

Dual representations of quasiconvex compositions with applications to systemic risk*

Çağın Ararat[†] Mücahit Aygün[‡]

August 29, 2021

Abstract

Motivated by the problem of finding dual representations for quasiconvex systemic risk measures in financial mathematics, we study quasiconvex compositions in an abstract infinite-dimensional setting. We calculate an explicit formula for the penalty function of the composition in terms of the penalty functions of the ingredient functions. The proof makes use of a nonstandard minimax inequality (rather than equality as in the standard case) that is available in the literature. In the second part of the paper, we apply our results in concrete probabilistic settings for systemic risk measures, in particular, in the context of Eisenberg-Noe clearing model. We also provide novel economic interpretations of the dual representations in these settings.

Keywords and phrases: quasiconvex function, penalty function, composition of functions, minimax inequality, systemic risk measure, dual representation

Mathematics Subject Classification (2020): 46N10, 91G45, 46A20, 52A01

1 Introduction

This paper is concerned with extended real-valued functions of the form $f \circ g$, where f and g are functions defined on some general preordered topological vector spaces. We look for minimal assumptions on f and g to ensure that their composition $f \circ g$ is a monotone, quasiconvex, and

*This work is supported by TÜBİTAK (Scientific & Technological Research Council of Turkey), Project No. 117F438.

[†]Bilkent University, Department of Industrial Engineering, Ankara, Turkey, cararat@bilkent.edu.tr.

[‡]University of Amsterdam, Department of Quantitative Economics, Amsterdam, The Netherlands, m.aygun@uva.nl.

lower semicontinuous function. In our main results, we provide novel duality formulae in which the dual function for $f \circ g$ is calculated in terms of the same type of functions for f and g .

In the literature, the study of $f \circ g$ from a duality point of view is not new in the convex case. For a single function, Fenchel-Moreau theorem provides a dual representation for a convex lower semicontinuous function in terms of its Legendre-Fenchel conjugate (Rockafellar [28, Thm. 12.2]). Then, it is natural to ask how and when we can have a dual representation for the composition of convex functions. This question has been answered in the literature, for instance, by Zălinescu [32, Thm. 2.8.10], Boţ et al. [6, Thm. 3]; see also the more recent work Burke et al. [7].

As a natural extension of the convex case, we look for dual representations of $f \circ g$ when it is guaranteed to be quasiconvex. This is an open problem to the best of our knowledge. For a single function, the quasiconvex duality theory in Penot and Volle [26] provides a suitable replacement of conjugate functions in convex duality. This is further explored in Cerreia-Vioglio et al. [9] within an abstract framework and also in Cerreia-Vioglio et al. [8], Drapeau and Kupper [12], Frittelli and Maggis [16] within the context of risk measures. In line with Drapeau and Kupper [12], the dual functions for quasiconvex duality will be referred to as *penalty functions* in this paper.

Our motivation for studying quasiconvex compositions also comes from financial mathematics, specifically, from the theory of systemic risk measures as we describe briefly next.

Initiated by Artzner et al. [4], risk measures have been studied extensively in the financial mathematics and operations research literature. In the original framework of Artzner et al. [4], *coherent risk measures* are defined as monotone, convex, translative and positively homogeneous functionals defined on a space of real-valued random variables. These random variables could be used to model the uncertain future worths of investments, and a risk measure assigns to each random variable its minimum deterministic capital requirement. Among the properties of coherent risk measures, *monotonicity* is a natural requirement which asserts that the risk of an investment with consistently higher future values should be lower. *Convexity* is related to diversification; under this property, the risk of a mixture, that is, convex combination, of two portfolios is not higher than the same type of mixture of the individual risks. *Positive homogeneity* is a scaling property that is relaxed for defining *convex risk measures* in Föllmer and Schied [15]. Finally, *translativity* asserts that a deterministic increase in the value of a portfolio decreases its risk by the same amount. This is indeed the property that justifies the interpretation of risk measure as capital requirement.

One might question whether convexity provides the correct encoding of the impact of diversification on risk. A weaker alternative is *quasiconvexity*, which bounds the risk of a mixture only by the maximum of the individual risks, hence the statement “Diversification does not increase risk.” is reflected properly. Under translativity, convexity is equivalent to quasiconvexity. Hence, the switch from convexity to quasiconvexity implies working with non-translative functionals in general. Indeed, the work Drapeau and Kupper [12] proposes a minimalist framework for risk measures in which only monotonicity and quasiconvexity are taken for granted, such functionals are called *quasiconvex risk measures*; see also Frittelli and Maggis [16]. For the use of quasiconvex risk measures in the context of financial optimization problems; see Mastrogiamomo and Rosazza Gianin [25], Källblad [22], Ararat [2].

The theory of risk measures outlined above is for univariate, that is, real-valued, random variables. In more complex settings such as markets with transaction costs (Hamel and Heyde [20], Hamel et al. [21]) and financial networks with interdependencies (Chen et al. [10], Feinstein et al. [14], Biagini et al. [5], Ararat and Rudloff [3]), it becomes necessary to evaluate the risks of random vectors. In this paper, we are particularly interested in the latter situation where the participating financial institutions are subject to correlated sources of risk, typically affecting the future values of their assets. Hence, the resulting future values are naturally modeled as correlated random vectors, explaining the multivariate nature of the problem. At the same time, the institutions form a network through mutual obligations and the aforementioned uncertainty affects the ability of the institutions to meet these obligations. Hence, the aim of a *systemic risk measure* is to quantify the overall risk associated to the financial network.

In the pioneering work Chen et al. [10], a systemic risk measure R is defined as the composition of a univariate risk measure ρ with a so-called *aggregation function* Λ , that is, $R = \rho \circ \Lambda$. The role of the aggregation function is to summarize the impact of the random shock vector X , on the economy (or society) as a scalar random quantity $\Lambda(X)$. The definition of Λ is made precise by the structure of the network and the accompanying clearing mechanism. For instance, one can consider a clearing system in the framework of Eisenberg and Noe [13] and define the aggregation function as the total payment made to society as in Ararat and Rudloff [3], in which case Λ is an increasing concave function. The output of Λ is further given as input to a convex risk measure ρ to calculate the value of $R(X)$. The resulting systemic risk measure R is a monotone convex functional that is

not translative in general. In Ararat and Rudloff [3], dual representations for *convex* systemic risk measures are studied in detail. The mathematical machinery used in that work is the conjugation formula in Zălinescu [32, Thm. 2.8.10] and Boț et al. [6, Thm. 3] for convex compositions.

When ρ is only assumed to be a quasiconvex risk measure, the resulting systemic risk measure R is also quasiconvex. Providing dual representations for this case is the starting point of this paper. However, we will first study the problem in greater generality. As stated at the beginning, we will explore the dual representation of a quasiconvex composition $f \circ g$, where the ingredients f, g are defined on general preordered topological vector spaces. To the best of our knowledge, the quasiconvex analogues of the conjugation results in Zălinescu [32] and Boț et al. [6] are not known in the literature. We provide a solution to this problem by proving a formula for the penalty function of $f \circ g$, roughly speaking, in terms of the penalty functions of f and g . More precisely, apart from the more technical continuity conditions, we will assume that f is an extended real-valued monotone, quasiconvex function. Since g is a vector-valued function (in a possibly infinite-dimensional space), choosing the right notion of quasiconvexity requires extra care. To this end, we will use the notion of *natural quasiconvexity*, which is introduced for vector-valued functions in Tanaka [31] and for set-valued functions in Kuroiwa [23]. When g is a monotone, naturally quasiconcave function, the resulting composition $f \circ g$ is a monotone, quasiconvex function.

For the proof of our main duality theorem (Theorem 4.6), we need a nonstandard minimax result since the assumptions of the standard minimax theorem in Sion [30] may not hold in our case. We are able to overcome this issue by using the minimax inequality in Liu [24] (see also Greco and Moschen [18], Cheng and Lin [11]), which works under weaker conditions. With additional arguments that use the properties of the involved functions, we are able to turn the inequality into an equality. Hence, the proof of the main theorem makes novel use of minimax theory.

After building the general theory, we go back to our motivating problem on systemic risk measures. Using a quasiconvex univariate risk measure ρ and a concave aggregation function Λ , we are able to provide a dual representation for the systemic risk measure $R = \rho \circ \Lambda$ in a probabilistic framework. We also discuss the economic interpretations of the dual variables and penalty functions in terms of the underlying financial network.

The rest of this paper is organized as follows. In Section 2, we review some basic notions and results about convex and quasiconvex functions. Section 3 is dedicated to some more technical

notions for vector-valued functions: natural quasiconvexity, regular monotonicity, and lower demi-continuity. In Section 4, we prove the main theorem on quasiconvex compositions together with some important special cases. This is followed by Section 5, where we discuss the validity of a compactness assumption in concrete settings. In Section 6, we apply the theory to obtain dual representations for systemic risk measures. Among the various examples that we study, Eisenberg-Noe model is discussed separately as it has a more sophisticated aggregation function. Some proofs of the results in Sections 4 and 6 are collected in Section 7, the appendix.

2 Convex and quasiconvex functions

2.1 Preliminaries

We begin with some basic notations and definitions that are used throughout the paper. We denote by $\overline{\mathbb{R}} := \mathbb{R} \cup \{+\infty, -\infty\}$ the extended real line. Given $a, b \in \overline{\mathbb{R}}$, we define $a \vee b := \max\{a, b\}$, $a \wedge b := \min\{a, b\}$. For each $n \in \mathbb{N} := \{1, 2, \dots\}$, we denote by \mathbb{R}^n the n -dimensional Euclidean space, by \mathbb{R}_+^n the set of all $z = (z_1, \dots, z_n)^\top \in \mathbb{R}^n$ with $z_i \geq 0$ for each $i \in \{1, \dots, n\}$, and by \mathbb{R}_{++}^n the set of all $z \in \mathbb{R}^n$ with $z_i > 0$ for each $i \in \{1, \dots, n\}$. We write $\mathbb{R}_+ = \mathbb{R}_+^1$ and $\mathbb{R}_{++} = \mathbb{R}_{++}^1$.

Let \mathcal{X} be a Hausdorff locally convex topological vector space. For a set $A \subseteq \mathcal{X}$, $\text{cl}(A)$ and $\text{conv}(A)$ denote the closure and convex hull of A , respectively. We denote by \mathcal{X}^* the topological dual space of \mathcal{X} , endowed with the weak* topology $\sigma(\mathcal{X}^*, \mathcal{X})$. The bilinear duality mapping on $\mathcal{X}^* \times \mathcal{X}$ is denoted by $\langle \cdot, \cdot \rangle$. For nonempty sets $A, B \subseteq \mathcal{X}$ and $\lambda \in \mathbb{R}$, we define the sum $A + B := \{x + y \mid x \in A, y \in B\}$ and the product $\lambda A := \{\lambda x \mid x \in A\}$ in the Minkowski sense. When $A = \{x\}$ for some $x \in \mathcal{X}$, we write $x + B := \{x\} + B$.

Throughout this section, let $f: \mathcal{X} \rightarrow \overline{\mathbb{R}}$ be a function. Given $m \in \mathbb{R}$, the m -sublevel set of f is defined as

$$S_m^f := \{x \in \mathcal{X} \mid f(x) \leq m\}.$$

A straightforward calculation yields that f can be recovered from its sublevel sets via

$$f(x) = \inf\{m \in \mathbb{R} \mid x \in S_m^f\}, \quad x \in \mathcal{X}. \quad (2.1)$$

The function f is called *positively homogeneous* if $f(\lambda x) = \lambda f(x)$ for every $\lambda > 0$ and $x \in \mathcal{X}$. It is

called *proper* if $f(x) > -\infty$ for every $x \in \mathcal{X}$ and $f(x) < +\infty$ for at least one $x \in \mathcal{X}$. The *conjugate function* or the *Legendre-Fenchel transform* $f^*: \mathcal{X}^* \rightarrow \overline{\mathbb{R}}$ of f is defined by

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

As an important special case, we may take $f = I_A$ for some $A \subseteq \mathcal{X}$, where I_A is the (convex analytic) *indicator function* of A defined by $I_A(x) := 0$ if $x \in A$, and by $I_A(x) = +\infty$ if $x \in \mathcal{X} \setminus A$. Then, the conjugate function of I_A is the *support function* of A given by

$$I_A^*(x^*) = \sup_{x \in A} \langle x^*, x \rangle, \quad x^* \in \mathcal{X}^*. \quad (2.2)$$

Definition 2.1. (i) The function f is called *quasiconvex* if $f(\lambda x + (1 - \lambda)y) \leq f(x) \vee f(y)$ for every $x, y \in \mathcal{X}$ and $\lambda \in [0, 1]$. It is called *quasiconcave* if $-f$ is *quasiconvex*.

(ii) Let $x \in \mathcal{X}$. The function f is called *lower semicontinuous* at x if $f(x) \leq \liminf_{i \in I} f(x_i)$ whenever $(x_i)_{i \in I}$ is a net in \mathcal{X} that converges to x . It is called *lower semicontinuous* if it is *lower semicontinuous* at each $x \in \mathcal{X}$. It is called *upper semicontinuous* (at x) if $-f$ is *lower semicontinuous* (at x).

Remark 2.2. It is well-known that f is quasiconvex if and only if the sublevel set S_m^f is convex for every $m \in \mathbb{R}$ (Zălinescu [32, Sect. 2.1, p. 41]), and f is lower semicontinuous if and only if S_m^f is closed for every $m \in \mathbb{R}$ (Aliprantis and Border [1, Lemma 2.39]). Moreover, every closed convex strict subset of \mathcal{X} can be written as the intersection of all closed halfspaces that contain it (Aliprantis and Border [1, Cor. 5.83]). Thus, when f is proper, lower semicontinuous and quasiconvex, S_m^f can be written as an intersection of closed halfspaces for each $m \in \mathbb{R}$.

2.2 The order structure

To be able to handle monotone functions, e.g., in the risk measure applications in Section 6, we introduce an order structure on \mathcal{X} . To that end, let $C \subseteq \mathcal{X}$ be a convex cone and define a relation \leq_C on \mathcal{X} by

$$x \leq_C y \quad \Leftrightarrow \quad y - x \in C \quad (2.3)$$

for each $x, y \in \mathcal{X}$. It follows that \leq_C is a *vector preorder*, that is, $x \leq_C y$ implies $x + z \leq_C y + z$ and $\lambda x \leq_C \lambda y$ for every $x, y, z \in \mathcal{X}$ and $\lambda > 0$.

Remark 2.3. It can be checked that every vector preorder \preceq on \mathcal{X} can be written as $\preceq = \leq_C$, where $C := \{x \in \mathcal{X} \mid 0 \preceq x\}$ is a convex cone. Hence, the assumption that C is a convex cone is not a restriction on the vector preorder of interest.

The elements of C are called *positive elements* of \mathcal{X} . We define the (*positive*) *dual cone* of C by

$$C^+ := \{x^* \in \mathcal{X}^* \mid \forall x \in C: \langle x^*, x \rangle \geq 0\},$$

which is a closed convex cone in \mathcal{X}^* . Then, we define the cone of *strictly positive elements* of \mathcal{X} by

$$C^\# = \{x \in C \mid \forall x^* \in C^+ \setminus \{0\}: \langle x^*, x \rangle > 0\}. \quad (2.4)$$

Finally, given $\pi \in C^\#$, we may scale the elements of C^+ and obtain the closed convex set

$$C_\pi^+ := \{x^* \in C^+ \mid \langle x^*, \pi \rangle = 1\}.$$

Remark 2.4. When \mathcal{X} is finite-dimensional, $C^\#$ coincides with the interior of C . However, in the infinite-dimensional setting, we prefer working with $C^\#$ since the interior of C can be empty for many important examples including Lebesgue spaces; see Glück and Weber [17, Ex. 2.12].

The next lemma shows that C^+ can be recovered from the (much) smaller set C_π^+ if $\pi \in C^\#$.

Lemma 2.5. *Assume that $C^\# \neq \emptyset$ and let $\pi \in C^\#$. Then, we have $C^+ \setminus \{0\} = \mathbb{R}_{++} C_\pi^+$.*

Proof. Let $\lambda > 0$ and $x^* \in C_\pi^+$. By definition, $C_\pi^+ \subseteq C^+ \setminus \{0\}$ and $C^+ \setminus \{0\}$ is a cone so that $\lambda x^* \in C^+ \setminus \{0\}$. Hence, $\mathbb{R}_{++} C_\pi^+ \subseteq C^+ \setminus \{0\}$. Conversely, let $x^* \in C^+ \setminus \{0\}$. We have $\langle x^*, \pi \rangle > 0$. Taking $z^* := \frac{x^*}{\langle x^*, \pi \rangle}$, we have $\langle z^*, \pi \rangle = 1$, which implies that $z^* \in C_\pi^+$. Moreover, taking $\lambda = \frac{1}{\langle x^*, \pi \rangle} > 0$, we have $x^* = \lambda z^* \in \mathbb{R}_{++} C_\pi^+$. Hence, $C^+ \setminus \{0\} \subseteq \mathbb{R}_{++} C_\pi^+$. \square

Thanks to the order structure provided by \leq_C , we may define the monotonicity of sets and functions. We say that a set $A \subseteq \mathcal{X}$ is *monotone* if $x \leq_C y$ and $x \in A$ imply $y \in A$, for every $x, y \in \mathcal{X}$. Similarly, we say that f is a *decreasing* function (with respect to C) if $x \leq_C y$ implies

$f(x) \geq f(y)$ for every $x, y \in \mathcal{X}$; we say that f is an *increasing* function (with respect to C) if it is decreasing with respect to $-C$.

Remark 2.6. It is easy to check that f is decreasing if and only if its sublevel sets are monotone.

2.3 Dual representations

In convex analysis, Fenchel-Moreau theorem provides a dual representation for a proper lower semicontinuous convex function f in terms of its conjugate function f^* :

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

One immediate consequence of this theorem is that a set $A \subseteq \mathcal{X}$ and its closed convex hull have the same support function, that is,

$$I_A^*(x^*) = I_{\text{cl}(\text{conv}(A))}^*(x^*), \quad x^* \in \mathcal{X}^*. \quad (2.5)$$

We will use (2.5) in the proof of Proposition 7.3, which is a significant tool for proving Theorem 4.6, the main theorem of the paper.

For monotone functions, the following refinement of Fenchel-Moreau theorem is possible.

