Multivariate Shortfall Risk Allocation and Systemic Risk

Yannick Armenti [a,1,†], Stéphane Crépey [a,2,*,§], Samuel Drapeau [b,3,¶], Antonis Papapantoleon [c,4,‡,§]

July 26, 2016

ABSTRACT

The ongoing concern about systemic risk since the outburst of the global financial crisis has highlighted the need for risk measures at the level of sets of interconnected financial components, such as portfolios, institutions or members of clearing houses. The two main issues in systemic risk measurement are the computation of an overall reserve level and its allocation to the different components according to their systemic relevance. We develop here a pragmatic approach to systemic risk measurement and allocation based on multivariate shortfall risk measures, where acceptable allocations are first computed and then aggregated so as to minimize costs. We analyze the sensitivity of the risk allocations to various factors and highlight its relevance as an indicator of systemic risk. In particular, we study the interplay between the loss function and the dependence structure of the components, that provides valuable insights into the properties of good loss functions. Moreover, we provide numerical schemes to assess the risk allocation in high dimensions.

KEYWORDS: Systemic risk, risk allocation, multivariate shortfall risk, sensitivities, numerical methods.
AMSCLASSIFICATION: 91G, 91B30, 91G60

AUTHORS INFO

[a] *Université d'Evry, 23 Boulevard de France, 91037 Evry, France*
[b] *School of Mathematical Sciences & Shanghai Advanced Institute for Finance (CAFR/CMAR), Shanghai Jiao Tong University, 211 West Huaihai Road, Shanghai, P.R. 200030 China*
[c] *Institute of Mathematics, Technical University Berlin, Straße des 17. Juni 136, 10623 Berlin, Germany*
[1] yannick.armenti@gmail.com
[2] stephane.crepey@univ-evry.fr
[3] sdrapeau@saif.sjtu.edu.cn
[4] papapan@math.tu-berlin.de

[*] *Financial support from the EIF grant "Collateral management in centrally cleared trading", from the Chair "Markets in Transition", Fédération Bancaire Française, and from the ANR 11-LABX-0019.*
[†] *Financial support from LCH.Clearnet Paris.*
[‡] *Financial support from the EIF grant "Post-crisis models for interest rate markets".*
[§] *Financial support from the DAAD PROCOPE project "Financial markets in transition: mathematical models and challenges".*
[¶] *Financial support from the National Science Foundation of China, "Research Fund for International Young Scientists", Grant number 11550110184.*

PAPER INFO

ArXiv ePrint: 1507.05351

1. Introduction

The ongoing concern about systemic risk since the onset of the global financial crisis has prompted intensive research on the design and properties of multivariate risk measures. In this paper, we study the risk assessment for financial systems with interconnected risky components, focusing on two major aspects, namely:

- The quantification of a monetary risk measure corresponding to an overall reserve of liquidity such that the whole system can overcome unexpected stress or default scenarios;
- The allocation of this overall amount between the different risk components in a way that reflects the systemic risk of each one.

Our goal is fourfold. First, designing a theoretically sound but numerically tractable class of systemic risk measures. Second, assessing the impact of the dependence structure of the system on the risk allocation. Third, studying the sensitivity of this allocation with respect to exogenous shocks. Finally, developing and testing efficient numerical schemes for the computation of the risk allocation.

Review of the Literature: Monetary risk measures have been the subject of intensive research since the seminal paper of Artzner et al. [4], which was further extended by Föllmer and Schied [19] and Frittelli and Rosazza Gianin [20], among others. The corresponding risk measures, including conditional value-at-risk by Artzner et al. [4], shortfall risk measures by Föllmer and Schied [19] or optimized certainty equivalents by Ben-Tal and Teboulle [5], can be applied in a multivariate framework that models the dependence of several financial risk components. Multivariate market data-based risk measures include the marginal expected shortfall of Acharya et al. [1], the systemic risk measure of Acharya et al. [2] and Brownlees and Engle [7], the delta conditional value-at-risk of Adrian and Brunnermeier [3] or the contagion index of Cont et al. [12]. In parallel, theoretical economical and mathematical considerations have led to multivalued and set-valued risk measures, in static or even dynamic setup; see for instance Cascos and Molchanov [9], Hamel et al. [23] and Jouini et al. [24].

More recently, the risk management of financial institutions raised concerns about the allocation of the overall risk among the different components of a financial system. A bank, for instance, for real time monitoring purposes, wants to channel to each trading desk a cost reflecting its responsibility in the overall capital requirement of the bank. A central clearing counterparty — CCP for short, also known as a clearing house — is interested in quantifying the size of the so-called default fund and allocating it in a meaningful way among the different clearing members. On a macroeconomic level, regulators are considering to require from financial institutions an amount of capital reflecting their systemic relevance. The aforementioned approaches can only address the allocation problem indirectly, through the sensitivity of the risk measure with respect to the different risk components. For instance, the so-called *Euler rule* allocates the total amount of risk according to the marginal impact of each risk factor. However, a practical limitation of the Euler rule is that it is based on Gâteaux derivatives. In addition, the Euler risk allocation does not add up to the total risk, unless the univariate risk measure that is used in the first place is sub-additive. In other words, the Euler rule does not automatically fulfill the so-called *full allocation* property.

In a recent work, Brunnemeier and Cheridito [8] address systematically the question of allocation of systemic risk with regard to certain economic properties:

- Full allocation: the sum of the components of the risk allocation is equal to the overall risk measure;
- Riskless allocation: if a risk factor is riskless, the corresponding component of the risk allocation is equal to it;
- Causal responsibility: any system component bears the entire additional costs of any additional risk that it takes.

More specifically, Brunnemeier and Cheridito [8] propose a framework where an overall capital requirement is first determined by utility indifference principles and then allocated according to a rule such that the above three properties are fulfilled, at least at a first order level of approximation. In fact, as far as dependence is concerned, whether the last two properties should hold is debatable. One may argue that each component in the system is not only responsible for its own risk taking but also for its relative exposure to other components. This is also what comes out from the present study, see Section 4.3.

Contribution and Outline of the Paper: Our approach addresses simultaneously the design of an overall risk measure regarding a financial system of interconnected components and the allocation of this risk measure among the different risk components; the emphasis lies on risk allocation. In contrast to [8, 10], we *first allocate* the monetary risk among the different risk components and *then aggregate* and minimizes the risk allocations in order to obtain the overall capital requirement. In two recent papers,

Feinstein et al. [17] and Biagini et al. [6] develop approaches in a similar spirit, covering allocation first followed by aggregation, in general frameworks with different aggregation procedures. They focus on the resulting risk measure, conducting systematic studies of their properties in terms of set valued functions, diversification and monotonicity, among others. Sharing with these references the "allocate first, then aggregate" perspective, our approach is restricted to a systemic extension of shortfall risk measures, see [19], based on multivariate loss functions. However, in contrast to the aforementioned references, we focus on the resulting risk allocation in terms of existence, uniqueness, sensitivities and numerical computation. In our view, the *systemic risk* is the risk that stems specifically from the intrinsic dependence structure of an interconnected system of risk components. In this perspective, the risk allocation and its properties provide a "cartography" of the systemic risk, see Section 6.2 for a numerical illustration thereof. It turns out that special care has to be given to the specifications of the loss function in order to stress the systemic risk. In [6], by allowing random allocations, the impact of the interdependence structure can be observed in the future. Such random allocations may be interesting in view of a posterior management of default allocations. By contrast, our deterministic allocation is sensitive to the dependence of the system already at the moment of the quantification, as suited for the vast majority of applications. In this sense, to the best of our knowledge, this paper is the first in the literature to make use of loss functions that emphasize the dependence structure of the system, see Section 4 and see *a contrario* Proposition 3.10. In addition, we present numerical schemes for the computation of multivariate shortfall risk measures and their risk allocations, with either Monte Carlo simulations or Fourier transform methods along the lines of the univariate case of Drapeau et al. [14], possibly combined with a Chebyshev polynomial interpolation, a method recently applied to option pricing in Gaß et al. [21].

The paper is organized as follows: Section 2 introduces the class of systemic loss functions, acceptance sets and risk measures that we use in this work. Section 3 establishes the existence and uniqueness of a risk allocation. Section 4 focuses on sensitivities. In Section 5, the algorithmic aspects of the problem are discussed and in Section 6, numerical results are presented. Appendices A and B gather classical facts from convex optimization and results on multivariate Orlicz spaces. Appendix C provides more details on the Fourier and Chebyshev schemes.

1.1. Basic Notation

Let x_k denote the generic coordinate of a vector $x \in \mathbb{R}^d$. By $\geqslant$ we denote the lattice order on $\mathbb{R}^d$, that is, $x \geqslant y$ if and only if $x_k \geq y_k$ for every $1 \leq k \leq d$. We denote by $\|\cdot\|$ the Euclidean norm and by $^{\pm}, \wedge, \vee, |\cdot|$ the lattice operations on $\mathbb{R}^d$. For $x, y \in \mathbb{R}^d$, we write $x > y$ for $x_k > y_k$ componentwise, $x \cdot y = \sum x_k y_k$, $xy = (x_1 y_1, \ldots, x_d y_d)$ and $x/y = (x_1/y_1, \ldots, x_d/y_d)$.

Let $(\Omega, \mathcal{F}, P)$ denote an atomless probability space, where P represents the objective probability measure, with related expectation denoted by E. We denote by $L^0(\mathbb{R}^d)$ the space of $\mathcal{F}$-measurable d-variate random variables on this space. The space $L^0(\mathbb{R}^d)$ inherits the lattice structure of $\mathbb{R}^d$, hence we can use the above notation in a P-almost sure sense. For instance, for $X, Y \in L^0(\mathbb{R}^d)$, we say that $X \geqslant Y$ or $X > Y$ if $P[X \geqslant Y] = 1$ or $P[X > Y] = 1$, respectively. Since we mainly deal with multivariate functions or random variables, to simplify notation we drop the reference to $\mathbb{R}^d$ in $L^0(\mathbb{R}^d)$, writing simply L^0 unless a particular dimension is meant, mainly for $L^0(\mathbb{R})$ in the case of univariate random variables.

2. Systemic Risk Measures

Let $X = (X_1, \dots, X_d) \in L^0$ be a random vector of financial losses, that is, negative values of X_k represent actual profits. We want to determine an overall monetary measure $R(X)$ of the risk of X as well as a sound risk allocation $RA_k(X), k = 1, \dots, d$, of $R(X)$ among the d risk components. We consider a flexible class of risk measures defined by means of loss functions and sets of acceptable monetary allocations. This class allows us to discuss in detail the properties of the resulting risk allocation as an indicator of systemic risk. Inspired by the shortfall risk measure introduced in [19] in the univariate case, we start with a loss function ℓ defined on $\mathbb{R}^d$, used to measure the expected loss $E[\ell(X)]$ of the financial loss vector X.

Definition 2.1. A function $\ell : \mathbb{R}^d \to (-\infty, \infty]$ is called a *loss function* if

(A1) ℓ is increasing, that is, $\ell(x) \geq \ell(y)$ if $x \geqslant y$;

(A2) ℓ is convex, lower semi-continuous and finite on some neighborhood of 0;

(A3) $\ell(0) = 0$ and $\ell(x) \geq \sum x_k$ on $\mathbb{R}^d$.

A risk neutral assessment of the losses corresponds to $E[\sum X_k] = \sum E[X_k]$. Thus, **(A3)** expresses a form of risk aversion, whereby the loss function puts more weight on high losses than a risk neutral evaluation. As for **(A1)** and **(A2)**, they express the respective normative facts about risk that "the more losses, the riskier" and "diversification should not increase risk"; see Drapeau and Kupper [13] for related discussions.

Remark 2.2. The considered risk components are often of the same type — banks, members of a clearing house or trading desks within a trading floor. In that case, the loss function should not discriminate some components against others. In other words, the loss function should be invariant under permutation of its variables. ⧫

Example 2.3. Let $h : \mathbb{R} \to (-\infty, \infty]$ be a one-dimensional loss function, that is, a convex, increasing, lower semi-continuous function such that $h(0) = 0$ and $h(x) \geq x$ for every $x \in \mathbb{R}$. Classical examples of loss functions[1] are

$$h(x) = \beta x^+, \ \beta > 1, \quad h(x) = x + (x^+)^2/2 \quad \text{or} \quad h(x) = e^x - 1.$$

Using these as building blocks, we obtain the following classes of multivariate loss functions, which will be used for illustrative purposes in the discussion of systemic risk, see Section 3.

(C1) $\ell(x) = h(\sum x_k)$;

(C2) $\ell(x) = \sum h(x_k)$;

(C3) $\ell(x) = \alpha h(\sum x_k) + (1 - \alpha) \sum h(x_k)$, where $0 \leq \alpha \leq 1$.

Moreover, in the numerical part we consider the following specific loss function:

(C4) $\ell(x) = \sum x_k + \frac{1}{2} \sum (x_k^+)^2 + \alpha \sum_{j<k} x_j^+ x_k^+$, where $0 \leq \alpha \leq 1$,[2]

[1] The second one, related to mean-variance penalization of the losses, is smoother than the first one, whilst being less explosive than the third one, hence yielding a good compromise for optimization routines.

[2] This is a loss function since $\frac{1}{2} \sum (x_k^+)^2 + \alpha \sum_{j<k} x_j^+ x_k^+ = \frac{1-\alpha}{2} \sum (x_k^+)^2 + \frac{\alpha}{2} (\sum x_k^+)^2$,

which is neither of the form **(C1)** nor **(C2)**. Note that each of these loss functions are permutation invariant, see Remark 2.2 for the economical motivation. ◊

Integrability and topological reasons lead us to consider loss vectors in the following multivariate Orlicz heart:

$$M^\theta = \left\{X \in L^0 \colon E\left[\theta\left(\lambda X\right)\right] < \infty \text{ for all } \lambda \in \mathbb{R}_+\right\},$$

where $\theta(x) = \ell(|x|)$, $x \in \mathbb{R}^d$; see Appendix B.

