho}(X, J_1 \cup J_2) \leq \widetilde{\rho}(X, J_1) + \widetilde{\rho}(X, J_2),$$

which reflects a diversification effect. This implies that it is always profitable to merge subportfolios. In order to have a fair allocation $k \in \mathbb{R}^d$ this diversification benefit should be shared among the different financial assets, i.e. $k_j \leq \widetilde{\rho}(X, \{j\})$ for $j = 1, ..., d$. Otherwise the investor would demerge this asset from the portfolio and would hold it separately. Therefore the allocated risk of every single financial asset should be less than its standalone risk. Similarly we can argue for subportfolios, which then results in (4.2.1) that is the main property of the core C^-.

The following lemma relates the cores of two subsystem risk measures where one is always more conservative than the other. It is a direct consequence of the definition of the core.

Lemma 4.2.5. *Let $\widetilde{\rho}_1$ be a subsystem risk measure. If $k \in C^-_{\widetilde{\rho}_1}(X)$, then $k \in C^-_{\widetilde{\rho}_2}(X)$ for all $\widetilde{\rho}_2$ with $\widetilde{\rho}_1(X, J) \leq \widetilde{\rho}_2(X, J)$ for all $J \in \mathcal{P}$ and $\rho_1(X) = \rho_2(X)$.*

For the construction of the core, we just considered subsystems of type $\mathcal{P}$, i.e. a risk factor of a financial institution can either be accounted for completely or not at all. But especially in the context of Example 4.2.4 a subsystem can also be created by taking fractional parts of the profits and losses. Thus we will now characterize a subsystem by a fractional participation level $\lambda \in [0, 1]^d$ or $\lambda \in \mathbb{R}^d_+$. For this purpose we need to generalize the notion of a subsystem risk measure to a function $\rho : \mathcal{X}^d \times \mathbb{R}^d_+ \to \mathbb{R}$, where $\rho(X, \lambda)$ is the risk of the system $X \in \mathcal{X}^d$ where bank j participates with λ_j and $\rho(X, \mathbf{1}_d) = \rho(X)$. Here $\lambda_j = 1$ means full participation and $\lambda_j = 0$ is the absence of bank j.

Definition 4.2.6 (Fuzzy core). *We say $k \in \mathbb{R}^d$ is in the fuzzy core $FC^-_\rho(X)$ if k is an allocation, i.e. $\rho(X, \mathbf{1}_d) = \mathbf{1}_d^\top k$ and for all $\lambda \in [0, 1]^d$ it holds that*

$$\rho(X, \lambda) \geq \lambda^\top k.$$

Since each subsystem risk measure with fractional participation ρ yields a subsystem risk measure $\widetilde{\rho} : \mathcal{X}^d \times \mathcal{P} \to \mathbb{R}$ via

$$\widetilde{\rho}(X, J) := \rho\left(X, \sum_{j \in J} e_j\right) \quad \text{for all } J \in \mathcal{P},$$

we have that $FC^-_\rho(X) \subseteq C^-_{\widetilde{\rho}}(X)$.

In the following theorem we recall the well-known result that under differentiability, convexity and positive homogeneity of a subsystem risk measure the fuzzy core is single-valued and equal to its gradient. For a proof see for instance Aubin (1979). In this case the fuzzy core is also called Euler allocation or Euler principle in the literature, c.f. Denault (2001), Tasche (2004).

Theorem 4.2.7. *Let $\rho : \mathcal{X}^d \times \mathbb{R}^d_+ \to \mathbb{R}$ be a subsystem risk measure which is positive homogeneous on the diagonal of its second argument, i.e. $\rho(\cdot, \alpha \mathbf{1}_d) = \alpha\rho(\cdot, \mathbf{1}_d)$ for all $\alpha \geq 0$. Then, the extended fuzzy core*

$$\overline{FC}^-_\rho(X) := \left\{k \in \mathbb{R}^d : \rho(X, \mathbf{1}_d) = \mathbf{1}_d^\top k \text{ and } \rho(X, \lambda) \geq \lambda^\top k, \forall \lambda \in \mathbb{R}^d_+\right\}$$

is equal to the subdifferential

$$\partial^- \rho(X, \mathbf{1}_d) := \left\{k \in \mathbb{R}^d : \rho(X, \lambda) \geq \rho(X, \mathbf{1}_d) + k^\top(\lambda - \mathbf{1}_d) \ \forall \lambda \in \mathbb{R}^d_+\right\}.$$

Thus, if the function $\lambda \mapsto \rho(X, \lambda)$ is additionally convex and differentiable in $\mathbf{1}_d$ the extended fuzzy core

$$\overline{FC}^-_\rho(X) = \nabla\rho(X, \mathbf{1}_d)$$

where $\nabla\rho(X, \cdot)$ is the gradient of ρ in its second argument.

Example 4.2.8 (Portfolio approach cont.)**.** A possible extension of (4.2.2) to allow for fractional participation is given by

$$\rho(X, \lambda) := \eta\left(\sum_{i=1}^d \lambda_i X_i\right) = \eta\left(\lambda^\top X\right), \tag{4.2.3}$$

where η is the coherent univariate risk measure from (4.2.2). Clearly, $\lambda \mapsto \rho(X, \lambda)$ is positively homogeneous and convex, thus the fuzzy core $FC^-_\rho(X)$ is non-empty. If $\lambda \mapsto \rho(X, \lambda)$ is also differentiable then the fuzzy core $FC^-_\rho(X)$ is even single valued. Therefore the fuzzy core $FC^-(X)$ or the larger core $C^-(X)$ seems to be a feasible allocation approach in the portfolio context.

4.3 Financial model with contagion

In this section we will investigate if the (fuzzy) core still yields fair allocations given that our financial network allows for feedback mechanisms among the financial institutions. For this

purpose we need to alter the aggregation function in (4.2.2) and (4.2.3) respectively from a simple sum to a more complex aggregation function which allows for the inclusion of a channel of contagion. The aggregation function we will use traces back to the seminal paper of Eisenberg and Noe (2001) and has been extended in many directions, e.g. to include multiple sources of contagion, c.f. Awiszus and Weber (2015) for a survey. In order to focus on the impact of the feedback mechanism on the allocations, we will consider a deterministic risk $x \in \mathbb{R}^d$. The treatment of random risk factors is discussed in the appendix.
As before we assume that $\mathcal{I} := \{1, ..., d\}$ represents a financial system. However, we now assume that only the first $d-1$ components are financial institutions and the last component represents the real economy. We suppose that the financial institutions have claims against each other which appear as interbank assets/liabilities on their balance sheets. The interbank assets/liabilities are summarized by the matrix $\mathbf{L} = (L_{i,j})_{i,j=1,...,d}$, where $L_{i,j}$ is the monetary amount of bank i which it owes to bank j. Furthermore the total amount of the interbank liabilities of bank i is denoted by $L_i := \sum_{j=1}^d L_{i,j}$ for all $i = 1, ..., d$. Three standing assumptions on the liability matrix will be that each bank does not have claims against itself and against the real economy, however the real economy has claims against each bank. In short, we assume that $L_{i,i} = 0, L_{d,i} = 0$ and $L_{i,d} > 0$ for all $i = 1, ..., d$. The first and the last assumption are more technical and not really restricting. The second assumption needs some further explanation. Of course, banks have claims against the real economy like households or industrial companies. However, we will not model these connections within the financial system, but the real economy can contribute losses to the banks via an initial shock. Another model assumptions is that in case of a default the debtors of the defaulting institution divide the remaining assets proportional to their claims, i.e. it will suffice to consider the relative liability matrix $\mathbf{\Pi} = (\Pi_{i,j})_{i,j=1,...,d}$ which is given by

$$\Pi_{i,j} := \begin{cases} \frac{L_{i,j}}{L_i} & \text{, if } L_i > 0 \\ 0 & \text{, if } L_i = 0 \end{cases} .$$

Moreover, the institutions are endowed with an initial capital/equity $c \in \mathbb{R}_+^d$. On the asset side the institutions have interbank assets as described above and some external assets, which also contains claims against the real economy. Therefore we have a full description of the balance sheet of each bank. Next we suppose that at a future point in time the external assets of the banks are hit by some adverse market event $y \leq \mathbf{0}_d$. Due to this market shock also the liability side of the balance sheets have to decrease by the same amount. As debt is senior to equity, the banks have to use their equity to buffer the shock. However, if there is not a sufficient amount of equity to dampen the shock, a bank is in default and pays out the remaining assets proportionally to its creditors. Since the creditors are not paid in full this creates a further loss on their balance sheets which can result in a default of one or more of the creditors. Finally these defaults can trigger other defaults, so that a large fraction of the system might be affected. This contagion is modeled by the following aggregation function

$$\begin{aligned} \Lambda(x) := \min_{a \in \mathbb{R}_+^d} \ & \mathbf{1}_d^\top a \\ \text{s.t. } a = & \left(\mathbf{\Pi}^\top a - x\right)^+ , \end{aligned} \tag{4.3.1}$$

where $x = c + y$ is the equity value of the financial institutions directly after the adverse market event y took place. Note that by monotonicity of the function $\mathbb{R}^d_+ \ni a \mapsto \big(\mathbf{\Pi}^\top a - x\big)^+$, we have that the optimal value for a in the optimization problem (4.3.1) of the aggregation function Λ can be found by iterating $a(n) = \big(\mathbf{\Pi}^\top a(n-1) - x\big)^+$ with $a(0) := \mathbf{0}_d$. The interpretation of this iteration procedure is as follows:
First, we suppose that no bank defaults which corresponds to $a(0) = \mathbf{0}_d$. In the first iteration we thus have that $a(1) = (-x)^+$ which is identical to the losses of the banks defaulting initially due to the adverse market event. Next $a(2) = \big(\mathbf{\Pi}^\top a(1) - x\big)^+ = \big(\mathbf{\Pi}^\top (-x)^+ - x\big)^+$, where $\mathbf{\Pi}^\top(-x)^+$ are the losses which the banks receive from the initially defaulting banks. Hence $a(2)$ contains the losses of the initially defaulting banks and of those banks which fail due to the losses transmitted by the defaulting banks. In each subsequent step these losses further spread into the system and we approach an equilibrium a fulfilling the constraint in (4.3.1).

In contrast to Eisenberg and Noe (2001) we do not cap the transmission of losses to other banks by the corresponding interbank liabilities. We did so in order to keep the model simple and thus for a better understanding of the contagion effects later on and second the inclusion of the real economy makes the events where the losses exceed the interbank liabilities rather unlikely. As there is not an upper bound for the transmitted losses, it could be that for a finite shock the contagion effects wind each other up more and more. However, we will see in Lemma 4.3.1 below that this is not possible in our framework, since a certain percentage of the losses is always transfered to the real economy, where the channel of contagion ends.

Lemma 4.3.1. *For each $x \in \mathbb{R}^d$ the aggregation function $\Lambda(x)$ is finite.*

Proof. Firstly, we observe that Λ is monotonically decreasing and that $\Lambda(x) = 0$ for all $x \geq \mathbf{0}_d$. Thus it suffices to consider $x \leq \mathbf{0}_d$, i.e. all institutions default initially. Then the constraint in (4.3.1) can be simplified to

$$a = \Big(\mathbf{\Pi}^\top a - x\Big)^+ = \mathbf{\Pi}^\top a - x$$

and thus if the matrix $\mathbf{I}_d - \mathbf{\Pi}^\top$ is invertible, then there exists a unique solution

$$a = -(\mathbf{I}_d - \mathbf{\Pi}^\top)^{-1} x,$$

where $\mathbf{I}_d$ the $d \times d$ dimensional identity matrix.
We denote by $(A_{j,k})_{j,k=1,\dots,d} = \mathbf{A} := \mathbf{I}_d - \mathbf{\Pi}^\top$. Moreover by $\widetilde{\mathbf{\Pi}}$ and $\widetilde{\mathbf{A}}$ we denote the $(d-1) \times (d-1)$ matrices which are obtained by erasing the last row and column from $\mathbf{\Pi}$ and $\mathbf{A}$ resp. Note that, since we assumed that every institution has liabilities to the real economy, i.e. $\Pi_{j,d} > 0$ for all $j = 1, \dots, d-1$, the row sums of $\widetilde{\mathbf{\Pi}}$ are less then 1 and thus the operator norm $\|\widetilde{\mathbf{\Pi}}^\top\|_1 = \max_{i=1,\dots,d-1} \sum_{j=1}^{d-1} |\Pi_{ij}| < 1$. Hence a classic result from functional analysis, see e.g. Werner (2011) Satz II.1.11, yields that the Neumann series $\sum_{i=0}^{n} (\widetilde{\mathbf{\Pi}}^\top)^i, n \in \mathbb{N}$, converges and that the inverse of $\widetilde{\mathbf{A}} = \mathbf{I}_{d-1} - \widetilde{\mathbf{\Pi}}^\top$ exists and is equal to the limit of the series. Furthermore, a Laplace expansion along the last column of $\mathbf{A}$ yields that $\det(\mathbf{A}) = \det(\widetilde{\mathbf{A}}) \neq 0$ and thus $\mathbf{A}$ is invertible.