Proposition 2.7. *Suppose that f is proper, decreasing, convex and lower semicontinuous. Then,*

$$f(x) = \sup_{x^* \in C^+} (\langle -x^*, x \rangle - f^*(-x^*)), \quad x \in \mathcal{X}. \quad (2.6)$$

Proof. We first prove that $f^*(x^*) = +\infty$ when $x^* \notin -C^+$. Note that, in this case, there exists $c \in C$ such that $\langle x^*, c \rangle > 0$. Let $x_0 \in \text{dom } f$ and $\lambda > 0$. Since f is decreasing, we have $f(x_0 + \lambda c) \leq f(x_0)$ so that $\langle x^*, x_0 + \lambda c \rangle - f(x_0 + \lambda c) \geq \lambda \langle x^*, c \rangle + \langle x^*, x_0 \rangle - f(x_0)$. Since $\langle x^*, c \rangle > 0$, letting $\lambda \rightarrow \infty$ implies that

$$f^*(x^*) \geq \sup_{\lambda > 0} (\langle x^*, x_0 + \lambda c \rangle - f(x_0 + \lambda c)) \geq +\infty.$$

Hence, $f^*(x^*) = +\infty$. Combining this with Fenchel-Moreau theorem yields (2.6). \square

For a quasiconvex function, a suitable generalization of conjugation is possible by the so-called

penalty function, which is defined in terms of the support function of the negative of sublevel sets. The precise definition is given next.

Definition 2.8. *The penalty function $\alpha_f: \mathcal{X}^* \times \mathbb{R} \rightarrow \overline{\mathbb{R}}$ associated with f is defined by*

$$\alpha_f(x^*, m) := I_{-S_m^f}^*(x^*) = I_{S_m^f}^*(-x^*) = \sup_{x \in S_m^f} \langle x^*, -x \rangle, \quad x^* \in \mathcal{X}^*, m \in \mathbb{R}.$$

Remark 2.9. When $x^* \neq 0$, we can extend Definition 2.8 for $m = +\infty$ and $m = -\infty$ by letting $\alpha_f(x^*, +\infty) := +\infty$ and $\alpha_f(x^*, -\infty) := -\infty$. These values are consistent with the original definition. For $m = +\infty$, we consider a supremum over the whole space $\{x \in \mathcal{X} \mid f(x) \leq \infty\} = \mathcal{X}$, which gives $+\infty$. Similarly, for $m = -\infty$, we consider a supremum over the empty set, which gives $-\infty$.

The next remark states some elementary properties of the penalty function α_f .

Remark 2.10. It is clear that the penalty function α_f is positively homogeneous in the first argument, that is, $\alpha_f(\lambda x^*, m) = \lambda \alpha_f(x^*, m)$ for every $x^* \in \mathcal{X}^*$, $m \in \mathbb{R}$. Moreover, α_f is increasing in the second argument. Indeed, by taking $m_1, m_2 \in \mathbb{R}$ with $m_1 \leq m_2$, we have $S_{m_1}^f \subseteq S_{m_2}^f$ so that $\alpha_f(x^*, m_1) \leq \alpha_f(x^*, m_2)$ for every $x^* \in \mathcal{X}^*$.

We continue with a remark which serves as a basis for dual representations.

Remark 2.11. Let $A \subseteq \mathcal{X}$ be a nonempty, closed, convex and monotone set. An immediate consequence of Hahn-Banach theorem is that, for every $x \in \mathcal{X}$, we have

$$x \in A \quad \Leftrightarrow \quad \forall x^* \in C^+ \setminus \{0\}: \langle x^*, -x \rangle \leq \sup_{y \in A} \langle x^*, -y \rangle.$$

In particular, if f is a decreasing, lower semicontinuous and quasiconvex function, then we have

$$x \in S_m^f \quad \Leftrightarrow \quad \forall x^* \in C^+ \setminus \{0\}: \langle x^*, -x \rangle \leq \alpha_f(x^*, m)$$

by Remarks 2.2, 2.6. Similarly, if f is an increasing, lower semicontinuous and quasiconvex function, then f is decreasing with respect to $-C$ so that

$$x \in S_m^f \quad \Leftrightarrow \quad \forall x^* \in C^- \setminus \{0\}: \langle x^*, -x \rangle \leq \alpha_f(x^*, m),$$

where $C^- := -C^+ = \{x^* \in \mathcal{X}^* \mid \langle x^*, x \rangle \leq 0 \text{ for all } x \in C\}$.

When f is lower semicontinuous and quasiconvex, its dual representation will be stated in terms of a special pseudoinverse of α_f , which we recall in the next definition.

Definition 2.12. (Drapeau and Kupper [12, App. B]) Let $\alpha: \mathcal{X}^* \times \mathbb{R} \rightarrow \overline{\mathbb{R}}$ be a function. We define its left inverse $\alpha^{-l}: \mathcal{X}^* \times \mathbb{R} \rightarrow \overline{\mathbb{R}}$ with respect to the second argument by

$$\alpha^{-l}(x^*, s) := \sup \{m \in \mathbb{R} \mid \alpha(x^*, m) < s\} = \inf \{m \in \mathbb{R} \mid \alpha(x^*, m) \geq s\}, \quad (2.7)$$

for each $x^* \in \mathcal{X}^*$ and $s \in \mathbb{R}$.

The next lemma provides simple strong duality results that will be useful in later calculations.

Lemma 2.13. Let $\alpha: \mathcal{X}^* \times \mathbb{R} \rightarrow \overline{\mathbb{R}}$ be a function that is increasing in its second argument.

(i) Let $r: \mathcal{X}^* \rightarrow \overline{\mathbb{R}}$ be a function and $A \subseteq \mathcal{X}^*$ a nonempty set. Then, we have

$$\inf \{m \in \mathbb{R} \mid \forall x^* \in A: r(x^*) \leq \alpha(x^*, m)\} = \sup_{x^* \in A} \alpha^{-l}(x^*, r(x^*)).$$

(ii) Let S be a nonempty set and $r: \mathcal{X}^* \times S \rightarrow \overline{\mathbb{R}}$. Then, for every $x^* \in \mathcal{X}^*$, we have

$$\inf \{m \in \mathbb{R} \mid \forall s \in S: r(x^*, s) \leq \alpha(x^*, m)\} = \sup_{s \in S} \alpha^{-l}(x^*, r(x^*, s)).$$

Proof. Let us prove (i). By the definition of left inverse, the claimed equality is equivalent to

$$\inf \{m \in \mathbb{R} \mid \forall x^* \in A: r(x^*) \leq \alpha(x^*, m)\} = \sup_{x^* \in A} \inf \{m \in \mathbb{R} \mid r(x^*) \leq \alpha(x^*, m)\}. \quad (2.8)$$

The \geq part is true by weak duality. For the other side, to get a contradiction, assume that there exists $\tilde{m} \in \mathbb{R}$ such that

$$\inf \{m \in \mathbb{R} \mid \forall x^* \in A: r(x^*) \leq \alpha_f(x^*, m)\} > \tilde{m} > \sup_{x^* \in A} \inf \{m \in \mathbb{R} \mid r(x^*) \leq \alpha_f(x^*, m)\}. \quad (2.9)$$

The first inequality in (2.9) implies that there exists $\tilde{x}^* \in A$ such that $r(\tilde{x}^*) > \alpha_f(\tilde{x}^*, \tilde{m})$. The second inequality in (2.9) implies that $\tilde{m} > \inf \{m \in \mathbb{R} \mid r(\tilde{x}^*) \leq \alpha_f(\tilde{x}^*, m)\}$. Hence, by the monotonicity of α_f , we must have $r(\tilde{x}^*) \leq \alpha_f(\tilde{x}^*, \tilde{m})$, a contradiction. So (2.8) follows.

The proof of (ii) is similar, we omit it for brevity. \square

We state the dual representation theorem for lower semicontinuous quasiconvex functions, which is a part of Drapeau and Kupper [12, Thm. 3]. It is formulated in terms of the left inverse of the penalty function. We provide the proof for completeness.

Theorem 2.14. *Suppose that $f: \mathcal{X} \rightarrow \overline{\mathbb{R}}$ is a decreasing, lower semicontinuous and quasiconvex function. Then, f has the dual representation*

$$f(x) = \sup_{x^* \in C^+ \setminus \{0\}} \alpha_f^{-l}(x^*, \langle x^*, -x \rangle), \quad x \in \mathcal{X}. \quad (2.10)$$

Proof. Let $x \in \mathbb{R}$. By (2.1) and Remark 2.11, we have

$$\begin{aligned} f(x) &= \inf\{m \in \mathbb{R} \mid x \in S_m^f\} = \inf\{m \in \mathbb{R} \mid \forall x^* \in C^+ \setminus \{0\}: \langle x^*, -x \rangle \leq \alpha_f(x^*, m)\} \\ &= \sup_{x^* \in C^+ \setminus \{0\}} \alpha_f^{-l}(x^*, \langle x^*, -x \rangle), \end{aligned}$$

where the last equality is a direct result of Lemma 2.13(i) since α_f is increasing by Remark 2.10. \square

In Drapeau and Kupper [12], a decreasing quasiconvex function on \mathcal{X} is called a *risk measure* as a generalization of convex and coherent risk measures studied in the financial mathematics literature; see Föllmer and Schied [15, Ch. 4], for instance. Hence, Theorem 2.14 provides a dual representation for a lower semicontinuous (quasiconvex) risk measure. For the current discussion, we keep using the general terminology of convex analysis and do not use the term *risk measure*. In Section 6, we will focus on applications in systemic risk measures, where we also introduce the financial background as necessary.

In applications, it might be necessary to consider a function that is defined on some subset of the vector space \mathcal{X} . The next corollary is for this purpose. To that end, let $\mathcal{K} \subseteq \mathcal{X}$ be a monotone convex set such that $C \subseteq \mathcal{K}$. In particular, we have $\mathcal{K} + C \subseteq \mathcal{K}$. Given a function $g: \mathcal{K} \rightarrow \overline{\mathbb{R}}$, we may extend g to \mathcal{X} as a function \bar{g} defined by $\bar{g}(x) = g(x)$ for $x \in \mathcal{K}$, and by $\bar{g}(x) = +\infty$ for $x \in \mathcal{X} \setminus \mathcal{K}$. Hence, the sublevel sets, penalty function, and algebraic properties (quasiconvexity, monotonicity, etc.) of g are defined as those of \bar{g} .

Corollary 2.15. *Let $g: \mathcal{K} \rightarrow \overline{\mathbb{R}}$ be a quasiconvex, decreasing and lower semicontinuous (with respect to the relative topology) function. Then, we have*

$$g(x) = \sup_{x^* \in C^+ \setminus \{0\}} \alpha_g^{-l}(x^*, \langle x^*, -x \rangle), \quad x \in \mathcal{K}. \quad (2.11)$$

Proof. Let us define a function $\tilde{g}: \mathcal{X} \rightarrow \mathbb{R}$ by

$$\tilde{g}(x) := \inf\{m \in \mathbb{R} \mid x \in \text{cl}(S_m^g)\}, \quad x \in \mathcal{X}.$$

Note that $S_m^{\tilde{g}} = \text{cl}(S_m^g)$ for each $m \in \mathbb{R}$. Let $m \in \mathbb{R}$. Since g is quasiconvex, it follows that $S_m^{\tilde{g}}$ is closed and convex. To show that it is also monotone, let $x \in S_m^{\tilde{g}} = \text{cl}(S_m^g)$, $c \in C$. Let $U \subseteq \mathcal{X}$ be a neighborhood of $x + c$. Since \mathcal{X} is a topological vector space, $(U - c)$ is an open set; hence, it is a neighborhood of x . Therefore, $(U - c) \cap S_m^g \neq \emptyset$. Let $z \in (U - c) \cap S_m^g$ so that $z + c \in U$. On the other hand, since g is decreasing, S_m^g is monotone, which yields that $z + c \in S_m^g$. It follows that $U \cap S_m^g \neq \emptyset$. Since U is an arbitrary neighborhood of $x + c$, we conclude that $x + c \in \text{cl}(S_m^g) = S_m^{\tilde{g}}$. Hence, $S_m^{\tilde{g}}$ is monotone. By Remarks 2.2, 2.6, it follows that \tilde{g} is decreasing, lower semicontinuous, and quasiconvex. By Theorem 2.14, we get

$$\tilde{g}(x) = \sup_{x^* \in C^+ \setminus \{0\}} \alpha_{\tilde{g}}^{-l}(x^*, \langle x^*, -x \rangle), \quad x \in \mathcal{X}. \quad (2.12)$$

By definition, $S_m^{\tilde{g}}$ is the closed convex hull of S_m^g for each $m \in \mathbb{R}$. Hence, (2.5) yields

$$\alpha_{\tilde{g}}(x^*, m) = \sup_{y \in S_m^{\tilde{g}}} \langle x^*, -y \rangle = \sup_{y \in S_m^g} \langle x^*, -y \rangle = \alpha_g(x^*, m), \quad x^* \in \mathcal{X}^*, m \in \mathbb{R}. \quad (2.13)$$

For $x \in \mathcal{K}$, by (2.1), we have

$$\tilde{g}(x) = \inf\{m \in \mathbb{R} \mid x \in S_m^{\tilde{g}}\} = \inf\{m \in \mathbb{R} \mid x \in S_m^g \cap \mathcal{K}\}. \quad (2.14)$$

We claim that $S_m^{\tilde{g}} \cap \mathcal{K} = S_m^g$. Indeed, it is clear that $S_m^{\tilde{g}} \cap \mathcal{K} = \text{cl}(S_m^g) \cap \mathcal{K} \supseteq S_m^g$. On the other hand, since g is lower semicontinuous with respect to the relative topology, we have $S_m^g = A \cap \mathcal{K}$ for some closed set $A \subseteq \mathcal{X}$. Since $S_m^g \subseteq A$, we have $\text{cl}(S_m^g) \subseteq A$. It follows that $\text{cl}(S_m^g) \cap \mathcal{K} \subseteq A \cap \mathcal{K} = S_m^g$.

Hence, the claim follows. Then, (2.14) yields $\tilde{g}(x) = \inf \{m \in \mathbb{R} \mid x \in S_m^g\} = g(x)$. After combining this result with (2.12) and (2.13), we obtain (2.11). \square

When f is a proper lower semicontinuous convex function, two dual representations are possible: the one provided by Fenchel-Moreau theorem, and the one provided by Theorem 2.14 since f is also quasiconvex. To establish the link between the two representations, we calculate the left inverse of the penalty function in terms of the conjugate function.

Proposition 2.16. *Suppose that $f: \mathcal{X} \rightarrow \mathbb{R}$ is convex and lower semicontinuous. If $m \in \mathbb{R}$ is such that the strict sublevel set $\{x \in \mathcal{X} \mid f(x) < m\}$ is nonempty, then*

$$\alpha_f(x^*, m) = \inf_{\lambda > 0} \left(\lambda m + \lambda f^* \left(-\frac{x^*}{\lambda} \right) \right), \quad x^* \in \mathcal{X}^*. \quad (2.15)$$

Moreover, for the left inverse, we have

$$\alpha_f^{-l}(x^*, s) = \sup_{\gamma \geq 0} (\gamma s - f^*(-\gamma x^*)), \quad x^* \in \mathcal{X}^*, s \in \mathbb{R}. \quad (2.16)$$

Proof. Let $x^* \in \mathcal{X}^*$ and $m \in \mathbb{R}$ such that $\{x \in \mathcal{X} \mid f(x) < m\} \neq \emptyset$. Note that $\alpha_f(x^*, m) = \sup_{x \in S_m^f} \langle x^*, -x \rangle$ can be seen as the optimal value of the following convex optimization problem:

$$\text{maximize } \langle x^*, -x \rangle \text{ subject to } f(x) \leq m, x \in \mathcal{X}.$$

By supposition, Slater's condition holds, that is, there exists $x_0 \in \mathcal{X}$ such that $f(x_0) < m$. Hence, we have strong duality for this problem, that is,

$$\alpha_f(x^*, m) = \inf_{\lambda \geq 0} \sup_{x \in \mathcal{X}} (\langle x^*, -x \rangle - \lambda(f(x) - m)).$$

When $\lambda = 0$, $\sup_{x \in \mathcal{X}} (\langle x^*, -x \rangle - \lambda(f(x) - m)) = \sup_{x \in \mathcal{X}} \langle x^*, -x \rangle = +\infty$. Hence, we can evaluate the infimum over $\lambda > 0$. Then,

$$\begin{aligned} \alpha_f(x^*, m) &= \inf_{\lambda > 0} \sup_{x \in \mathcal{X}} (\langle x^*, -x \rangle - \lambda(f(x) - m)) \\ &= \inf_{\lambda > 0} \left(\lambda m + \sup_{x \in \mathcal{X}} (\langle x^*, -x \rangle - \lambda f(x)) \right) = \inf_{\lambda > 0} \left(\lambda m + \lambda f^* \left(-\frac{x^*}{\lambda} \right) \right), \end{aligned}$$

which proves (2.15). For $m > \inf_{x \in \mathcal{X}} f(x)$, the strict sublevel set $\{x \in \mathcal{X} \mid f(x) < m\}$ is nonempty. Let us define $F := (\inf_{x \in \mathcal{X}} f(x), +\infty)$. For $m < \inf_{x \in \mathcal{X}} f(x)$, we have $\alpha_f(x^*, m) = -\infty$. Then, to prove (2.16), for each $s \in \mathbb{R}$, we have

$$\begin{aligned} \alpha_f^{-l}(x^*, s) &= \inf \{m \in \mathbb{R} \mid \alpha_f(x^*, m) \geq s\} \\ &= \inf_{m \in F} m \vee \inf \left\{ m \in \mathbb{R} \mid \forall \lambda > 0: \lambda m + \lambda f^* \left(-\frac{x^*}{\lambda} \right) \geq s \right\} \\ &= \inf_{x \in \mathcal{X}} f(x) \vee \inf \left\{ m \in \mathbb{R} \mid \forall \lambda > 0: m \geq \frac{s}{\lambda} - f^* \left(-\frac{x^*}{\lambda} \right) \right\}. \end{aligned}$$

Also, note that $\inf_{x \in \mathcal{X}} f(x) = -\sup_{x \in \mathcal{X}} (0 - f(x)) = -f^*(0)$. Hence,

$$\begin{aligned} \alpha_f^{-l}(x^*, s) &= -f^*(0) \vee \sup_{\lambda > 0} \left(\frac{s}{\lambda} - f^* \left(-\frac{x^*}{\lambda} \right) \right) \\ &= -f^*(0) \vee \sup_{\gamma > 0} (\gamma s - f^*(-\gamma x^*)) = \sup_{\gamma \geq 0} (\gamma s - f^*(-\gamma x^*)), \end{aligned}$$

which completes the proof. □

Remark 2.17. Under the assumptions of Proposition 2.16, we may rewrite the dual representation in Theorem 2.14 using Proposition 2.16 and the fact that C^+ is a cone, which gives

$$\begin{aligned} f(x) &= \sup_{x^* \in C^+ \setminus \{0\}} \alpha_f^{-l}(x^*, \langle x^*, -x \rangle) = \sup_{x^* \in C^+ \setminus \{0\}} \sup_{\gamma \geq 0} (\langle \gamma x^*, -x \rangle - f^*(-\gamma x^*)) \\ &= \sup_{x^* \in C^+} (\langle x^*, -x \rangle - f^*(-x^*)) \end{aligned}$$

for each $x \in \mathcal{X}$. Hence, in the convex case, the representation in Theorem 2.14 reproduces the standard Fenchel-Moreau-type representation in Proposition 2.7.