Definition 2.4. A monetary allocation $m \in \mathbb{R}^d$ is *acceptable* for X at the loss level $c > 0$ if

$$E\left[\ell\left(X - m\right)\right] \leq c.$$

We denote by

$$A(X) := \left\{m \in \mathbb{R}^d : E\left[\ell\left(X - m\right)\right] \leq c\right\}$$

the corresponding set of *acceptable monetary allocations*.

Example 2.5. In a centrally cleared trading setup, each clearing member k is required to post a default fund contribution m_k in order to make the risk of the clearing house acceptable with respect to a risk measure accounting for extreme and systemic risk. The default fund is a pooled resource of the clearing house, in the sense that the default fund contribution of a given member can be used by the clearing house not only in case the liquidation of this member requires it, but also in case the liquidation of another member requires it. For the determination of the default fund contributions, the methodology of this paper can be applied, to the vector X defined as the vector of stressed losses-and-profits of the clearing members. According to the findings of Section 3 and 4, a "systemic" loss function such as **(C3)** or **(C4)** with $\alpha > 0$ would be consistent with the purpose of the default fund. ◊

The next proposition gathers the main properties of the sets of acceptable monetary allocations. The convexity property in *(i)* means that a diversification between two acceptable monetary allocations remains acceptable. If a monetary allocation is acceptable, then any greater amount of money should also be acceptable, which is the monotonicity property in *(i)*. As for *(ii)*, it says that, if the losses X are less than Y almost surely, then any monetary allocation that is acceptable for Y is also for X. Next, *(iii)* means that a convex combination of allocations acceptable in two markets is still acceptable in the diversified market. In particular, the acceptability concept pushes towards a greater diversification among the different risk components. Finally, *(iv)* means that acceptable positions translate with cash in the sense of scalar monetary risk measures à la [4, 19, 20]. As an immediate consequence of these properties, $X \mapsto A(X)$ defines a monetary set-valued risk measure in the sense of Hamel et al. [23], that is, a set-valued map A from M^θ into the set of monotone, closed and convex subsets of $\mathbb{R}^d$.

Proposition 2.6. *For X, Y in M^θ, it holds:*

(i) $A(X)$ *is convex, monotone and closed;*

(ii) $A(X) \supseteq A(Y)$ *whenever* $X \leqslant Y$*;*

(iii) $A(\alpha X + (1-\alpha)Y) \supseteq \alpha A(X) + (1-\alpha)A(Y)$*, for any* $\alpha \in (0,1)$*;*

(iv) $A(X + m) = A(X) + m$*, for any* $m \in \mathbb{R}^d$*;*

(v) $\emptyset \neq A(X) \neq \mathbb{R}^d$.

Proof. Since ℓ is convex, increasing and lower semi-continuous, it follows that $(m, X) \mapsto E[\ell(X - m)]$ is convex and lower semi-continuous, decreasing in m and increasing in X. This implies the properties *(i)* through *(iii)* by Definition 2.4 of $A(X)$. Regarding *(iv)*, a change of variables yields

$$A(X + m) = \{n \in \mathbb{R}^d \colon E[\ell(X + m - n) \leq c]\} = \{n + m \in \mathbb{R}^d \colon E[\ell(X - n)] \leq c\} = A(X) + m.$$

As for *(v)*, on the one hand, $\ell(X - m) \searrow \ell(-\infty) \leq \ell(0) = 0 < c$ as $m \to \infty$ component-wise. Since $X \in M^\theta$ it follows that $\ell(X) \in L^1$, thus monotone convergence yields $E[\ell(X - m)] \searrow \ell(-\infty) < c$ and in turn the existence of $m \in \mathbb{R}^d$ such that $E[\ell(X - m)] \leq c$, showing that $A(X) \neq \emptyset$. On the other hand, ℓ being increasing and such that $\ell(x) \geq \sum x_k$, it implies that $\ell(X - m) \geq \sum X_k - \sum m_k \nearrow \infty$ as $m \to -\infty$, component-wise. Hence, monotone convergence yields $E[\ell(X - m)] \nearrow \infty > c$, therefore there exists $m \in \mathbb{R}^d$ such that $E[\ell(X - m)] > c$, that is, $m \notin A(X)$. □

Figure 1 shows sets of acceptable monetary allocations for a bivariate normal distribution with varying correlation coefficient. The location and shape of these sets change with the correlation: the higher the correlation, the more costly the acceptable monetary allocations, as expected in terms of systemic risk. As discussed in Sections 3 and 4, this systemic feature is not always immediate, and depends on the specification of the loss function. Indeed, the present plot stems from a loss function that emphasises the interdependence structure of the risk components, see Section 4.2.

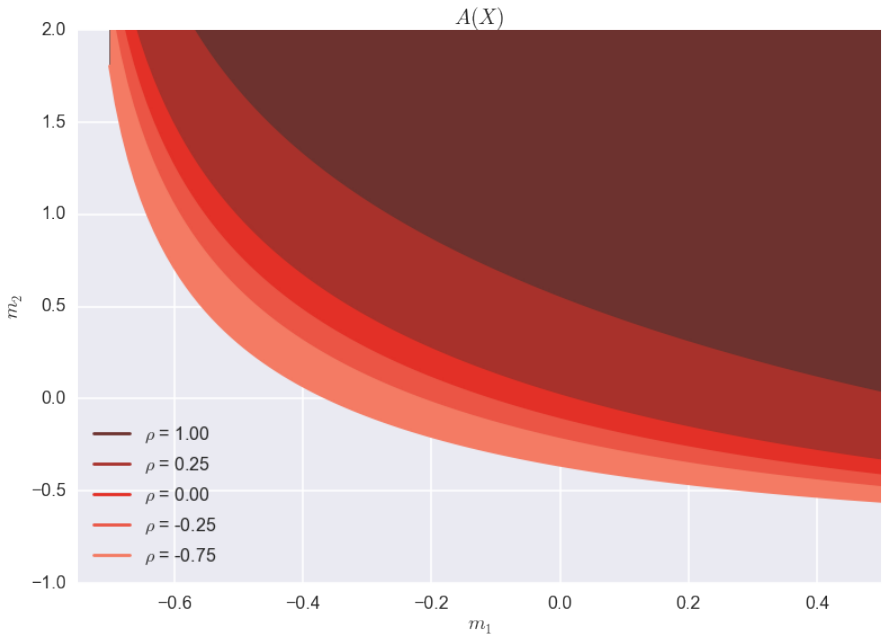

Figure 1: Acceptance sets $A(X)$ corresponding to the case study of Section 4.2 for different correlations.

Given an acceptable monetary allocation $m \in A(X)$, its aggregated liquidity cost is $\sum m_k$. The smaller the cost, the better, which motivates the following.

Definition 2.7. The *multivariate shortfall risk* of $X \in M^\theta$ is

$$R(X) := \inf\left\{\sum m_k \colon m \in A(X)\right\} = \inf\left\{\sum m_k \colon E[\ell(X - m)] \leq c\right\}.$$

Example 2.8. Following up on the central clearing house Example 2.5, any acceptable allocation $m \in A(X)$ yields a corresponding value for the default fund. Clearing houses are in competition with each other, hence they are looking for the cheapest acceptable allocation to require from their members. $\Diamond$

When $d = 1$, the above definition corresponds exactly to the shortfall risk measure in [19], of which this paper can be viewed as a multivariate extension. Our next result, which uses the concepts and notation of Appendix B, shows that all the classical properties of the shortfall risk measure, including its dual representation, can be extended to the multivariate case.

Theorem 2.9. *The function*

$$R(X) = \inf\left\{\sum m_k \colon m \in A(X)\right\}, \quad X \in M^\theta,$$

is real valued, convex, monotone and translation invariant.[3] *In particular, it is continuous and sub-differentiable. Moreover, it admits the dual representation*

$$R(X) = \max_{Q \in \mathcal{Q}^{\theta^*}} \left\{E_Q[X] - \alpha(Q)\right\}, \quad X \in M^\theta, \tag{2.1}$$

where $\mathcal{Q}^{\theta^*}$ is the set of measures Q on the product space $\Omega \times \{1, \dots, d\}$ with density Y in L^{θ^*} normalized to d, in the sense that $E[1 \cdot dQ/dP] = d$, and where the penalty function is given by

$$\alpha(Q) = \inf_{\lambda > 0} \frac{1}{\lambda}\left(c + E\left[\ell^*\left(\lambda \frac{dQ}{dP}\right)\right]\right), \quad Q \in \mathcal{Q}^{\theta^*}. \tag{2.2}$$

Remark 2.10. This robust representation can also be inferred from the general results of [16]. However, for the sake of completeness and since the multivariate shortfall risk measure is closely related to a multidimensional version of the optimized certainty equivalent, we give a self contained proof tailored to this particular context. $\blacklozenge$

Proof. By Proposition 2.6 *(v)*, we have $A(X) \neq \emptyset$ and in turn $R(X) < \infty$. If $R(X) = -\infty$ for some $X \in M^\theta$, then there exists a sequence $(m^n) \subseteq A(X)$ such that $\sum m_k^n \to -\infty$, in contradiction with $\infty > c \geq E[\ell(X - m^n)] \geq E[\sum X_k] - \sum m_k^n$. Hence, $R(X) > -\infty$. Monotonicity, convexity and translation invariance readily follow from Proposition 2.6 *(ii)*, *(iii)* and *(iv)*, respectively. In particular, R is a convex, real-valued and increasing functional on the Banach lattice M^θ. Hence, by Cheridito and Li [11, Theorem 4.1], R is continuous and sub-differentiable. Therefore, the results recalled in Appendix B and the Fenchel-Moreau theorem imply

$$R(X) = \sup_{Y \in L^{\theta^*}} \left\{E[X \cdot Y] - R^*(Y)\right\} = \max_{Y \in L^{\theta^*}} \left\{E[X \cdot Y] - R^*(Y)\right\}, \tag{2.3}$$

where

$$R^*(Y) = \sup_{X \in M^\theta} \left\{E[X \cdot Y] - R(X)\right\}, \quad Y \in L^{\theta^*}.$$

By the bipolar theorem, for $Y \not\geqslant 0$, there exists $K \in M_\theta$, $K \geqslant 0$ with $E[Y \cdot K] < -\varepsilon < 0$ for some $\varepsilon > 0$. By monotonicity of R, it follows that $R(-\lambda K) \leq R(0) < \infty$ for every $\lambda > 0$. Hence

$$R^*(Y) = \sup_{X \in M^\theta} \left\{E[Y \cdot X] - R(X)\right\} \geq \sup_{\lambda > 0} \left\{-\lambda E[Y \cdot K] - R(-\lambda K)\right\} \geq \sup_{\lambda} \lambda\varepsilon - R(0) = \infty,$$

[3] In the sense that $R(X + m) = R(X) + \sum m_k$.

showing that the supremum and maximum in (2.3) can be attained on the set of those $Y \in L_\theta$ such that $Y \geq 0$. Furthermore, by translation invariance, setting $X = (r, \dots, r)$ for $r \in \mathbb{R}$, it follows that

$$R^*(Y) \geq rE\left[1 \cdot Y\right] - R(0) - rd = r\left(E\left[1 \cdot Y\right] - d\right) - R(0),$$

where the right hand side can be made arbitrarily large whenever $E\left[1 \cdot Y\right] \neq d$. In order to obtain a more explicit expression of the penalty function $\alpha(Q) := R^*(dQ/dP) = R^*(Y)$, we set

$$\begin{aligned} L(m, \lambda, X) &= \sum m_k + \lambda E\left[\ell\left(X - m\right) - c\right] \\ S(\lambda, X) &= \inf_{m \in \mathbb{R}^d} L(m, \lambda, X) = \inf_{m \in \mathbb{R}^d} \left\{\sum m_k + \lambda E\left[\ell\left(X - m\right) - c\right]\right\}. \end{aligned}$$

The functional $X \mapsto S(\lambda, X)$ is a multivariate version of the so called optimized certainty equivalent, see [5]. Clearly,

$$R(X) = \inf_{m \in \mathbb{R}^d} \sup_{\lambda > 0} L(m, \lambda, X) \geq \sup_{\lambda > 0} \inf_{m \in \mathbb{R}^d} L(m, \lambda, X) = \sup_{\lambda > 0} S(\lambda, X).$$

Since $A(X)$ is nonempty and monotone, there exists $m \in \mathrm{Int}(A(X))$, that is, the so called Slater condition is fulfilled. As a consequence of Rockafellar [26, Theorem 28.2], there is no duality gap. Namely, $R(X) = \sup_{\lambda > 0} S(\lambda, X)$. Via the first part of the proof, an easy multivariate adaptation of [5, 14] yields

$$S(\lambda, X) = \sup_{Q \in \mathcal{Q}^{\theta^*}} \left\{E_Q\left[X\right] - E\left[\left(\ell_\lambda\right)^*\left(\frac{dQ}{dP}\right)\right]\right\},$$

where $\ell_\lambda(m) = \lambda(\ell(m) - c)$, hence $(\ell_\lambda)^*(m^*) = \lambda(c + \ell^*(m^*/\lambda))$. Combining this with $R(X) = \sup_{\lambda > 0} S(\lambda, X)$, the dual representation (2.2) follows. □

Remark 2.11. Let us emphasize the dependence on c by writing $S_c(\lambda, L) = \inf_m\{\sum m_k + \lambda E[\ell(X - m) - c]\}$. The proof of Theorem 2.9 shows that

$$R(X) = \sup_{\lambda > 0} S_c(\lambda, X) \geq S_c(1, X) = S_0(1, X) - c = \inf_m \left\{\sum m_k + E\left[\ell\left(X - m\right)\right]\right\} - c.$$

Note that $S_0(1, X)$ is a multivariate version of the so-called optimized certainty equivalent in [5], which, depending on the specification of the loss function, covers popular risk measures such as the conditional Value-at-Risk and the entropic risk measure, see [5, 14]. Hence this relation shows that R is more conservative than its optimized certainty equivalent counterpart, modulo the constant c. In the particular case corresponding to $CV@R_\beta$, that is, when $\ell(x) = (\sum x_k)^+/\beta$, $\beta \in (0, 1)$, it follows that

$$R(X) \geq CV@R_\beta\left(\sum X_k\right) - c. \qquad \blacklozenge$$

3. Risk Allocation

We have established in Theorem 2.9 that the infimum over all allocations $m \in \mathbb{R}^d$ used for defining $R(X)$ is real valued and has the desired properties of a risk measure. Beyond the question of the overall liquidity reserve, the allocation of this amount between the different risk components is key for systemic risk purposes. Hence, existence and uniqueness of a risk allocation are crucial questions.