□

In the next lemma we derive an element in the (fuzzy) core for the subsystem construction scheme $\Lambda^0(x,\lambda) := \Lambda(\lambda * x)$ by making use of Theorem 4.2.7. As already pointed out earlier this subsystem construction scheme parallels the portfolio approach (4.2.3), where the sum as aggregation function is replaced by Λ. For the result we need to identify the institutions which default after all possible contagion has taken place. We denote the set of these institutions for a given $x \in \mathbb{R}^d$ by

$$\mathcal{D}(x) = \{p_1, ..., p_{|\mathcal{D}|}\} := \{i \in \mathcal{I} : (\mathbf{\Pi}^\top a - x)_i \geq 0\},$$

where a is the limit of the sequence $a(n) = (\mathbf{\Pi}^\top a(n-1) - x)^+$ with $a(0) = \mathbf{0}_d$. If it is clear from the context we will mostly drop the reference to the risk factor x.

Lemma 4.3.2. *Let $x \in \mathbb{R}^d$ and define $k \in \mathbb{R}^d$ by*

$$k_{p_i} := -\sum_{j=1}^{|\mathcal{D}|} \left(\left(\mathbf{I}_{|\mathcal{D}|} - \mathbf{\Pi}_{\mathcal{D},\mathcal{D}}^\top \right)^{-1} \right)_{j,i} x_{p_i}, \quad i = 1, ..., |\mathcal{D}|,$$

and $k_i := 0$ for $i \notin \mathcal{D}$. Here $\mathbf{\Pi}_{\mathcal{D},\mathcal{D}} = (\Pi_{i,j})_{i,j\in\mathcal{D}}$. Then

$$k \in FC^-_{\Lambda^0}(x).$$

Proof. That the matrix $\mathbf{I}_{|\mathcal{D}|} - \mathbf{\Pi}_{\mathcal{D},\mathcal{D}}^\top$ is invertible can be shown analogously to the considerations made in the proof of Lemma 4.3.1.
Denote by a the limit of the sequence $a(n) = (\mathbf{\Pi}^\top a(n-1) - x)^+$ with $a(0) = \mathbf{0}_d$. Since $(\mathbf{\Pi}^\top a - x)_i < 0$ for all non-defaulting institutions $i \notin \mathcal{D}$ we have that $a_i = 0$. Therefore we obtain for the vector of losses of the defaulting institutions $a_{\mathcal{D}} := (a_j)_{j\in\mathcal{D}}$ that

$$a_{\mathcal{D}} = \left((\mathbf{\Pi}^\top a - x)_j^+ \right)_{j\in\mathcal{D}} = \mathbf{\Pi}_{\mathcal{D},\mathcal{D}}^\top a_{\mathcal{D}} - x_{\mathcal{D}},$$

where and $x_{\mathcal{D}} := (x_j)_{j\in\mathcal{D}}$. Finally, we obtain that

$$\Lambda(x) = \sum_{i\in\mathcal{D}} a_i = \sum_{i=1}^{|\mathcal{D}|} -\mathbf{1}_{|\mathcal{D}|}^\top \left(\mathbf{I}_{|\mathcal{D}|} - \mathbf{\Pi}_{\mathcal{D},\mathcal{D}}^\top \right)^{-1} \widetilde{e}_i x_{p_i} = \sum_{i=1}^{|\mathcal{D}|} k_{p_i} = \sum_{i=1}^{d} k_i, \tag{4.3.2}$$

where $\widetilde{e}_i$ is the i-th unit vector in $\mathbb{R}^{|\mathcal{D}|}$. Thus k is an allocation.
Moreover, suppose that $a \in \mathbb{R}^d_+$ is such that $a = (\mathbf{\Pi}^\top a - x)^+$ and $\Lambda(x) = \mathbf{1}_d^\top a$. Then we have for each $\lambda > 0$ that $\lambda a = (\mathbf{\Pi}^\top (\lambda a) - \lambda x)^+$. Hence

$$\Lambda^0(x, \lambda\mathbf{1}_d) = \Lambda(\lambda x) \leq \mathbf{1}_d^\top (\lambda a) = \lambda\Lambda(x) = \lambda\Lambda^0(x, \mathbf{1}_d).$$

On the other hand we obtain by a similar argumentation for λx that

$$\begin{aligned}\lambda\Lambda^0(x, \mathbf{1}_d) &= \lambda\Lambda^0\left(x, \frac{1}{\lambda}\lambda\mathbf{1}_d\right) = \lambda\Lambda^0\left(\lambda x, \frac{1}{\lambda}\mathbf{1}_d\right) \\ &\leq \Lambda^0(\lambda x, \mathbf{1}_d) = \Lambda^0(x, \lambda\mathbf{1}_d).\end{aligned}$$

Combining both results yields the positive homogeneity on the diagonal of the second argument of Λ^0.

Furthermore, it can be easily shown that Λ is a convex function, from which it immediately follows that the function $z \mapsto \Lambda^0(x, z)$ is also convex. Hence, we obtain from Theorem 4.2.7 that $FC^-_{\Lambda^0}(x)$ is equal to the subdifferential of $z \mapsto \Lambda^0(x, z)$ at $\mathbf{1}_d$.

Firstly, we suppose that there exists a neighborhood N of x such that $\mathcal{D}(z) = \mathcal{D}(x)$ for all $z \in N$. Then $\Lambda(x)$ is the linear function (4.3.2) on N and thus differentiable in x. Therefore k is the gradient of $\Lambda^0(x, \cdot)$ at $\mathbf{1}_d$ and hence $\{k\} = FC^-_{\Lambda^0}(x)$.
If no such neighborhood N exists the subdifferential might not be single valued. However, it can still be shown that k is a member of the subdifferential.
Since the function $z \mapsto |\mathcal{D}(z)|$ is left-continuous with values in $\{0, ..., d\}$, we can always find $\widetilde{x} \leq x$ with $\|x - \widetilde{x}\|_\infty > \varepsilon$ for some $\varepsilon > 0$ such that $\mathcal{D}(\widetilde{x}) = \mathcal{D}(x)$ and $\lambda \mapsto \Lambda^0(\widetilde{x}, \lambda)$ is differentiable in $\mathbf{1}_d$. Note that $\widetilde{x}$ can also be chosen such that $\widetilde{x}_i \neq 0$ for all $i = 1, ..., d$ and thus the componentwise quotient $u \in \mathbb{R}^d_+$ of x and $\widetilde{x}$, i.e. $u_i = \frac{x_i}{\widetilde{x}_i}$ for all $i = 1, ..., d$, is well-defined. Since Λ is linear between $\widetilde{x}$ and x, we have that

$$\Lambda^0(\widetilde{x}, \mathbf{1}_d) + \nabla\Lambda^0(\widetilde{x}, \mathbf{1}_d)^\top (u - \mathbf{1}_d) = \Lambda^0(\widetilde{x}, u) = \Lambda^0(x, \mathbf{1}_d), \tag{4.3.3}$$

where $\nabla\Lambda^0(\widetilde{x}, \mathbf{1}_d)$ denotes the gradient of the function $\lambda \mapsto \Lambda^0(\widetilde{x}, \lambda)$ at $\mathbf{1}_d$. From this we can immediately infer that for all $\lambda \in \mathbb{R}^d_+$

$$\begin{aligned} \Lambda^0(x, \lambda) &= \Lambda^0(\widetilde{x}, \lambda * u) \\ &\geq \Lambda^0(\widetilde{x}, \mathbf{1}_d) + \nabla\Lambda^0(\widetilde{x}, \mathbf{1}_d)^\top (\lambda * u - \mathbf{1}_d) \\ &= \Lambda^0(x, \mathbf{1}_d) + \left(\nabla\Lambda^0(\widetilde{x}, \mathbf{1}_d) * u\right)^\top (\lambda - \mathbf{1}_d), \end{aligned}$$

where we used (4.3.3) in the last step. Hence $\nabla\Lambda^0(\widetilde{x}, \mathbf{1}_d) * u$ is a subdifferential of $\lambda \mapsto \Lambda^0(x, \lambda)$ at $\mathbf{1}_d$. By using (4.3.2) a simple calculation shows that $k = \nabla\Lambda^0(\widetilde{x}, \mathbf{1}_d) * u$ and the result follows. □

Lemma 4.3.3. *Let $x \in \mathbb{R}^d$ and $k \in FC^-_{\Lambda^0}(x)$ be the allocation from Lemma 4.3.2. Moreover, denote by $\mathcal{D}_0 := \{i \in \mathcal{I} : x_i \leq 0\}$ the set of initially defaulting institutions. Then we have that the allocations $k_i, i \in \mathcal{D}\backslash\mathcal{D}_0$ of the institutions which default due to contagion are non-positive.*

Proof. First, we prove that $\left(\mathbf{I}_{|\mathcal{D}|} - \mathbf{\Pi}^\top_{\mathcal{D},\mathcal{D}}\right)^{-1} = \sum_{i=0}^\infty \left(\mathbf{\Pi}^\top_{\mathcal{D},\mathcal{D}}\right)^i$. We have already seen in the proof of Lemma 4.3.1 that this holds true if $\|\mathbf{\Pi}^\top_{\mathcal{D},\mathcal{D}}\|_1 < 1$. Therefore we assume that $\|\mathbf{\Pi}^\top_{\mathcal{D},\mathcal{D}}\|_1 = 1$. In particular this implies that the real economy $d \in \mathcal{D}$. We consider the matrix $\widetilde{\mathbf{\Pi}} \in \mathbb{R}^{(|\mathcal{D}|-1)\times(|\mathcal{D}|-1)}$ which we obtain from $\mathbf{\Pi}_{\mathcal{D},\mathcal{D}}$ by erasing the last row and column and the vector $\mathbf{\Pi}_d = (\Pi_{p_i,d})_{i=1,...,|\mathcal{D}|-1} \in \mathbb{R}^{(|\mathcal{D}|-1)}$ containing the relative liabilities of the defaulting banks to the real economy. Then

$$\mathbf{\Pi}_{\mathcal{D},\mathcal{D}} = \begin{pmatrix} \widetilde{\mathbf{\Pi}} & \mathbf{\Pi}_d \\ \mathbf{0}^\top_{|\mathcal{D}|-1} & 0 \end{pmatrix}$$

and it can be easily shown that for all $n \in \mathbb{N}$

$$\sum_{i=0}^{n} \left(\mathbf{\Pi}_{\mathcal{D},\mathcal{D}}^{\top}\right)^i = \begin{pmatrix} \sum_{i=0}^{n}(\widetilde{\mathbf{\Pi}}^{\top})^i & \mathbf{0}_{|\mathcal{D}|-1} \\ \mathbf{\Pi}_d^{\top} \sum_{i=0}^{n-1}(\widetilde{\mathbf{\Pi}}^{\top})^i & 0 \end{pmatrix}.$$

Since $\|\widetilde{\mathbf{\Pi}}\|_1 < 1$ the Neumann series $\sum_{i=0}^{n}(\widetilde{\mathbf{\Pi}}^{\top})^i$ converges and hence also $\sum_{i=0}^{n}(\mathbf{\Pi}_{\mathcal{D},\mathcal{D}}^{\top})^i$. From Werner (2011) it thus follows that the limit $\sum_{i=0}^{\infty}(\mathbf{\Pi}_{\mathcal{D},\mathcal{D}}^{\top})^i = \left(\mathbf{I}_{|\mathcal{D}|} - \mathbf{\Pi}_{\mathcal{D},\mathcal{D}}^{\top}\right)^{-1}$. Therefore all entries of the inverse of $\mathbf{I}_{|\mathcal{D}|} - \mathbf{\Pi}_{\mathcal{D},\mathcal{D}}^{\top}$ must be positive. Finally this implies that for all $i \in \mathcal{D}$ the allocation k from Lemma 4.3.2 can be rewritten as $k_i = -w_i x_i$ for the positive weighting factor $w_i := -\sum_{j=1}^{|\mathcal{D}|} \left(\left(\mathbf{I}_{|\mathcal{D}|} - \mathbf{\Pi}_{\mathcal{D},\mathcal{D}}^{\top}\right)^{-1} \right)_{j,i}$. Therefore, we have for all $i \in \mathcal{D}\backslash\mathcal{D}_0$ that $k_i \leq 0$, since $x_i > 0$. □

In summary the allocation k from Lemma 4.3.2 seems to be reasonable for the initially defaulting banks, since it is a combination of the severity of the loss x_i and of how much the loss propagates further into the system which is specified by the weighting factor w_i. Moreover, those banks which do not default at all get an allocation of zero which could also be declared as fair. However, those banks which have enough equity at the beginning $x_i > 0$ but which default due to contagion, get an allocation which is strictly negative. Compendiously, this allocation creates an incentive to control the standalone risk factor x_i, but to ignore (or even to increase) the systemic risk which originates from the network effects w_i.