3 Naturally quasiconvex vector-valued functions

Throughout this section, let \mathcal{X}, \mathcal{Y} be Hausdorff locally convex topological vector spaces with vector preorders \leq_C, \leq_D , where $C \subseteq \mathcal{X}$ and $D \subseteq \mathcal{Y}$ are closed convex cones. We denote by $2^{\mathcal{Y}}$ the power set of \mathcal{Y} . Let $f: \mathcal{Y} \rightarrow \overline{\mathbb{R}}$ and $g: \mathcal{X} \rightarrow \mathcal{Y}$ be functions. The main purpose of this paper is to provide a dual representation for a quasiconvex composition of the form $f \circ g$. While Section 2 provides the

background for the study of extended real-valued function, we dedicate this section to the study of vector-valued functions.

We start by recalling some generalized notions of convexity for vector-valued functions.

Definition 3.1. *Consider the following notions for $g: \mathcal{X} \rightarrow \mathcal{Y}$.*

(i) *g is called D -convex if $g(\lambda x_1 + (1 - \lambda)x_2) \leq_D \lambda g(x_1) + (1 - \lambda)g(x_2)$ for every $x_1, x_2 \in \mathcal{X}$ and $\lambda \in (0, 1)$. It is called D -concave if $-g$ is D -convex.*

(ii) *g is called D -naturally quasiconvex if, for every $x_1, x_2 \in \mathcal{X}$ and $\lambda \in [0, 1]$, there exists $\mu \in [0, 1]$ such that $g(\lambda x_1 + (1 - \lambda)x_2) \leq_D \mu g(x_1) + (1 - \mu)g(x_2)$. It is called D -naturally quasiconcave if $-g$ is naturally D -quasiconvex.*

From Definition 3.1, it is clear that D -convexity implies D -natural quasiconvexity. For real-valued functions with $D = \mathbb{R}_+$, D -natural quasiconvexity coincides with quasiconvexity; see the notes after Kuroiwa [23, Def. 2.1].

For the function $g: \mathcal{X} \rightarrow \mathcal{Y}$, let us consider the *scalarization* $h_{y^*}^g(x): \mathcal{X} \rightarrow \mathbb{R}$ defined by

$$h_{y^*}^g(x) := \langle y^*, g(x) \rangle, \quad x \in \mathcal{X}, \quad (3.1)$$

for each $y^* \in D^+ \setminus \{0\}$. The next proposition provides useful characterizations of the convexity and D -natural quasiconvexity of g in terms of the analogous properties of the family of scalarizations. It can be seen as a modified version of Kuroiwa [23, Prop. 2.2, Theorem 2.1], which are stated in a set-valued setting.

Proposition 3.2. *We have the following equivalences for g and its scalarizations.*

(i) *g is D -convex if and only if $h_{y^*}^g$ is convex for each $y^* \in D^+ \setminus \{0\}$.*

(ii) *g is D -naturally quasiconvex if and only if $h_{y^*}^g$ is quasiconvex for every $y^* \in D^+ \setminus \{0\}$.*

Proof. We prove (i) first. Assume that g is D -convex and take $y^* \in D^+ \setminus \{0\}$. Let $x_1, x_2 \in \mathcal{X}$ and $\lambda \in (0, 1)$. Then, the D -convexity of g implies that $g(\lambda x_1 + (1 - \lambda)x_2) \leq_D \lambda g(x_1) + (1 - \lambda)g(x_2)$, that is, $\lambda g(x_1) + (1 - \lambda)g(x_2) - g(\lambda x_1 + (1 - \lambda)x_2) \in D$. Hence,

$$\langle y^*, \lambda g(x_1) + (1 - \lambda)g(x_2) - g(\lambda x_1 + (1 - \lambda)x_2) \rangle \geq 0$$

so that

$$\begin{aligned} h_{y^*}^g(\lambda x_1 + (1 - \lambda)x_2) &= \langle y^*, g(\lambda x_1 + (1 - \lambda)x_2) \rangle \\ &\leq \langle y^*, \lambda g(x_1) + (1 - \lambda)g(x_2) \rangle = \lambda h_{y^*}^g(x_1) + (1 - \lambda)h_{y^*}^g(x_2). \end{aligned}$$

Therefore, $h_{y^*}^g$ is convex.

Conversely, assume that $h_{y^*}^g$ is convex for each $y^* \in D^+ \setminus \{0\}$. Let $x_1, x_2 \in \mathcal{X}$ and $\lambda \in (0, 1)$. For each $y^* \in D^+ \setminus \{0\}$, since $h_{y^*}^g$ is convex, we have

$$\begin{aligned} \langle y^*, g(\lambda x_1 + (1 - \lambda)x_2) \rangle &= h_{y^*}^g(\lambda x_1 + (1 - \lambda)x_2) \\ &\leq \lambda h_{y^*}^g(x_1) + (1 - \lambda)h_{y^*}^g(x_2) = \langle y^*, \lambda g(x_1) + (1 - \lambda)g(x_2) \rangle. \end{aligned}$$

Hence, $\langle y^*, \lambda g(x_1) + (1 - \lambda)g(x_2) - g(\lambda x_1 + (1 - \lambda)x_2) \rangle \geq 0$ for every $y^* \in D^+ \setminus \{0\}$, that is, $\lambda g(x_1) + (1 - \lambda)g(x_2) - g(\lambda x_1 + (1 - \lambda)x_2) \in D$, that is, $g(\lambda x_1 + (1 - \lambda)x_2) \leq_D \lambda g(x_1) + (1 - \lambda)g(x_2)$. Therefore, g is D -convex, which completes the proof of (i).

Next, we prove (ii). Assume that g is D -naturally quasiconvex. Let $y^* \in D^+ \setminus \{0\}$ and consider $h_{y^*}^g$. Let $x_1, x_2 \in \mathcal{X}$ and $\lambda \in [0, 1]$. Since g is D -naturally quasiconvex, there exists $\mu \in [0, 1]$ such that $g(\lambda x_1 + (1 - \lambda)x_2) \leq_D \mu g(x_1) + (1 - \mu)g(x_2)$. Hence,

$$\begin{aligned} h_{y^*}^g(\lambda x_1 + (1 - \lambda)x_2) &= \langle y^*, g(\lambda x_1 + (1 - \lambda)x_2) \rangle \leq \langle y^*, \mu g(x_1) + (1 - \mu)g(x_2) \rangle \\ &\leq \langle y^*, g(x_1) \rangle \vee \langle y^*, g(x_2) \rangle = h_{y^*}^g(x_1) \vee h_{y^*}^g(x_2). \end{aligned}$$

Therefore, $h_{y^*}^g$ is quasiconvex.

Conversely, assume that $h_{y^*}^g$ is quasiconvex for each $y^* \in D^+ \setminus \{0\}$. To get a contradiction, suppose that g is not D -naturally quasiconvex. Hence, there exist $x_1, x_2 \in \mathcal{X}$, $\lambda \in [0, 1]$ such that

$$(\text{conv}(\{g(x_1), g(x_2)\}) - g(\lambda x_1 + (1 - \lambda)x_2)) \cap D = \emptyset.$$

Since D is closed and convex, and the (shifted) line segment $\text{conv}(\{g(x_1), g(x_2)\}) - g(\lambda x_1 + (1 - \lambda)x_2)$

is compact and convex, by Hahn-Banach strong separation theorem, there exists $y_0^* \in \mathcal{Y}^* \setminus \{0\}$ with

$$\inf_{d \in D} \langle y_0^*, d \rangle > \sup_{y \in \text{conv}(\{g(x_1), g(x_2)\}) - g(\lambda x_1 + (1-\lambda)x_2)} \langle y_0^*, y \rangle \quad (3.2)$$

Since D is a cone, $\inf_{d \in D} \langle y_0^*, d \rangle$ is either 0 or $-\infty$. However, the term on the right of (3.2) is finite. Hence, we must have $\inf_{d \in D} \langle y_0^*, d \rangle = 0$ so that $y_0^* \in D^+$. Using this information in (3.2) implies $\langle y_0^*, \mu g(x_1) + (1-\mu)g(x_2) \rangle < \langle y_0^*, g(\lambda x_1 + (1-\lambda)x_2) \rangle$ for every $\mu \in [0, 1]$. It follows that

$$\langle y_0^*, g(x_1) \rangle \vee \langle y_0^*, g(x_2) \rangle < \langle y_0^*, g(\lambda x_1 + (1-\lambda)x_2) \rangle,$$

which contradicts the quasiconvexity of $h_{y_0^*}^g$. Hence, g is D -naturally quasiconvex. \square

Remark 3.3. The equivalent condition in Proposition 3.2(ii) is sometimes called **-quasiconvexity*; see, for instance, Kuroiwa [23, Def. 2.1].

If $f: \mathcal{Y} \rightarrow \overline{\mathbb{R}}$ is a decreasing convex function, and $g: \mathcal{X} \rightarrow \mathcal{Y}$ is a D -concave function, then it is easy to check that the composition $f \circ g: \mathcal{X} \rightarrow \overline{\mathbb{R}}$ is a convex function. The following proposition provides an analogue of this observation when the resulting composition is quasiconvex.

Proposition 3.4. *Suppose that f is quasiconvex and decreasing, and g is D -naturally quasiconcave. Then, $f \circ g: \mathcal{X} \rightarrow \overline{\mathbb{R}}$ is quasiconvex.*

Proof. Let $x_1, x_2 \in \mathcal{X}$ and $\lambda \in [0, 1]$. Since g is naturally D -quasiconcave, there exists $\mu \in [0, 1]$ such that $\mu g(x_1) + (1-\mu)g(x_2) \leq_D g(\lambda x_1 + (1-\lambda)x_2)$. Using the monotonicity and quasiconvexity of f , we obtain $f(g(\lambda x_1 + (1-\lambda)x_2)) \leq f(\mu g(x_1) + (1-\mu)g(x_2)) \leq f(g(x_1)) \vee f(g(x_2))$. Hence, $f \circ g$ is quasiconvex. \square

When $f: \mathcal{Y} \rightarrow \overline{\mathbb{R}}$ is quasiconvex and decreasing, Proposition 3.4 gives a sufficient condition on $g: \mathcal{X} \rightarrow \mathcal{Y}$ that is weaker than D -convexity so that $f \circ g$ is quasiconvex. In the rest of the section, we investigate further properties of g that will help us obtain a dual representation for $f \circ g$ in Section 4. To that end, we start by studying monotonicity for the vector-valued case.

Definition 3.5. *The function $g: \mathcal{X} \rightarrow \mathcal{Y}$ is called decreasing if $x_1 \leq_C x_2$ implies $g(x_2) \leq_D g(x_1)$ for every $x_1, x_2 \in \mathcal{X}$; it is called increasing if $x_1 \leq_C x_2$ implies $g(x_1) \leq_D g(x_2)$ for every $x_1, x_2 \in \mathcal{X}$.*

The preservation of monotonicity under compositions is formulated next.

Proposition 3.6. *Suppose that f is decreasing and g is increasing. Then, $f \circ g$ is decreasing.*

Proof. Let $x_1, x_2 \in \mathcal{X}$ such that $x_1 \leq_C x_2$. Since g is increasing, we have $g(x_1) \leq_D g(x_2)$. Since f is decreasing, we obtain $f(g(x_1)) \geq f(g(x_2))$. Hence, $f \circ g$ is decreasing. \square

The next proposition connects the monotonicity of g and that of its scalarizations.

Proposition 3.7. *Suppose that g is decreasing. Then, $h_{y^*}^g$ is decreasing for every $y^* \in D^+ \setminus \{0\}$.*

Proof. Let $y^* \in D^+ \setminus \{0\}$. Let $x_1, x_2 \in \mathcal{X}$ such that $x_1 \leq_C x_2$. Since g is decreasing, we have $g(x_2) \leq_D g(x_1)$, that is, $g(x_1) - g(x_2) \in D$. Hence, $\langle y^*, g(x_1) - g(x_2) \rangle \geq 0$, that is, $h_{y^*}^g(x_1) \geq h_{y^*}^g(x_2)$. Therefore, $h_{y^*}^g$ is decreasing. \square

In Section 4, we will also need a notion of strict monotonicity for a vector-valued function. The next definition gives one that suits our purposes. Recall that $C^\#$ and $D^\#$ are the (convex) cones of strictly positive elements in \mathcal{X} and \mathcal{Y} , respectively; see (2.4). Although these cones are not closed in general, we define their induced preorders $\leq_{C^\#}$ and $\leq_{D^\#}$ as in (2.3).

Definition 3.8. *The function g is called regularly increasing if it is increasing, and $x_1 \leq_{C^\#} x_2$ implies $g(x_1) \leq_{D^\#} g(x_2)$ for every $x_1, x_2 \in \mathcal{X}$; it is called regularly decreasing if it is decreasing, and $x_1 \leq_{C^\#} x_2$ implies $g(x_2) \leq_{D^\#} g(x_1)$ for every $x_1, x_2 \in \mathcal{X}$.*

To be able to employ Definition 3.8, we need to work under the following assumption.

Assumption 3.9. *The cones $C^\#$ and $D^\#$ are nonempty.*

We proceed with a continuity concept for g , which is defined through its set-valued extension $G: \mathcal{X} \rightarrow 2^{\mathcal{Y}}$ given by

$$G(x) := g(x) + D, \quad x \in \mathcal{X}.$$

Given $M \subseteq \mathcal{Y}$, the sets

$$G^L(M) := \{x \in \mathcal{X} \mid G(x) \cap M \neq \emptyset\}, \quad G^U(M) := \{x \in \mathcal{X} \mid G(x) \subseteq M\}$$

are called the *lower inverse image* and *upper inverse image* of M under G , respectively. It is easy to check that $(G^U(M))^c = G^L(M^c)$ and $(G^L(M))^c = G^U(M^c)$.

Definition 3.10. (Ha [19, Def. 2.1]) The function g is called D -lower demicontinuous if the lower inverse image $G^L(M)$ is open for every open halfspace $M \subseteq \mathcal{Y}$.

When $\mathcal{Y} = \mathbb{R}$ and $D = \mathbb{R}_+$, note that Definition 3.10 coincides with the usual notion of lower semicontinuity, see Remark 2.2.

Remark 3.11. Note that g is D -lower demicontinuous if and only if the upper inverse image $G^U(M)$ is closed for every closed halfspace $M \subseteq \mathcal{Y}$. This follows from the observations that M is a closed halfspace if and only if M^c is an open halfspace, and that $G^U(M) = (G^L(M^c))^c$.

We conclude this section by relating the D -lower demicontinuity of g with the upper semicontinuity of its scalarizations.

Proposition 3.12. The function g is D -lower demicontinuous if and only if $h_{y^*}^g$ is upper semicontinuous for every $y^* \in D^+ \setminus \{0\}$.

Proof. Let $m \in \mathbb{R}$ and $y^* \in D^+ \setminus \{0\}$. Let us define the sets

$$A_{m,y^*} := \{x \in \mathcal{X} \mid h_{y^*}^g(x) \geq m\}, \quad B_{m,y^*} := G^U(M_{m,y^*}) = \{x \in \mathcal{X} \mid g(x) + D \subseteq M_{m,y^*}\},$$

where $M_{m,y^*} := \{y \in \mathcal{Y} \mid \langle y^*, y \rangle \geq m\}$. We claim that $A_{m,y^*} = B_{m,y^*}$. First, let $x \in A_{m,y^*}$ and take $d \in D$. Hence, $\langle y^*, g(x) \rangle \geq m$ and $\langle y^*, d \rangle \geq 0$. Combining these two inequalities, we get $\langle y^*, g(x) + d \rangle \geq m$, that is, $g(x) + d \in M_{m,y^*}$. Since $d \in D$ is arbitrary, we have $g(x) + D \subseteq M_{m,y^*}$. Hence, $x \in B_{m,y^*}$. Conversely, let $x \in B_{m,y^*}$. In particular, $g(x) \in M_{m,y^*}$, that is, $h_{y^*}^g(x) = \langle y^*, g(x) \rangle \geq m$. Hence, $x \in A_{m,y^*}$, which completes the proof of the claim. By this claim and Remark 3.11, the statement of the proposition follows. \square

Let $y^* \in D^+ \setminus \{0\}$. In view of Propositions 3.2 and 3.7, when g is D -naturally quasiconcave increasing and D -lower demicontinuous, the function $-h_{y^*}^g$ is quasiconvex, decreasing and lower semicontinuous. In this case, we may apply Theorem 2.14 for $-h_{y^*}^g$ to get

$$-h_{y^*}^g(x) = \sup_{x^* \in C^+ \setminus \{0\}} \alpha_{-h_{y^*}^g}^{-l}(x^*, \langle x^*, -x \rangle), \quad x \in \mathcal{X}. \quad (3.3)$$

The availability of (3.3) will be useful in Section 4 when obtaining dual representations for quasiconvex compositions.

4 Quasiconvex compositions

The aim of this section is to establish dual representations for quasiconvex compositions. We continue working in the framework of Section 3, where we have locally convex topological vector spaces \mathcal{X}, \mathcal{Y} with respective preorders \leq_C, \leq_D .

4.1 The main theorem

Let us fix two functions $f: \mathcal{Y} \rightarrow \overline{\mathbb{R}}$ and $g: \mathcal{X} \rightarrow \mathcal{Y}$. To motivate the discussion, we make the following simple observation: if f is decreasing and quasiconvex, and g is increasing and D -naturally quasiconcave, then $f \circ g$ is decreasing and quasiconvex by Propositions 3.4 and 3.6. Hence, in view of Theorem 2.14, a dual representation for $f \circ g$ is readily available once $f \circ g$ is guaranteed to be lower semicontinuous. This is achieved in the next proposition by suitable continuity assumptions.