Definition 3.1. A *risk allocation* is an acceptable monetary allocation $m \in A(X)$ such that $R(X) = \sum m_k$. When a risk allocation is uniquely determined, we denote it by $RA(X)$.

Remark 3.2. By definition, if a risk allocation exists, then the full allocation property automatically holds; see also Section 4.3. ⧫

In contrast to the univariate case, where the unique risk allocation is given by $m = R(X)$, existence and uniqueness are no longer straightforward in the multivariate case. The following example shows that existence may fail.

Example 3.3. Consider the loss function

$$\ell(x,y) = \begin{cases} x + y + (x+y)^+/(1-y), & y < 1 \\ \infty, & \text{otherwise.} \end{cases}$$

For $c = 1$, it follows that

$$A(0) = \left\{ m \in \mathbb{R}^2 \colon m_2 > -1 \text{ and } 1 \geq -m_1 - m_2 + \frac{(-m_1 - m_2)^+}{1 + m_2} \right\}.$$

Computations yield

$$R(0) = \inf_{m_2 > -1} \left\{ m_2 - \frac{m_2^2 + 3m_2 + 1}{m_2 + 2} \right\} = \inf_{m_2 > -1} -\frac{m_2 + 1}{m_2 + 2} = -1.$$

However, the infimum is not attained. ◊

Our next result introduces mathematically and economically sound conditions towards the existence and uniqueness of a risk allocation. A *zero-sum allocation* is a monetary allocation $u \in \mathbb{R}^d$ such that $\sum u_k = 0$. Note that, for any zero-sum allocation u,

$$\ell(0) = 0 = \sum u_k \leq \ell(u).$$

Definition 3.4. We call a loss function ℓ *unbiased* if, for every zero-sum allocation u, $\ell(ru) = 0$ for any $r > 0$ implies that $\ell(-ru) = 0$ for any $r > 0$.

Remark 3.5. Note that any permutation invariant loss function is unbiased, in particular those of Example 2.3. By contrast, the loss function used in Example 3.3 is biased. In fact, taking the zero-sum allocation $u = (1, -1)$ yields $\ell(r, -r) = 0$ for every $r \geq 0$. However, $\ell(-r, r) = \infty$ as soon as $r \geq 1$. ⧫

Theorem 3.6. *If ℓ is an unbiased loss function, then, for every $X \in M^\theta$, risk allocations m^* exist. They are characterized by the first order conditions*

$$1 \in \lambda^* E\left[\nabla \ell\left(X - m^*\right)\right] \quad \text{and} \quad E\left[\ell\left(X - m^*\right)\right] = c, \tag{3.1}$$

where λ^* is a Lagrange multiplier. In particular, when ℓ has no zero-sum direction of recession[4] except 0, the set of the solutions (m^*, λ^*) to the first order conditions (3.1) is bounded.

If ℓ is strictly convex outside $\mathbb{R}^d_-$ along zero-sums allocations,[5] then the risk allocation is unique.

[4] We refer the reader to Appendix A regarding the notions and properties of recession cones and functions. In particular, if ℓ has no zero-sum direction of recession except 0, then ℓ is an unbiased loss function.

[5] That is, for $x, y \in \mathbb{R}^d$ such that $x - y$ is a zero-sum allocation and $\ell(\alpha x + (1-\alpha)y) = \alpha\ell(x) + (1-\alpha)\ell(y)$ for every $0 \leq \alpha \leq 1$, it follows that $\{\alpha x + (1-\alpha)y \colon 0 \leq \alpha \leq 1\} \subseteq \mathbb{R}^d_-$.

Proof. Given a set $C \subseteq \mathbb{R}^d$, we define $m \mapsto \delta(m, C)$ as 0 if $m \in C$ and ∞ otherwise. The function $f(m) = \sum m_k + \delta(m, A(X))$ is increasing, convex, lower semi-continuous, proper and such that $R(X) = \inf\{f(m)\colon m \in \mathbb{R}^d\}$. Hence, by [26, Theorem 27.1 (b)], in order to prove the existence of a risk allocation, it suffices to show that f is constant along its directions of recession. By Theorem A.1, this is equivalent to showing that $u \in 0^+ f$ implies $(-u) \in 0^+ f$ where $0^+ f$ is the recession cone of f, see Appendix A for the exact definition. Let $\gamma \in \mathbb{R}$ be such that $B := \{x \in \mathbb{R}^d\colon f(x) \leq \gamma\} \neq \emptyset$, so that $0^+ B = 0^+ f$, by Theorem A.1. By means of Theorem 2.9, $R(X) > -\infty$. Fix $b \in B$. Hence, $u \in 0^+ B$ if and only if

$$-\infty < R(X) \leq \sum b_k + r \sum u_k \leq \gamma < \infty \quad \text{and} \quad b + ru \in A(X)$$

for every $r \geq 0$. As a consequence, $u \in 0^+ B$ if and only if $u \in Z \cap 0^+ A(X)$, where Z denotes the set of zero-sum allocations in $\mathbb{R}^d$. This reduces the proof of existence to showing that if $u \in Z \cap 0^+ A(X)$, then $(-u) \in 0^+ A(X)$. Since $A(X)$ is a non-empty lower level set of $m \mapsto g(m) = E[\ell(X - m)]$, Theorem A.1 implies that $u \in 0^+ A(X)$ if and only if $0 \geq (g0^+)(u)$. As b is in $A(X)$, therefore in $\mathrm{dom}(g)$, Theorem A.1 yields

$$\begin{aligned} (g0^+)(u) &= \sup_{r>0} E\left[\frac{\ell(X - b - ru) - \ell(X - b)}{r}\right] \\ &= E\left[\sup_{r>0} \frac{\ell(X - b - ru) - \ell(X - b)}{r}\right] \\ &= (\ell 0)^+(-u), \end{aligned} \tag{3.2}$$

for any $u \in \mathbb{R}^d$, where the second equality follows by means of the monotone convergence theorem. It follows that $u \in 0^+ A(X)$ if and only if $0 \geq (\ell 0^+)(-u)$. Hence, if $u \in Z \cap 0^+ A(X)$, then, on the one hand, $0 \geq (\ell 0^+)(-u) = \sup_r \ell(-ru)/r$, by Theorem A.1. But, on the other hand, $u \in Z$ implies that $\ell(-ru)/r \geq -r \sum u_k / r = 0$. In conclusion, if $u \in Z \cap 0^+ A(X)$, then $\ell(-ru) = 0$, for every $r \geq 0$, which in turn yields that $\ell(ru) = 0$, since ℓ is unbiased. In particular, $0 = (\ell 0)^+(u) = (g0)^+(-u)$, by (3.2) applied to $-u$. Hence, $-u \in 0^+ A(X)$, which proves the existence of a risk allocation m^*.

Since $E[\ell(X - m)] - c < 0$ for some m large enough, the Slater condition for the convex optimization problem $R(X) = \inf_m f(m)$ is fulfilled. Hence, according to [26, Theorems 28.1, 28.2 and 28.3], optimal solutions m^* are characterized by (3.1).

In the case where ℓ has no zero-sum direction of recession except 0, it follows from the previous computations that $0^+ f = Z \cap 0^+ A(X) = Z \cap 0^+ g = Z \cap (-0^+ \ell) = \{0\}$. Hence, by [26, Theorem 27.1, (d)], the set of risk allocations is non-empty and bounded.

Finally, let $m \neq n$ be two risk allocations. It follows that $\alpha m + (1 - \alpha) n$ is a risk allocation as well for every $\alpha \in [0, 1]$. Furthermore, $(m - n)$ is a zero sum allocation. By convexity, it follows that $c = E[\ell(X - \alpha m - (1-\alpha)n)] \leq \alpha E[\ell(X - m)] + (1-\alpha) E[\ell(X - n)] = c$ for every $0 \leq \alpha \leq 1$, which shows that $\alpha \ell(X - m) + (1-\alpha)\ell(X - n) = \ell(X - \alpha m - (1-\alpha)n)$ P-almost surely for every $0 \leq \alpha \leq 1$. Since ℓ is strictly convex along zero-sum allocations outside $\mathbb{R}^d_-$, it follows that $\{X - \alpha m - (1-\alpha)n\colon 0 \leq \alpha \leq 1\} \subseteq \mathbb{R}^d_-$. In particular $X - m \leqslant 0$, which implies $E[\ell(X - m)] \leq \ell(0) = 0 < c$, a contradiction. □

Corollary 3.7. *Let ℓ be an unbiased loss function, strictly convex outside $\mathbb{R}^d_-$. It holds*

$$RA(X + r) = RA(X) + r \quad \textit{for every } X \in M^\theta \textit{ and } r \in \mathbb{R}^d.$$

Proof. From Theorem 3.6, the assumptions on ℓ ensure the existence and uniqueness of a risk allocation uniquely characterized, together with the Lagrange multiplier, by the first order conditions. Let $m =$

$RA(X+r)$, for which there exists a unique λ such that

$$\lambda E\left[\nabla \ell\left(X+r-m\right)\right]=1 \quad \text{and} \quad E\left[\ell\left(X+r-m\right)\right]=c.$$

Hence, $n=m-r$ and λ satisfy the first order conditions

$$\lambda E\left[\nabla \ell\left(X-n\right)\right]=1 \quad \text{and} \quad E\left[\ell\left(X-n\right)\right]=c,$$

which by uniqueness shows that $n=RA(X)=m-r=RA(X+r)-r$. □

Remark 3.8. In general, the positivity of the risk allocation is not required. If positivity or any other convex constraint is imposed, for instance by regulators, it can easily be embedded in our setup. In case of positivity, this would modify the definition of $R(X)$ into

$$R(X)=\inf\left\{\sum m_k \colon E\left[\ell(X-m)\right]\leq c \text{ and } m_k\geq 0 \text{ for every } k\right\},$$

with accordingly modified first order conditions. ⧫

The question of existence and uniqueness of a risk allocation having been addressed, the following example shows the economic importance of the uniqueness.

Example 3.9. Any loss function of class **(C1)**, that is, $\ell(x)=h(\sum x_k)$, is permutation invariant and therefore unbiased, see Remark 3.5. Thus, a risk allocation $m^*\in A(X)$ exists by means of Theorem 3.6. However, for any zero-sum allocation u, we have $R(X)=\sum m_k^*+u_k=\sum m_k^*$ and $E[h(\sum X_k-(m_k^*+u_k))]=E[h(\sum X_k-m_k^*)]\leq c$, so that m^*+u is another risk allocation.

Economically, this is not a sound situation. Indeed, suppose that we have two banks as risk components and we require from them 110 M € and 500 M €, respectively, as risk allocation. In such a case, we could equally well require 610 M € from the first bank and nothing from the second. Such arbitrariness is unlikely to be accepted by any economic agent. ◊

Example 3.9 shows that loss functions of the class **(C1)** are economically unsuitable, by lack of uniqueness of a risk allocation. By contrast, for loss functions of class **(C2)**, that is, $\ell(x)=\sum h(x_k)$, the following proposition shows that, while there exists a unique risk allocation under very mild conditions, the risk allocation only depends on the marginal distributions of the loss vector $X=(X_1,\dots,X_d)$. In other words, the risk measure and the risk allocation do not account for the dependence structure of the risk components. This makes the loss functions of class **(C2)** equally unsuitable from a systemic risk point of view.

Proposition 3.10. *Let $\ell(x):=\sum h_k(x_k)$ for univariate loss functions $h_k:\mathbb{R}\to(-\infty,\infty]$ strictly convex on $\mathbb{R}_+$, $k=1,\dots,d$. For every $X\in M^\theta$, there exists a unique optimal risk allocation $RA(X)$ and we have $RA(X)=RA(Y)$, for every $Y\in M^\theta$ such that Y_k has the same distribution as X_k, $k=1,\dots,d$.*

Proof. Let x,y be such that $\alpha x+(1-\alpha)y\notin\mathbb{R}^d_-$ for every $\alpha\in(0,1)$. It follows that $\ell(\alpha x+(1-\alpha)y)=\sum h_k(\alpha x_k+(1-\alpha)y_k)<\sum \alpha h_k(x_k)+(1-\alpha)h_k(y_k)=\alpha\ell(x)+(1-\alpha)\ell(y)$. The loss function ℓ is furthermore unbiased. Indeed, for every zero-sum allocation u, assuming without loss of generality $u_1>0$, it follows that

$$\ell 0^+(u)\geq \lim_{r\to\infty} h_1(ru_1)/r+\sum_{k\geq 2} h_k(ru_k)/r\geq \lim_{r\to\infty} h_1(ru_1)/r+\sum_{k\geq 2} u_k=\infty$$

since h_1 is strictly convex and $h_1(t) \geq t$. Hence, ℓ has no zero-sum direction of recession other than 0. The strict convexity of h_k yields, according to Theorem 3.6, the existence of a unique risk allocation for every $X \in M^\theta$. The first order conditions (3.1) are written as

$$1 \in \lambda E\left[\partial h_k(X_k - m_k)\right], \quad k = 1, \ldots, d, \quad \text{and} \quad \sum E\left[h_k\left(X_k - m_k\right)\right] = c,$$

which only depend on the marginal distributions of X. □

4. Systemic Sensitivity of Shortfall Risk and its Allocation

The previous results emphasize the importance of using a loss function that adequately captures the systemic risk inherent to the system. This motivates the study of the sensitivity of shortfall risk and its allocation so as to identify the systemic features of a loss function.