The main problem of this allocation is that it is based on the subsystem construction scheme Λ^0. Whilst in the portfolio framework the entities outside of a subsystem had no influence on the risk evaluation of the subsystem, the subsystem construction scheme Λ^0 just sets the equity of the neighboring entities to zero. However, this does not imply that they have no impact on the subsystem anymore, since the network linkages have not changed at all. Even worse the banks outside of the subsystem are assumed to be already in default which means that they transmit all the losses. Thus this construction scheme corresponds in some sense to a worst-case view on how much a subsystem is able to spread its losses within the whole system. This interpretation is also in line with the definition of the core, i.e. that the construction scheme is always an upper bound of the subsystems allocation.

Another problem with the core allocations in this interaction model is that each entity does not only act as a spreader of risk as in the portfolio approach, but can also function as a transmitter of the losses of some other entities. This perspective is exactly the crucial part for the fuzzy core element from above. Namely the banks which are in $\mathcal{D}\backslash\mathcal{D}_0$, do not contribute losses to the system and thus any core allocation must be bounded by zero. Nonetheless in the complete system they face losses from other institutions and transmit them further into the system. However, they can not be charged for this loss transmission as their share is already capped by zero.

Contrarily, we now want to find an appropriate subsystem construction scheme such that the causality of the risk of a subsystem can solely be explained by the subsystem itself. Moreover, we also want that the feedback effects within the subsystem remain intact. The most intuitive choice for such a subsystem construction scheme is equipping all banks outside of the subsystem with a very high amount of capital such that these banks can never face a default, i.e.

$$\Lambda^b(x, \lambda) = \Lambda(\lambda * x + (\mathbf{1}_d - \lambda) * b), \tag{4.3.4}$$

with $b \in \mathbb{R}^d_+$ sufficiently large.

Note that, whilst in the financial network without feedback mechanisms we had that joining two subgroups of banks always resulted in a risk reduction compared to the sum of the single risks, it might now happen that two single subsystems are not able to trigger a default of a bank but they can in a combined subsystem. That is the diversification benefit can turn into a cost. Moreover, in contrast to the prior subsystem construction scheme Λ^0, we now have a best-case view as we suppose that the external system is capable of covering all losses.

4.4 The reverse core

Because of the change of the perspective towards a best-case view, we also need to change the definition of a fair allocation in a way that the allocation of a subsystem should at least cover the risk of the subsystem. Moreover, for this new subsystem construction, we have that the risk of a single financial institution is just a measure of the adverse market event, since we assume that no loss can spread to the other institutions. This is in line with the current market practice of measuring the risk on a standalone basis. Therefore we should demand that the allocation of the systemic risk to this bank does not fall below this threshold in order to cover it own losses. For this reason we introduce the notion of the reverse (fuzzy) core.

Definition 4.4.1 (Reverse (fuzzy) core)**.** *Let $X \in \mathcal{X}^d$ and $\widetilde{\rho} : \mathcal{X}^d \times \mathcal{P} \to \mathbb{R}$ a subsystem risk measure. We say that $k \in \mathbb{R}^d$ is in the reverse core $C^+_{\widetilde{\rho}}(X)$ if $\sum_{i=1}^d k_j = \rho(X, \mathcal{I})$ and for all $J \in \mathcal{P}$*

$$\widetilde{\rho}(X, J) \leq \sum_{j \in J} k_j.$$

Similarly, we say that k is in the reverse fuzzy core $FC^+_\rho(X)$ for a subsystem risk measure $\rho : \mathcal{X}^d \times \mathbb{R}^d_+ \to \mathbb{R}$ if $\sum_{i=1}^d k_i = \rho(X, \mathbf{1}_d)$ and for all $\lambda \in [0,1]^d$ we have that

$$\rho(X, \lambda) \leq \lambda^\top k.$$

Next we investigate the relationship between the core and the reverse core which are generated by the same subsystem risk measure. Note that a similar result also holds for the fuzzy cores.

Lemma 4.4.2. *Let $\widetilde{\rho} : \mathcal{X}^d \times \mathcal{P} \to \mathbb{R}$ be a subsystem risk measure. The core and the reverse core are related in one of the following ways*

- *$C^-_{\widetilde{\rho}}(X) = C^+_{\widetilde{\rho}}(X) = \emptyset$;*
- *$C^-_{\widetilde{\rho}}(X) = C^+_{\widetilde{\rho}}(X) = \{k\}$ with $\widetilde{\rho}(X, J) = \sum_{j \in J} k_j$ for all $J \in \mathcal{P}$;*
- *One core contains only allocations where the inequality is strict for at least one $J \in \mathcal{P}$ and the other core is empty.*

In particular, if $C^\pm_{\widetilde{\rho}}(X) \neq \emptyset$, then $|C^\mp_{\widetilde{\rho}}(X)| \leq 1$.

Proof. Suppose $C^-_{\widetilde{\rho}}(X)$ as well as $C^+_{\widetilde{\rho}}(X)$ are non-empty. Let $k^- \in C^-_{\widetilde{\rho}}(X)$ and $k^+ \in C^+_{\widetilde{\rho}}(X)$ from which it follows that

$$\sum_{i=1}^d k_i^- = \sum_{i=1}^d k_i^+ \quad \text{and} \quad \sum_{j\in J} k_j^- \leq \sum_{j\in J} k_j^+ \text{ for all } J \in \mathcal{P}.$$

We define further $K := \{i \in \mathcal{I} : k_i^- < k_i^+\}$ and $G := \{i \in \mathcal{I} : k_i^- > k_i^+\}$. We assume that $G \cup K \neq \emptyset$. If $K \neq \emptyset$, then $\sum_{i=1}^d k_i^- = \sum_{i=1}^d k_i^+$ implies that

$$\sum_{i\in G} k_i^+ - k_i^- = \sum_{i\in K} k_i^- - k_i^+ < 0.$$

Thus also $G \neq \emptyset$, which contradicts $\sum_{j\in J} k_j^- \leq \sum_{j\in J} k_j^+$ for all $J \in \mathcal{P}$. Hence $k^- = k^+$ and we can deduce that both cores are equal and single valued.
Finally, if one core is empty, we clearly have that

$$\left\{ k \in \mathbb{R}^d : \widetilde{\rho}(X, J) = \sum_{j\in J} k_j, \forall J \in \mathcal{P} \right\} = \emptyset$$

and thus the other core has to be empty as well or the inequality has to be strict for at least one $J \in \mathcal{P}$ for each allocation. □

In the prior lemma we studied the connection of the two core concepts for the same subsystem risk measure. In contrast to this, we will see in the next lemma that we can also translate one core concept to the other by changing the underlying subsystem risk measure.

Lemma 4.4.3. *Let $\widetilde{\rho} : \mathcal{X}^d \times \mathcal{P} \to \mathbb{R}$ be a subsystem risk measure with $\widetilde{\rho}(X, \emptyset) = 0$ for all $X \in \mathcal{X}^d$. Then*

$$C^+_{\widetilde{\rho}}(X) \subseteq \left\{ k \in \mathbb{R}^d : \sum_{j\in J} k_j \leq \rho(X) - \widetilde{\rho}\left(X, J^C\right) \ \forall J \in \mathcal{P} \right\},$$

where $J^C := \mathcal{I}\backslash J$ is again the complement w.r.t. the complete system. The interpretation is that each element of the reverse core must "undercut" the "with and without risk". In particular by defining the subsystem risk measure $\bar{\rho}$ via

$$\bar{\rho}(X, J) := \rho(X) - \widetilde{\rho}(X, J^C)$$

we obtain from the result above and from $\rho(X) - \bar{\rho}(X, J^C) = \widetilde{\rho}(X, J)$ that

$$C^+_{\widetilde{\rho}}(X) = C^-_{\bar{\rho}}(X).$$

Proof. Let $k \in C^+_{\widetilde{\rho}}(X)$. Then for each $J \in \mathcal{P}$ it holds that

$$\widetilde{\rho}(X, J^C) - \sum_{j\in J^C} k_j \leq 0 = \rho(X, \mathcal{I}) - \sum_{i=1}^d k_i,$$

which is equivalent to

$$\sum_{j\in J} k_j \leq \rho(X) - \widetilde{\rho}\left(X, J^C\right).$$

□

We remark that a similar result also holds for the fuzzy and reverse fuzzy cores.

4.5 The reverse core in the financial model with contagion

Equipped with this new core concept, we come back to our interaction model from Section 4.3. Recall that we are interested in a risk factor $x = c + y$, where $c \in \mathbb{R}^d_+$ is the vector of some initial capital endowments of the financial institutions and $-y \in \mathbb{R}^d_+$ is a negative shock. For the subsystem construction scheme Λ^b defined in (4.3.4) we vaguely demanded that $b \in \mathbb{R}^d_+$ should be sufficiently large. In Lemma 4.5.2 we will show that it is already sufficient to consider Λ^c in order that the reverse fuzzy core $FC^+_{\Lambda^c}(x)$ is non-empty. Moreover, similar to Lemma 4.2.5 we derive that $FC^+_{\Lambda^c}(x) \subseteq FC^+_{\Lambda^b}(x)$ for all $b \geq c$.
Note that the subsystem construction scheme Λ^b is independent of the specific decomposition of the risk factor x into a positive capital amount c and a shock y. Therefore, if we allow also for positive shocks, i.e. $x = c + y$ with $y \in \mathbb{R}^d$, then we can choose the decomposition $x = \widetilde{c} + \widetilde{y}$ with $\widetilde{c} = c + \max\{y, \mathbf{0}_d\}$ and $\widetilde{y} = \min\{y, \mathbf{0}_d\}$. Since $\widetilde{c}$ and $-\widetilde{y}$ are again positive, we have that $FC^+_{\Lambda^{\widetilde{c}}}(x) \subseteq FC^+_{\Lambda^b}(x)$. Thus, if we are interested in the non-emptiness of the reverse core of Λ^b, assuming a negative shock is essentially not a restriction.
Before we identify an element of the reverse fuzzy core of Λ^c, we need the following preparatory lemma:

Lemma 4.5.1. *Let* $\mathbf{A} = (A_{i,j})_{i,j=1,\dots,d} \in \mathbb{R}^{d\times d}_+$ *and* $b \in \mathbb{R}^d_+$. *Then there exists a* $\mathbf{B} = (B_{i,j})_{i,j=1,\dots,d} \in \mathbb{R}^{d\times d}_+$ *such that*

$$\left(\sum_{j=1}^{d} A_{\cdot,j} - b\right)^+ = \sum_{j=1}^{d} \left(A_{\cdot,j} - B_{\cdot,j}\right)^+, \tag{4.5.1}$$

where $\sum_{j=1}^d B_{i,j} = b_i$ *and* $B_{i,i}$ *is either equal to* $A_{i,i}$ *or* b_i *for all* $i = 1, \dots, d$.