Proposition 4.1. *Suppose that f is decreasing, lower semicontinuous, and quasiconvex; and that g is increasing, D -lower demicontinuous, and D -naturally quasiconcave. Then, $f \circ g$ is a decreasing, lower semicontinuous, and quasiconvex function. Moreover, for every $x \in \mathcal{X}$, we have*

$$f \circ g(x) = \sup_{y^* \in D^+ \setminus \{0\}} \alpha_f^{-l} \left(y^*, \sup_{x^* \in C^+ \setminus \{0\}} \alpha_{-h_{y^*}^g}^{-l} (x^*, \langle x^*, -x \rangle) \right) = \sup_{x^* \in C^+ \setminus \{0\}} \alpha_{f \circ g}^{-l} (x^*, \langle x^*, -x \rangle). \quad (4.1)$$

Proof. By Propositions 3.4 and 3.6, $f \circ g$ is decreasing and quasiconvex. Let us show that it is also lower semicontinuous. To that end, let $m \in \mathbb{R}$. Note that

$$S_m^{f \circ g} = \{x \in \mathcal{X} \mid f \circ g(x) \leq m\} = \{x \in \mathcal{X} \mid g(x) \in S_m^f\} = \{x \in \mathcal{X} \mid G(x) \subseteq S_m^f\} = G^U(S_m^f). \quad (4.2)$$

Here, only the third equality needs a proof. Since f is decreasing, S_m^f is monotone. Let $x \in \mathcal{X}$ with $g(x) \in S_m^f$, and $d \in D$. Since S_m^f is monotone, we have $g(x) + d \in S_m^f$. As this is true for every $d \in D$, we have $G(x) = g(x) + D \subseteq S_m^f$. Conversely, let $x \in \mathcal{X}$ with $G(x) \subseteq S_m^f$. Since $0 \in D$, we have $g(x) \in g(x) + D = G(x) \subseteq S_m^f$. These observations verify the third equality in (4.2).

By Remark 2.2, we may write $S_m^f = \bigcap_{M \in \mathcal{M}} M$, where \mathcal{M} is the collection of all closed halfspaces

M such that $S_m^f \subseteq M$. Therefore,

$$G^U(S_m^f) = G^U\left(\bigcap_{M \in \mathcal{M}} M\right) = \bigcap_{M \in \mathcal{M}} G^U(M).$$

Since g is D -lower demicontinuous, $G^U(M)$ is closed for each $M \in \mathcal{M}$. By (4.2), it follows that $S_m^{f \circ g} = G^U(S_m^f)$ is closed. Therefore, $f \circ g$ is lower semicontinuous by Remark 3.11.

By Theorem 2.14, we obtain the second dual representation in (4.1). Finally, we show the first equality in (4.1). Let $x \in \mathcal{X}$. By applying Theorem 2.14 for f at the point $g(x)$, we get

$$f(g(x)) = \sup_{y^* \in D^+ \setminus \{0\}} \alpha_f^{-l}(y^*, \langle y^*, -g(x) \rangle).$$

On the other hand, by (3.3), we have

$$\langle y^*, -g(x) \rangle = -h_{y^*}^g(x) = \sup_{x^* \in C^+ \setminus \{0\}} \alpha_{-h_{y^*}^g}^{-l}(x^*, \langle x^*, -x \rangle), \quad y^* \in D^+ \setminus \{0\}.$$

Combining the last two observations gives the first equality in (4.1). \square

Under the assumptions of Proposition 4.1, $f \circ g$ has a dual representation in the sense of Theorem 2.14. We have a more explicit dual representation for $f \circ g$ in the next theorem.

Theorem 4.2. *Suppose that f is decreasing, lower semicontinuous, and quasiconvex; and that g is increasing, D -lower demicontinuous, and D -naturally quasiconcave. We have*

$$f \circ g(x) = \sup_{x^* \in C^+ \setminus \{0\}} \sup_{y^* \in D^+ \setminus \{0\}} \alpha_f^{-l}\left(y^*, \alpha_{-h_{y^*}^g}^{-l}(x^*, \langle x^*, -x \rangle)\right), \quad x \in \mathcal{X}.$$

Proof. By (2.1), Remark 2.11 and Lemma 2.13(i), we have

$$\begin{aligned} f \circ g(x) &= \inf\{m \in \mathbb{R} \mid g(x) \in S_m^f\} \\ &= \inf\{m \in \mathbb{R} \mid \forall y^* \in D^+ \setminus \{0\}: \langle y^*, -g(x) \rangle \leq \alpha_f(y^*, m)\} \\ &= \sup_{y^* \in D^+ \setminus \{0\}} \inf\{m \in \mathbb{R} \mid \langle y^*, -g(x) \rangle \leq \alpha_f(y^*, m)\}. \end{aligned} \tag{4.3}$$

By using (3.3) and then applying Lemma 2.13(i), we obtain

$$\begin{aligned}
f \circ g(x) &= \sup_{y^* \in D^+ \setminus \{0\}} \inf \{m \in \mathbb{R} \mid \langle y^*, -g(x) \rangle \leq \alpha_f(y^*, m)\} \\
&= \sup_{y^* \in D^+ \setminus \{0\}} \inf \left\{ m \in \mathbb{R} \mid \sup_{x^* \in C^+ \setminus \{0\}} \alpha_{-h_{y^*}^g}^{-l}(x^*, \langle -x^*, x \rangle) \leq \alpha_f(y^*, m) \right\} \\
&= \sup_{y^* \in D^+ \setminus \{0\}} \inf \left\{ m \in \mathbb{R} \mid \forall x^* \in C^+ \setminus \{0\}, \alpha_{-h_{y^*}^g}^{-l}(x^*, \langle -x^*, x \rangle) \leq \alpha_f(y^*, m) \right\} \\
&= \sup_{y^* \in D^+ \setminus \{0\}} \sup_{x^* \in C^+ \setminus \{0\}} \inf \left\{ m \in \mathbb{R} \mid \alpha_{-h_{y^*}^g}^{-l}(x^*, \langle -x^*, x \rangle) \leq \alpha_f(y^*, m) \right\} \\
&= \sup_{x^* \in C^+ \setminus \{0\}} \sup_{y^* \in D^+ \setminus \{0\}} \alpha_f^{-l} \left(y^*, \alpha_{-h_{y^*}^g}^{-l}(x^*, \langle x^*, -x \rangle) \right),
\end{aligned}$$

which gives the conclusion of the theorem. \square

The main problem is to calculate the penalty function $\alpha_{f \circ g}$ as well as its left inverse $\alpha_{f \circ g}^{-l}$ in terms of the same type of functions for f and g (more precisely, the scalarizations of g). The solution of this problem will be provided by Theorem 4.6 and Corollary 4.8. It turns out that these results work under a mild compactness assumption on D^+ as we describe next.

Definition 4.3. A set $\bar{D}^+ \subseteq D^+$ is called a cone generator for D^+ if every $y^* \in D^+ \setminus \{0\}$ can be written as $y^* = \lambda \bar{y}^*$ for some $\lambda > 0$ and $\bar{y}^* \in \bar{D}^+$.

It is clear that if \bar{D}^+ is a cone generator for D^+ , then D^+ is the conic hull of \bar{D}^+ .

Remark 4.4. Suppose that $D^\# \neq \emptyset$ and let $\pi \in D^\#$. Then, D_π^+ is a closed convex cone generator for D^+ thanks to Lemma 2.5.

In Section 5, we will discuss the existence and compactness of cone generators for several examples that show up frequently in applications. For the theoretical development of this section, we work under the following assumption.

Assumption 4.5. There exists a convex and compact cone generator \bar{D}^+ for D^+ .

Next, we state the main theorem of the paper, which provides a formula for the penalty function of $f \circ g$. Its proof is presented separately in Section 7.1. The proof consists of several auxiliary results together with the use of a minimax inequality in Liu [24] for two functions. Assumption 4.5 will be crucial in applying this inequality.

Theorem 4.6. *Suppose that Assumptions 3.9 and 4.5 hold. In addition, suppose that f is decreasing, lower semicontinuous, and quasiconvex; and that g is regularly increasing, D -lower demicontinuous, and D -naturally quasiconcave. Then, for every $x^* \in C^+ \setminus \{0\}$ and $m \in \mathbb{R}$, we have*

$$\alpha_{f \circ g}(x^*, m) = \inf_{y^* \in D^+ \setminus \{0\}} \alpha_{-h_{y^*}^g}(x^*, \alpha_f(y^*, m)) = \inf_{y^* \in D_\pi^+ \setminus \{0\}} \alpha_{-h_{y^*}^g}(x^*, \alpha_f(y^*, m)). \quad (4.4)$$

Remark 4.7. It should be noted that \bar{D}^+ does not have to be the same as D_π^+ but the second equality in (4.4) still holds.

The next corollary complements Theorem 4.6 by providing a formula for the left inverse of the penalty function of $f \circ g$, which is the actual function that shows up in the dual representation of $f \circ g$ in Proposition 4.1. Its proof is given in Section 7.1.

Corollary 4.8. *In the setting of Theorem 4.6, for every $x^* \in C^+ \setminus \{0\}$ and $s \in \mathbb{R}$, we have*

$$\alpha_{f \circ g}^{-l}(x^*, s) = \sup_{y^* \in D^+ \setminus \{0\}} \alpha_f^{-l}\left(y^*, \alpha_{-h_{y^*}^g}^{-l}(x^*, s)\right) \quad (4.5)$$

and

$$f \circ g(x) = \sup_{x^* \in C^+ \setminus \{0\}} \sup_{y^* \in D^+ \setminus \{0\}} \alpha_f^{-l}\left(y^*, \alpha_{-h_{y^*}^g}^{-l}(x^*, \langle x^*, -x \rangle)\right), \quad x \in \mathcal{X}. \quad (4.6)$$

Remark 4.9. The second part Corollary 4.8 gives the same result as Theorem 4.2. In the proof of Corollary 4.8, we use Theorem 4.6, a stronger result.

4.2 Two important special cases

We consider special cases of the setting in Section 4.1 where at least one of the functions in the composition is convex/concave. In these cases, we can obtain simplified formulae for the penalty function of the composition. As before, we work with two functions $f: \mathcal{Y} \rightarrow \overline{\mathbb{R}}$, $g: \mathcal{X} \rightarrow \mathcal{Y}$.

We first work on the case where both f and g satisfy a stronger convexity assumption so that $f \circ g$ becomes convex. As the next corollary shows, the reduced form of the dual representation is consistent with the ones available for convex compositions in the literature; see, for instance, Zălinescu [32, Thm. 2.8.10] and Boţ et al. [6, Thm. 3].

Corollary 4.10. *Suppose that $f: \mathcal{Y} \rightarrow \overline{\mathbb{R}}$ is convex, decreasing and lower semicontinuous; and that g is increasing, D -lower demicontinuous, and D -concave. Then, we have*

$$f \circ g(x) = \sup_{x^* \in C^+} \sup_{y^* \in D^+} \left(\langle x^*, -x \rangle - (-h_{y^*}^g)^*(-x^*) - f^*(-y^*) \right), \quad x \in \mathcal{X}.$$

Proof. Let $x \in \mathcal{X}$. First, we prove the following scaling property for arbitrary $\gamma \geq 0$, $x^* \in C^+ \setminus \{0\}$:

$$\gamma \alpha_{-h_{y^*}^g}^{-l} (x^*, \langle x^*, -x \rangle) = \alpha_{-h_{\gamma y^*}^g}^{-l} (x^*, \langle x^*, -x \rangle). \quad (4.7)$$

Let us consider the case $\gamma > 0$. By Definition 2.12, we have

$$\begin{aligned} \alpha_{-h_{\gamma y^*}^g}^{-l} (x^*, \langle x^*, -x \rangle) &= \inf \left\{ m \in \mathbb{R} \mid \alpha_{-h_{\gamma y^*}^g}^{-l} (x^*, m) \geq \langle x^*, -x \rangle \right\} \\ &= \inf \left\{ m \in \mathbb{R} \mid \sup_{z \in S_{-h_{\gamma y^*}^g}^m} \langle x^*, -z \rangle \geq \langle x^*, -x \rangle \right\}. \end{aligned}$$

We have the following relations:

$$z \in S_{-h_{\gamma y^*}^m} \Leftrightarrow \langle \gamma y^*, -g(z) \rangle \leq m \Leftrightarrow \langle y^*, -g(z) \rangle \leq \frac{m}{\gamma} \Leftrightarrow z \in S_{-h_{y^*}^{\frac{m}{\gamma}}}.$$

Therefore, we get

$$\begin{aligned} \alpha_{-h_{\gamma y^*}^g}^{-l} (x^*, \langle x^*, -x \rangle) &= \inf \left\{ m \in \mathbb{R} \mid \sup_{z \in S_{-h_{\gamma y^*}^m}^m} \langle x^*, -z \rangle \geq \langle x^*, -x \rangle \right\} \\ &= \inf \left\{ m \in \mathbb{R} \mid \sup_{z \in S_{-h_{y^*}^{\frac{m}{\gamma}}}^{\frac{m}{\gamma}}} \langle x^*, -z \rangle \geq \langle x^*, -x \rangle \right\} \\ &= \gamma \inf \left\{ n \in \mathbb{R} \mid \sup_{z \in S_{-h_{y^*}^n}^n} \langle x^*, -z \rangle \geq \langle x^*, -x \rangle \right\} = \gamma \alpha_{-h_{y^*}^g}^{-l} (x^*, \langle x^*, -x \rangle), \end{aligned}$$

where the third equality comes from the change-of-variables $\frac{m}{\gamma} = n$ and the last equality is by Definition 2.12. For the case $\gamma = 0$, we will prove that $\alpha_{-h_{\gamma y^*}^g}^{-l} (x^*, \langle x^*, -x \rangle) = 0$. Observe that $S_{-h_{\gamma y^*}^m}^m = \{z \in \mathcal{X} \mid \langle 0, g(z) \rangle \leq m\}$. Therefore, $S_{-h_{\gamma y^*}^m}^m = \mathcal{X}$ if $m \geq 0$, and $S_{-h_{\gamma y^*}^m}^m = \emptyset$ if $m < 0$.

Hence, for $\gamma = 0$,

$$\begin{aligned}\alpha_{-h_{\gamma y^*}^g}^{-l}(x^*, \langle x^*, -x \rangle) &= \inf \left\{ m \in \mathbb{R} \mid \sup_{z \in S_{-h_{\gamma y^*}^g}^m} \langle x^*, -z \rangle \geq \langle x^*, -x \rangle \right\} \\ &= \inf \left\{ m \geq 0 \mid \sup_{z \in \mathcal{X}} \langle x^*, -z \rangle \geq \langle x^*, -x \rangle \right\} = 0.\end{aligned}$$

We have proved (4.7), now we continue with the main result. By Theorem 4.2, we have

$$f \circ g(x) = \sup_{x^* \in C^+ \setminus \{0\}} \sup_{y^* \in D^+ \setminus \{0\}} \alpha_f^{-l} \left(y^*, \alpha_{-h_{y^*}^g}^{-l}(x^*, \langle x^*, -x \rangle) \right).$$

By applying the second part of Proposition 2.16 to f and using (4.7), we get

$$\begin{aligned}f \circ g(x) &= \sup_{x^* \in C^+ \setminus \{0\}} \sup_{y^* \in D^+ \setminus \{0\}} \sup_{\gamma \geq 0} \left(\gamma \alpha_{-h_{y^*}^g}^{-l}(x^*, \langle x^*, -x \rangle) - f^*(-\gamma y^*) \right) \\ &= \sup_{x^* \in C^+ \setminus \{0\}} \sup_{y^* \in D^+ \setminus \{0\}} \sup_{\gamma \geq 0} \left(\alpha_{-h_{\gamma y^*}^g}^{-l}(x^*, \langle x^*, -x \rangle) - f^*(-\gamma y^*) \right) \\ &= \sup_{x^* \in C^+ \setminus \{0\}} \sup_{\tilde{y}^* \in D^+} \left(\alpha_{-h_{\tilde{y}^*}^g}^{-l}(x^*, \langle x^*, -x \rangle) - f^*(-\tilde{y}^*) \right),\end{aligned}$$

where the last equality comes from the change-of-variables $\gamma y^* = \tilde{y}^*$ since D^+ is a cone. Let us apply Proposition 2.16 to $-h_{\tilde{y}^*}^g$ and obtain

$$\begin{aligned}f \circ g(x) &= \sup_{x^* \in C^+ \setminus \{0\}} \sup_{\tilde{y}^* \in D^+} \left(\alpha_{-h_{\tilde{y}^*}^g}^{-l}(x^*, \langle x^*, -x \rangle) - f^*(-\tilde{y}^*) \right) \\ &= \sup_{x^* \in C^+ \setminus \{0\}} \sup_{\tilde{y}^* \in D^+} \left(\sup_{\beta \geq 0} \left(\beta \langle x^*, -x \rangle - (-h_{\tilde{y}^*}^g)^*(-\beta x^*) \right) - f^*(-\tilde{y}^*) \right) \\ &= \sup_{\tilde{y}^* \in D^+} \sup_{x^* \in C^+ \setminus \{0\}} \sup_{\beta \geq 0} \left(\langle \beta x^*, -x \rangle - (-h_{\tilde{y}^*}^g)^*(-\beta x^*) - f^*(-\tilde{y}^*) \right) \\ &= \sup_{\tilde{y}^* \in D^+} \sup_{\tilde{x}^* \in C^+} \left(\langle \tilde{x}^*, -x \rangle - (-h_{\tilde{y}^*}^g)^*(-\tilde{x}^*) - f^*(-\tilde{y}^*) \right),\end{aligned}$$

where the last equality is by the change-of-variables $\beta x^* = \tilde{x}^*$ since C^+ is a cone. \square

Next, we work on the case where only one of the functions in the composition has a stronger convexity assumption. While Corollary 4.10 reproduces earlier results in the literature, the next result is novel to this work to the best of our knowledge. In Section 6.1, we will use this result to

obtain new dual representations for quasiconvex systemic risk measures.

Proposition 4.11. *Suppose that f is decreasing, lower semicontinuous, and quasiconvex; and that g is increasing, D -lower demicontinuous, and D -concave. Then, $f \circ g$ is an decreasing, lower semicontinuous, and quasiconvex function; and the following dual representation holds:*

$$f \circ g(x) = \sup_{x^* \in C^+ \setminus \{0\}} \sup_{y^* \in D^+ \setminus \{0\}} \alpha_f^{-l} \left(y^*, \langle x^*, -x \rangle - (-h_{y^*}^g)^*(-x^*) \right), \quad x \in \mathcal{X}. \quad (4.8)$$

Assume further that g is also regularly increasing and Assumption 4.5 holds. We have the following:

(i) Let $x^* \in C^+ \setminus \{0\}$, $m \in \mathbb{R}$ with $\alpha_f(y^*, m) \in \mathbb{R}$ and $S_{-h_{y^*}^g}^{\alpha_f(y^*, m)} \neq \emptyset$ for all $y^* \in D^+ \setminus \{0\}$. Then,

$$\alpha_{f \circ g}(x^*, m) = \inf_{y^* \in D^+ \setminus \{0\}} \left((-h_{y^*}^g)^*(-x^*) + \alpha_f(y^*, m) \right).$$

(ii) For every $x^* \in C^+ \setminus \{0\}$ and $s \in \mathbb{R}$,

$$\alpha_{f \circ g}^{-l}(x^*, s) = \sup_{y^* \in D^+ \setminus \{0\}} \alpha_f^{-l} \left(y^*, -(-h_{y^*}^g)^*(0) \vee (s - (-h_{y^*}^g)^*(-x^*)) \right).$$

Proof. Note that $x \mapsto \langle y^*, -g(x) \rangle$ is convex and lower semicontinuous by Propositions 3.2 and 3.12. By (4.3) in the proof of Theorem 4.2 and Fenchel-Moreau theorem, we have

$$\begin{aligned} f \circ g(x) &= \sup_{y^* \in D^+ \setminus \{0\}} \inf \{ m \in \mathbb{R} \mid \langle y^*, -g(x) \rangle \leq \alpha_f(y^*, m) \} \\ &= \sup_{y^* \in D^+ \setminus \{0\}} \inf \left\{ m \in \mathbb{R} \mid \sup_{x^* \in C^+ \setminus \{0\}} \left(\langle -x^*, x \rangle - (-h_{y^*}^g)^*(-x^*) \right) \leq \alpha_f(y^*, m) \right\} \\ &= \sup_{y^* \in D^+ \setminus \{0\}} \sup_{x^* \in C^+ \setminus \{0\}} \inf \left\{ m \in \mathbb{R} \mid \langle -x^*, x \rangle - (-h_{y^*}^g)^*(-x^*) \leq \alpha_f(y^*, m) \right\} \\ &= \sup_{x^* \in C^+ \setminus \{0\}} \sup_{y^* \in D^+ \setminus \{0\}} \alpha_f^{-l} \left(y^*, \langle x^*, -x \rangle - (-h_{y^*}^g)^*(-x^*) \right), \end{aligned}$$

where the third equality comes from Lemma 2.13(ii). Hence, (4.8) follows.