Definition 4.1. The *marginal risk contribution* of $Y \in M^\theta$ to $X \in M^\theta$ is defined as the sensitivity of the risk of X with respect to the impact of Y, that is

$$R(X;Y) := \limsup_{t \searrow 0} \frac{R(X+tY) - R(X)}{t}.$$

In the case where $R(X+tY)$ admits a unique risk allocation $RA(X+tY)$ for every t, the *risk allocation marginals* of the risk of X with respect to the impact of Y are given by

$$RA_k(X;Y) = \limsup_{t \searrow 0} \frac{RA_k(X+tY) - RA_k(X)}{t}, \quad k = 1, \ldots, d.$$

Theorem 2.9 and its proof show that the determination of the risk measure $R(X)$ reduces to the saddle point problem

$$R(X) = \min_m \max_{\lambda > 0} L(m, \lambda, X) = \max_{\lambda > 0} \min_m L(m, \lambda, X).$$

Using [26], the "argminmax" set of saddle points (m^*, λ^*) is a product set that we denote by $B(X) \times C(X)$.

Theorem 4.2. *Assuming that ℓ has no zero-sum direction of recession, then*

$$R(X;Y) = \min_{m \in B(X)} \max_{\lambda \in C(X)} \lambda E\left[Y \cdot \nabla \ell\left(X - m\right)\right].$$

Supposing further that ℓ is twice differentiable and that $(m, \lambda) \in B(X) \times C(X)$ is such that

$$M = \begin{bmatrix} \lambda E\left[\nabla^2 \ell(X-m)\right] & -1/\lambda \\ 1 & 0 \end{bmatrix}$$

is non-singular, then

- *there exists $t_0 > 0$ such that $B(X+tY) \times C(X+tY)$ is a singleton, for every $0 \leq t \leq t_0$;*
- *the corresponding unique saddle point $(m_t, \lambda_t) = (RA(X+tY), \lambda_t)$ is differentiable as a function of t and we have*

$$\begin{bmatrix} RA(X;Y) \\ \lambda(X;Y) \end{bmatrix} = M^{-1} V,$$

where $\lambda(X;Y) = \limsup_{t \searrow 0} (\lambda_t - \lambda_0)/t$ and

$$V = \begin{bmatrix} \lambda E\left[\nabla^2 \ell(X-m) Y\right] \\ R(X;Y) \end{bmatrix}.$$

Proof. Let $L(m,\lambda,t) = \sum m_k + \lambda E[\ell(X+tY-m)-c]$. Theorem 2.9 yields

$$R(X+tY) = \min_m \max_\lambda L(m,\lambda,t) = \max_\lambda \min_m L(m,\lambda,t) = L(m_t,\lambda_t,t),$$

for every selection $(m_t,\lambda_t) \in B(X+tY)\times C(t+tY)$.

Regarding the first assertion of the theorem, since ℓ has no zero-sum direction of recession other than 0, it follows from Theorem 3.6 that $B(X)\times C(X)$ is non empty and bounded. Hence, the assumptions of Golshtein's Theorem on the perturbation of saddle values, see Rockafellar and Wets [27, Theorem 11.52], are satisfied and the first assertion follows.

As for the second assertion, the assumptions of Fiacco and McCormick [18, Theorem 6, pp. 34–45] are fulfilled. The Jacobian of the vector

$$\begin{bmatrix} \nabla_m L(m,\lambda,0) \\ \lambda E\left[\ell\left(X-m\right)-c\right] \end{bmatrix}$$

that is used to specify the first order conditions is given by the matrix M. Hence, the second assertion follows from [18, Theorem 6, pp. 34–35]. □

4.1. Sensitivity with Respect to an Exogenous Shock

We want to study the impact of an additional loss on the first risk components in a bivariate framework. To this end we consider the loss function

$$\ell(x_1,x_2) = \frac{(x_1^+)^2+(x_2^+)^2}{2} + \alpha x_1^+ x_2^+ + x_1 + x_2, \quad 0\le\alpha\le 1,$$

which gives rise to a unique risk allocation by virtue of Theorem 3.6; this allocation is denoted by $(m_1,m_2) = (RA_1(X), RA_2(X))$. The first order conditions (3.1) yield

$$\begin{aligned}
\frac{1}{\lambda} - 1 &= E\left[(X_1-m_1)^+\right] + \alpha E\left[(X_2-m_2)^+ 1_{\{X_1\ge m_1\}}\right], \\
\frac{1}{\lambda} - 1 &= E\left[(X_2-m_2)^+\right] + \alpha E\left[(X_1-m_1)^+ 1_{\{X_2\ge m_2\}}\right], \\
c &= \frac{1}{2}E\left[((X_1-m_1)^+)^2 + ((X_2-m_2)^+)^2\right] + \alpha E\left[(X_1-m_1)^+(X_2-m_2)^+\right] \\
&\quad + E[X_1] + E[X_2] - m_1 - m_2.
\end{aligned}$$

These conditions show that the first component m_1 of the risk allocation primarily depends

- first, on the loss X_1 in excess above the liquidity requirement m_1 itself, via the term

$$E\left[(X_1-m_1)^+\right];$$

- second, weighted by α, on the loss X_2 in excess over the liquidity requirement m_2, conditioned on the fact that the loss of X_1 also exceeds its own liquidity requirement m_1, via the term

$$\alpha E\left[(X_2-m_2)^+ 1_{\{X_1\ge m_1\}}\right],$$

  which clearly involves the dependence structure between X_1 and X_2.

Let then $Y = (Y_1, 0)$ be an exogenous shock impacting the first risk component.

4.1.1. Marginal Risk Contribution

According to Theorem 4.2, we have

$$R(X;Y) = \lambda E\left[Y_1\right] + \lambda E\left[Y_1\left(X_1 - m_1\right)^+\right] + \alpha\lambda E\left[Y_1\left(X_2 - m_2\right)^+ 1_{\{X_1 \geq m_1\}}\right].$$

This equation shows that the impact of Y_1 on the first risk component X_1 decomposes into three parts in the marginal risk contribution $R(X;Y)$:

- an impact of Y_1 on the whole system via the term $E[Y_1]$;
- an impact of Y_1 related to the concerned risk component, that is, conditioned on the fact that X_1 exceeds its own liquidity requirement, via the term $E[Y_1\left(X_1 - m_1\right)^+]$;
- a systemic impact of Y_1 onto the dependent risk component, that is, conditioned on the fact that both components exceed their respective liquidity requirements, via the term $\alpha E[Y_1\left(X_2 - m_2\right)^+ 1_{\{X_1 \geq m_1\}}]$.

The last effect takes into account the dependence structure of the system, provided α is strictly positive. In particular, if α is non-zero, then the systemic risk impact increases with the "loss dependence" between X_1 and X_2.

4.1.2. Risk Allocation Sensitivity

We write $p = P[X_1 \geq m_1]$, $q = P[X_2 \geq m_2]$ and $r = P[X_1 \geq m_1; X_2 \geq m_2]$. In the notation of Theorem 4.2, we have:

$$M = \begin{bmatrix} \lambda p & \lambda\alpha r & -1/\lambda \\ \lambda\alpha r & \lambda q & -1/\lambda \\ 1 & 1 & 0 \end{bmatrix} \quad \text{and} \quad V = \begin{bmatrix} \lambda E\left[Y_1 | X_1 \geq m_1\right] p \\ \lambda\alpha E\left[Y_1 | X_1 \geq m_1; X_2 \geq m_2\right] r \\ R(X;Y) \end{bmatrix},$$

which by Theorem 4.2 yields

$$RA_1(X;Y) = \frac{(q - \alpha r)R(X;Y) + E\left[Y_1 | X_1 \geq m_1\right] p - \alpha E\left[Y_1 \middle| X_1 \geq m_1; X_2 \geq m_2\right] r}{p + q - 2\alpha r},$$
$$RA_2(X;Y) = \frac{(p - \alpha r)R(X;Y) - E\left[Y_1 | X_1 \geq m_1\right] p + \alpha E\left[Y_1 \middle| X_1 \geq m_1; X_2 \geq m_2\right] r}{p + q - 2\alpha r},$$

consistent with the general identity $RA_1(X;Y) + RA_2(X;Y) = R(X;Y)$. Observe that

- the first risk component, impacted by the exogenous effect, is taking away from the second component the non-correlated share of additional risk $E[Y_1|X_1 \geq m_1]p$;
- in the presence of the systemic weight $\alpha > 0$, the first risk component offloads a share of its additional liquidity requirement to the second component according to $\alpha E[Y_1|X_1 \geq m_1; X_2 \geq m_2]r$.

4.2. Sensitivity with Respect to the Interdependence Structure

In the following example, we analyse the sensitivity of the risk allocation with respect to the interdependence structure in a bivariate Gaussian setup. To this end, we consider the following loss function

$$\ell(x_1, x_2) = \frac{1}{1+\alpha}\left[\frac{1}{2}e^{2x_1} + \frac{1}{2}e^{2x_2} + \alpha e^{x_1}e^{x_2}\right] - 1.$$

Let $X = (X_1, X_2)$ be a normally distributed random vector with zero mean and covariance matrix

$$\Sigma = \begin{bmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{bmatrix}.$$

We denote

$$s_k := E\left[e^{2X_k}\right] = e^{2\sigma_k^2}, \quad k = 1, 2 \quad \text{and} \quad r := E\left[e^{X_1}e^{X_2}\right] = e^{\frac{1}{2}\left(\sigma_1^2+\sigma_2^2+2\rho\sigma_1\sigma_2\right)}.$$

The first order conditions yield

$$\begin{aligned} 1/\lambda &= \frac{\alpha}{1+\alpha} r e^{-m_1} e^{-m_2} + \frac{1}{1+\alpha} s_1 e^{-2m_1}, \\ 1/\lambda &= \frac{\alpha}{1+\alpha} r e^{-m_1} e^{-m_2} + \frac{1}{1+\alpha} s_2 e^{-2m_2}, \\ \tilde{c} := c + 1 &= \frac{\alpha}{1+\alpha} r e^{-m_1} e^{-m_2} + \frac{1}{2(1+\alpha)} \left(s_1 e^{-2m_1} + s_2 e^{-2m_2}\right), \end{aligned}$$

with solution

$$\begin{aligned} RA_1(X) &= \sigma_1^2 + \frac{1}{2}\ln\left(1 + \alpha e^{\rho\sigma_1\sigma_2 - \frac{1}{2}(\sigma_1^2+\sigma_2^2)}\right) - \frac{1}{2}\ln\left(\tilde{c}(1+\alpha)\right), \\ RA_2(X) &= \sigma_2^2 + \frac{1}{2}\ln\left(1 + \alpha e^{\rho\sigma_1\sigma_2 - \frac{1}{2}(\sigma_1^2+\sigma_2^2)}\right) - \frac{1}{2}\ln\left(\tilde{c}(1+\alpha)\right), \\ R(X) &= \sigma_1^2 + \sigma_2^2 + \ln\left(1 + \alpha e^{\rho\sigma_1\sigma_2 - \frac{1}{2}(\sigma_1^2+\sigma_2^2)}\right) - \ln\left(\tilde{c}(1+\alpha)\right). \end{aligned}$$

These equations show that the risk allocations are disentangled into the respective individual contributions σ_i^2, $i = 1, 2$, and a *systemic risk contribution*

$$SRC = \frac{1}{2}\ln\left(1 + \alpha e^{\rho\sigma_1\sigma_2 - \frac{1}{2}(\sigma_1^2+\sigma_2^2)}\right) - \frac{1}{2}\ln\left(\tilde{c}(1+\alpha)\right), \tag{4.1}$$

which depends on the correlation parameter ρ and on the systemic weight α of the loss function. Figure 2 shows the value of this systemic risk contribution as a function of ρ and σ_1.

We obtain the following sensitivities of the systemic risk contribution

$$\begin{aligned} \frac{\partial SRC}{\partial \sigma_1} &= \frac{\alpha\left(\rho\sigma_2 - \sigma_1\right)}{2} \frac{e^{\rho\sigma_1\sigma_2 - \frac{1}{2}(\sigma_1^2+\sigma_2^2)}}{1 + \alpha e^{\rho\sigma_1\sigma_2 - \frac{1}{2}(\sigma_1^2+\sigma_2^2)}}, \\ \frac{\partial SRC}{\partial \sigma_2} &= \frac{\alpha\left(\rho\sigma_1 - \sigma_2\right)}{2} \frac{e^{\rho\sigma_1\sigma_2 - \frac{1}{2}(\sigma_1^2+\sigma_2^2)}}{1 + \alpha e^{\rho\sigma_1\sigma_2 - \frac{1}{2}(\sigma_1^2+\sigma_2^2)}}, \\ \frac{\partial SRC}{\partial \rho} &= \frac{\alpha\sigma_1\sigma_2}{2} \frac{e^{\rho\sigma_1\sigma_2 - \frac{1}{2}(\sigma_1^2+\sigma_2^2)}}{1 + \alpha e^{\rho\sigma_1\sigma_2 - \frac{1}{2}(\sigma_1^2+\sigma_2^2)}}. \end{aligned}$$

In particular, the systemic risk contribution is

- increasing with respect to the correlation ρ;
- decreasing with respect to σ_1 if the correlation is negative;
- increasing up to $\rho\sigma_2$ and then decreasing with respect to σ_1 if the correlation is positive.