Proof. Let $i \in \mathcal{I}$ be fixed. We denote by $\pi : \mathcal{I} \to \mathcal{I}$ the permutation which exchanges the first and the i-th entry, i.e. $\pi(1) = i$, $\pi(i) = 1$ and $\pi(j) = j$ for all $j \notin \{1, i\}$. We distinguish the following two cases:

- If $\sum_{j=1}^d A_{i,j} \leq b_i$, then set $B_{i,\pi(j)} := A_{i,\pi(j)}$ for all $j = 1, \dots, d-1$ and $B_{i,\pi(d)} := b_i - \sum_{j=1}^{d-1} A_{i,\pi(j)}$.

- If $\sum_{j=1}^d A_{i,j} > b_i$, then define for $j = 1, \dots, d$

$$B_{i,\pi(j)} := A_{i,\pi(j)} \mathbb{1}_{\{\sum_{k=1}^j A_{i,\pi(k)} \leq b_i\}}$$

$$+\left(b_i - \sum_{k=1}^{j-1} A_{i,\pi(k)}\right) \mathbb{1}_{\{\sum_{k=1}^{j-1} A_{i,\pi(k)} \leq b_i, \sum_{k=1}^{j} A_{i,\pi(k)} > b_i\}}.$$

Obviously, (4.5.1) is fulfilled and we have for this choice of $B_{i,\cdot}$ that $\sum_{j=1}^{d} B_{i,j} = b_i$. Depending on the size of $A_{i,i}$, we have either $B_{i,i} = A_{i,i}$ or $B_{i,i} = b_i$. □

Lemma 4.5.2. *Let $x = y + c$ with $y \leq \mathbf{0}_d$ and $c \in \mathbb{R}^d_+$. Then the reverse fuzzy core $FC^+_{\Lambda^c}(x)$ with the subsystem construction scheme*

$$\Lambda^c(x, \lambda) := \Lambda\left(\lambda * x + (\mathbf{1}_d - \lambda) * c\right) = \Lambda(\lambda * y + c),$$

is non-empty. Moreover, there exists an allocation $k \in FC^+_{\Lambda^c}(x)$ such that for all $i = 1, ..., d$

$$k_i \leq \Lambda(x_i e_i),$$

that is k also fulfills the property of the core for the single financial institutions and the subsystem construction scheme $\Lambda^0(x, \lambda) = \Lambda(\lambda * x)$.

Proof. As in Lemma 4.3.2, we will use the fact that the optimal value for a in the optimization problem (4.3.1) of the aggregation function Λ can be found by iterating $a(n) = \left(\mathbf{\Pi}^\top a(n-1) - x\right)^+$, $n \in \mathbb{N}$ with $a(0) := \mathbf{0}_d$.
First, we iteratively define a non-negative partition $A_i(n) \in \mathbb{R}^d_+, i = 1, ..., d$ of $a(n)$, that is $\sum_{i=1}^{d} A_i(n) = a(n)$. Clearly, $A_i(0) := \mathbf{0}_d, i = 1, ..., d$ is a partition of $a(0) = \mathbf{0}_d$. Note that, if we have found a partition of $a(n-1)$ for some $n \in \mathbb{N}$, then

$$\begin{aligned} a(n) &= \left(\mathbf{\Pi}^\top a(n-1) - x\right)^+ \\ &= \left(\sum_{i=1}^{d} \mathbf{\Pi}^\top A_i(n-1) - y - c\right)^+ \\ &= \left(\sum_{i=1}^{d} \left(\mathbf{\Pi}^\top A_i(n-1) - y_i e_i\right) - c\right)^+. \end{aligned}$$

Since $\mathbf{\Pi}^\top A_i(n-1) - y_i e_i \geq \mathbf{0}_d$ for all $i = 1, ..., d$ and $c \geq \mathbf{0}_d$, we can apply Lemma 4.5.1 in order to obtain the existence of $C_i(n) \in \mathbb{R}^d_+, i = 1, ..., d$ with $\sum_{i=1}^{d} C_i(n) = c$ and

$$a(n) = \sum_{i=1}^{d} \left(\mathbf{\Pi}^\top A_i(n-1) - y_i e_i - C_i(n)\right)^+.$$

Hence, $A_i(n) := \left(\mathbf{\Pi}^\top A_i(n-1) - y_i e_i - C_i(n)\right)^+ \in \mathbb{R}^d_+, i = 1, ..., d$ is a non-negative partition of $a(n)$.

Since $(a(n))_{n\in\mathbb{N}}$ is an increasing sequence, we have that $A_i(n)$ is also bounded from above by $a = \lim_{n\to\infty} a(n)$ for all $i = 1, ..., d$. Recall that a is finite by Lemma 4.3.1. Thus $((A_i(n))_{i=1,...,d})_{n\in\mathbb{N}}$ is a bounded sequence and we obtain by applying the Bolzano-Weierstrass theorem, that there exists a converging subsequence $\big((A_i(n_k))_{i=1,...,d}\big)_{k\in\mathbb{N}}$. We denote the limit of this subsequence by $(A_i)_{i=1,...,d}$. Thus, we have that

$$a = \lim_{k\to\infty} a(n_k) = \lim_{k\to\infty} \sum_{i=1}^{d} A_i(n_k) = \sum_{i=1}^{d} A_i$$

and that $A_i \geq \mathbf{0}_d$ for all $i = 1, ..., d$.

Next we define for a level of participation $\lambda \in [0,1]^d$ the sequence

$$a(\lambda, n) := \Big(\mathbf{\Pi}^\top a(\lambda, n-1) - \lambda * y - c\Big)^+ \quad \text{for all } n \in \mathbb{N}$$

and $a(\lambda, 0) = \mathbf{0}_d$ which corresponds to the fixpoint iteration for $\Lambda^c(x,\lambda)$. In the following we prove that

$$\sum_{i=1}^{d} \lambda_i A_i(n) \geq a(\lambda, n), \qquad \text{for all } n \in \mathbb{N}. \tag{4.5.2}$$

Obviously (4.5.2) holds true for $n = 0$. Suppose now that (4.5.2) is valid for $n-1$, with $n \in \mathbb{N}$. Then, we have that

$$\begin{aligned}
\sum_{i=1}^{d} \lambda_i A_i(n) &= \sum_{i=1}^{d} \lambda_i \Big(\mathbf{\Pi}^\top A_i(n-1) - y_i e_i - C_i(n)\Big)^+ \\
&\geq \left(\mathbf{\Pi}^\top \sum_{i=1}^{d} \lambda_i A_i(n-1) - \sum_{i=1}^{d} \lambda_i y_i e_i - \sum_{i=1}^{d} \lambda_i C_i(n)\right)^+ \\
&\geq \left(\mathbf{\Pi}^\top \sum_{i=1}^{d} \lambda_i A_i(n-1) - \lambda * y - c\right)^+ \\
&\geq \Big(\mathbf{\Pi}^\top a(\lambda, n-1) - \lambda * y - c\Big)^+ \\
&= a(\lambda, n),
\end{aligned}$$

where we used the induction hypothesis in the penultimate step and that $\sum_{i=1}^{d} \lambda_i C_i(n) \leq c$ in the third. By taking the limit we obtain

$$\sum_{i=1}^{d} \lambda_i A_i = \lim_{k\to\infty} \sum_{i=1}^{d} \lambda_i A_i(n_k) \geq \lim_{k\to\infty} a(\lambda, n_k) =: a(\lambda).$$

Finally, by defining $k = (k_1, ..., k_d)$ with $k_i := \mathbf{1}_d^\top A_i, i = 1, ..., d$, we have for all $\lambda \in [0,1]^d$ that

$$\lambda^\top k = \mathbf{1}_d^\top \sum_{i=1}^{d} \lambda_i A_i \geq \mathbf{1}_d^\top a(\lambda) = \Lambda^c(x, \lambda)$$

and

$$\mathbf{1}_d^\top k = \mathbf{1}_d^\top \sum_{i=1}^{d} A_i = \mathbf{1}_d^\top a = \Lambda^c(x, \mathbf{1}_d) = \Lambda(x).$$

Thus $k = (k_1, ..., k_d) \in FC^+_{\Lambda^c}(x)$.

In the end we still have to show that $k_i \leq \Lambda^0(x, e_i)$ for all $i = 1, ..., d$. For this purpose denote by

$$a(i,n) := \left(\mathbf{\Pi}^\top a(i,n-1) - x_i e_i\right)^+, \quad a(i,0) = \mathbf{0}_d,$$

the fixpoint iteration for $\Lambda^0(x, e_i)$. Again we will use induction to show that

$$A_i(n) \leq a(i,n), \quad \text{for all } n \in \mathbb{N}. \tag{4.5.3}$$

Then, obviously $A_i(0) \leq a(i,0)$. Thus suppose that (4.5.3) holds for $n-1$. Before we proceed with the induction step recall that by construction of $\big(C_i(n)\big)_i = (\mathbf{\Pi}^\top A_i(n-1) - y_i e_i)_i$ or $\big(C_i(n)\big)_i = c_i$ and in both cases we obtain

$$\left(\big(\mathbf{\Pi}^\top A_i(n-1)\big)_i - y_i - \big(C_i(n)\big)_i\right)^+ \leq \left(\big(\mathbf{\Pi}^\top A_i(n-1)\big)_i - y_i - c_i\right)^+.$$

Moreover, since $C_i(n) \geq \mathbf{0}_d$ for all $i = 1, ..., d$ we have that

$$\begin{aligned} A_i(n) &= \left(\mathbf{\Pi}^\top A_i(n-1) - y_i e_i - C_i(n)\right)^+ \\ &\leq \left(\mathbf{\Pi}^\top A_i(n-1) - y_i e_i - c_i e_i\right)^+ \\ &\leq \left(\mathbf{\Pi}^\top a(i,n-1) - y_i e_i - c_i e_i\right)^+ = a(i,n). \end{aligned}$$

Hence (4.5.3) holds and we can conclude that

$$k_i = \mathbf{1}_d^\top A_i \leq \mathbf{1}_d^\top a(i) = \Lambda^0(x, e_i),$$

where $a(i) := \lim_{k\to\infty} a(i, n_k)$.

□

We have seen in Lemma 4.5.2 that the reverse fuzzy core for the subsystem construction scheme Λ^c is non-empty and that there exists an element in the reverse fuzzy core which additionally fulfills the essential property (4.2.1) of the core of Λ^0 at least for the single institutions. As we have seen that the usual core might not be a useful allocation in a financial model with contagion, we want to investigate if there is also an element in the intersection of the two cores $C^+_{\Lambda^c}$ and $C^-_{\Lambda^0}$. In Lemma 4.5.4 it will be shown that under a rather weak assumption on the risk factor x the intersection of the cores is empty. In order to put this assumption into context, we precede the following lemma.