From now on, we assume that g is regularly increasing and Assumption 4.5 holds. To prove (i), let $x^* \in C^+ \setminus \{0\}$ and $m \in \mathbb{R}$ with $S_{-h_{y^*}^g}^{\alpha_f(y^*, m)} \neq \emptyset$. By Theorem 4.6, we have

$$\alpha_{f \circ g}(x^*, m) = \inf_{y^* \in D^+ \setminus \{0\}} \alpha_{-h_{y^*}^g}(x^*, \alpha_f(y^*, m)).$$

Also, take $x \in S_{-h_{y^*}^g}^{\alpha_f(y^*, m)}$ and let $c \in C^\#$. Then, there exists $d \in D^\#$ such that $g(x + c) = g(x) + d$ since g is regularly increasing. Therefore, by using the definition of $D^\#$, we get

$$\langle y^*, -g(x + c) \rangle = \langle -y^*, g(x) + d \rangle = \langle -y^*, g(x) \rangle + \langle -y^*, d \rangle < \langle -y^*, g(x) \rangle \leq \alpha_f(y^*, m),$$

which gives that $\{x \in \mathcal{X} \mid -h_{y^*}^g(x) < \alpha_f(y^*, m)\} \neq \emptyset$. Hence, by Proposition 2.16, we have

$$\alpha_{f \circ g}(x^*, m) = \inf_{y^* \in D^+ \setminus \{0\}} \inf_{\beta > 0} \left(\beta (-h_{y^*}^g)^* \left(-\frac{x^*}{\beta} \right) + \beta \alpha_f(y^*, m) \right).$$

Then, by Zălinescu [32, Thm. 2.3.1] on the elementary rules of conjugation, we have

$$\alpha_{f \circ g}(x^*, m) = \inf_{y^* \in D^+ \setminus \{0\}} \inf_{\beta > 0} \left((-\beta h_{y^*}^g)^*(-x^*) + \beta \alpha_f(y^*, m) \right).$$

By the positive homogeneity of $y^* \mapsto \alpha_f(y^*, m)$ and that of $y^* \mapsto h_{y^*}^g(x)$ for each $x \in \mathcal{X}$, we get

$$\alpha_{f \circ g}(x^*, m) = \inf_{y^* \in D^+ \setminus \{0\}} \inf_{\beta > 0} \left((-h_{\beta y^*}^g)^*(-x^*) + \alpha_f(\beta y^*, m) \right).$$

Finally, since D^+ is cone, we can make a change of variables and obtain (i).

We prove (ii) next. By Corollary 4.8, the second part of Proposition 2.16, and the definition of left inverse, we have

$$\begin{aligned} \alpha_{f \circ g}^{-l}(x^*, s) &= \sup_{y^* \in D^+ \setminus \{0\}} \alpha_f^{-l}\left(y^*, \alpha_{-h_{y^*}^g}^{-l}(x^*, s)\right) \\ &= \sup_{y^* \in D^+ \setminus \{0\}} \alpha_f^{-l}\left(y^*, \sup_{\gamma \geq 0} (\gamma s - (-h_{y^*}^g)^*(-\gamma x^*))\right) \\ &= \sup_{y^* \in D^+ \setminus \{0\}} \inf \left\{ m \in \mathbb{R} \mid \sup_{\gamma \geq 0} (\gamma s - (-h_{y^*}^g)^*(-\gamma x^*)) \leq \alpha_f(y^*, m) \right\} \\ &= \sup_{y^* \in D^+ \setminus \{0\}} \sup_{\gamma \geq 0} \inf \left\{ m \in \mathbb{R} \mid \gamma s - (-h_{y^*}^g)^*(-\gamma x^*) \leq \alpha_f(y^*, m) \right\}, \end{aligned}$$

where the last equality comes from Lemma 2.13(ii). By the conjugation formula, for $\gamma > 0$,

$$(-h_{y^*}^g)^*(-\gamma x^*) = \sup_{x \in \mathcal{X}} (\langle -\gamma x^*, x \rangle + \langle y^*, g(x) \rangle) = \gamma \sup_{x \in \mathcal{X}} \left(\langle -x^*, x \rangle + \left\langle \frac{y^*}{\gamma}, g(x) \right\rangle \right) = \gamma (-h_{\frac{y^*}{\gamma}}^g)^*(-x^*).$$

For $\gamma = 0$, we have

$$\begin{aligned} \inf \left\{ m \in \mathbb{R} \mid \gamma s - (-h_{y^*}^g)^*(-\gamma x^*) \leq \alpha_f(y^*, m) \right\} &= \inf \left\{ m \in \mathbb{R} \mid -(-h_{y^*}^g)^*(0) \leq \alpha_f(y^*, m) \right\} \\ &= \alpha_f^{-l}(y^*, -(-h_{y^*}^g)^*(0)). \end{aligned}$$

Therefore, by using the previous two equations and the positive homogeneity of α_f , we get

$$\begin{aligned} \alpha_{f \circ g}^{-l}(x^*, s) &= \sup_{y^* \in D^+ \setminus \{0\}} \sup_{\gamma \geq 0} \inf \left\{ m \in \mathbb{R} \mid \gamma s - (-h_{y^*}^g)^*(-\gamma x^*) \leq \alpha_f(y^*, m) \right\} \\ &= \sup_{y^* \in D^+ \setminus \{0\}} \left(\alpha_f^{-l}(y^*, -(-h_{y^*}^g)^*(0)) \vee \sup_{\gamma > 0} \inf \left\{ m \in \mathbb{R} \mid \gamma s - \gamma (-h_{\frac{y^*}{\gamma}}^g)^*(-x^*) \leq \alpha_f(y^*, m) \right\} \right) \\ &= \sup_{y^* \in D^+ \setminus \{0\}} \left(\alpha_f^{-l}(y^*, -(-h_{y^*}^g)^*(0)) \vee \sup_{\gamma > 0} \inf \left\{ m \in \mathbb{R} \mid s - (-h_{\frac{y^*}{\gamma}}^g)^*(-x^*) \leq \alpha_f \left(\frac{y^*}{\gamma}, m \right) \right\} \right) \\ &= \sup_{y^* \in D^+ \setminus \{0\}} \alpha_f^{-l}(y^*, -(-h_{y^*}^g)^*(0)) \vee \sup_{y^* \in D^+ \setminus \{0\}, \gamma > 0} \inf \left\{ m \in \mathbb{R} \mid s - (-h_{\frac{y^*}{\gamma}}^g)^*(-x^*) \leq \alpha_f \left(\frac{y^*}{\gamma}, m \right) \right\}. \end{aligned}$$

Hence,

$$\begin{aligned} \alpha_{f \circ g}^{-l}(x^*, s) &= \sup_{y^* \in D^+ \setminus \{0\}} \alpha_f^{-l}(y^*, -(-h_{y^*}^g)^*(0)) \vee \sup_{y^* \in D^+ \setminus \{0\}} \inf \left\{ m \in \mathbb{R} \mid (s - (-h_{y^*}^g)^*(-x^*)) \leq \alpha_f(y^*, m) \right\} \\ &= \sup_{y^* \in D^+ \setminus \{0\}} \alpha_f^{-l}(y^*, -(-h_{y^*}^g)^*(0)) \vee \sup_{y^* \in D^+ \setminus \{0\}} \alpha_f^{-l}(y^*, s - (-h_{y^*}^g)^*(-x^*)) \\ &= \sup_{y^* \in D^+ \setminus \{0\}} \left(\alpha_f^{-l}(y^*, -(-h_{y^*}^g)^*(0)) \vee \alpha_f^{-l}(y^*, s - (-h_{y^*}^g)^*(-x^*)) \right). \end{aligned}$$

By the monotonicity of α_f^{-l} , we can also write the last line as

$$\sup_{y^* \in D^+ \setminus \{0\}} \alpha_f^{-l}(y^*, -(-h_{y^*}^g)^*(0) \vee (s - (-h_{y^*}^g)^*(-x^*))),$$

which completes the proof. \square

4.3 Quasiconvex composition on a convex set

We turn our attention to the case where the composition is considered on a monotone convex set $\mathcal{K} \subseteq \mathcal{X}$ with $C \subseteq \mathcal{K}$, see Corollary 2.15, the analogous result for a single function. The treatment

here will be relevant for some applications in Section 6.

We work with two functions $f: \mathcal{Y} \rightarrow \overline{\mathbb{R}}$ and $g: \mathcal{K} \rightarrow \mathcal{Y}$. The following results extend Theorem 4.6 and Theorem 4.2. Their proofs are given in Section 7.1.

Corollary 4.12. *Suppose that f is decreasing, lower semicontinuous, and quasiconvex; and that g is regularly increasing, D -lower demicontinuous (with respect to the relative topology), and D -naturally quasiconcave. Then, $f \circ g$ is an decreasing, lower semicontinuous, and quasiconvex function. Moreover, for each $x^* \in C^+ \setminus \{0\}$ and $m \in \mathbb{R}$, we have*

$$\alpha_{f \circ g}(x^*, m) = \inf_{y^* \in D^+ \setminus \{0\}} \alpha_{-h_{y^*}^g}(x^*, \alpha_f(y^*, m)).$$

Proposition 4.13. *Suppose that f is decreasing, lower semicontinuous, and quasiconvex; and that g is increasing, D -lower demicontinuous (with respect to the relative topology), and D -naturally quasiconcave. Then, we have*

$$f \circ g(x) = \sup_{x^* \in C^+ \setminus \{0\}} \alpha_{f \circ g}^{-l}(x^*, \langle x^*, -x \rangle), \quad x \in \mathcal{K}, \quad (4.9)$$

and

$$f \circ g(x) = \sup_{x^* \in C^+ \setminus \{0\}} \sup_{y^* \in D^+ \setminus \{0\}} \alpha_f^{-l}\left(y^*, \alpha_{-h_{y^*}^g}^{-l}(x^*, \langle x^*, -x \rangle)\right), \quad x \in \mathcal{K}. \quad (4.10)$$

For a more specific case, we have the following proposition.

Proposition 4.14. *Suppose that f is decreasing, lower semicontinuous, and quasiconvex; and that g is increasing, D -lower demicontinuous (with respect to the relative topology), and concave. Then,*

$$f \circ g(x) = \sup_{x^* \in C^+ \setminus \{0\}} \sup_{y^* \in D^+ \setminus \{0\}} \alpha_f^{-l}\left(y^*, \langle x^*, -x \rangle - (-h_{y^*}^g)^*(-x^*)\right), \quad x \in \mathcal{K}. \quad (4.11)$$

5 Compact cone generators

In this section, we will discuss the existence of compact convex cone generators in some concrete spaces and show that Theorem 4.6 is applicable in these examples. This will justify the use of our results in the context of systemic risk measures in Section 6.

As noted in Remark 4.4, D_π^+ is a closed convex generator but it is not always compact. However,

we do not have to restrict ourselves to this generator and can search for other compact generators because after guaranteeing the existence of a compact convex cone generator \bar{D}^+ , we can still work with D_π^+ thanks to the second part of (4.4) in Theorem 4.6.

5.1 Finite-dimensional spaces

Let us take $\mathcal{Y} = \mathbb{R}^n$ with the Euclidean norm $\|\cdot\|$. As a natural consequence, $\mathcal{Y}^* = \mathbb{R}^n$ with the same norm $\|\cdot\|$. Let us choose a convex cone D and denote the unit ball by $B = \{y \in \mathbb{R}^n : \|y\| \leq 1\}$. We show the existence of a compact convex generator for D^+ so that we can use Theorem 4.6 for the case $\mathcal{Y} = \mathbb{R}^n$.

Proposition 5.1. *The set $\bar{D}^+ := D^+ \cap B$ is a compact and convex cone generator for D^+ .*

Proof. Since D^+ and B are closed and convex sets, their intersection is also closed and convex. Moreover, B is compact since it is closed and bounded. By using this fact and that \bar{D}^+ is a closed subset of B , we conclude that \bar{D}^+ is compact. To show that \bar{D}^+ generates D^+ , let us take $y^* \in D^+ \setminus \{0\}$. We have $\frac{y^*}{\|y^*\|} \in D^+$ since D^+ is a cone and $\left\| \frac{y^*}{\|y^*\|} \right\| = 1$, which implies that $\frac{y^*}{\|y^*\|} \in B$ and hence $\frac{y^*}{\|y^*\|} \in \bar{D}^+$. We can write $y^* = \|y^*\| \frac{y^*}{\|y^*\|}$ where $\|y^*\| > 0$ and $\frac{y^*}{\|y^*\|} \in \bar{D}^+$; hence, \bar{D}^+ is a cone generator for D^+ . \square

5.2 Lebesgue spaces

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, and let $p \in [1, +\infty]$, $n \in \mathbb{N}$. We denote by $L^0(\mathbb{R}^n)$ the space of all n -dimensional random vectors that are identified up to \mathbb{P} -almost sure equality. We denote by $L^p(\mathbb{R}^n)$ the space of all $X \in L^0(\mathbb{R}^n)$ such that $\|X\|_p < +\infty$, where $\|X\|_p := (\mathbb{E}[\|X\|^p])^{1/p}$ for $p < +\infty$ and $\|X\|_p := \inf\{c > 0 \mid \mathbb{P}\{\|X\| \leq c\} = 1\}$ for $p = +\infty$. For $p \in \{0\} \cup [1, +\infty]$ and a set $A \subseteq \mathbb{R}^n$, we denote by $L^p(A)$ the set of all $X \in L^p(\mathbb{R}^n)$ such that $\mathbb{P}\{X \in A\} = 1$.

In this section, we fix $p \in [1, +\infty)$ and consider the case $\mathcal{Y} = L^p(\mathbb{R}^n)$, which is equipped with the norm $\|\cdot\|_p$ and the induced topology. Then, $\mathcal{Y}^* = L^q(\mathbb{R}^n)$ with the norm $\|\cdot\|_q$ and we consider it with the topology $\sigma(\mathcal{Y}^*, \mathcal{Y})$, where $q \in (1, +\infty]$ is defined by $\frac{1}{p} + \frac{1}{q} = 1$. Let $D \subseteq \mathcal{Y}$ be a closed convex cone and denote the unit ball in $L^q(\mathbb{R}^n)$ by $B_q^n = \{Y^* \in L^q(\mathbb{R}^n) \mid \|Y^*\|_q \leq 1\}$. We show the existence of a compact convex cone generator for D^+ next.

Proposition 5.2. *The set $\bar{D}^+ := D^+ \cap B_q^n$ is a compact and convex cone generator for D^+ .*

Proof. Since D^+ and B_q^n are closed convex sets, so is their intersection \bar{D}^+ . Also, B_q^n is (weakly) compact by Banach-Alaoglu Theorem (Reed and Simon [27, Thm. IV.21]). By using this fact and that \bar{D}^+ is a closed subset of B_q , we conclude that \bar{D}^+ is also compact. The proof of the claim that \bar{D}^+ is a cone generator for D^+ is similar to the proof of Proposition 5.1, hence omitted. \square

Let us consider the special case $n = 1$ and take $D = L^p(\mathbb{R}_+)$ which is the set of all almost surely positive elements of $L^p(\mathbb{R})$. Then, $D^+ = L^q(\mathbb{R}_+)$. Also, we can take $\pi \equiv 1$ and get $D_1^+ \cong \mathcal{M}_1^q(\mathbb{P})$, where $\mathcal{M}_1^q(\mathbb{P})$ denotes the set of all probability measures \mathbb{Q} on (Ω, \mathcal{F}) that are absolutely continuous with respect to \mathbb{P} with Radon-Nikodym derivatives $\frac{d\mathbb{Q}}{d\mathbb{P}}$ in $L^q(\mathbb{R}_+)$. Therefore, the formula in Theorem 4.6 can be rewritten as

$$\alpha_{f \circ g}(X^*, m) = \inf_{\mathbb{Q} \in \mathcal{M}_1^q(\mathbb{P})} \alpha_{-h^g} \left(X^*, \alpha_f \left(\frac{d\mathbb{Q}}{d\mathbb{P}}, m \right) \right). \quad (5.1)$$

We can work with any closed convex cone generator after guaranteeing the existence of a compact convex cone generator since we do not need the compactness of D_π^+ for the second equality in Theorem 4.6. Therefore, the dual representation in Theorem 4.6 can be written as in (5.1) since D_1^+ is a closed convex cone generator by Remark 4.4.

6 Applications to systemic risk measures

In this section, we will explore the implications of the general theory developed in Section 4 on some quasiconvex risk measures for interconnected financial systems. Such risk measures are referred to as *systemic risk measures*, which are of recent interest in financial mathematics. We refer the reader to Chen et al. [10], Biagini et al. [5], Feinstein et al. [14], Ararat and Rudloff [3] for detailed discussions on this subject.

Throughout this section, we fix a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. The proofs of the results in this section are given in Section 7.2.

6.1 General results on quasiconvex systemic risk measures

We consider an interconnected financial system with $n \in \mathbb{N}$ institutions in a static setting. Due to their financial activities, the assets of the institutions are subject to uncertainty. Consequently,

the future values of the assets of all institutions can be modeled as a random vector $X \in L^0(\mathbb{R}^n)$, which is sometimes called a *random shock*. A systemic risk measure quantifies the overall risk of the system by taking into account the correlations between the components of the random shock as well as the underlying structure of the system. In line with Chen et al. [10] and Biagini et al. [5], we study systemic risk measures of the form

$$R(X) = \rho(\tilde{\Lambda} \circ X), \quad (6.1)$$

where $\tilde{\Lambda}: \mathbb{R}^n \rightarrow \mathbb{R}$ is an aggregation function and ρ is a risk measure, see Definition 6.1 below for the precise description of these term. The aggregation function produces a univariate quantity $\tilde{\Lambda} \circ X \in L^0(\mathbb{R})$ that summarizes the impact of the random shock on the economy (or society), which can be seen as an external entity of the system. The risk of this aggregate quantity is then evaluated through the univariate functional ρ and the output $\rho(\tilde{\Lambda} \circ X)$ is the risk associated to the overall system when it faces random shock X .