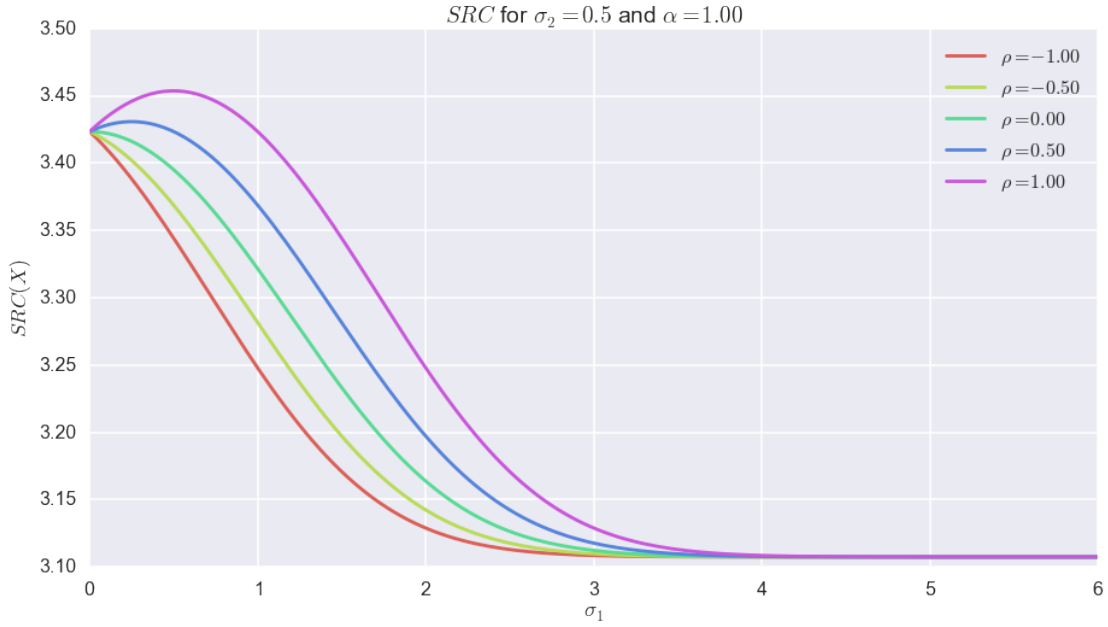

Figure 2: Systemic risk contribution (4.1) as a function of σ_1 for different values of the correlation ρ in the case where $\alpha = 1$ and $c = 1$.

4.3. Riskless Allocation, Causal Responsibility and Additivity

We conclude this section regarding risk allocation and its sensitivity by a discussion of their properties in light of the following economic features of risk allocations introduced in [8].

(FA) Full Allocation: $\sum RA_k(X) = R(X)$;

(RA) Riskless Allocation: $RA_k(X) = X_k$ if X_k is deterministic;

(CR) Causal Responsibility: $R(X + \Delta X_k) - R(X) = RA_k(X + \Delta X_k) - RA_k(X)$, where ΔX_k is a loss increment of the k-th risk component;

As mentioned before, per design, shortfall risk allocations always satisfy the full allocation property **(FA)**. As visible from the above case studies, riskless allocation **(RA)** and causal responsibility **(CR)** are not satisfied in general. In fact, from a systemic risk point of view, we think that **(RA)** and **(CR)** are not desirable properties. Indeed, both imply that risk taking, or non-taking, should only impact the concerned risk component. However, the risk components are interdependent and any move in one of them bears consequences to the rest of the system. The search for an optimal allocation is a non-cooperative game between the different system components, each of them respectively looking for its own minimal risk allocation while impacting the others by doing so. In other words, everyone is responsible for its own risk but also for its relative exposure with respect to the others. The sensitivity analysis of this section shows that external shocks are primarily born by the risk component that is hit. Then a correction appears and a fraction of the shock is offloaded to the other risk components according to their relative exposure to the concerned component.

5. Numerical Schemes

The aim of this section is to present and compare numerical schemes for the computation of multivariate shortfall risk measures and the corresponding risk allocations. They are provided, according to Theorem 3.6, by the following system of equations:

$$1 \in \lambda^* E[\nabla \ell(X - m^*)] \quad \text{and} \quad E[\ell(X - m^*)] = c,$$

where the random variable X, the loss function ℓ and the threshold level c are given. This can be viewed as a deterministic optimization problem, as soon as the expectations have been computed by some numerical scheme. In the sequel, we will compare three schemes for this computation: the Fourier transform method, the Monte Carlo method and the Chebyshev interpolation method. We will focus on the loss function **(C4)** in order to be able to make explicit computations and compare numerical results. However, we want to point out that these methods are general, apply to any (unbiased) loss function and are well-suited for high-dimensional problems. Indeed, in the next section we will discuss a 10- and a 30-dimensional example.

We consider the loss function **(C4)** of Example 2.3, that is,

$$\ell(x) = \sum_{k=1}^{d} x_k + \frac{1}{2} \sum_{k=1}^{d} (x_k^+)^2 + \alpha \sum_{1 \le j < k \le d} x_j^+ x_k^+, \tag{5.1}$$

for $\alpha = 0$ or 1, whose partial derivative with respect to x_k equals:

$$\partial_k \ell(x) = 1 + x_k^+ + \alpha \sum_{j=1, j \neq k}^{d} x_j^+ 1_{\{x_k \ge 0\}}.$$

Let us define the following functions:

$$\begin{aligned} e_k(m) &= E[(X_k - m)] & h_{jk}(m,n) &= E\left[(X_j - m)^+ (X_k - n)^+\right] \\ f_k(m) &= E\left[(X_k - m)^+\right] & l_{jk}(m,n) &= E\left[(X_j - m)^+ 1_{\{X_k \ge n\}}\right]. \\ g_k(m) &= E\left[\left((X_k - m)^+\right)^2\right] \end{aligned} \tag{5.2}$$

According to Theorem 3.6, the risk allocation is determined by the first order conditions (3.1), which read in this case:

$$\begin{cases} \lambda f_k(m_k) + \alpha\lambda \displaystyle\sum_{j=1, j\neq k}^{d} l_{jk}(m_j, m_k) = 1 - \lambda, \quad \text{for } k = 1, \dots, d; \\ \displaystyle\sum_{k=1}^{d} \left\{ e_k(m_k) + \frac{1}{2} g_k(m_k) \right\} + \alpha \sum_{1 \le j < k \le d} h_{jk}(m_j, m_k) = c. \end{cases} \tag{5.3}$$

In order to solve this optimization problem, we will rely on standard packages for root-finding algorithms and combine them with the different methods for computing the expectations. There are two major challenges here: On the one hand, the number of bivariate expectations to be evaluated increases quadratically with the dimension d of the system. On the other hand, the number of steps needed to ensure the convergence to the root increases linearly with the dimension d of the system; for instance, the Broyden method requires $2d$-steps to convergence quadratically, see Gay [22].[6]

Below we provide a brief overview of the numerical methods for the computation of the expectations, while more details are postponed for Appendix C.

[6] Under some regularity conditions, the Broyden method yields a convergence $\left\|m^* - m^{2d+n}\right\|_2 \le \gamma \left\|m^* - m^n\right\|_2^2$, where n denotes the current step in the algorithm.

Fourier transform method: Assuming that the moment generating functions M_{X_k}, M_{X_k,X_j}, for all $1 \leq k, j \leq d$, are available, Fourier methods allow us to exactly compute the functions e, f, g, h, l in (5.2). In particular, we have the following result:

Proposition 5.1. *Let $X \in M^\theta$ and assume that the following conditions hold:*

$$\begin{aligned} \mathcal{I}'_k &= \{\eta \in \mathbb{R}_+ : M_{X_k}(\eta) < \infty \text{ and } M_{X_k}(\eta + i\cdot) \in L^1\} \neq \emptyset, \\ \mathcal{I}'_{j,k} &= \{\eta \in \mathbb{R}^2_+ : M_{X_k,X_j}(\eta) < \infty \text{ and } M_{X_k,X_j}(\eta + i\cdot) \in L^1\} \neq \emptyset, \end{aligned} \tag{5.4}$$

for all $1 \leq k, j \leq d$. Then, the functions f_k, g_k, h_{jk}, l_{jk} admit the following representations:

$$\begin{aligned} f_k(m) &= -\frac{1}{2\pi} \int_{\mathbb{R}} e^{(iu_k - \eta_k)m} \frac{M_{X_k}(\eta_k - iu_k)}{(u_k + i\eta_k)^2} du_k, \\ g_k(m) &= \frac{1}{\pi} \int_{\mathbb{R}} e^{(iu_k - \eta_k)m} \frac{M_{X_k}(\eta_k - iu_k)}{i(u_k + i\eta_k)^3} du_k, \\ h_{jk}(m,n) &= \frac{1}{(2\pi)^2} \int_{\mathbb{R}^2} e^{\langle iu - \eta, (m,n)^\intercal \rangle} \frac{M_{X_j,X_k}(\eta - iu)}{(u_j + i\eta_j)^2 (u_k + i\eta_k)^2} du_j du_k, \\ l_{jk}(m,n) &= \frac{1}{(2\pi)^2} \int_{\mathbb{R}^2} e^{\langle iu - \eta, (m,n)^\intercal \rangle} \frac{M_{X_j,X_k}(\eta - iu)}{(u_j + i\eta_j)^2 (iu_k - \eta_k)} du_j du_k. \end{aligned} \tag{5.5}$$

Here $\eta_k \in \mathcal{I}'_k$ and $\eta = (\eta_j, \eta_k) \in \mathcal{I}'_{j,k}$, for all $1 \leq j, k \leq d$.

Remark 5.2. Assuming that the moment generating function of X is known in an (open) set, then we can easily compute the expectations of its marginals X_k by differentiation. Thus, we will not further discuss the computation of $e_k(m)$ in (5.2). ♦

See Appendix C.1 for more details and the proof of this result. The main advantage of this method is that it is theoretically possible to compute the value of the integrals at any level of precision, while the basic computational time is roughly doubled for every additional digit of accuracy. Moreover, it is very efficient for several applications in mathematical finance, for instance model calibration. However, as shown in Table 1, this method suffers from several drawbacks in the present context. The basic computational time for four digits of accuracy varies from one to over thirty seconds. This depends on the values of m, n and on the dampening factor η. In addition, it may happen that for given m, n, several different values of the dampening factor η have to be tried at random before the integral converges in the first place. Taking the large number of double integrals to be computed into account, see Table 2, the computational time can become prohibitively long.

Monte Carlo method: We can also use Monte Carlo simulations for the estimation of the functions e, f, g, h, l in (5.2). Indeed, let $(X^i)_{1 \leq i \leq M}$ be M independent realizations of the random vector $X = (X_1, \ldots, X_d)$, then the estimators for e_k and h_{jk} are provided by

$$I^{MC}_M[e_k](m) = \frac{1}{M} \sum_{i=1}^{M} (X^i_k - m) \quad \text{and} \quad I^{MC}_M[h_{jk}](m,n) = \frac{1}{M} \sum_{i=1}^{M} (X^i_k - m)^+ (X^i_j - n)^+,$$

while the estimators for f, g anf l are analogous. An important observation here is that we can generate and store all M realizations in advance, and then use them for the estimation of the functions e, f, g, h, l for different m's and n's in every step of the root-finding procedure.

The main advantage of Monte Carlo relative to Fourier methods is that a wider variety of models can be considered; think, for example, of models with copulas or of random variables with Pareto type distributions. The main disadvantage is the slow statistical convergence of the scheme: For an additional digit of accuracy, one hundred times more samples are required. In our context, though, a four digit accuracy is reasonable. In addition, the time to generate, once and for all, the samples, as well as to compute the Monte Carlo averages, is very fast, see Table 1, and is independent of the value of m and n.

	Fourier		Monte Carlo	
	$m = n = 1$	$m = n = -10$	Generation of M	$\sum_M$
Computational Time	1190 ms	27150 ms	138 ms	45 ms

Table 1: Computation of h_{jk} for a bivariate Gaussian distribution with zero mean, $\sigma_1 = \sigma_2 = 1$ and $\rho = 0.8$. Accuracy 10^{-4} for Fourier and $M = 10^6$ samples for Monte Carlo. The result of 27150 milliseconds includes two non-convergent trials of $\eta = (\eta_j, \eta_k)$ before converging.

The expectations in every step of the root-finding routine can be computed either directly using Fourier or Monte Carlo methods, or in combination with the following interpolation scheme.