Lemma 4.5.3. *We have for all $x \in \mathbb{R}^d$ that*

$$\sum_{i\in\mathcal{D}_0} \Lambda\left(\sum_{j\in\mathcal{D}_0^C\cup\{i\}} e_j x_j\right) \leq \Lambda(x), \tag{4.5.4}$$

where $\mathcal{D}_0 := \{i \in \mathcal{I} : x_i \leq 0\}$ denotes the set of institutes which default initially.
Moreover, if (4.5.4) *is strict, then there is at least one institution which defaults due to contagion, i.e. $\mathcal{D}\backslash\mathcal{D}_0 \neq \emptyset$.*

Proof. Similar to Lemma 4.5.2 we consider the sequences

$$a(n) := \left(\mathbf{\Pi}^\top a(n-1) - x\right)^+, \quad a(0) = \mathbf{0}_d,$$

and for all $i \in \mathcal{D}_0$

$$a(i,n) := \left(\mathbf{\Pi}^\top a(i,n-1) - \sum_{j \in \mathcal{D}_0^C \cup \{i\}} e_j x_j\right)^+, \quad a(i,0) = \mathbf{0}_d.$$

By construction $a(0) = \sum_{i \in \mathcal{D}_0} a(i,0)$. Thus suppose that

$$a(n-1) \geq \sum_{i \in \mathcal{D}_0} a(i,n-1)$$

for some $n \in \mathbb{N}$, then

$$a(n) = \left(\mathbf{\Pi}^\top a(n-1) - x\right)^+ \geq \left(\mathbf{\Pi}^\top \sum_{i \in \mathcal{D}_0} a(i,n-1) - x\right)^+. \tag{4.5.5}$$

Now we look at the single entries of the vector on the right hand side of (4.5.5). For $l \in \mathcal{D}_0$ we have

$$\begin{aligned}
\left(\mathbf{\Pi}^\top \sum_{i \in \mathcal{D}_0} a(i,n-1) - x\right)_l^+ &= \left(\sum_{i \in \mathcal{D}_0} \left(\mathbf{\Pi}^\top a(i,n-1)\right)_l - x_l\right)^+ \\
&= \sum_{i \in \mathcal{D}_0} \left(\mathbf{\Pi}^\top a(i,n-1)\right)_l - x_l \\
&= \left(\left(\mathbf{\Pi}^\top a(l,n-1)\right)_l - x_l\right)^+ + \sum_{i \in \mathcal{D}_0 \setminus \{l\}} \left(\left(\mathbf{\Pi}^\top a(i,n-1)\right)_l\right)^+ \\
&= \sum_{i \in \mathcal{D}_0} \left(\mathbf{\Pi}^\top a(i,n-1) - \sum_{j \in \mathcal{D}_0^C \cup \{i\}} e_j x_j\right)_l^+
\end{aligned}$$

and for $l \in \mathcal{D}_0^C$

$$\begin{aligned}
\left(\mathbf{\Pi}^\top \sum_{i \in \mathcal{D}_0} a(i,n-1) - x\right)_l^+ &= \left(\sum_{i \in \mathcal{D}_0} \left(\mathbf{\Pi}^\top a(i,n-1)\right)_l - x_l\right)^+ \\
&= \sum_{i \in \mathcal{D}_0} \left(\left(\mathbf{\Pi}^\top a(i,n-1)\right)_l - X_{l,i}\right)^+ \\
&\geq \sum_{i \in \mathcal{D}_0} \left(\left(\mathbf{\Pi}^\top a(i,n-1)\right)_l - x_l\right)^+
\end{aligned}$$

$$
\begin{aligned}
&= \sum_{i \in \mathcal{D}_0} \left(\mathbf{\Pi}^\top a(i, n-1) - x \right)_l^+ \\
&= \sum_{i \in \mathcal{D}_0} \left(\mathbf{\Pi}^\top a(i, n-1) - \sum_{j \in \mathcal{D}_0^C \cup \{i\}} e_j x_j - \sum_{j \in \mathcal{D}_0 \setminus \{i\}} e_j x_j \right)_l^+ \\
&\geq \sum_{i \in \mathcal{D}_0} \left(\mathbf{\Pi}^\top a(i, n-1) - \sum_{j \in \mathcal{D}_0^C \cup \{i\}} e_j x_j \right)_l^+
\end{aligned}
$$

where $(X_{l,i})_{l \in \mathcal{D}_0^C, i \in \mathcal{D}_0}$ is specified by Lemma 4.5.1. Note that, since $X_{l,i} \geq 0$ and $\sum_{i \in \mathcal{D}_0} X_{l,i} = x_l$ we have that $X_{l,i} \leq x_l$ for all $l \in \mathcal{D}_0^C$ and $i \in \mathcal{D}_0$ which we used in the third step. Now we can continue with (4.5.5)

$$
a(n) \geq \sum_{i \in \mathcal{D}_0} \left(\mathbf{\Pi}^\top a(i, n-1) - \sum_{j \in \mathcal{D}_0^C \cup \{i\}} e_j x_j \right)^+ = \sum_{i \in \mathcal{D}_0} a(i, n)
$$

and thus we have shown that $a(n) \geq \sum_{i \in \mathcal{D}_0} a(i, n)$ for all $n \in \mathbb{N}$. Finally by considering the limit for $n \to \infty$ we obtain (4.5.4).

Next we show the second claim. For this we will prove that $\mathcal{D} \setminus \mathcal{D}_0 = \emptyset$ implies that (4.5.4) holds with equality. Obviously this is true if $\mathcal{D}_0 = \emptyset$. Thus we suppose that $\mathcal{D}_0 \neq \emptyset$, i.e. at least one bank defaults initially. It can be readily seen that $\mathcal{D}_0 \subseteq \mathcal{D}$ and thus, since $\mathcal{D} \setminus \mathcal{D}_0 = \emptyset$, we have that $\mathcal{D}_0 = \mathcal{D}$. For each $i \in \mathcal{D}$ we will also need the sets of banks which default initially and after all possible contagion took place for the subsystem with corresponding risk factor $\sum_{j \in \mathcal{D}_0^C \cup \{i\}} e_j x_j$. We denote these sets by $\mathcal{D}_0(i)$ and $\mathcal{D}(i)$ respectively. Since for all $i \in \mathcal{D}$ we have that $\sum_{j \in \mathcal{D}_0^C \cup \{i\}} e_j x_j \geq x$, it follows directly that

$$
\mathcal{D}(i) \subseteq \mathcal{D}.
$$

Contrarily, due to the fact that $\left(\sum_{j \in \mathcal{D}_0^C \cup \{i\}} e_j x_j \right)_l \leq 0$ for all $l \in \mathcal{D}_0$, we also have that

$$
\mathcal{D}(i) \supseteq \mathcal{D}_0(i) = \mathcal{D}_0 = \mathcal{D}
$$

and thus $\mathcal{D}(i) = \mathcal{D}$ for all $i \in \mathcal{D}$.

As in Lemma 4.3.2 let $\mathcal{D} = \{p_1, ..., p_{|\mathcal{D}|}\}$ and denote by $\mathbf{\Pi}_{\mathcal{D},\mathcal{D}} := (\Pi_{i,j})_{i,j \in \mathcal{D}} \in \mathbb{R}^{|\mathcal{D}| \times |\mathcal{D}|}$ the matrix $\mathbf{\Pi}$ where the rows and columns which are not in $\mathcal{D}$ have been erased. Similar to Lemma 4.3.2 we get that

$$
\begin{aligned}
\Lambda(x) &= -\mathbf{1}_{|\mathcal{D}|}^\top \left(\mathbf{I}_{|\mathcal{D}|} - \mathbf{\Pi}_{\mathcal{D},\mathcal{D}}^\top \right)^{-1} \sum_{i=1}^{|\mathcal{D}|} \widetilde{e}_i x_{p_i} \\
&= \sum_{i=1}^{|\mathcal{D}|} -\mathbf{1}_{|\mathcal{D}|}^\top \left(\mathbf{I}_{|\mathcal{D}|} - \mathbf{\Pi}_{\mathcal{D},\mathcal{D}}^\top \right)^{-1} \widetilde{e}_i x_{p_i}
\end{aligned}
$$

$$= \sum_{i \in \mathcal{D}_0} \Lambda \left(\sum_{j \in \mathcal{D}_0^C \cup \{i\}} e_j x_j \right),$$

where $\widetilde{e}_i \in \mathbb{R}^{|\mathcal{D}|}$ is the i-th $|\mathcal{D}|$-dimensional unit vector. □

It is obvious that the reverse implication of Lemma 4.5.3 does not hold. As a counterexample take for instance a financial network comprising two banks, where the first bank defaults initially and the second fails as a consequence of this default. Then we have a contagious default, but (4.5.4) holds with equality.
Therefore, (4.5.4) is strict if there is a contagious default and this default must be triggered by more than one defaulted bank. Thus (4.5.4) being strict can be interpreted as a scenario of a high level of interactions in the network.

Lemma 4.5.4. *If* (4.5.4) *is strict for some* $x \in \mathbb{R}^d$*, i.e.*

$$\sum_{i \in \mathcal{D}_0} \Lambda \left(\sum_{j \in \mathcal{D}_0^C \cup \{i\}} e_j x_j \right) < \Lambda(x),$$

then

$$C^+_{\widetilde{\Lambda}^c}(x) \cap C^-_{\widetilde{\Lambda}^0}(x) = \emptyset,$$

where $\widetilde{\Lambda}^b(x, J) := \Lambda\left(\sum_{j \in J} x_j e_j + \sum_{j \in J^C} b_j e_j\right)$ *for all* $J \in \mathcal{P}$ *and* $b \in \mathbb{R}^d$*.*

Proof. Assume there exists an $k \in C^+_{\widetilde{\Lambda}^c}(x) \cap C^-_{\widetilde{\Lambda}^0}(x)$. We have for all $i \in \mathcal{D}_0^C$ that

$$0 \leq \Lambda \left(e_i x_i + \sum_{j \neq i} e_j c_j \right) \leq \Lambda(e_i x_i) = 0,$$

and thus the respective core properties imply that $k_i = 0$. Hence

$$\sum_{i=1}^d k_i = \sum_{i \in \mathcal{D}_0} k_i = \sum_{i \in \mathcal{D}_0} \sum_{j \in \mathcal{D}_0^C \cup \{i\}} k_i \leq \sum_{i \in \mathcal{D}_0} \Lambda \left(\sum_{j \in \mathcal{D}_0^C \cup \{i\}} e_j x_j \right) < \Lambda(x) = \sum_{i=1}^d k_i,$$

which is a contradiction. □

We finish this section with a small but concrete calculation of the core and the reverse core in order to exemplify their differences. We consider a financial network with the following specifications:

$$\mathbf{\Pi} = \begin{pmatrix} 0 & 1/2 & 0 & 0 & 1/2 \\ 0 & 0 & 1/2 & 0 & 1/2 \\ 0 & 1/4 & 0 & 1/2 & 1/4 \\ 0 & 1/3 & 0 & 0 & 2/3 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \text{ and } x = c + y = \begin{pmatrix} 0 - 5 \\ 5 - 3 \\ 10 - 12 \\ 5 - 3 \\ 0 - 0 \end{pmatrix} = \begin{pmatrix} -5 \\ 2 \\ -2 \\ 2 \\ 0 \end{pmatrix}.$$

The corresponding network is depicted in Figure 4.5.1 and the values of the subsystem construction schemes can be found in Table 4.5.1. Note that, since the inclusion of the real economy does not change the risk of a subsystem, we omit the results in Table 4.5.1. Clearly, the initially defaulting banks are $\mathcal{D}_0 = \{1, 3\}$. Moreover, we observe that the initial default of bank 1 triggers a contagious default of bank 2 and that even after all possible defaults took place, bank 4 remains solvent. Applying Lemma 4.3.2 yields that

$$(13.57, -4.86, 3.71, 0, 0)^\top = FC^-_{\Lambda^0}(x)$$

and by using the partition from the proof of Lemma 4.5.2 we obtain that

$$(8.71, 0, 3.71, 0, 0)^\top \in FC^+_{\Lambda^c}(x).$$

As bank 4 is not participating in the contagion process, it gets in both allocations a share of zero which can be considered as fair. However, here we see clearly that bank 2 is a transmitter of losses in the system and by the allocation of the fuzzy core it is rewarded for this position with a negative share compared to the solvent bank 4. Contrarily, bank 2 also gets a share of zero for the allocation which is in the reverse fuzzy core. Since, bank 2 does not default initially this allocation can barely be considered as fair. However, since bank 2 is also in a channel of contagion later on, it would also be fair that bank 2 gets a strictly positive share. Using the Table 4.5.1, it can be readily seen that this holds for all other allocations in the reverse core, which is given by

$$\begin{aligned} C^+_{\widetilde{\Lambda}^c}(x) = \Big\{ (k_1, k_2, k_3, 0, 0) \in \mathbb{R}^5_+ : \sum_{i=1}^{3} k_i = 12.43, \\ k_1 \geq 7.5, k_3 \geq 2.5, k_1 + k_2 \geq 8.25 \Big\}. \end{aligned}$$

The largest share of the systemic risk for bank 2 in the reverse core is attained for the allocation

$$(7.5, 2.43, 2.5, 0, 0)^\top.$$

Furthermore, since $\mathcal{D}_0 = \{1, 3\}$ and

$$\begin{aligned} \Lambda(x) &= 12.43 > 11.21 = 8.71 + 2.5 \\ &= \widetilde{\Lambda}^0(x, \{1, 2, 4, 5\}) + \widetilde{\Lambda}^0(x, \{2, 3, 4, 5\}) \\ &= \Lambda\left(\sum_{j \in \mathcal{D}_0^C \cup \{1\}} x_j e_j \right) + \Lambda\left(\sum_{j \in \mathcal{D}_0^C \cup \{3\}} x_j e_j \right), \end{aligned}$$

(4.5.4) is strict and thus the reverse core of $\widetilde{\Lambda}^c$ and the core of $\widetilde{\Lambda}^0$ do not have a common element.