To view the structure of R in (6.1) as a composition of two functions, we may simply define the functional version $\Lambda: L^0(\mathbb{R}^n) \rightarrow L^0(\mathbb{R})$ of the aggregation function via $\Lambda(X) := \tilde{\Lambda} \circ X$, that is,

$$\Lambda(X)(\omega) := \tilde{\Lambda}(X(\omega)), \quad \omega \in \Omega. \quad (6.2)$$

Then, (6.1) can be rewritten as

$$R = \rho \circ \Lambda. \quad (6.3)$$

To obtain dual representations for systemic risk measures of the (6.3), we will consider random shocks that are sufficiently integrable. As in Section 5.2, we choose $\mathcal{X} = L^p(\mathbb{R}^n)$ and $\mathcal{Y} = L^p(\mathbb{R})$, where $p \in [1, +\infty]$. These spaces are equipped with their norm topologies when $p < +\infty$ and with weak* topologies when $p = +\infty$. In all cases, we have $\mathcal{X}^* = L^q(\mathbb{R}^n)$ and $\mathcal{Y}^* = L^q(\mathbb{R})$, with their weak topologies, where $q \in [1, +\infty]$ is determined by $\frac{1}{p} + \frac{1}{q} = 1$. We denote by $\mathcal{M}_n^q(\mathbb{P})$ the set of all vectors $\mathbb{S} = (\mathbb{S}_1, \dots, \mathbb{S}_n)$, where \mathbb{S}_i is a probability measure on (Ω, \mathcal{F}) that is absolutely continuous with respect to \mathbb{P} and $\frac{d\mathbb{S}_i}{d\mathbb{P}} \in L^q(\mathbb{R}_+)$ for each $i \in \{1, \dots, n\}$. We take $C = L^p(\mathbb{R}_+^n)$ and $D = L^p(\mathbb{R}_+)$; hence, the dual cones are given by $C^+ = L^q(\mathbb{R}_+^n)$ and $D^+ = L^q(\mathbb{R}_+)$. With this choice of D , for the sake of convenience, we remove D from the terminology; for instance, we simply call a function

concave if it is D -concave.

The formal definitions of aggregation function and risk measure are given next.

Definition 6.1. (i) A function $\tilde{\Lambda}: \mathbb{R}^n \rightarrow \mathbb{R}$ is called an aggregation function if it is increasing (with respect to \mathbb{R}_+^n and \mathbb{R}_+) and its functional version Λ defined by (6.2) satisfies the following condition: $\Lambda(X) \in L^p(\mathbb{R})$ for every $X \in L^p(\mathbb{R})$. (ii) A function $\rho: L^p(\mathbb{R}) \rightarrow \overline{\mathbb{R}}$ is called a quasiconvex risk measure if it is quasiconvex and decreasing. (iii) A function $R: L^p(\mathbb{R}^n) \rightarrow \overline{\mathbb{R}}$ is called a systemic risk measure if it is of the form (6.1), where $\tilde{\Lambda}$ is an aggregation function and ρ is a quasiconvex risk measure.

Consider a systemic risk measure $R = \rho \circ \Lambda$ as in Definition 6.1. In view of Proposition 3.4, R is quasiconvex whenever Λ is naturally quasiconcave. We are particularly interested in the special case where Λ is concave. As we will illustrate in Section 6.2, such aggregation functions appear frequently in concrete examples. On the other hand, to ensure the lower demicontinuity of Λ , we need to impose sufficient regularity on $\tilde{\Lambda}$. This is done in the following lemma.

Lemma 6.2. Let $\tilde{\Lambda}: \mathbb{R}^n \rightarrow \mathbb{R}$ be an aggregation function and define Λ by (6.2).

- (i) If $\tilde{\Lambda}$ is concave and bounded from above, then Λ is concave and lower demicontinuous.
- (ii) If $\tilde{\Lambda}$ is linear, then Λ is linear and lower demicontinuous.
- (iii) If $\tilde{\Lambda}$ is regularly increasing (with respect to \mathbb{R}_+^n and \mathbb{R}_+), then Λ is regularly increasing.

In the next proposition, we calculate the penalty function of a systemic risk measure when the aggregation function is concave and regularly increasing, and the univariate risk measure is quasiconvex and lower semicontinuous. It should be noted that, in Ararat and Rudloff [3], dual representations are provided for convex systemic risk measures, where ρ is further assumed to be a convex (translative) risk measure. Hence, our results will extend these representations to the quasiconvex case. For convenience, we define the conjugate function $\tilde{\Phi}$ by

$$\tilde{\Phi}(x^*) := (-\tilde{\Lambda})^*(-x^*) = \sup_{x \in \mathbb{R}^n} \left(\Lambda(x) - (x^*)^\top x \right), \quad x^* \in \mathbb{R}^n, \quad (6.4)$$

Similar to (6.2), we also define the functional version Φ of $\tilde{\Phi}$ by

$$\Phi(X^*) := \tilde{\Phi} \circ X^*, \quad X^* \in L^q(\mathbb{R}^n). \quad (6.5)$$

Moreover, for each $X^* \in L^q(\mathbb{R}^n)$, we introduce the set

$$T_{X^*} := \{Y^* \in L^q(\mathbb{R}_+) \mid \mathbb{P}\{X^* \neq 0, Y^* = 0\} = 0\}. \quad (6.6)$$

Proposition 6.3. *Assume that $p \in [1, +\infty)$. Let $\tilde{\Lambda}: \mathbb{R}^n \rightarrow \mathbb{R}$ be a concave, regularly increasing aggregation function that is either bounded from above or linear. Let Λ be defined by (6.2). Let ρ be a lower semicontinuous quasiconvex risk measure. Let $X^* \in L^q(\mathbb{R}^n)$ and $m \in \mathbb{R}$ such that the strict sublevel set $\{X \in L^p(\mathbb{R}^n) \mid \mathbb{E}[-Y^*\Lambda(X)] < m\}$ is nonempty for every $Y^* \in L^q(\mathbb{R}_+) \setminus \{0\}$. Then,*

$$\alpha_{\rho \circ \Lambda}(X^*, m) = \inf_{Y^* \in T_{X^*}} \left(\mathbb{E} \left[Y^* \Phi \left(\frac{X^*}{Y^*} \right) 1_{\{Y^* > 0\}} \right] + \alpha_\rho(Y^*, m) \right).$$

Next, we aim to rewrite the formula in Proposition 6.3 in terms of probability measures. This reformulation will make it possible to provide economic interpretations of the dual representation in view of model uncertainty. Since $D_1^+ = \{\frac{d\mathbb{Q}}{d\mathbb{P}} \mid \mathbb{Q} \in \mathcal{M}_1^q(\mathbb{P})\}$ is a closed convex cone generator for $D^+ = L^q(\mathbb{R}_+)$, we can write every $Y^* \in L^q(\mathbb{R}_+) \setminus \{0\}$ as $Y^* = \lambda \frac{d\mathbb{Q}}{d\mathbb{P}}$ for some $\lambda > 0$ and $\mathbb{Q} \in \mathcal{M}_1^q(\mathbb{P})$ by Remark 4.4. Similarly, every $X^* \in C^+ = L^q(\mathbb{R}_+^n)$ can be written as $X^* = w \cdot \frac{d\mathbb{S}}{d\mathbb{P}}$, where $w \in \mathbb{R}_+^n \setminus \{0\}$, $\mathbb{S} = (\mathbb{S}_1, \dots, \mathbb{S}_n) \in \mathcal{M}_n^q(\mathbb{P})$, and $w \cdot \frac{d\mathbb{S}}{d\mathbb{P}} := (w_1 \frac{d\mathbb{S}_1}{d\mathbb{P}}, \dots, w_n \frac{d\mathbb{S}_n}{d\mathbb{P}})$. The interpretation of these dual variables is as follows. In the presence of model uncertainty, we consider \mathbb{Q} as a probability measure that is assigned to an external entity, e.g., society, and, for each $i \in \{1, \dots, n\}$, \mathbb{S}_i is a probability measure that is assigned to internal entity i , e.g., a bank in the network, with corresponding weight w_i . Moreover, since we consider X^* and Y^* satisfying the condition $\mathbb{P}\{X^* \neq 0, Y^* = 0\} = 0$ in Proposition 6.3, it follows from Ararat and Rudloff [3, Lem. 6.3] that $w_i \mathbb{S}_i$ is a finite measure that is absolutely continuous with respect to \mathbb{Q} , and we can write

$$\frac{w \cdot d\mathbb{S}}{d\mathbb{P}} = \frac{w \cdot d\mathbb{S}}{d\mathbb{Q}},$$

where all Radon-Nikodym derivatives are well-defined. Therefore, in probabilistic terms, the formula in Proposition 6.3 can be rewritten as

$$\alpha_{\rho \circ \Lambda} \left(w \cdot \frac{d\mathbb{S}}{d\mathbb{P}}, m \right) = \inf_{\substack{\lambda > 0, \mathbb{Q} \in \mathcal{M}_1^q(\mathbb{P}), \\ w_i \mathbb{S}_i \ll \mathbb{Q}}} \left(\mathbb{E}_{\mathbb{Q}} \left[\lambda \Phi \left(\frac{w \cdot d\mathbb{S}}{\lambda d\mathbb{Q}} \right) \right] + \lambda \alpha_\rho \left(\frac{d\mathbb{Q}}{d\mathbb{P}}, m \right) \right). \quad (6.7)$$

According to (6.7), the total penalty of choosing probability vector \mathbb{S} and weight vector w for the financial institutions is calculated by considering all possible choices of society's probability measure \mathbb{Q} and an associated weight λ . As in the convex case studied in Ararat and Rudloff [3], \mathbb{Q} is chosen from an absolute continuity interval $w_i \mathbb{S}_i \ll \mathbb{Q} \ll \mathbb{P}$, $i \in \{1, \dots, n\}$, determined by $w \cdot \mathbb{S}$. The objective function of minimization in (6.7) can be seen as a directed distance from $w \cdot \mathbb{S}$ to \mathbb{P} that is calculated through society's measure \mathbb{Q} . The first term is the multivariate divergence of $w \cdot \mathbb{S}$ relative to \mathbb{Q} . The divergence function is determined by the structure of the network, see Section 6.2 and Section 6.3 for concrete calculations, Moreover, this function is scaled by the weight $\lambda > 0$ through $\lambda \Phi(\frac{\cdot}{\lambda})$, which is the conjugate function corresponding to $\lambda \Lambda(\cdot)$. In other words, society's weight λ amplifies/shrinks the impact of the shock to society as a factor. The second term of the objective function is the penalty of choosing \mathbb{Q} with respect to the physical measure \mathbb{P} in the presence of model uncertainty, which is quantified by the choice of the univariate risk measure ρ . Hence, the overall penalty is calculated as the least possible sum of these two distance terms. It is notable that the objective function of the penalty function has an additive structure in our quasiconvex framework, which generalizes the observations in Ararat and Rudloff [3] for the convex case.

As a continuation of Proposition 6.3, we calculate the inverse of the penalty function in the next proposition.

Proposition 6.4. *Assume that $p \in [1, +\infty)$. Let $\tilde{\Lambda}: \mathbb{R}^n \rightarrow \mathbb{R}$ be a concave, regularly increasing aggregation function. Let Λ be defined by (6.2). Let ρ be a lower semicontinuous quasiconvex risk measure.*

(i) *Suppose that $\tilde{\Lambda}$ is bounded from above, that is, $\tilde{\Phi}(0) < +\infty$. Then, we have*

$$\alpha_{\rho \circ \Lambda}^{-l}(X^*, s) = \sup_{Y^* \in L_+^q(\mathbb{R}) \setminus \{0\}} \alpha_{\rho}^{-l}(Y^*, -\Phi(0)\mathbb{E}[Y^*]) \vee \sup_{Y^* \in T_{X^*}} \alpha_{\rho}^{-l}\left(Y^*, s - \mathbb{E}\left[Y^* \Phi\left(\frac{X^*}{Y^*}\right) 1_{\{Y^* > 0\}}\right]\right),$$

where T_{X^*} is defined by (6.6). In particular, when we transform the variables into the probabilistic setting, we get

$$\alpha_{\rho \circ \Lambda}^{-l}\left(w \cdot \frac{d\mathbb{S}}{d\mathbb{P}}, s\right) = \sup_{\mathbb{Q} \in \mathcal{M}_1^q(\mathbb{P})} \alpha_{\rho}^{-l}\left(\frac{d\mathbb{Q}}{d\mathbb{P}}, -\Phi(0)\right) \vee \sup_{\substack{\mathbb{Q} \in \mathcal{M}_1^q(\mathbb{P}), \lambda > 0 \\ w_i \mathbb{S}_i \ll \mathbb{Q}}} \alpha_{\rho}^{-l}\left(\frac{d\mathbb{Q}}{d\mathbb{P}}, \frac{s}{\lambda} - \mathbb{E}_{\mathbb{Q}}\left[\Phi\left(\frac{w \cdot d\mathbb{S}}{\lambda d\mathbb{Q}}\right)\right]\right).$$

(ii) Suppose that $\tilde{\Lambda}$ is linear and it is unbounded from above, that is, $\tilde{\Phi}(0) = +\infty$. Then, we have

$$\alpha_{\rho \circ \Lambda}^{-l}(X^*, s) = \sup_{Y^* \in T_{X^*}} \alpha_{\rho}^{-l} \left(Y^*, s - \mathbb{E} \left[Y^* \Phi \left(\frac{X^*}{Y^*} \right) 1_{\{Y^* > 0\}} \right] \right),$$

and

$$\alpha_{\rho \circ \Lambda}^{-l} \left(w \cdot \frac{d\mathbb{S}}{d\mathbb{P}}, s \right) = \sup_{\substack{\mathbb{Q} \in \mathcal{M}_1^q(\mathbb{P}), \lambda > 0 \\ w_i \mathbb{S}_i \ll \mathbb{Q}}} \alpha_{\rho}^{-l} \left(\frac{d\mathbb{Q}}{d\mathbb{P}}, \frac{s}{\lambda} - \mathbb{E}_{\mathbb{Q}} \left[\Phi \left(\frac{w \cdot d\mathbb{S}}{\lambda d\mathbb{Q}} \right) \right] \right).$$

In the next proposition, we give a dual representation for quasiconvex systemic risk measures. Unlike Propositions 6.3 and 6.4, we allow for $p = +\infty$ here as we do not rely on the expression for the penalty function (hence not on the existence of a compact cone generator).

Proposition 6.5. *Assume that $p \in [1, +\infty)$. Let $\tilde{\Lambda}: \mathbb{R}^n \rightarrow \mathbb{R}$ be a concave aggregation function that is either bounded from above or linear. Let Λ be defined by (6.2). Let ρ be a lower semicontinuous quasiconvex risk measure. Then, we have*

$$R(X) = \rho \circ \Lambda(X) = \sup_{\substack{w \in \mathbb{R}_+^n \setminus \{0\}, \mathbb{S} \in \mathcal{M}_n^q(\mathbb{P}) \\ \mathbb{Q} \in \mathcal{M}_1^q(\mathbb{P}), w_i \mathbb{S}_i \ll \mathbb{Q}}} \alpha_{\rho}^{-l} \left(\frac{d\mathbb{Q}}{d\mathbb{P}}, -\mathbb{E}_{\mathbb{Q}} \left[\Phi \left(\frac{w \cdot d\mathbb{S}}{d\mathbb{Q}} \right) \right] - w^{\top} \mathbb{E}_{\mathbb{S}}[X] \right) \quad (6.8)$$

for every $X \in L^p(\mathbb{R}^n)$.

While the objective function of the penalty function has an additive structure in Proposition 6.3, we see in Proposition 6.4 that this might not be the case for its inverse. In other words, the inverse penalty function of ρ and the divergence term including Φ might interact in a non-additive way. We will see such cases in Section 6.2. Consequently, due to Proposition 6.5, the same structure also shows up in the final dual representation of the systemic risk measure. This is contrary to the convex framework of Ararat and Rudloff [3], where the penalty function directly appears in the dual representation of a convex systemic risk measure. Hence, our results shed light on a new feature of quasiconvex systemic risk measures that does not exist in convex systemic risk measures.

6.2 Examples

In this section, we first recall some examples of quasiconvex risk measures and concave aggregation functions studied in the literature. Then, we will combine some choices of these two functions and

illustrate the forms of the penalty functions and dual representations of the resulting systemic risk measures.

We start by recalling two families of quasiconvex lower semicontinuous risk measures studied in Drapeau and Kupper [12]. The first family consists of functionals of the form

$$\rho(Y) = \ell^{-1}(\mathbb{E}[\ell \circ (-Y)]), \quad Y \in L^p(\mathbb{R}),$$

where $p \in [1, +\infty]$, and $\ell: \mathbb{R} \rightarrow (-\infty, \infty]$ is a proper lower semicontinuous convex increasing function, called a *loss function*. For simplicity, we assume that ℓ is differentiable. Such ρ is called the *certainty equivalent* associated to ℓ . It is calculated in Drapeau and Kupper [12] that

$$\alpha_\rho\left(\frac{d\mathbb{Q}}{d\mathbb{P}}, m\right) = \mathbb{E}_{\mathbb{Q}}\left[h \circ \left(\beta \frac{d\mathbb{Q}}{d\mathbb{P}}\right)\right], \quad \mathbb{Q} \in \mathcal{M}_1^q(\mathbb{P}), m \in \mathbb{R},$$

where h is the right inverse of the derivative ℓ' , and $\beta = \beta(\mathbb{Q}, m)$ is the solution of the equation $\mathbb{E}[\ell \circ h \circ (\beta \frac{d\mathbb{Q}}{d\mathbb{P}})] = \ell^+(m)$ under some integrability and positivity conditions.

Let us provide some concrete examples of the loss function ℓ and recall the penalty functions for the corresponding certainty equivalents, already calculated in Drapeau and Kupper [12, Ex. 8].