Chebyshev interpolation method: A numerical scheme to approximate the functions e, f, g, h, l in (5.2) that might be more amenable to root-finding routines is the Chebyshev interpolation method, recently applied to option pricing by Gaß et al. [21]. This method can be summarized as follows: Suppose you want to evaluate quickly a function $F(m)$, of one or several variables, for a large number of m's. The first step of the Chebyshev method is to evaluate the function $F(m)$ on a given set of nodes m^i, $1 \leq i \leq N$. These evaluations can be computed by e.g. Fourier or Monte Carlo schemes, are independent of each other and can thus be realized in parallel. The next step, in order to compute $F(m)$ for an m outside the nodes m^i, is to perform a polynomial interpolation of the $F(m^i)$'s using the Chebyshev coefficients. In other words, the Chebyshev method provides a polynomial approximation $\hat{F}(m)$ of $F(m)$. As an example, having evaluated $f(m^i)$ at the Chebyshev nodes m^i for $0 \leq i \leq N$, the approximation of the function f using the Chebyshev method reads

$$I_N[f_k](m) = \sum_{i=0}^{N} c_i T_i(m),$$

where

$$c_i = \frac{2^{1_{\{i>0\}}}}{N+1} \sum_{r=0}^{N} f_k(m^r) \cos\left(i\pi \frac{2r+1}{2N+1}\right) \quad \text{and} \quad T_i(m) = \cos(i \arccos(m)).$$

See Appendix C.2 for more details and formulas in the two-dimensional case. One drawback of the method tough, is that the m^i lie in bounded intervals for which a good initial guess is not trivial.

The Chebyshev method can be combined with a root-finding algorithm for solving the equation $F(m) = c$ as follows: Compute first $F(m^i)$ on the Chebyshev nodes m^i. At eact step of the root-finding procedure approximate $F(m)$ with $\hat{F}(m)$, and proceed iteratively. If the computation of $F(m)$ is expensive in

computational time, then also the iterative computation is also very costly[7]. On the contrary, the Chebyshev method offers a very fast and accurate approximation of F—the convergence is (sub)exponential in the number of nodes m—while the computationally most expensive part can be performed outside the root-finding loop.

In the subsequent discussion and the following chapter with numerical results, we will use the following shorthand notation:

- CTP: Computational time for the preprocessing (i.e. before the root-finding starts);
- CTR: Computational time of the root finding method;
- CT: Computational time in total;
- MC: Monte Carlo scheme;
- F: Fourier scheme;
- Cheby: Chebyshev scheme;
- d: Number of risk components in the system;
- N: Number of nodes for the Chebyshev scheme;
- $I(d)$: Number of iterations for the root-finding method;
- M: Number of samples for the Monte Carlo scheme.

Discussion: We want now to compare the computational cost for the Fourier and Monte Carlo methods, with or without the use of Chebychev interpolation. Whether it is advantageous to use the Chebyshev interpolation or not, is a matter of two competing factors that affect the computational time: On the one hand, the number of iterations $I(d)$ needed to find the root of the system, which is as mentioned beforehand roughly of order d, and on the other hand, the size of the grid N^2 used in the Chebyshev interpolation. Table 2 summarizes the computational burden for the Fourier and Monte Carlo methods, with and without the use of Chebyshev interpolation as a function of $I(d)$ and N, in the case where $\alpha = 1$.[8]

This table reveals that, rather counterintuitively, the Monte Carlo schemes are better than the Fourier schemes in the range of our accuracy requirements, since they require the least amount of work during each step of the root-finding procedure or for the pre-processing computations in the Chebyshev method. Only when the dimension is low, less than three, and the accuracy requirement is high can the Fourier methods be faster. Next, the choice between Chebyshev or not is a matter of comparison between $I(d)$ and N^2. In high dimensions, when $I(d)$ dominates N^2, with $I(d)$ being in principle of order d and N usually between 10 and 20, then the Chebyshev method is less costly. Furthermore, the Chebyshev method can intensively benefit from parallel computing as the pre-processing step is not sequential.

6. Numerical Results

In this section, we present computational results based on the different numerical schemes discussed in the previous section for the loss function (5.1). The implementation was done on standard desktop computers in the Python programming language, mainly based on the numpy and scipy packages. In particular, all resolutions of system (5.3) were realized using the fsolve function of the scipy.optimize package. As for

[7] We make use of the SciPy fsolve which is a wrapper of MINPACK's hybrd and hybrj algorithms. Quoting MINPACK's official documentation, "Unless FCN [the system of functions] can be evaluated quickly, the timing of HYBRJ1 will be strongly influenced by the time spent in FCN."

[8] In the case where $\alpha = 0$ there are no bivariate integrals to be computed and all the schemes are very fast, with advantage for Fourier together with Chebyshev.

Method		Computational Costs	
		Pre-Processing	Root Finding
w/o Chebyshev	Fourier	-	$O\left(I(d)d^2 \times \iint\right)$
	MC (*)	-	$O\left(I(d)d^2 \times \sum_M\right)$
w/ Chebyshev	Fourier	$O\left(N^2d^2 \times \iint\right)$	$O\left(I(d) \times \hat{F}\right)$
	MC (*)	$O\left(N^2d^2 \times \sum_M\right)$	

Table 2: Comparison of the computational work of the numerical schemes in leading terms, as a function of the number of iterations $I(d)$ for the root finding and of the number of nodes N in the Chebyshev method in the case where $\alpha \neq 0$. The (*) for the Monte Carlo methods means that, even though the generation of samples takes time, see Table 1, as it is done once and for all, it is negligible with respect to the other computational times. The symbols $\iint$ and $\sum_M$ denote the computational time for a Fourier double integral and a Monte Carlo average over M samples, respectively.

the distribution of X, we mainly use Gaussian distributions with mean vector μ and variance-covariance matrix Σ. In the bi- and tri-variate cases the variance-covariance matrix is parameterized by a single correlation factor ρ and the variances σ_k^2 of X_k for all k. In other words, $\Sigma_{ij} = \rho\sigma_i\sigma_j$ for $i \neq j$. In this case, using the notation of (5.2), the functions f_k and g_k have the following analytical forms

$$
\begin{aligned}
f_k(x) &= E\left[(X_k - x)^+\right] = \frac{\sigma_k}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2\sigma_k^2}\right) - x\Phi\left(-\frac{x}{\sigma_k}\right) \\
g_k(x) &= E\left[\left[(X_k - x)^+\right]^2\right] \\
&= \left(x^2 + \sigma_k^2\right)\Phi\left(-\frac{x}{\sigma_k}\right) - \frac{x\sigma_k}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2\sigma_k^2}\right).
\end{aligned}
$$

6.1. Bivariate case

We suppose that $d = 2$ and consider a bivariate Gaussian distribution with zero mean, $\sigma_1 = \sigma_2 = 1$ and correlation ρ in $\{-0.9, -0.5, -0.2, 0, 0.2, 0.5, 0.9\}$. Without systemic risk weight, that is, setting $\alpha = 0$ in the loss function (5.1), the results do not depend on the correlation value. See Table 3. Since $\sigma_1 = \sigma_2 = 1$, the allocation is symmetric. Hence, we only provide the first component of m^*. Table 3 shows that the analytical scheme is naturally the fastest. However, Fourier combined with Chebychev with 15 nodes provides an accuracy of 10^{-3} with a total computational time slightly less than pure Fourier. Indeed, since there are no bivariate integrals for $\alpha = 0$, there are 15 times 6 univariate integrals to compute for the Chebychev preprocessing, whereas Fourier needs 64 times 6 univariate integrals.

Table 4 shows the same results for different levels of correlation with the systemic factor in the case where $\alpha = 1$. As expected, the values of the risk allocation are increasing in ρ. For $\alpha = 1$, the presence of bivariate integrals makes the Fourier methods, with or without Chebyshev, considerably less efficient than the pure Monte Carlo method. However, since we are in low dimension, the pure Fourier method is more efficient than coupled with Chebyshev with 10 nodes. Indeed, d is definitely smaller than N in that case, therefore, requiring more computations of bivariate integrals in the preprocessing than during the

Scheme	m_1^*	$I(2)$	CTP	CTR	CT
Analytical	-0.173	59	0.00 s	0.06 s	0.06 s
Fourier	-0.173	64	0.00 s	1.03 s	1.03 s
Fourier + Chebychev 3 nodes	-0.140	84	0.14 s	0.08 s	0.22 s
Fourier + Chebychev 5 nodes	-0.158	66	0.24 s	0.07 s	0.31 s
Fourier + Chebychev 10 nodes	-0.163	56	0.40 s	0.09 s	0.49 s
Fourier + Chebychev 15 nodes	-0.173	49	0.57 s	0.09 s	0.66 s
Monte Carlo 1 Mio	-0.174	5	0.19 s	9.38 s	9.57 s
Monte Carlo 10 Mio	-0.173	5	1.74 s	73.3 s	75.04 s

Table 3: Bivariate case without systemic weight, that is, for $\alpha = 0$.

	Fourier					Fourier + Chebychev 10 nodes					Monte Carlo 2 Mio				
ρ	m_1^*	$I(2)$	CTP	CTR	CT	m_1^*	$I(2)$	CTP	CTR	CT	m_1^*	$I(2)$	CTP	CTR	CT
-0.9	-0.167	140	0.00 s	898.26 s	898.26 s	-0.151	59	2552.94 s	0.70 s	2553.64 s	-0.168	86	0.35 s	80.29 s	80.64 s
-0.5	-0.143	66	0.00 s	259.48 s	259.48 s	-0.131	60	1926.83 s	0.73 s	1927.56 s	-0.144	82	0.35 s	71.65 s	72.00 s
-0.2	-0.120	93	0.00 s	410.01 s	410.01 s	-0.113	84	1590.02 s	0.91 s	1590.93 s	-0.121	86	0.34 s	75.60 s	75.94 s
0	-0.103	72	0.00 s	250.33 s	250.33 s	-0.098	76	1635.34 s	0.79 s	1636.13 s	-0.103	75	0.46 s	66.83 s	67.29 s
0.2	-0.086	94	0.00 s	342.27 s	342.27 s	-0.082	94	1751.79 s	1.28 s	1753.07 s	-0.085	73	0.39 s	64.38 s	64.77 s
0.5	-0.057	124	0.00 s	559.06 s	559.06 s	-0.055	73	2154.19 s	0,74 s	2154.93 s	-0.057	103	0.36 s	90.55 s	90.91 s
0.9	-0.013	44	0.00 s	932.66 s	932.66 s	-0.013	84	2162.33 s	0,54 s	2162.87 s	-0.013	96	0.41 s	85.27 s	85.68 s

Table 4: Bivariate case with systemic weight, that is, for $\alpha = 1$.

iterations of the root finding.

6.2. Trivariate Case

In this section, we illustrate the systemic contribution of the loss function with three risk components and study the impact of the interdependence of two components with respect to the third one. We start with a Gaussian vector with the following variance-covariance matrix

$$\Sigma = \begin{bmatrix} 0.5 & 0.5\rho & 0 \\ 0.5\rho & 0.5 & 0 \\ 0 & 0 & 0.6 \end{bmatrix},$$

for different correlations $\rho \in \{-0.9, -0.5, -0.2, 0, 0.2, 0.5, 0.9\}$. Here the third risk component has a higher marginal risk than the first two so that, in the absence of systemic component, it should contribute most to the overall risk. When $\alpha = 0$, this is indeed the case. The result is independent of the correlation and is typically overall lower, charging the risk component with the highest variance more than the other two, as shown in Table 5.

However, with systemic risk weights, the contribution of the first two overcomes the third one for high correlation, as shown in Table 6. These results illustrate that the systemic risk weights corrects the risk allocation as the correlation between the first two risk components increases. Without Chebychev method, the Monte Carlo scheme in this trivariate case is radically faster than Fourier, from 10 times up to 40 times more efficient.

$m_1^* = m_2^*$		m_3^*	$R(X)$	$I(3)$	CTP	CTR	CT
-0.166	$\leq$	0.120	-0.212	148	0 s	8.05 s	8.05 s

Table 5: Trivariate case without systemic weight, that is $\alpha = 0$. Computed by Fourier.

	Fourier Method								Monte Carlo 2 Mio							
ρ	$m_1^* = m_2^*$		m_3^*	$R(X)$	$I(3)$	CTP	CTR	CT	$m_1^* = m_2^*$		m_3^*	$R(X)$	$I(3)$	CTP	CTR	CT
-0.9	-0.189	$\leq$	0.096	-0.282	157	0 s	2154.30 s	2154.30 s	-0.188	$\leq$	0.096	-0.281	98	0.43 s	156.62 s	157.05 s
-0.5	-0.135	$\leq$	0.017	-0.253	131	0 s	2286.76 s	2286.76 s	-0.135	$\leq$	0.018	-0.252	123	0.35 s	216.27 s	216.62 s
-0.2	-0.099	$\leq$	-0.030	-0.229	80	0 s	4711.94 s	4711.94 s	-0.100	$\leq$	-0.030	-0.230	97	0.36 s	154.86 s	155.22 s
0	-0.076	$\leq$	-0.059	-0.212	9	0 s	3062.09 s	3062.09 s	-0.078	$\leq$	-0.059	-0.215	89	0.37 s	140.80 s	141.17 s
0.2	-0.054	$\leq$	-0.087	-0.194	155	0 s	2277.04 s	2277.04 s	-0.054	$\leq$	-0.087	-0.194	107	0.37 s	175.17 s	175.54 s
0.5	-0.020	$\geq$	-0.125	-0.165	35	0 s	4678.19 s	4678.19 s	-0.022	$\geq$	-0.124	-0.168	139	0.34 s	217.90 s	218.24 s
0.9	0.026	$\geq$	-0.173	-0.122	202	0 s	6000.57 s	6000.57 s	0.026	$\geq$	-0.173	-0.122	103	0.34 s	164.17 s	164.51 s

Table 6: Trivariate case with systemic weight, that is $\alpha = 1$. Computed by Fourier.