Finally, we also observe that not only the fuzzy core, but also all core elements do not respect a fair ordering in the sense that $k_u \geq k_v \geq k_w$ for all $u \in \mathcal{D}_0, v \in \mathcal{D} \backslash \mathcal{D}_0$ and $w \in \mathcal{D}^C$. Recall that bank 2 defaults due to contagion, but not initially, and thus a core allocation k must fulfill that $k_2 \leq 0$. Since, this bank participates in the contagion process later on we want that its allocation should be non-negative. Now we assume that there exists an allocation $k \in C^{-}_{\widetilde{\Lambda}^0}(x)$ which respects our notion of a fair ordering, i.e. $k = (k_1, 0, k_3, k_4, 0)$ such that $k_1, k_3 \geq 0$ and $k_4 \leq 0$. Then,

$$\begin{aligned} k_4 &= (k_1 + k_4) + (k_3 + k_4) - \sum_{i=1}^{5} k_i \\ &\leq \widetilde{\Lambda}^0(x, \{1,2,4,5\}) + \widetilde{\Lambda}^0(x, \{2,3,4,5\}) - \Lambda(x) = -1.22. \end{aligned}$$

Moreover, we have that

$$12.43 = \Lambda(x) = \sum_{i=1}^{5} k_i \leq \widetilde{\Lambda}^0(x, \{1,2\}) + \widetilde{\Lambda}^0(x, \{2,3\}) + k_4 = 13.28 + k_4,$$

which immediately yields the contradiction that $k_4 \geq -0.85$. Hence there does not exist a core element which respects the fair ordering from above.

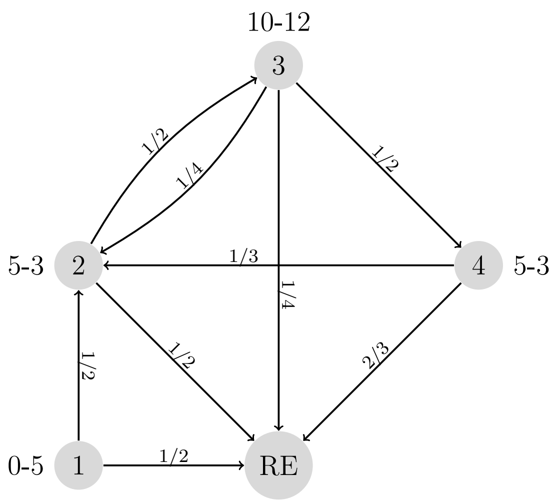


Figure 4.5.1: Exemplary system.

J	$\widetilde{\Lambda}^c(x, J)$	$\widetilde{\Lambda}^0(x, J)$
$\{1,2,3,4\}$	12.43	12.43
$\{1,2,3\}$	12.43	15.95
$\{1,2,4\}$	8.25	8.71
$\{1,3,4\}$	10.00	17.29
$\{2,3,4\}$	2.50	2.50
$\{1,2\}$	8.25	9.11
$\{1,3\}$	10.00	22.37
$\{1,4\}$	7.50	13.57
$\{2,3\}$	2.50	4.17
$\{2,4\}$	0	0
$\{3,4\}$	2.50	3.71
$\{1\}$	7.50	15.53
$\{2\}$	0	0
$\{3\}$	2.50	6.84
$\{4\}$	0	0

Table 4.5.1: Risks of the subsystems.

4.A Random risks

This section is devoted to the discussion on how we can derive the non-emptiness of the core also for random risks. For this purpose, we first recall the well-known Bondareva-Shapley theorem which gives an alternative characterization of the non-emptiness of the core. For this we need the notion of a balanced collection of weights.

Definition 4.A.1. *We say $(\alpha_J)_{J\in\mathcal{P}}$ is a balanced collection of weights if $\alpha_J \geq 0$ for all $J \in \mathcal{P}$ and $\sum_{J\in\mathcal{P}_i} \alpha_J = 1$ for all $i = 1, ..., d$. Here $\mathcal{P}_i := \{J \in \mathcal{P} : i \in J\}$ denotes the set of all subgroups containing the i-th financial institution.*

Theorem 4.A.2 (Bondareva-Shapley)**.** *The core $C^-_{\widetilde{\rho}}(X)$ of the subsystem risk measure $\widetilde{\rho}$ is not empty if and only if for all balanced collections of weights $(\alpha_J)_{J\in\mathcal{P}}$ it holds that*

$$\rho(X) \leq \sum_{J\in\mathcal{P}} \alpha_J \widetilde{\rho}(X, J).$$

For a proof see for instance Shapley (1967).

In the following we suppose that $\widetilde{\rho} : \mathcal{X}^d \times \mathcal{P} \to \mathbb{R}$ is given by

$$\widetilde{\rho}(X, J) := \eta(-\widetilde{\Lambda}(X, J)),$$

where $\widetilde{\Lambda} : \mathbb{R}^d \times \mathcal{P} \to \mathbb{R}$ is a subsystem construction scheme and η is a univariate risk measure.

Lemma 4.A.3. *If the subsystem risk measure $\widetilde{\rho}$ is given by $\widetilde{\rho}(X, J) = \eta(-\widetilde{\Lambda}(X, J))$ for all $X \in \mathcal{X}^d$ and $J \in \mathcal{P}$, where η is a positive homogeneous and subadditive univariate risk measure and $\widetilde{\Lambda}$ is a subsystem construction scheme which is additive with respect to the subsystems, i.e. for all disjoint sets $J_1, J_2 \in \mathcal{P}$ and $X \in \mathcal{X}^d$*

$$\widetilde{\Lambda}(X, J_1 \cup J_2) = \widetilde{\Lambda}(X, J_1) + \widetilde{\Lambda}(X, J_2), \tag{4.A.1}$$

then there exists a core allocation $k \in C^-_{\widetilde{\rho}}(X)$.

Proof. In order to prove the lemma we will utilize Theorem $4.A.2$. Let $(\alpha_J)_{J\in\mathcal{P}}$ be a balanced collection of weights, then we obtain by additivity of $\widetilde{\Lambda}$ that

$$\begin{aligned}\widetilde{\Lambda}(X) &= \sum_{i=1}^{d} \widetilde{\Lambda}\big(X, \{i\}\big) = \sum_{i=1}^{d} \sum_{J\in\mathcal{P}_i} \alpha_J \widetilde{\Lambda}\big(X, \{i\}\big) \\ &= \sum_{J\in\mathcal{P}} \alpha_J \sum_{i\in J} \widetilde{\Lambda}\big(X, \{i\}\big) = \sum_{J\in\mathcal{P}} \alpha_J \widetilde{\Lambda}\big(X, J\big)\end{aligned}$$

and thus by subadditivity and positive homogeneity that

$$\eta(-\widetilde{\Lambda}(X)) = \eta\left(-\sum_{J\in\mathcal{P}} \alpha_J \widetilde{\Lambda}(X, J)\right) \leq \sum_{J\in\mathcal{P}} \alpha_J \eta\big(-\widetilde{\Lambda}(X, J)\big).$$

□

Remark 4.A.4. Note that in order to prove Lemma $4.A.3$ it would be sufficient to show that $\widetilde{\Lambda}(X) \leq \sum_{J\in\mathcal{P}} \alpha_J \widetilde{\Lambda}\big(X, J\big)$ for all balanced collection of weights $(\alpha_J)_{J\in\mathcal{P}}$.

Example 4.A.5. The additivity over subsystems (4.A.1) is clearly satisfied by the subsystem construction scheme $\widetilde{\Lambda}(x, J) = -\sum_{j\in J} x_j$ which we already know as a suitable aggregation for financial systems without contagion. Furthermore, the additivity still holds if we just consider the losses of the financial institutions in this model, i.e. if $\widetilde{\Lambda}(x, J) = \sum_{j\in J} x_j^-$.

Note that the additivity property (4.A.1) in Lemma 4.A.3 directly implies that the core of the subsystem construction scheme is always non-empty. In the following lemma we show that this weaker property is already sufficient for the core of the subsystem risk measure to be non-empty.

Lemma 4.A.6. *Let $\widetilde{\rho}(X,J) = \eta\big(-\widetilde{\Lambda}(X,J)\big)$ be a subsystem risk measure where η is a positive homogeneous and subadditive univariate risk measure and $\widetilde{\Lambda}$ is a subsystem construction scheme such that the functions $x \mapsto \widetilde{\Lambda}(x,J)$ are continuous for all $J \in \mathcal{P}$. Then we have that $C^-_{\widetilde{\rho}}(X) \neq \emptyset$ for all $X \in \mathcal{X}^d$ with $C^-_{\widetilde{\Lambda}}(X(\omega)) \neq \emptyset$ for all $\omega \in \Omega$.*

Proof. Let $X \in \mathcal{X}^d$ such that $C^-_{\widetilde{\Lambda}}(X(\omega)) \neq \emptyset$ for all $\omega \in \Omega$. It is well-known that the set-valued function $\mathcal{C}^-$ mapping all possible $\nu : \mathcal{P} \to \mathbb{R}$ to its core is upper hemicontinuous, see for instance Delbaen (1974). Since $x \mapsto \widetilde{\Lambda}(x,\cdot)$ are continuous, we get that the set-valued composition $C^-_{\widetilde{\Lambda}}(x) = \mathcal{C}^- \circ \widetilde{\Lambda}(x,\cdot)$ is also upper-hemicontinuous, i.e. for all open $A \subset \mathbb{R}^d$, we have that $\{x \in \mathbb{R}^d : C^-_{\widetilde{\Lambda}}(x) \subset A\}$ is open. Moreover this implies that $C^-_{\widetilde{\Lambda}}$ is measurable and thus according to Theorem 8.1.3 in Aubin and Frankowska (2009) there exists a Borel measurable selection of $C^-_{\widetilde{\Lambda}}$. Therefore there also exists a measurable selection $K \in \mathcal{X}^d$ of $C^-_{\widetilde{\Lambda}}(X)$, i.e. $K(\omega) \in C^-_{\widetilde{\Lambda}}(X(\omega))$ for each $\omega \in \Omega$. Now, define the subsystem risk measure

$$\bar{\rho} : \mathcal{X}^d \times \mathcal{P} \to \mathbb{R}; (K,J) \mapsto \eta\left(\sum_{j\in J} K_j\right).$$

By applying Lemma 4.A.3 we obtain that $C^-_{\bar{\rho}}(K) \neq \emptyset$. The monotonicity of η yields

$$\widetilde{\rho}(X) = \eta\big(-\widetilde{\Lambda}(X)\big) = \eta\left(\sum_{j=1}^{d} K_j\right) = \bar{\rho}(K)$$

as well as for all $J \in \mathcal{P}$

$$\widetilde{\rho}(X,J) = \eta\big(-\widetilde{\Lambda}(X,J)\big) \geq \eta\left(\sum_{j\in J} K_j\right) = \bar{\rho}(K,J)$$

and it immediately follows from Lemma 4.2.5 that also $C^+_{\widetilde{\rho}}(X) \neq \emptyset$ □

In particular Lemma 4.A.6 implies that for every coherent risk measure $\eta : \mathcal{X} \to \mathbb{R}$ the core of the subsystem risk measure $\eta \circ \Lambda^0$ is always non-empty. Note that this is possible since both the coherent risk measure η and the subsystem construction scheme Λ^0 share the same perspective towards diversification, that is joining to subgroups always results in a risk reduction. Unfortunately, for subsystem construction schemes for which the reverse core is non-empty like Λ^c this is no longer true. Thus it is more problematic to obtain a similar result as in Lemma 4.A.6, i.e. that the reverse core of a random risk is non-empty if the reverse cores for the corresponding scenario-wise deterministic risks are non-empty. For instance this would hold if we ask that the univariate risk measure η is superadditive instead of subadditive, i.e.

$$\eta(F+G) \geq \eta(F) + \eta(G) \quad \text{for all } F, G \in \mathcal{X}.$$

However, the requirement of superadditivity is less clear on the level of the univariate risk measure compared to the level of aggregation. Clearly, a compromise in this context would be a linear risk measure.