Example 6.6. (i) (Quadratic loss function) Let us take $p = 2$, and $\ell(s) = s^2/2 + s$ for $s \geq -1$, $\ell(s) = -\frac{1}{2}$ for $s < -1$. Then, for each $\mathbb{Q} \in \mathcal{M}_1^2(\mathbb{P})$, we have $\alpha_\rho(\frac{d\mathbb{Q}}{d\mathbb{P}}, m) = -1$ for $m \geq -1$ and

$$\alpha_\rho\left(\frac{d\mathbb{Q}}{d\mathbb{P}}, m\right) = (1+m) \left\| \frac{d\mathbb{Q}}{d\mathbb{P}} \right\|_2 - 1, \quad m < -1, \quad \alpha_\rho^{-l}\left(\frac{d\mathbb{Q}}{d\mathbb{P}}, s\right) = \frac{s+1}{\left\| \frac{d\mathbb{Q}}{d\mathbb{P}} \right\|_2} - 1, \quad s < -1.$$

(ii) (Logarithmic loss function) Let us take $p = 1$ or $p = +\infty$, and $\ell(s) = -\ln(-s)$ for $s < 0$, $\ell(s) = +\infty$ for $s \geq 0$. Then, for each $\mathbb{Q} \in \mathcal{M}_1^q(\mathbb{P})$,

$$\alpha_\rho\left(\frac{d\mathbb{Q}}{d\mathbb{P}}, m\right) = m e^{\mathbb{E}[\ln(\frac{d\mathbb{Q}}{d\mathbb{P}})]}, \quad m < 0, \quad \alpha_\rho^{-l}\left(\frac{d\mathbb{Q}}{d\mathbb{P}}, s\right) = s e^{-\mathbb{E}[\ln(\frac{d\mathbb{Q}}{d\mathbb{P}})]}, \quad s < 0.$$

(iii) (Power loss function) Let us take $p = 1$ or $p = +\infty$, and fix some $\gamma \in (0, 1)$. Take $\ell(s) =$

$-\frac{(-s)^{1-\gamma}}{1-\gamma}$ for $s \leq 0$, $\ell(s) = \infty$ for $s > 0$. Then, for each $\mathbb{Q} \in \mathcal{M}_1^q(\mathbb{P})$,

$$\alpha_\rho \left(\frac{d\mathbb{Q}}{d\mathbb{P}}, m \right) = \frac{m}{\left\| \frac{d\mathbb{Q}}{d\mathbb{P}} \right\|_{\frac{\gamma-1}{\gamma}}}, \quad m < 0, \quad \alpha_\rho^{-1} \left(\frac{d\mathbb{Q}}{d\mathbb{P}}, s \right) = s \left\| \frac{d\mathbb{Q}}{d\mathbb{P}} \right\|_{\frac{\gamma-1}{\gamma}}, \quad s < 0.$$

Here, for $Y^* \in L^1(\mathbb{R})$, we use the notation $\|Y^*\|_a := (\mathbb{E}[|Y^*|^a])^{\frac{1}{a}}$ for $a < 1$ as well, although $\|\cdot\|_a$ is not a norm in general.

We also revisit the *economic index of riskiness* as another example of a quasiconvex risk measure. Based on a loss function ℓ as before, this risk measure is defined by

$$\rho(Y) = \frac{1}{\sup\{\lambda > 0 \mid \mathbb{E}[\ell \circ (-\lambda Y)] \leq c_0\}}, \quad Y \in L^p(\mathbb{R}),$$

where $c_0 \in \mathbb{R}$ is a fixed threshold for expected loss levels. To make this risk measure well-defined, ℓ is usually assumed to have the superlinear growth condition $\lim_{s \rightarrow \infty} \ell(s)/s = \infty$ and p is chosen in accordance with ℓ . Following the arguments in Drapeau and Kupper [12], it can be shown that

$$\alpha_\rho \left(\frac{d\mathbb{Q}}{d\mathbb{P}}, m \right) = \mathbb{E}_{\mathbb{Q}} \left[mh \circ \left(m\beta \frac{d\mathbb{Q}}{d\mathbb{P}} \right) \right], \quad \mathbb{Q} \in \mathcal{M}_1^q(\mathbb{P}), m \in \mathbb{R},$$

where $\beta = \beta(\mathbb{Q}, m)$ is the solution of the equation $\mathbb{E}[\ell \circ h \circ (m\beta \frac{d\mathbb{Q}}{d\mathbb{P}})] = c_0$.

The following example is the analogue of Example 6.6(ii) for the economic index of riskiness; see Drapeau and Kupper [12, Ex. 3, 9] for more details.

Example 6.7. Let us take $p = 1$ and $c_0 > 0$, and consider $\ell(s) = -\ln(1-s)$ for $s < 1$, $\ell(s) = +\infty$ for $s \geq 1$. Then, for each $\mathbb{Q} \in \mathcal{M}_1^\infty(\mathbb{P})$, $m < 0$, $s < 0$, we have

$$\alpha_\rho \left(\frac{d\mathbb{Q}}{d\mathbb{P}}, m \right) = m \left(1 - \exp \left(\mathbb{E} \left[\ln \left(\frac{d\mathbb{Q}}{d\mathbb{P}} \right) \right] - c_0 \right) \right), \quad \alpha_\rho^{-1} \left(\frac{d\mathbb{Q}}{d\mathbb{P}}, s \right) = \frac{s}{1 - \exp \left(\mathbb{E} \left[\ln \left(\frac{d\mathbb{Q}}{d\mathbb{P}} \right) \right] - c_0 \right)},$$

where $\exp(x) = e^x$ for $x \in \mathbb{R}$.

Next, we recall some examples of concave aggregation functions from Ararat and Rudloff [3, Sect. 4]. In each example, we calculate the conjugate function $\tilde{\Phi}$ given by (6.4). A more sophisticated aggregation function based on a clearing mechanism will be discussed separately in Section 6.3.

Example 6.8. (i) (Total profit-loss model) Let us take $\tilde{\Lambda}(x) = \sum_{i=1}^n x_i$ for each $x \in \mathbb{R}^n$. Then,

$$\tilde{\Phi}(x^*) = \begin{cases} 0 & \text{if } x^* = \mathbf{1}, \\ \infty & \text{else.} \end{cases}$$

The condition that $\Lambda(X) \in L^p(\mathbb{R})$ for every $X \in L^p(\mathbb{R}^n)$ is satisfied for every $p \in [1, +\infty]$.

(ii) (Total loss model) Let us take $\tilde{\Lambda}(x) = -\sum_{i=1}^n x_i^-$ for each $x \in \mathbb{R}^n$. Then,

$$\tilde{\Phi}(x^*) = \begin{cases} 0 & \text{if } x_i^* \in [0, 1] \text{ for every } i \in \{1, \dots, n\}, \\ \infty & \text{else.} \end{cases}$$

As in (i), for every choice of $p \in [1, +\infty]$, we have $\Lambda(X) \in L^p(\mathbb{R})$ for every $X \in L^p(\mathbb{R}^n)$.

(iii) (Exponential model) Let us take $\tilde{\Lambda}(x) = -\sum_{i=1}^n e^{-x_i-1}$ for each $x \in \mathbb{R}^n$. Then,

$$\tilde{\Phi}(x^*) = \sum_{i=1}^n x_i^* \ln(x_i^*),$$

where $\ln(0) := -\infty$ and $0 \ln(0) := 0$ as conventions. The condition that $\Lambda(X) \in L^p(\mathbb{R})$ for every $X \in L^p(\mathbb{R}^n)$ is satisfied only for $p = +\infty$. As a result, Propositions 6.3 and 6.4 is not applicable. However, we can still use the dual representation in Proposition 6.5.

Thanks to Lemma 6.2, each aggregation function $\tilde{\Lambda}$ above yields a lower demicontinuous concave functional version Λ via (6.2). In (i) and (iii), the aggregation function is also regularly increasing.

By combining Examples 6.6 and 6.7 with Example 6.8, we will consider some examples of quasiconvex systemic risk measures and provide their penalty functions and dual representations in view of Propositions 6.3 and 6.5.

Example 6.9. (Total profit-loss model with economic index of riskiness) Take $\tilde{\Lambda}(x) = \sum_{i=1}^n x_i$ and $p \in [1, +\infty)$. By (6.7), we have

$$\alpha_{\rho \circ \Lambda} \left(w \cdot \frac{\mathbf{dS}}{d\mathbb{P}}, m \right) = \inf_{\substack{\lambda > 0, \mathbb{Q} \in \mathcal{M}_1^q(\mathbb{P}), \\ w_i \mathbb{S}_i \ll \mathbb{Q}}} \left(\lambda \alpha_{\rho} \left(\frac{d\mathbb{Q}}{d\mathbb{P}}, m \right) + \mathbb{E}_{\mathbb{Q}} \left[\lambda \Phi \left(\frac{w \cdot \mathbf{dS}}{\lambda d\mathbb{Q}} \right) \right] \right).$$

Thanks to the calculation in Example 6.8(i), it is enough to consider only the case where $\frac{w \cdot \mathbf{dS}}{\lambda d\mathbb{Q}} = \mathbf{1}$

almost surely, that is, $w_1 = \dots = w_n = \lambda$ and $\mathbb{S}_1 = \dots = \mathbb{S}_n = \mathbb{Q}$. Therefore,

$$\alpha_{\rho \circ \Lambda} \left(w \cdot \frac{d\mathbb{S}}{d\mathbb{P}}, m \right) = \lambda \alpha_{\rho} \left(\frac{d\mathbb{Q}}{d\mathbb{P}}, m \right)$$

if $\frac{w \cdot d\mathbb{S}}{d\mathbb{Q}} = \lambda \mathbf{1}$ for some $\mathbb{Q} \in \mathcal{M}_1^q(\mathbb{P})$, $\lambda > 0$, and $\alpha_{\rho \circ \Lambda} \left(w \cdot \frac{d\mathbb{S}}{d\mathbb{P}}, m \right) = +\infty$ otherwise. As a further special case, let us assume that ρ is the economic index of riskiness in Example 6.7 corresponding to the logarithmic loss function with $p = 1$. In this case, we obtain

$$\alpha_{\rho \circ \Lambda} \left(w \cdot \frac{d\mathbb{S}}{d\mathbb{P}}, m \right) = m\lambda \left(1 - \exp \left(\mathbb{E} \left[\ln \left(\frac{d\mathbb{Q}}{d\mathbb{P}} \right) \right] - c_0 \right) \right)$$

if $\frac{w \cdot d\mathbb{S}}{d\mathbb{Q}} = \lambda \mathbf{1}$ for some $\mathbb{Q} \in \mathcal{M}_1^\infty(\mathbb{P})$ and $\lambda > 0$, and $\alpha_{\rho \circ \Lambda} \left(w \cdot \frac{d\mathbb{S}}{d\mathbb{P}}, m \right) = +\infty$ otherwise.

Example 6.10. (i) Let $\tilde{\Lambda}(x) = \sum_{i=1}^n x_i$ be the aggregation function in Example 6.8(i) and $p \in [1, +\infty]$. Then, by Proposition 6.5 and Example 6.8,

$$\rho \circ \Lambda(X) = \sup_{\mathbb{Q} \in \mathcal{M}_1^q(\mathbb{P})} \alpha_{\rho}^{-l} \left(\frac{d\mathbb{Q}}{d\mathbb{P}}, -\sum_{i=1}^n \mathbb{E}_{\mathbb{Q}}[X_i] \right).$$

In particular, if we take ρ as the certainty equivalent corresponding to the power loss function (Example 6.6(iii)) and $p = 1$, then by Proposition 6.5 and Example 6.6, we get

$$\rho \circ \Lambda(X) = \sup_{\mathbb{Q} \in \mathcal{M}_1^\infty(\mathbb{P})} - \left\| \frac{d\mathbb{Q}}{d\mathbb{P}} \right\|_{\frac{\gamma-1}{\gamma}} \sum_{i=1}^n \mathbb{E}_{\mathbb{Q}}[X_i].$$

(ii) Let us take the total loss model in Example 6.8 and $p \in [1, +\infty]$. Then, we have the following dual representation by Proposition 6.5:

$$R(X) = \rho \circ \Lambda(X) = \sup_{\substack{w \in \mathbb{R}_+^n \setminus \{0\}, \mathbb{S} \in \mathcal{M}_n^q(\mathbb{P}) \\ \frac{w_i d\mathbb{S}_i}{d\mathbb{P}} \leq 1, \\ \mathbb{Q} \in \mathcal{M}_1^q(\mathbb{P}), w_i \mathbb{S}_i \ll \mathbb{Q}}} \alpha_{\rho}^{-l} \left(\frac{d\mathbb{Q}}{d\mathbb{P}}, -w^\top \mathbb{E}_{\mathbb{S}}[X] \right). \quad (6.9)$$

As a special case, let us take $p = 2$ and consider the quadratic loss function in Example 6.6(i), which gives

$$R(X) = \rho \circ \Lambda(X) = \sup_{\substack{w \in \mathbb{R}_+^n \setminus \{0\}, \mathbb{S} \in \mathcal{M}_n^2(\mathbb{P}) \\ \frac{w_i d\mathbb{S}_i}{d\mathbb{P}} \leq 1, w^\top \mathbb{E}_{\mathbb{S}}[X] < 1 \\ \mathbb{Q} \in \mathcal{M}_1^2(\mathbb{P}), w_i \mathbb{S}_i \ll \mathbb{Q}}} \frac{-w^\top \mathbb{E}_{\mathbb{S}}[X] + 1}{\left\| \frac{d\mathbb{Q}}{d\mathbb{P}} \right\|_2} - 1. \quad (6.10)$$

(iii) Let us suppose that ρ is the certainty equivalent corresponding to the logarithmic loss function in Example 6.6(ii) with $p = +\infty$. Then, by Proposition 6.5 and Example 6.6, we have

$$\rho \circ \Lambda(X) = \sup_{\substack{w \in \mathbb{R}_+^n \setminus \{0\}, \mathbb{S} \in \mathcal{M}_n^1(\mathbb{P}) \\ \mathbb{Q} \in \mathcal{M}_1^1(\mathbb{P}), w_i \mathbb{S}_i \ll \mathbb{Q}}} - \frac{\mathbb{E}_{\mathbb{Q}} \left[\Phi \left(\frac{w \cdot d\mathbb{S}}{d\mathbb{Q}} \right) \right] + w^\top \mathbb{E}_{\mathbb{S}} [X]}{e^{\mathbb{E}[\ln(\frac{d\mathbb{Q}}{d\mathbb{P}})]}}.$$

In particular, let us assume that $\tilde{\Lambda}$ is the exponential aggregation function in Example 6.7(iii). Then, (6.11) simplifies as

$$\rho \circ \Lambda(X) = \sup_{\substack{w \in \mathbb{R}_+^n \setminus \{0\}, \mathbb{S} \in \mathcal{M}_n^1(\mathbb{P}) \\ \mathbb{Q} \in \mathcal{M}_1^1(\mathbb{P}), w_i \mathbb{S}_i \ll \mathbb{Q}}} \frac{w^\top \mathbb{E}_{\mathbb{S}} [-X] - \sum_{i=1}^n \mathcal{H}(w_i \mathbb{S}_i || \mathbb{Q})}{e^{\mathbb{E}[\ln(\frac{d\mathbb{Q}}{d\mathbb{P}})]}}, \quad (6.11)$$

where $\mathcal{H}(w_i \mathbb{S}_i || \mathbb{Q}) := w_i \mathbb{E}_{\mathbb{S}_i} [\ln(\frac{w_i d\mathbb{S}_i}{d\mathbb{Q}})]$ is the relative entropy of the finite measure $w_i \mathbb{S}_i$ with respect to society's probability measure \mathbb{Q} .

We conclude this section by providing an economic interpretation of the dual representation in (6.11). For given choices of the network's probability vector \mathbb{S} and weight vector w , and society's probability \mathbb{Q} , the risk of the random shock X is first calculated linearly as $w^\top \mathbb{E}_{\mathbb{S}} [-X]$. This linear evaluation is adjusted by the relative entropy term $\sum_{i=1}^n \mathcal{H}(w_i \mathbb{S}_i || \mathbb{Q})$, which is a multivariate directed distance from $w \cdot \mathbb{S}$ to \mathbb{Q} . In the presence of model uncertainty for society, further adjustment of risk by the directed distance $e^{\mathbb{E}[\ln(\frac{d\mathbb{Q}}{d\mathbb{P}})]}$ from society's measure \mathbb{Q} to the physical measure \mathbb{P} . The nonlinear interaction between the numerator and the denominator is due to the quasiconvex (but not convex) choice of ρ , as discussed in Section 6.1. Finally, the systemic risk measure is calculated as the most conservative evaluation of the ratio over all choices of $w, \mathbb{S}, \mathbb{Q}$. Similar interpretations can be made for the other instances of systemic risk measures discussed above.

6.3 Eisenberg-Noe model

In some applications, random shocks might take values only in a certain subset of \mathbb{R}^n . In such cases, the aggregation function is naturally defined on this subset instead of the whole space. In this section, we will discuss the Eisenberg-Noe clearing model for which the aggregation function is of the form $\tilde{\Lambda}: \mathbb{R}_+^n \rightarrow \mathbb{R}$. Before describing this model in detail, as a preparation, we first state slightly different versions of Propositions 6.3 and 6.5 for a generic aggregation function $\tilde{\Lambda}: \mathbb{R}_+^n \rightarrow \mathbb{R}$.

Accordingly, we modify the definition of $\tilde{\Phi}$ in (6.4) as

$$\tilde{\Phi}(x^*) = \sup_{x \in \mathbb{R}_+^n} (\Lambda(x) - (x^*)^\top x), \quad x^* \in \mathbb{R}^n,$$

and we define the functional version Φ by (6.5) as before.