6.3. Higher Dimensions

We performed the computations in a 10- as well as in a 30-variate case. In the 10-variate case without systemic weight, that is for $\alpha = 0$, we used the Fourier and Monte Carlo schemes, see Table 7. Since there is no systemic risk weight involved, the Fourier scheme outperforms slightly the Monte Carlo method. In the case with systemic weight, that is for $\alpha = 1$, we used the Monte Carlo scheme without and with

Scheme	m_1^*	m_2^*	m_3^*	m_4^*	m_5^*	m_6^*	m_7^*	m_8^*	m_9^*	m_{10}^*	$R(X)$	$l(10)$	CTP	CTR	CT
F	0.408	0.295	-0.049	-0.341	1.352	-0.193	-0.037	0.214	0.130	-0.047	1.733	1245	0 s	71.80 s	71.80 s
MC 1 Mio	0.410	0.298	-0.049	-0.345	1.344	-0.192	-0.034	0.214	0.131	-0.045	1.733	180	0.77 s	138.59 s	139.36 s
MC 10 Mio	0.410	0.295	-0.050	-0.344	1.348	-0.193	-0.034	0.213	0.132	-0.045	1.733	165	1.57 s	220.41 s	221.98 s

Table 7: 10-variate case without systemic weight, that is, for $\alpha = 0$.

Chebyshev, the Fourier scheme being too slow to run in a reasonable amount of time, see Table 8. The pure Monte Carlo scheme requires 236 steps times 55 sample summations for a total of 12980 sample summations, whereas the Monte Carlo together with Chebyshev with 15 nodes needed in preprocessing 15^2 times 55 sample summations for a total of 12375 sample summations, resulting in a roughly similar total computational time.

Scheme	m_1^*	m_2^*	m_3^*	m_4^*	m_5^*	m_6^*	m_7^*	m_8^*	m_9^*	m_{10}^*	$R(X)$	$l(10)$	CTP	CTR	CT
MC 2 Mio	0.344	0.254	0.260	0.113	1.324	0.180	0.430	0.703	0.584	0.448	4.639	236	1.58 s	6976.70 s	6978.28 s
MC 2 Mio + Cheby 10	0.182	0.106	0.272	0.211	1.188	0.198	0.529	0.776	0.596	0.477	4.536	120	3533.13 s	51.47 s	3584.60 s
MC 2 Mio + Cheby 15	0.373	0.297	0.310	0.213	1.298	0.251	0.432	0.690	0.583	0.447	4.896	158	6183.01 s	82.02 s	6265.03 s
MC 2 Mio + Cheby 20	0.312	0.229	0.266	0.179	1.315	0.193	0.439	0.698	0.572	0.430	4.632	127	6689.62 s	164.19 s	6853.81 s

Table 8: 10-variate case with systemic weight, that is, for $\alpha = 1$.

In the 30-variate case with systemic weight, that is for $\alpha = 1$, the Monte Carlo with 10^6 samples did not deliver a result in a reasonable amount of time – we stopped after more than three days. Using however Monte Carlo with 2×10^6 samples together with Chebyshev yields the computational results in

Table 9. Note moreover that the preprocessing in the Chebyshev case could be considerably reduced by making use of parallel processing.

Scheme	$l(10)$	CTP	CTR	CT
MC 2 Mio + Cheby 15 for $\alpha = 0$	216	1.87 s	439.15 s	441.02 s
MC 2 Mio + Cheby 15 for $\alpha = 1$	465	21925.33 s	3035.83 s	25426.16 s

Table 9: 30-variate case without and with systemic weight, that is, for $\alpha = 0, 1$.

Figures 3 and 4 show the variance-covariance matrix and the resulting risk allocation in the 10- and 30-variate case, respectively. The plots show that the risk allocation depends not only on the variance of the different risk components, but also, in the case where $\alpha = 1$, on the corresponding dependence stucture (in this case, the column of covariances). For instance, compare components 2 and 3 in the 10-variate case in Figure 3. In the first case we observe that when $\alpha = 0$, component 2 contributes more than 3, and conversely when $\alpha = 1$. The reason is that even if component 2 has a slightly higher variance, unlike member 3, it is negatively correlated with component 5 which has the highest variance, and thus is the most 'dangerous' from the systemic point of view. Hence, component 3 is more exposed than 2 in case of a systemic event. The same pattern can be observed between components 28 and 29 in the 30-variate case in Figure 4.

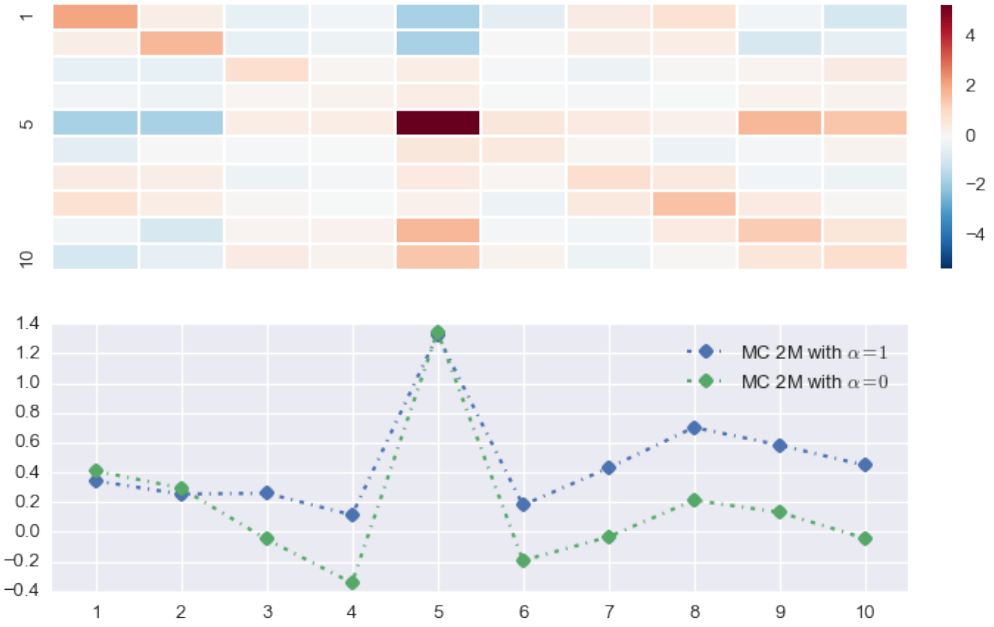

Figure 3: Plot showing the variance-covariance matrix together with the respective allocation in the 10-variate case for $\alpha = 0, 1$.

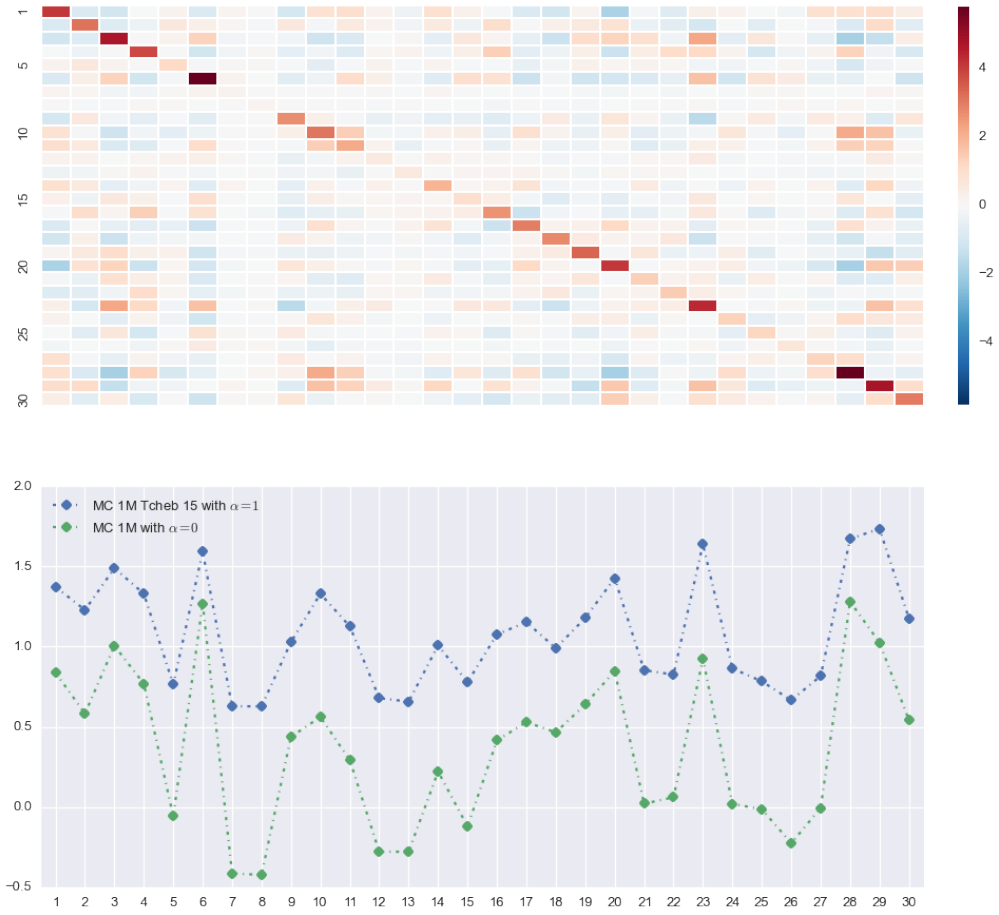

Figure 4: Plot showing the variance-covariance matrix together with the respective allocation in the 30-variate case for $\alpha = 0, 1$.

A. Some Classical Facts in Convex Optimization

For an extended real valued function f on a locally convex topological vector space X, its convex conjugate is defined as

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

where X^* is the topological dual of X. The Fenchel–Moreau theorem states that if f is lower semi-continuous, convex and proper, then so is f^*, and it holds

$$f(x) = f^{**}(x) = \sup_{x^* \in X^*} \{\langle x^*, x\rangle - f^*(x^*)\}, \quad x \in X.$$

Following Rockafellar [26], for any non-empty set $C \subseteq \mathbb{R}^d$, we define its *recession cone*

$$0^+C := \left\{y \in \mathbb{R}^d \colon x + \lambda y \in C \text{ for every } x \in C \text{ and } \lambda \in \mathbb{R}_+\right\}.$$

By [26, Theorem 8.3], if C is non-empty, closed and convex, then

$$0^+C = \left\{y \in \mathbb{R}^d \colon \text{ there exists } x \in C \text{ such that } x + \lambda y \in C \text{ for every } \lambda \in \mathbb{R}_+\right\}. \tag{A.1}$$

By [26, Theorem 8.4], a non-empty, closed and convex set C is compact if and only if $0^+C = \{0\}$.

Given a proper, convex and lower semi-continuous function f on $\mathbb{R}^d$, we call $y \in \mathbb{R}^d$ a *direction of recession* of f if there exists $x \in \mathrm{dom}(f)$ such that the map $\lambda \mapsto f(x + \lambda y)$ is decreasing on $\mathbb{R}_+$. We denote by $f0^+$ the *recession function* of f, that is, the function with epigraph given as the recession cone of the epigraph of f, and we call

$$0^+ f := \big\{ y \in \mathbb{R}^d \colon (f0^+)(y) \leq 0 \big\}$$

the *recession cone of* f. The following theorem gathers results from [26, Theorems 8.5, 8.6, 8.7 and Corollaries pp. 66–70].

Theorem A.1. *Let f be a proper, closed and convex function on $\mathbb{R}^d$.*

1. *Given x, y in $\mathbb{R}^d$, if $\liminf_{\lambda\to\infty} f(x + \lambda y) < \infty$, then $\lambda \mapsto f(x + \lambda y)$ is decreasing.*
2. *All the non-empty level sets $B := \{x \in \mathbb{R}^d \colon f(x) \leq \gamma\} \neq \emptyset$ of f have the same recession cone, namely the recession cone of f. That is:*

   $$0^+ f = 0^+ B, \text{ for every } \gamma \in \mathbb{R} \text{ such that } B \neq \emptyset.$$

3. *$f0^+$ is a positively homogeneous, proper, closed and convex function, such that*

   $$(f0^+)(y) = \sup_{\lambda>0} \frac{f(x + \lambda y) - f(x)}{\lambda} = \lim_{\lambda\to\infty} \frac{f(x + \lambda y) - f(x)}{\lambda}, \quad y \in \mathbb{R}^d,$$

   for every $x \in \mathrm{dom}(f)$.
4. *There exists $x \in \mathrm{dom}(f)$ such that the map $\lambda \mapsto f(x + \lambda y)$ is decreasing on $\mathbb{R}_+$, that is, y is a direction of recession of f, if and only if this map is decreasing for every $x \in \mathrm{dom}(f)$, which in turn is equivalent to $(f0^+)(y) \leq 0$.*
5. *The map $\lambda \mapsto f(x + \lambda y)$ is constant on $\mathbb{R}_+$ for every $x \in \mathrm{dom}(f)$ if and only if $(f0^+)(y) \leq 0$ and $(f0^+)(-y) \leq 0$.*

B. Multivariate Orlicz Spaces

In this appendix we briefly sketch how the classical theory of univariate Orlicz spaces carries over to the d-variate case without any significant change. We follow the lecture notes by Léonard [25], only providing the proofs that differ structurally from the univariate case.

A function $\theta : \mathbb{R}^d \to [0, \infty]$ is called a Young function if it is

- convex and lower semi-continuous;
- such that $\theta(x) = \theta(|x|)$ and $\theta(0) = 0$;
- non trivial, that is, $\mathrm{dom}(\theta)$ contains a neighborhood of 0 and $\theta(x) \geq a \|x\| - b$ for some $a > 0$.

In particular, θ achieves its minimum at 0 and is increasing on $\mathbb{R}^d_+$. It is said to be finite if $\mathrm{dom}(\theta) = \mathbb{R}^d$ and strict if $\lim_{x\to\infty} \theta(x)/\|x\| = \infty$.