Moreover, if we have a scenario-wise non-emptiness of the reverse core, but we insist upon a subadditive univariate risk measure, then a possible workaround is to consider the transition to the equivalent core. That is, if we suppose that there exists a $k(\omega) \in C^+_\Lambda(X(\omega))$ for all $\omega \in \Omega$, then it follows by Lemma 4.4.3 that $k(\omega) \in C^-_{\bar{\Lambda}}(X(\omega))$ for all $\omega \in \Omega$, where

$$\bar{\Lambda}(X(\omega), J) := \Lambda(X(\omega)) - \Lambda\left(X(\omega), J^C\right).$$

Now, define

$$\bar{\rho}(X, J) := \eta\left(-\bar{\Lambda}(X, J)\right),$$

where η is positive homogeneous and subadditive univariate risk measure. By Lemma $4.A.6$ we obtain that $C^-_{\bar{\rho}}(X) \neq \emptyset$. However, we remark that this is in general not equivalent to the reverse core of $\rho(X, J) := \eta(-\Lambda(X, J))$. To be more precise we only have that $C^+_{\hat{\rho}}(X) \neq \emptyset$ with

$$\hat{\rho}(X, J) := \bar{\rho}(X, \mathcal{I}) - \bar{\rho}(X, J^C) = \rho(X) - \eta\big(\Lambda(X, J) - \Lambda(X)\big).$$

In the special case of a linear univariate risk measure η, we also obtain that $\hat{\rho}(X, J) = \eta(-\Lambda(X, J)) = \rho(X, J)$.

List of Symbols

$\mathbb{N}$	positive integers
$\mathbb{Q}$	rational numbers
$\mathbb{R}, \mathbb{R}^+$	real numbers, positive real numbers
$\emptyset$	empty set
$\mathcal{I}$	financial system $\mathcal{I} = \{1, ..., d\}$
$\mathcal{P}$	powerset of $\mathcal{I}$
d	dimension of the systems considered
$\uparrow, \downarrow$	convergence from below, convergence from above
$\liminf, \limsup$	limes inferior and limes superior
$\wedge, \vee$	minimum and maximum
$\inf, \sup$	infimum and supremum
$\operatorname{essinf}, \operatorname{esssup}$	essential infimum and essential supremum
$*$	Hadamard product (componentwise multiplication)
$\circ$	composition of two functions
$(\cdot)^+, (\cdot)^-$	positive and negative part; $x^+ = \max\{x, 0\}$, $x^- = -\min\{x, 0\}$
$(\cdot)^\top$	transpose
id	identity function
$\left\Vert\cdot\right\Vert_\infty, \left\Vert\cdot\right\Vert_{\infty,d}$	supremum norm, $\left\Vert X\right\Vert_{\infty,d} = \max_{i=1,\dots,d} \left\Vert X_i\right\Vert_\infty$
$\mathbb{E}_\mathbb{P}\left[\cdot \mid \mathcal{G}\right]$	$\mathcal{G}$-conditional expectation
$\mu_X(\cdot \vert \mathcal{G})$	$\mathcal{G}$-conditional distribution of X
$\prec, \preceq$	preference order
δ_x	Dirac measure at $x \in \mathbb{R}^d$
$\mathbb{1}_A$	indicator function; $\mathbb{1}_A(\omega) = 1$ if $\omega \in A$ and $\mathbb{1}_A(\omega) = 0$ if $\omega \notin A$
$\mathcal{F}, \mathcal{G}, \mathcal{H}, \mathcal{T}$	σ-algebras
$\sigma(A)$	smallest σ-algebra generated by A
$\mathcal{I}_1, \mathcal{I}_2$	sets of σ-algebras
$\mathcal{E}$	$\mathcal{E} = \{(\mathcal{G}, \mathcal{T}) \in \mathcal{I}_1 \times \mathcal{I}_2 : \mathcal{G} \subseteq \mathcal{T}\}$
$\mathcal{E}(\mathcal{T})$	$\mathcal{E}(T) = \{\mathcal{G} \in \mathcal{I}_1 : \mathcal{G} \subseteq \mathcal{T}\}$
$\mathcal{B}(\mathbb{R}), \mathcal{B}(\mathbb{R}^d)$	Borel-σ-algebra on $\mathbb{R}, \mathbb{R}^d$
$(\Omega, \mathcal{F}, \mathbb{P})$	probability space
$(\Omega, \widehat{\mathcal{F}}, \widehat{\mathbb{P}})$	complete probability space
$(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$	filtered probability space
N	$\mathbb{P}$-nullset
$\mathscr{L}^\infty(\mathcal{F})$	space of $\mathcal{F}$-measurable, $\mathbb{P}$-a.s. bounded random variables
$\mathscr{L}^\infty_d(\mathcal{F})$	d-fold Cartesian product of $\mathscr{L}^\infty(\mathcal{F})$

$L_d^\infty(\mathcal{F})$ .... quotient space $\mathscr{L}_d^\infty(\mathcal{F})/\{X : X = \mathbf{0}_d \ \mathbb{P}\text{-a.s.}\}$
$\mathcal{S}$ .... set of $\mathcal{F}$-measurable simple random vectors, i.e. $\{X : X = \sum_{i=1}^k x_i \mathbb{1}_{A_i}, k \in \mathbb{N}, x_i \in \mathbb{R}^d$ and $A_i \in \mathcal{F} \ \forall i\}$
$\mathcal{X}$ .... domain of a conditional (base) risk measure
$\operatorname{Im}\Lambda_{\mathcal{G}}$ .... Image of the conditional aggregation function $\Lambda_{\mathcal{G}}$
ρ .... multivariate risk measure; $\rho : L_d^\infty(\mathcal{F}) \to \mathbb{R}$
subsystem risk measure; $\rho : L_d^\infty(\mathcal{F}) \times \mathbb{R}_+^d \to \mathbb{R}$
$\widetilde{\rho}$ .... subsystem risk measure; $\rho : L_d^\infty(\mathcal{F}) \times \mathcal{P} \to \mathbb{R}$
$\rho_{\mathcal{G}}$ .... risk-consistent conditional systemic risk measure; $\rho_{\mathcal{G}} : L_d^\infty(\mathcal{F}) \to L^\infty(\mathcal{G})$
multivariate conditional risk measure;$\rho_{\mathcal{G}} : L_d^\infty(\mathcal{F}) \to L^\infty(\mathcal{G})$
$\widetilde{\rho}_{\mathcal{G}}$ .... realization of $\rho_{\mathcal{G}}$
$\bar{\rho}_{\mathcal{G}}$ .... normalized multivariate conditional risk measure, i.e. $f_{\bar{\rho}_{\mathcal{G}}} = -\operatorname{id}$
$\rho_{\mathcal{G},\mathcal{T}}$ .... multivariate conditional risk measure;$\rho_{\mathcal{G},\mathcal{T}} : L_d^\infty(\mathcal{T}) \to L^\infty(\mathcal{G})$
η .... univariate risk measure
$\eta_{\mathcal{G}}$ .... conditional (base) risk measure; $\eta_{\mathcal{G}} : \mathcal{X} \to L^\infty(\mathcal{G})$
Λ .... conditional aggregation function; $\Lambda : L_d^\infty(\mathcal{F}) \to L^\infty(\mathcal{F})$
subsystem construction scheme $\Lambda : \mathbb{R}^d \times \mathbb{R}_+^d \to \mathbb{R}$
Λ^0 .... $\Lambda^0(x,\lambda) = \Lambda(\lambda * x)$
Λ^b .... $\Lambda^b(x,\lambda) = \Lambda(\lambda * x + (\mathbf{1}_d - \lambda) * b)$
$\Lambda_{\mathcal{G}}$ .... (extended) conditional aggregation function; $\Lambda_{\mathcal{G}} : L_d^\infty(\mathcal{F}) \to L^\infty(\mathcal{F})$
$\widetilde{\Lambda}$ .... deterministic aggregation function; $\widetilde{\Lambda} : \mathbb{R}^d \to \mathbb{R}$
subsystem construction scheme $\widetilde{\Lambda} : \mathbb{R}^d \times \mathcal{P} \to \mathbb{R}$
$\widetilde{\Lambda}_{\mathcal{G}}$ .... conditional aggregation function; $\widetilde{\Lambda}_{\mathcal{G}} : \mathbb{R}^d \times \Omega \to \mathbb{R}$
$f_{\rho_{\mathcal{G}}}$ .... $f_{\rho_{\mathcal{G}}} : L^\infty(\mathcal{G}) \to L^\infty(\mathcal{G}); \alpha \mapsto \rho_{\mathcal{G}}(\alpha \mathbf{1}_d)$
$\bar{f}_{\rho_{\mathcal{G}}}$ .... extension of $f_{\rho_{\mathcal{G}}}$; $\bar{f}_{\rho_{\mathcal{G}}} : L^\infty(\mathcal{F}) \to L^\infty(\mathcal{F})$
$f_{\Lambda_{\mathcal{G}}}$ .... $f_{\Lambda_{\mathcal{G}}} : L^\infty(\mathcal{F}) \to L^\infty(\mathcal{F}); F \mapsto \Lambda_{\mathcal{G}}(F\mathbf{1}_d)$
$\mathbf{0}_d$ .... real vector whose each entries are equal to 0
$\mathbf{1}_d$ .... real vector whose entries are equal to 1
e_i .... i-th unit vector
$\mathbf{L}, \mathbf{\Pi}$ .... liability matrix, relative liability matrix
L .... total liabilities
$\mathbf{I}_d$ .... d-dimensional identity matrix
$\mathbf{\Pi}_{\mathcal{D},\mathcal{D}}$ .... restriction of the matrix $\mathbf{\Pi}$; $\mathbf{\Pi}_{\mathcal{D},\mathcal{D}} = (\Pi_{i,j})_{i,j\in\mathcal{D}}$
$C_{\widetilde{\rho}}^-, C_{\widetilde{\rho}}^+$ .... core and reverse core of $\widetilde{\rho}$
$FC_{\widetilde{\rho}}^-, FC_{\widetilde{\rho}}^+$ .... fuzzy core and fuzzy reverse core of $\widetilde{\rho}$
$\partial^-\rho(X,\cdot)$ .... subdifferential of $\rho(X,\cdot)$
$\nabla\rho(X,\cdot)$ .... gradient of $\rho(X,\cdot)$
$\mathrm{VaR}_q, \mathrm{VaR}_q(\cdot|\mathcal{G})$ .... Value at Risk at level $q \in (0,1)$, conditional Value at Risk
$\mathrm{AVaR}_q, \mathrm{AVaR}_q(\cdot|\mathcal{G})$ .... Average Value at Risk at level $q \in (0,1)$, conditional Average Value at Risk

Bibliography

Acciaio, B. and V. Goldammer (2013). Optimal portfolio selection via conditional convex risk measures on lp. *Decisions in Economics and Finance 36*(1), 1–21.

Acciaio, B. and I. Penner (2011). Dynamic risk measures. In J. Di Nunno and B. Øksendal (Eds.), *Advanced Mathematical Methods for Finance*, Chapter 1., pp. 11–44. Springer.

Acharya, V., R. Engle, and M. Richardson (2012). Capital shortfall: A new approach to ranking and regulating systemic risks. *The American Economic Review 102*(3), 59–64.

Acharya, V. V., L. H. Pedersen, T. Philippon, and M. Richardson (2017). Measuring systemic risk. *Review of Financial Studies 30*(1), 2–47.

Adrian, T. and M. K. Brunnermeier (2016). CoVaR. *The American Economic Review 106*(7), 1705–1741.