Lemma B.1. *The function θ is Young if and only if θ^* is Young. Furthermore, θ is strict if and only if θ^* is strict if and only if θ and θ^* are both finite.*

Proof. This follows by application of the Fenchel-Moreau theorem and from the relation $x \cdot y \leq \theta(x) + \theta^*(y)$. □

For $X \in L^0$, the Luxembourg norm of X is given as

$$\|X\|_\theta = \inf \{\lambda \in \mathbb{R} \colon \lambda > 0 \text{ and } E[\theta(X/\lambda)] \leq 1\},$$

where $\inf \emptyset = \infty$. The Orlicz space and heart are respectively defined as

$$L^\theta := \left\{X \in L^0 \colon \|X\|_\theta < \infty\right\} = \left\{X \in L^0 \colon E[\theta(X/\lambda)] < \infty \text{ for some } \lambda \in \mathbb{R}, \lambda > 0\right\}$$
$$M^\theta := \left\{X \in L^0 \colon E[\theta(X/\lambda)] < \infty \text{ for all } \lambda \in \mathbb{R}, \lambda > 0\right\}.$$

Lemma B.2. *1. We have $\|X\|_\theta = 0$ if and only if $X = 0$.*

2. *If $0 < \|X\|_\theta < \infty$, then $E[\theta(X/\|X\|_\theta)] \leq 1$. In particular, $B := \{X \colon \|X\|_\theta \leq 1\} = \{X \colon E[\theta(X)] \leq 1\}$.*

3. *The gauge $\|\cdot\|_\theta$ is a norm both on the Orlicz space L^θ and on the Orlicz heart M^θ.*

4. *The following Hölder Inequality holds:*

$$E[|X \cdot Y|] \leq \|X\|_\theta \|Y\|_{\theta^*}.$$

5. *L^θ is continuously embedded into L^1, the space of integrable random variables on $\Omega \times \{1, \ldots, d\}$ for the product measure $P \otimes Unif_{\{1,\ldots,d\}}$.* [9]

6. *The normed spaces $(L^\theta, \|\cdot\|_\theta)$ and $(M^\theta, \|\cdot\|_\theta)$ are Banach spaces.*

Proof. These results can be established along the same lines as in the univariate case [See 25, Lemmas 1.8 and 1.10 and Propositions 1.11, 1.14, 1.15 and 1.18], using the Fenchel-Moreau Theorem in $\mathbb{R}^d_+$. □

Theorem B.3. *If θ is finite, then the topological dual of M^θ is L^{θ^*}.*

Proof. Again, the proof follows the univariate case [see 25, Proposition 1.20, Theorem 2.2 and Lemmas 2.4 and 2.5]. □

C. Numerical Methods

C.1. Fourier Transform

Let ℓ be a loss function and denote by ℓ_η the *dampened* loss function, defined by $\ell_\eta(x) := e^{-\langle \eta, x\rangle}\ell(x)$, for $\eta \in \mathbb{R}^d$. Moreover, let $\widehat{\ell}$ denote the Fourier transform of the function ℓ, that is, $\widehat{\ell}(u) = \int_{\mathbb{R}^d} e^{i\langle u,x\rangle}\ell(x)dx$, and let M_X denote the (extended) moment generating function of the random variable X, that is, $M_X(u) = E[e^{\langle u,X\rangle}]$, for suitable $u \in \mathbb{C}^d$. Let L^1, respectively L^1_{bc}, denote the set of measurable functions on $\mathbb{R}^d$ which are integrable, respectively bounded, continuous and integrable, with respect to the Lebesgue measure. We also denote by $\Im(z)$ the imaginary part of the complex number z.

[9] The case where $L^\theta = L^1$ corresponds to $\theta(x) = \sum |x_k|$.

Consider the following sets:

$$\mathcal{I} := \left\{\eta \in \mathbb{R}^d \colon M_X(\eta) < \infty\right\}, \qquad \mathcal{I}' := \left\{\eta \in \mathbb{R}^d \colon M_X(\eta) < \infty \text{ and } M_X(\eta + i\cdot) \in L^1\right\},$$
$$\mathcal{J} := \left\{\eta \in \mathbb{R}^d \colon \ell_\eta \in L^1_{\text{bc}} \text{ and } \widehat{\ell}_\eta \in L^1\right\}, \quad \mathcal{J}' := \left\{\eta \in \mathbb{R}^d \colon \ell_\eta \in L^1\right\}.$$

Assuming that $\mathcal{I} \cap \mathcal{J} \neq \emptyset$ or $\mathcal{I}' \cap \mathcal{J}' \neq \emptyset$ and using Eberlein et al. [15, Theorems 2.2 and 3.2], we have that

$$E\big[\ell(X - m)\big] = \frac{1}{(2\pi)^d} \int_{\mathbb{R}^d} e^{\langle iu - \eta, m\rangle} M_X(\eta - iu)\widehat{\ell}(u + i\eta)du, \tag{C.1}$$

where $\eta \in \mathcal{I} \cap \mathcal{J}$ or $\eta \in \mathcal{I}' \cap \mathcal{J}'$.

Proof (of Proposition 5.1). Starting from (5.2), we define the functions

$$\ell_1(x_k) = x_k^+, \qquad \ell_2(x_k) = (x_k^+)^2, \qquad \ell_3(x_j, x_k) = x_j^+ 1_{\{x_k \geq 0\}} \qquad \ell_4(x_j, x_k) = x_j^+ x_k^+,$$

whose Fourier transforms are provided by

$$\widehat{\ell}_1(z_k) = -\frac{1}{z_k^2}, \qquad \text{for } z_k \in \mathbb{C} \text{ with } \Im(z_k) > 0; \tag{C.2}$$
$$\widehat{\ell}_2(z_k) = \frac{2}{iz_k^3}, \qquad \text{for } z_k \in \mathbb{C} \text{ with } \Im(z_k) > 0; \tag{C.3}$$
$$\widehat{\ell}_3(z_j, z_k) = \frac{1}{iz_k z_j^2}, \qquad \text{for } z_j, z_k \in \mathbb{C} \text{ with } \Im(z_j), \Im(z_k) > 0; \tag{C.4}$$
$$\widehat{\ell}_4(z_j, z_k) = \frac{1}{z_j^2 z_k^2}, \qquad \text{for } z_j, z_k \in \mathbb{C} \text{ with } \Im(z_j), \Im(z_k) > 0. \tag{C.5}$$

The result now follows by combining (C.1) with (5.4) and (C.2)–(C.5), and observing that the analogue of assumption $\mathcal{I}' \cap \mathcal{J}' \neq \emptyset$ is satisfied in each case due to (5.4) and (C.2)–(C.5). □

C.2. Chebyshev Interpolation

We can use the Chebyshev interpolation method in order to compute values of $f_k(m)$, $g_k(m)$, $h_{jk}(m, n)$ and $l_{jk}(m, n)$ for a large number of m's and n's, as follows:

Step 1: Select upper and lower bounds for the parameter m, that is, $m \in [\underline{m}, \overline{m}]$, and rescale this by a linear interpolation such that $m \in [-1, 1]$. Then, compute the Chebyshev nodes via

$$m^i = \cos\left(\pi \frac{2i + 1}{2N + 1}\right), \quad i = 1, \ldots, N, \tag{C.6}$$

where N denotes the degree of the Chebyshev interpolation. Next, compute the values $f_k(m^i)$ and $g_k(m^i)$, at each node $i \in \{1, \ldots, N\}$ and for each risk factor $k \in \{1, \ldots, d\}$, and store them. This operation can be performed in parallel both in the nodes i and in the factors k.

Step 2: Compute approximate values of $f_k(m)$ and $g_k(m)$ for some arbitrary $m \in [-1, 1]$, or $m \in [\underline{m}, \overline{m}]$, using the Chebyshev interpolation:

$$I_N[f_k](m) = \sum_{i=0}^{N} c_i T_i(m),$$

with weights

$$c_i = \frac{2^{1_{\{i>0\}}}}{N+1} \sum_{r=0}^{N} f_k(m^r) \cos\left(i\pi \frac{2r+1}{2N+1}\right), \quad 0 \le i \le N,$$

and basis functions

$$T_i(m) = \cos(i \arccos(m)), \qquad 0 \le i \le N.$$

The interpolation function $I_N[g_k]$ for g_k is defined analogously.

Step 3: In case of the bivariate functions $h_{jk}(m,n)$ and $l_{jk}(m,n)$, the Chebyshev interpolation takes the following form:

$$I_{\bar{N}}[h_{jk}](m,n) = \sum_{i_1=0}^{N_1} \sum_{i_2=0}^{N_2} c_{i_1,i_2} T_{i_1,i_2}(m,n),$$

where $\bar{N} = (N_1, N_2)$, with weights

$$c_{i_1,i_2} = \prod_{j=1}^{2} \frac{2^{1_{\{i_j>0\}}}}{N_j+1} \sum_{r_1=0}^{N_1} \sum_{r_2=0}^{N_2} h_{jk}(m^{r_1}, n^{r_2}) \prod_{k=1}^{2} \cos\left(i_k \pi \frac{2r_k+1}{2N_k+1}\right),$$

for all $0 \le i_1 \le N_1$, $0 \le i_2 \le N_2$, and basis functions

$$T_{i_1,i_2}(m,n) = T_{i_1}(m) T_{i_2}(n).$$

Here, the Chebyshev nodes m^{i_1}, n^{i_2} are selected according to (C.6) for each of them, while the values of $h_{jk}(m^{i_1}, n^{i_2})$ are computed over the tensorized parameter space $(i_1, i_2) \in \{1, \ldots, N_1\} \times \{1, \ldots, N_2\}$. The interpolation function $I_{\bar{N}}[l_{jk}]$ for l_{jk} is defined analogously.

References

[1] V. Acharya, T. P. L. Pedersen, and M. Richardson. Measuring systemic risk. *SSRN: 1573171*, 2010.

[2] V. Acharya, R. Engle, and M. Richardson. Capital shortfall: A new approach to ranking and regulating systemic risks. *American Economic Review: Papers & Proceedings*, 102(3):59–64, 2012.

[3] T. Adrian and M. Brunnermeier. CoVaR. *National Bureau of Economic Research Working Paper*, 1745, 2011.

[4] P. Artzner, F. Delbaen, J. M. Eber, and D. Heath. Coherent measures of risk. *Mathematical Finance*, 9:203–228, 1999.

[5] A. Ben-Tal and M. Teboulle. An old-new concept of convex risk measures: the optimized certainty equivalent. *Mathematical Finance*, 17(3):449–476, 2007.

[6] F. Biagini, J.-P. Fouque, M. Frittelli, and T. Meyer-Brandis. A unified approach to systemic risk measures via acceptance sets. *ArXiV:1503.06354*, 2015.

[7] C. Brownlees and R. Engle. Volatility, correlation and tails for systemic risk measurement. *SSRN: 1611229*, 2012.

[8] M. K. Brunnemeier and P. Cheridito. Measuring and allocating systemic risk. *SSRN: 2372472*, 2014.

[9] I. Cascos and I. Molchanov. Multivariate risk measures: a constructive approach based on selections. *Mathematical Finance*, 2014. Forthcoming.

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

[11] P. Cheridito and T. Li. Risk measures on Orlicz hearts. *Mathematical Finance*, 19(2):189–214, 2009.

[12] R. Cont, E. Santos, and A. Moussa. Network structure and systemic risk in banking systems. In J.-P. Fouque and J. Langsam, editors, *Handbook of Systemic Risk*. Cambridge University Press, 2013.

[13] S. Drapeau and M. Kupper. Risk preferences and their robust representation. *Mathematics of Operations Research*, 28(1):28–62, 2013.

[14] S. Drapeau, M. Kupper, and A. Papapantoleon. A Fourier approach to the computation of CV@R and optimized certainty equivalents. *Journal of Risk*, 16(6):3–29, 2014.

[15] E. Eberlein, K. Glau, and A. Papapantoleon. Analysis of Fourier transform valuation formulas and applications. *Applied Mathematical Finance*, 17:211–240, 2010.

[16] W. Farkas and P. Koch-Medina. Measuring risk with multiple eligible assets. *Mathematics and Financial Economics*, 9(1):3–27, 2015.

[17] Z. Feinstein, B. Rudlof, and S. Weber. Measures of systemic risk. *ArXiV:1502.07961*, 2015.

[18] A. Fiacco and G. McCormick. *Nonlinear Programming: Sequential Unconstrained Minimization Techniques*. Classics in Applied Mathematics. Society for Industrial and Applied Mathematics, 1990. ISBN 9780898712544.

[19] H. Föllmer and A. Schied. Convex measures of risk and trading constraint. *Finance and Stochastics*, 6(4): 429–447, 2002.

[20] M. Frittelli and E. Rosazza Gianin. Putting order in risk measures. *Journal of Banking & Finance*, 26(7): 1473–1486, July 2002.

[21] M. Gaß, K. Glau, M. Mahlstedt, and M. Mair. Chebyshev interpolation for parametric option pricing. *ArXiV:1505.04648*, 2015.

[22] D. M. Gay. Some convergence properties of Broyden's method. *SIAM Journal on Numerical Analysis*, 16(4): pp. 623–630, 1979.

[23] A. Hamel, F. Heyde, and B. Rudloff. Set-valued risk measures for conical market models. *Mathematics and Financial Economics*, 5(1):1–28, 2011.

[24] E. Jouini, M. Meddeb, and N. Touzi. Vector-valued coherent risk measures. *Finance and Stochastics*, 8: 531–552, 2004.

[25] C. Léonard. Some notes on Orlicz spaces. 2007. URL http://www.cmap.polytechnique.fr/~leonard/papers/orlicz.pdf.

[26] R. T. Rockafellar. *Convex Analysis*. Princeton Mathematical Series, No. 28. Princeton University Press, Princeton, N.J., 1970.

[27] R. T. Rockafellar and R. J.-B. Wets. *Variational Analysis*. Springer, Berlin, New York, 3rd edition, 2009.