Amini, H., R. Cont, and A. Minca (2016). Resilience to contagion in financial networks. *Mathematical Finance 26*(2), 329–365.

Amini, H., D. Filipović, and A. Minca (2015). Systemic risk and central clearing counterparty design.

Amini, H., D. Filipović, and A. Minca (2016). Uniqueness of equilibrium in a payment system with liquidation costs. *Operations Research Letters 44*(1), 1–5.

Amini, H. and A. Minca (2013). Mathematical modeling of systemic risk. In *Advances in Network Analysis and its Applications*, pp. 3–26. Springer.

Armenti, Y., S. Crepey, S. Drapeau, and A. Papapantoleon (2015). Multivariate shortfall risk allocation. *arXiv preprint arXiv:1507.05351*.

Artzner, P., F. Delbaen, J.-M. Eber, and D. Heath (1999). Coherent measures of risk. *Mathematical finance 9*(3), 203–228.

Aubin, J. (1979). *Mathematical Methods of Game and Economic Theory*. North-Holland, Amsterdam.

Aubin, J.-P. and H. Frankowska (2009). *Set-valued analysis*. Springer Science & Business Media.

Aumann, R. J. and L. S. Shapley (1974). *Values of non-atomic games*. Princeton University Press.

Awiszus, K. and S. Weber (2015). The joint impact of bankruptcy costs, fire sales and cross-holdings on systemic risk in financial networks.

Battiston, S., D. Delli Gatti, M. Gallegati, B. Greenwald, and J. E. Stiglitz (2012). Liaisons dangereuses: Increasing connectivity, risk sharing, and systemic risk. *Journal of Economic Dynamics and Control 36*(8), 1121–1141.

Battiston, S., M. Puliga, R. Kaushik, P. Tasca, and G. Caldarelli (2012). DebtRank: Too central to fail? financial networks, the fed and systemic risk. *Scientific reports 2*, 541.

Bernard, C., E. C. Brechmann, and C. Czado (2013). Statistical assessments of systemic risk measures. In J.-P. Fouque and J. A. Langsam (Eds.), *Handbook on Systemic Risk*, Chapter 6, pp. 165. Cambridge University Press.

Biagini, F., J.-P. Fouque, M. Frittelli, and T. Meyer-Brandis (2015). A unified approach to systemic risk measures via acceptance sets. *arXiv preprint arXiv:1503.06354*.

Bion-Nadal, J. (2004). Conditional risk measure and robust representation of convex conditional risk measures. *CMAP Preprint 557*.

Bisias, D., M. Flood, A. W. Lo, and S. Valavanis (2012). A survey of systemic risk analytics. *Annu. Rev. Financ. Econ. 4*(1), 255–296.

Brownlees, C. and R. F. Engle (2016). SRISK: A conditional capital shortfall measure of systemic risk. *Review of Financial Studies*, 59–64.

Brunnermeier, M. K. and P. Cheridito (2014). Measuring and allocating systemic risk. Available at SSRN 2372472.

Buch, A. and G. Dorfleitner (2008). Coherent risk measures, coherent capital allocations and the gradient allocation principle. *Insurance: Mathematics and Economics 42*(1), 235–242.

Carmona, R., J.-P. Fouque, S. M. Mousavi, and L.-H. Sun (2016). Systemic risk and stochastic games with delay. *arXiv preprint arXiv:1607.06373*.

Carmona, R., J.-P. Fouque, and L.-H. Sun (2015). Mean field games and systemic risk. *Communications in Mathematical Sciences 13*(4), 911–933.

Cerreia-Vioglio, S., F. Maccheroni, M. Marinacci, and L. Montrucchio (2011). Risk measures: rationality and diversification. *Mathematical Finance 21*(4), 743–774.

Chen, C., G. Iyengar, and C. C. Moallemi (2013). An axiomatic approach to systemic risk. *Management Science 59*(6), 1373–1388.

Cheridito, P., F. Delbaen, and M. Kupper (2006). Dynamic monetary risk measures for bounded discrete-time processes. *Electronic Journal of Probability 11*(3), 57–106.

Cheridito, P. and M. Kupper (2011). Composition of time-consistent dynamic monetary risk measures in discrete time. *International Journal of Theoretical and Applied Finance 14*(01), 137–162.

Cifuentes, R., G. Ferrucci, and H. S. Shin (2005). Liquidity risk and contagion. *Journal of the European Economic Association 3*(2/3), pp. 556–566.

Cont, R., A. Moussa, and E. B. Santos (2013). Network structure and systemic risk in banking systems. In J.-P. Fouque and J. A. Langsam (Eds.), *Handbook on Systemic Risk*, Chapter 13., pp. 327–368. Cambridge University Press.

Delbaen, F. (1974). Convex games and extreme points. *Journal of Mathematical Analysis and Applications 45*(1), 210–233.

Delbaen, F. (2000). *Coherent Risk Measures*. Lecture Notes of the Cattedra Galileiana at the Scuola Normale Superiore di Pisa.

Denault, M. (2001). Coherent allocation of risk capital. *Journal of risk 4*, 1–34.

Detering, N., T. Meyer-Brandis, K. Panagiotou, and D. Ritter (2016). Managing systemic risk in financial networks. *arXiv preprint arXiv:1610.09542*.

Detlefsen, K. and G. Scandolo (2005). Conditional and dynamic convex risk measures. *Finance and Stochastics 9*(4), 539–561.

Drapeau, S. and M. Kupper (2013). Risk preferences and their robust representation. *Mathematics of Operations Research 38*(1), 28–62.

Drehmann, M. and N. Tarashev (2013). Measuring the systemic importance of interconnected banks. *Journal of Financial Intermediation 22*(4), 586–607.

Eisenberg, L. and T. H. Noe (2001). Systemic risk in financial systems. *Management Science 47*(2), 236–249.

Elliott, M., B. Golub, and M. O. Jackson (2014). Financial networks and contagion. *The American economic review 104*(10), 3115–3153.

Elsinger, H. (2011). Financial networks, cross holdings, and limited liability. Technical report, Österreichische Nationalbank Working Paper 156.

Engle, R., E. Jondeau, and M. Rockinger (2015). Systemic risk in Europe. *Review of Finance 19*(1), 145–190.

Feinstein, Z. and F. El-Masri (2015). The effects of leverage requirements and fire sales on financial contagion via asset liquidation strategies in financial networks. *arXiv preprint arXiv:1507.01847*.

Feinstein, Z., B. Rudloff, and S. Weber (2015). Measures of systemic risk. *arXiv preprint arXiv:1502.07961*.

Filipović, D., M. Kupper, and N. Vogelpoth (2012). Approaches to conditional risk. *SIAM Journal on Financial Mathematics 3*(1), 402–432.

Föllmer, H. (2014). Spatial risk measures and their local specification: The locally law-invariant case. *Statistics & Risk Modeling 31*(1), 79–103.

Föllmer, H. and C. Klüppelberg (2014). Spatial risk measures: Local specification and boundary risk. In *Crisan, D., Hambly, B. and Zariphopoulou, T.: Stochastic Analysis and Applications 2014 - In Honour of Terry Lyons*. Springer.

Föllmer, H. and A. Schied (2002). Convex measures of risk and trading constraints. *Finance and stochastics 6*(4), 429–447.

Föllmer, H. and A. Schied (2011). *Stochastic Finance: An introduction in discrete time* (3rd ed.). De Gruyter.

Fouque, J.-P. and T. Ichiba (2013). Stability in a model of interbank lending. *SIAM Journal on Financial Mathematics 4*(1), 784–803.

Fouque, J.-P. and L.-H. Sun (2013). Systemic risk illustrated. In J.-P. Fouque and J. A. Langsam (Eds.), *Handbook on Systemic Risk*, Chapter 17, pp. 444–452. Cambridge University Press.

Frittelli, M. and M. Maggis (2011). Conditional certainty equivalent. *International Journal of Theoretical and Applied Finance 14*(01), 41–59.

Frittelli, M. and E. Rosazza Gianin (2002). Putting order in risk measures. *Journal of Banking & Finance 26*(7), 1473–1486.

Furfine, C. H. (2003). Interbank exposures: Quantifying the risk of contagion. *Journal of Money, Credit & Banking (Ohio State University Press) 35*(1).

Gai, P. and S. Kapadia (2010). Contagion in financial networks. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Science 466*(2120), 2401–2423.

Girardi, G. and A. T. Ergün (2013). Systemic risk measurement: Multivariate GARCH estimation of CoVaR. *Journal of Banking & Finance 37*(8), 3169–3180.

Glasserman, P. and H. P. Young (2015). How likely is contagion in financial networks? *Journal of Banking & Finance 50*, 383–399.

Hoffmann, H., T. Meyer-Brandis, and G. Svindland (2016). Risk-consistent conditional systemic risk measures. *Stochastic Processes and their Applications 126*(7), 2014–2037.

Huang, X., H. Zhou, and H. Zhu (2012). Systemic risk contributions. *Journal of financial services research 42*(1-2), 55–83.

Kalkbrener, M. (2005). An axiomatic approach to capital allocation. *Mathematical Finance 15*(3), 425–437.

Kley, O., C. Klüppelberg, and L. Reichel (2014). Systemic risk through contagion in a core-periphery structured banking network. *arXiv preprint arXiv:1406.6575*.

Kromer, E., L. Overbeck, and K. Zilch (2016). Systemic risk measures on general measurable spaces. *Mathematical Methods of Operations Research*, 1–35.

Kromer, E., L. Overbeck, and K. A. Zilch (2014). Dynamic systemic risk measures for bounded discrete-time processes. Available at SSRN 2469475.

Kupper, M. and W. Schachermayer (2009). Representation results for law invariant time consistent functions. *Mathematics and Financial Economics 2*(3), 189–210.

Penner, I. (2007). *Dynamic convex risk measures: time consistency, prudence, and sustainability*. Ph. D. thesis, Humboldt-Universität zu Berlin, Mathematisch-Naturwissenschaftliche Fakultät II.

Revuz, D. and M. Yor (1999). *Continuous martingales and Brownian motion*, Volume 293. Springer.

Rogers, L. and L. A. Veraart (2013). Failure and rescue in an interbank network. *Management Science 59*(4), 882–898.

Roorda, B. and H. Schumacher (2013). Membership conditions for consistent families of monetary valuations. *Statistics & risk modeling 30*(3), 255–280.

Roorda, B. and J. M. Schumacher (2016). Weakly time consistent concave valuations and their dual representations. *Finance and stochastics 20*(1), 123–151.

Shapley, L. (1953). A value for n-person games. *Contributions to the Theory of Games 2*, 307–318.

Shapley, L. S. (1967). On balanced sets and cores. *Naval research logistics quarterly 14*(4), 453–460.

Staum, J. (2012). Systemic risk components and deposit insurance premia. *Quantitative Finance 12*(4), 651–662.

Staum, J., M. Feng, and M. Liu (2016). Systemic risk components in a network model of contagion. *IIE Transactions 48*(6), 501–510.

Tarashev, N., K. Tsatsaronis, and C. Borio (2016). Risk attribution using the shapley value: methodology and policy applications. *Review of Finance 20*(3), 1189–1213.

Tasche, D. (2004). Allocating portfolio economic capital to sub-portfolios. In *Economic Capital: A Practitioner Guide, Risk Books*, pp. 275–302.

Tasche, D. (2007). Capital allocation to business units and sub-portfolios: The Euler principle. *arXiv preprint arXiv:0708.2542*.

Tutsch, S. (2007). *Konsistente und konsequente dynamische Risikomaße und das Problem der Aktualisierung*. Ph. D. thesis, Humboldt-Universität zu Berlin, Mathematisch-Naturwissenschaftliche Fakultät II.

Tutsch, S. (2008). Update rules for convex risk measures. *Quantitative Finance 8*(8), 833–843.

Upper, C. and A. Worms (2004). Estimating bilateral exposures in the German interbank market: Is there a danger of contagion? *European Economic Review 48*(4), 827–849.

Weber, S. (2006). Distribution-invariant risk measures, information, and dynamic consistency. *Mathematical Finance 16*(2), 419–441.

Werner, D. (2011). *Funktionalanalysis*. Springer.