Proposition 6.11. *Assume that $p \in [1, +\infty)$. Let $\tilde{\Lambda}: \mathbb{R}_+^n \rightarrow \mathbb{R}$ be a concave, regularly increasing and increasing function that is bounded from above. Let Λ be defined by (6.2) and suppose that $\Lambda(X) \in L^p(\mathbb{R})$ for every $X \in L^p(\mathbb{R}_+^n)$. Let ρ be a lower semicontinuous quasiconvex risk measure. Let $X^* \in L^q(\mathbb{R}_+^n)$ and $m \in \mathbb{R}$ such that the strict sublevel set $\{X \in L^p(\mathbb{R}_+^n) \mid \mathbb{E}[-Y^* \Lambda(X)] < m\}$ is nonempty for every $Y^* \in L^q(\mathbb{R}_+) \setminus \{0\}$. Then,*

$$\alpha_{\rho \circ \Lambda}(X^*, m) = 0 \wedge \inf_{Y^* \in L^q(\mathbb{R}_{++})} \left(\mathbb{E} \left[Y^* \Phi \left(\frac{X^*}{Y^*} \right) \right] + \alpha_\rho(Y^*, m) \right).$$

Proposition 6.12. *Assume that $p \in [1, +\infty]$. Let $\tilde{\Lambda}: \mathbb{R}_+^n \rightarrow \mathbb{R}$ be a concave increasing function that is either bounded from above or linear. Let Λ be defined by (6.2) and suppose that $\Lambda(X) \in L^p(\mathbb{R})$ for every $X \in L^p(\mathbb{R}_+^n)$. Let $\rho: L^p(\mathbb{R}) \rightarrow \bar{\mathbb{R}}$ be a lower semicontinuous quasiconvex risk measure. Then, for every $X \in L^p(\mathbb{R}_+^n)$,*

$$\rho \circ \Lambda(X) = \sup_{\substack{X^* \in L^q(\mathbb{R}_+^n) \setminus \{0\}, \\ Y^* \in L^q(\mathbb{R}_{++})}} \alpha_\rho^{-l} \left(Y^*, -\mathbb{E} \left[(X^*)^\top X + Y^* \Phi \left(\frac{X^*}{Y^*} \right) \right] \right).$$

As in Section 6.1, we may switch to probability measures by writing $X^* = w \cdot \frac{d\mathbb{S}}{d\mathbb{P}}$ and $Y^* = \lambda \frac{d\mathbb{Q}}{d\mathbb{P}}$, where $w \in \mathbb{R}_+^n \setminus \{0\}$, $\lambda > 0$, $\mathbb{Q} \in \mathcal{M}_1^q(\mathbb{P})$, and $\mathbb{S} \in \mathcal{M}_n^q(\mathbb{P})$. Again, by Ararat and Rudloff [3, Lem. 6.3], we have $w_i \mathbb{S}_i \ll \mathbb{Q}$ if $Y^* \in L^q(\mathbb{R}_{++})$. Hence, the representation in Proposition 6.12 can be rewritten as

$$\rho \circ \Lambda(X) = \sup_{\substack{w \in \mathbb{R}_+^n \setminus \{0\}, \mathbb{S} \in \mathcal{M}_n^q(\mathbb{P}) \\ \mathbb{Q} \in \mathcal{M}_1^q(\mathbb{P}), w_i \mathbb{S}_i \ll \mathbb{Q}}} \alpha_\rho^{-l} \left(\frac{d\mathbb{Q}}{d\mathbb{P}}, -\mathbb{E}_\mathbb{Q} \left[\Phi \left(\frac{w \cdot d\mathbb{S}}{d\mathbb{Q}} \right) \right] - w^\top \mathbb{E}_\mathbb{S}[X] \right). \quad (6.12)$$

Next, we review the clearing model in Eisenberg and Noe [13], which takes into account the liabilities between the members of the financial network, hence the structure of the network. In this model, financial institutions are considered as the nodes of a graph, and their liabilities are

considered as the corresponding arcs. More precisely, let $\mathcal{N} = \{0, 1, \dots, n\}$ denote the nodes, where nodes $1, \dots, n$ typically represent the banks and node 0 represents society. For each $i, j \in \mathcal{N}$, let $\ell_{ij} \geq 0$ denote the nominal liability of member i to member j . Naturally, we assume no self-liabilities, that is, $\ell_{ii} = 0$ for each $i \in \mathcal{N}$; and society has no liabilities to banks, that is, $\ell_{0i} = 0$ for every $i \in \mathcal{N}$. We also assume that every bank has nonzero liability to society, that is, $\ell_{i0} > 0$ for every $i \in \mathcal{N} \setminus \{0\}$. Then, the relative liability of member i to member j is defined by

$$a_{ij} := \frac{\ell_{ij}}{\bar{p}_i},$$

where $\bar{p}_i := \sum_{j=0}^n \ell_{ij}$ is the total liability of member i . Finally, let $x \in \mathbb{R}_+^n$ denote a possible realization of the uncertain value of the assets of the banks. A clearing payment vector $p(x) \in \mathbb{R}^n$ is defined as a solution of the following fixed point problem:

$$p_i(x) = \min \left\{ \bar{p}_i, \sum_{j=1}^n a_{ji} p_j(x) \right\} \text{ for } i \in \mathcal{N} \setminus \{0\}.$$

In words, at clearing, each bank either pays in full what it owes or it partially meets its obligations by paying what it receives from other banks. Obviously, every clearing payment vector $p = p(x)$ is a feasible solution for the following linear programming problem.

$$\begin{aligned} & \text{maximize} && \sum_{i=1}^n a_{i0} p_i && (6.13) \\ & \text{subject to} && p_i \leq x_i + \sum_{j=1}^n a_{ji} p_j \quad \forall i \in \{1, \dots, n\}, \\ & && p_i \in [0, \bar{p}_i] \quad \forall i \in \{1, \dots, n\}. \end{aligned}$$

It is shown in Eisenberg and Noe [13, Lem. 4] that every optimal solution of this problem is a clearing payment vector for the system. In addition, it is shown in Eisenberg and Noe [13] that, for every $x \in \mathbb{R}_+^n$, the above linear programming problem is feasible, and hence it has an optimal solution; let us denote the optimal value by $\tilde{\Lambda}(x)$. It should be noted that $\tilde{\Lambda}(x) \in \mathbb{R}_+$ since $a_{i0} > 0$ by definition and $p_i \in [0, \bar{p}_i]$. $\tilde{\Lambda}$ calculates the effect of the realized values of the assets on society. Therefore, $\tilde{\Lambda}$ can be considered as an aggregation function. Let us take $D = L^p(\mathbb{R}_+)$ and

$D^+ = L^q(\mathbb{R}_+)$. Then, $\tilde{\Lambda}$ is concave and increasing as it is stated in Ararat and Rudloff [3, Sect. 4.4]; it is also bounded by $\sum_{i=1}^n a_{i0}\bar{p}_i$. Hence, the assumptions of Lemma 6.2 are satisfied.

Let us calculate the conjugate function $\tilde{\Phi}$: for every $x^* \in \mathbb{R}_+^n$, by (6.13), we have

$$\begin{aligned}\tilde{\Phi}(x^*) &= \sup_{x \in \mathbb{R}_+^n} \left(-x^\top x^* + \tilde{\Lambda}(x) \right) = \sup_{0 \leq p \leq \bar{p}} \left(\sum_{i=1}^n a_{i0}p_i - \inf_{\substack{x \geq 0 \\ x \geq p - A^\top p}} \sum_{i=1}^n x_i^* x_i \right) \\ &= \sup_{0 \leq p \leq \bar{p}} \sum_{i=1}^n \left(a_{i0}p_i - x_i^* \left(p_i - \sum_{j=1}^n a_{ji}p_j \right)^+ \right).\end{aligned}$$

Then, by Proposition 6.12, we have

$$\rho \circ \Lambda(X) = \sup_{\substack{X^* \in L^q(\mathbb{R}_+^n) \setminus \{0\} \\ Y^* \in L^q(\mathbb{R}_{++})}} \alpha_\rho^{-l} \left(Y^*, -\mathbb{E} \left[X^\top X^* + Y^* \Phi \left(\frac{X^*}{Y^*} \right) \right] \right).$$

We can pass to the probabilistic setting by using (6.12) as follows:

$$\rho \circ \Lambda(X) = \sup_{\substack{w \in \mathbb{R}_+^n \setminus \{0\}, \mathbb{S} \in \mathcal{M}_n^q(\mathbb{P}) \\ \mathbb{Q} \in \mathcal{M}_1^q(\mathbb{P}), w_i \mathbb{S}_i \ll \mathbb{Q}}} \alpha_\rho^{-l} \left(\frac{d\mathbb{Q}}{d\mathbb{P}}, -\mathbb{E}_{\mathbb{Q}} \left[\Phi \left(w \cdot \frac{d\mathbb{S}}{d\mathbb{Q}} \right) \right] - w^\top \mathbb{E}_{\mathbb{S}}[X] \right).$$

As a special case, let us assume that ρ is the certainty equivalent associated to the logarithmic loss function (see Example 6.6(ii)) for the case $p = 1$. Then, the dual representation simplifies as

$$\rho \circ \Lambda(X) = \sup_{\substack{w \in \mathbb{R}_+^n \setminus \{0\}, \mathbb{S} \in \mathcal{M}_n^\infty(\mathbb{P}) \\ \mathbb{Q} \in \mathcal{M}_1^\infty(\mathbb{P}), w_i \mathbb{S}_i \ll \mathbb{Q}}} \frac{w^\top \mathbb{E}_{\mathbb{S}}[-X] - \mathbb{E}_{\mathbb{Q}} \left[\Phi \left(w \cdot \frac{d\mathbb{S}}{d\mathbb{Q}} \right) \right]}{e^{\mathbb{E} \left[\ln \left(\frac{d\mathbb{Q}}{d\mathbb{P}} \right) \right]}}. \quad (6.14)$$

The economic interpretation of (6.14) is similar to the one at the end of Section 6.2. Different from the examples in Section 6.2, the multivariate divergence term here is specific to the Eisenberg-Noe model. Hence, we focus on the interpretation of this term. With the help of Rockafellar and

Wets [29, Thm. 14.60], we can calculate the divergence term more explicitly as

$$\begin{aligned} \mathbb{E}_{\mathbb{Q}} \left[\Phi \left(w \cdot \frac{d\mathbb{S}}{d\mathbb{Q}} \right) \right] &= \mathbb{E}_{\mathbb{Q}} \left[\sup_{0 \leq p \leq \bar{p}} \sum_{i=1}^n \left(a_{i0} p_i - w_i \frac{d\mathbb{S}_i}{d\mathbb{Q}} \left(p_i - \sum_{j=1}^n a_{ji} p_j \right)^+ \right) \right] \\ &= \sup_{P \in L^1(\mathbb{Q}, [0, \bar{p}])} \left(\mathbb{E}_{\mathbb{Q}} \left[\sum_{i=1}^n a_{i0} P_i \right] - \sum_{i=1}^n w_i \mathbb{E}_{\mathbb{S}_i} \left[\left(P_i - \sum_{j=1}^n a_{ji} P_j \right)^+ \right] \right), \end{aligned}$$

where $L^1(\mathbb{Q}, [0, \bar{p}])$ denotes the space of random vectors of the probability space $(\Omega, \mathcal{F}, \mathbb{Q})$ that take values in the rectangle $[0, \bar{p}]$. Hence, under the supremum, we consider a scenario-dependent payment vector P . The term $\sum_{i=1}^n a_{i0} P_i$ represents the total payment received by society. Therefore, we calculate its expectation with respect to \mathbb{Q} , that is, with respect to society's own perspective. Let us fix a bank $i \in \{1, \dots, n\}$. Then, $(P_i - \sum_{j=1}^n a_{ji} P_j)^+$ is the net equity of bank i ; we calculate its expectation with respect to \mathbb{S}_i , that is, with respect to the bank's own perspective. Hence, the weighted sum $\sum_{i=1}^n w_i \mathbb{E}_{\mathbb{S}_i} [(P_i - \sum_{j=1}^n a_{ji} P_j)^+]$ can be seen as the expected net equity from the perspective of the overall network (besides society). Then, the difference $\mathbb{E}_{\mathbb{Q}} [\sum_{i=1}^n a_{i0} P_i] - \sum_{i=1}^n w_i \mathbb{E}_{\mathbb{S}_i} [(P_i - \sum_{j=1}^n a_{ji} P_j)^+]$ is a measure of the mismatch between society's expectation and the network's overall expectation for the payments. Finally, the multivariate divergence term, as a directed distance from $w \cdot \mathbb{S}$ to \mathbb{Q} , is calculated as the largest possible value of this mismatch over all choices of the random payment vector P .

7 Appendix

7.1 Proof of some results in Section 4

The main purpose of this section is to prove Theorem 4.6. As a preparation for the proof, we will establish a sequence of technical results. In particular, these results will ensure that we may apply the minimax inequality in Liu [24].

We work in the setting of Section 4: we consider two functions $f: \mathcal{Y} \rightarrow \overline{\mathbb{R}}$ and $g: \mathcal{X} \rightarrow \mathcal{Y}$. We also suppose that Assumption 4.5 holds, that is, \bar{D}^+ is a convex and compact cone generator for D^+ . Given $m \in \mathbb{R}$ and $y^* \in D^+$, let us define the sets

$$A_{y^*}^m := \{x \in \mathcal{X} \mid \langle y^*, -g(x) \rangle \leq \alpha_f(y^*, m)\}, \quad \tilde{A}_{y^*}^m := \{x \in \mathcal{X} \mid \langle y^*, -g(x) \rangle < \alpha_f(y^*, m)\}.$$

Clearly, $\tilde{A}_{y^*}^m \subseteq A_{y^*}^m$. Also, observe that $A_{y^*}^m$ is actually the sublevel set of $-h_{y^*}^g$; see (3.1). Therefore, when the function g is D -naturally quasiconcave, increasing and D -lower demicontinuous, the set $A_{y^*}^m$ is a closed, convex and monotone set by Propositions 3.2, 3.7, 3.12. We give the precise relationship between the sets $\tilde{A}_{y^*}^m$ and $A_{y^*}^m$ in the following proposition.

Proposition 7.1. *Suppose that Assumption 3.9 holds. In addition, suppose that $g: \mathcal{X} \rightarrow \mathcal{Y}$ is D -naturally quasiconcave, regularly increasing and D -lower demicontinuous; and let $m \in \mathbb{R}$, $y^* \in D^+ \setminus \{0\}$. Then,*

$$A_{y^*}^m = \text{cl } \tilde{A}_{y^*}^m = \text{cl conv } \tilde{A}_{y^*}^m. \quad (7.1)$$

Proof. If $A_{y^*}^m = \emptyset$, then the result is obvious. Let us assume that $A_{y^*}^m \neq \emptyset$ and prove that $A_{y^*}^m$ is the closure of $\tilde{A}_{y^*}^m$. Since $\tilde{A}_{y^*}^m \subseteq A_{y^*}^m$ and $A_{y^*}^m$ is closed, we have $\text{cl } \tilde{A}_{y^*}^m \subseteq A_{y^*}^m$.

Now let us take $x \in A_{y^*}^m$, and fix some $c \in C^\#$ and $\lambda > 0$. It is clear that $\lambda c \in C^\#$ since $C^\#$ is a cone. Moreover, since g is regularly increasing, we have $g(x + \lambda c) - g(x) \in D^\#$. In particular, since $y^* \in D^+ \setminus \{0\}$, we have $\langle y^*, g(x + \lambda c) - g(x) \rangle > 0$. Therefore,

$$\begin{aligned} \langle y^*, -g(x + \lambda c) \rangle &= \langle y^*, -g(x) \rangle - \langle y^*, g(x + \lambda c) - g(x) \rangle \\ &\leq \alpha_f(y^*, m) - \langle y^*, g(x + \lambda c) - g(x) \rangle < \alpha_f(y^*, m). \end{aligned}$$

Hence, $x + \lambda c \in \tilde{A}_{y^*}^m$. The net $(x + \lambda c)_{\lambda > 0} \subseteq \tilde{A}_{y^*}^m$ converges to x as $\lambda \rightarrow 0$, which implies that $x \in \text{cl } \tilde{A}_{y^*}^m$. Hence, $A_{y^*}^m \subseteq \text{cl } \tilde{A}_{y^*}^m$ and the first equality in (7.1).

Finally, since $A_{y^*}^m$ is convex, we have $A_{y^*}^m = \text{conv}(\text{cl } \tilde{A}_{y^*}^m) \subseteq \text{cl}(\text{conv } \tilde{A}_{y^*}^m) \subseteq A_{y^*}^m$. Hence, the second equality in (7.1) follows as well. \square

Remark 7.2. Let $m \in \mathbb{R}$, $y^* \in D^+ \setminus \{0\}$. We may write $y^* = \lambda \bar{y}^*$ for some $\lambda > 0$ and $\bar{y}^* \in \bar{D}^+$. Then, it is easy to see that $x \in A_{y^*}^m$ if and only if $x \in A_{\bar{y}^*}^m$ for each $x \in \mathcal{X}$. Hence, $A_{y^*}^m = A_{\bar{y}^*}^m$.

Next, given $m \in \mathbb{R}$ and $x^* \in C^+$, we define two auxiliary functions $K_{x^*}^m, \tilde{K}_{x^*}^m: \mathcal{X} \times \bar{D}^+ \rightarrow \bar{\mathbb{R}}$ by

$$K_{x^*}^m(x, y^*) = \langle x^*, -x \rangle - I_{A_{y^*}^m}(x), \quad \tilde{K}_{x^*}^m(x, y^*) = \langle x^*, -x \rangle - I_{\tilde{A}_{y^*}^m}(x), \quad (7.2)$$

for each $(x, y^*) \in \mathcal{X} \times \bar{D}^+$. The next proposition shows the relation between these two functions.

Proposition 7.3. *Let $m \in \mathbb{R}$ and $x^* \in C^+$. Suppose that g is D -naturally quasiconcave, regularly increasing and D -lower demicontinuous. Then, for each $y^* \in \bar{D}^+$, we have*

$$\sup_{x \in \mathcal{X}} \tilde{K}_{x^*}^m(x, y^*) = \sup_{x \in \mathcal{X}} K_{x^*}^m(x, y^*).$$

Proof. Let $y^* \in \bar{D}^+$. By definition, we have

$$\sup_{x \in \mathcal{X}} \tilde{K}_{x^*}^m(x, y^*) = \sup_{x \in \mathcal{X}} (\langle -x^*, x \rangle - I_{\tilde{A}_{y^*}^m}(x)) = I_{\tilde{A}_{y^*}^m}^*(-x^*). \quad (7.3)$$

Moreover, by (2.5) and Proposition 7.1, we have $I_{\tilde{A}_{y^*}^m}^*(-x^*) = \sup_{x \in A_{y^*}^m} \langle -x^*, x \rangle$. Similar to (7.3), we also have

$$\sup_{x \in \mathcal{X}} K_{x^*}^m(x, y^*) = \sup_{x \in A_{y^*}^m} \langle -x^*, x \rangle.$$

Combining these, we obtain the desired result. \square

We will use a minimax theorem in the proof of Theorem 4.6. As a preparation, we check some properties of the functions defined in (7.2). These properties are necessary for the application of the minimax theorem.

Proposition 7.4. *Let $m \in \mathbb{R}$ and $x^* \in C^+$. Suppose that g is D -naturally quasiconcave. The following properties hold.*

(i) *Suppose further that g is D -lower demicontinuous. Then, $K_{x^*}^m$ is concave and upper semicontinuous in its first argument, and quasiconvex in its second argument.*

(ii) *$\tilde{K}_{x^*}^m$ is concave in its first argument, and quasiconvex and lower semicontinuous in its second argument.*

Proof. We prove (i) first. Let $y^* \in \bar{D}^+$. Since $A_{y^*}^m$ is a closed convex set, $I_{A_{y^*}^m}$ is a lower semicontinuous convex function. Hence, $x \mapsto K_{x^*}^m(x, y^*)$ is an upper semicontinuous concave function.

Next, let us fix $x \in \mathcal{X}$. We claim that $y^* \mapsto I_{A_{y^*}^m}(x)$ is a quasiconvex function. Indeed, let $y_1^*, y_2^* \in \bar{D}^+$ and $\lambda \in [0, 1]$. Since \bar{D}^+ is convex, $\lambda y_1^* + (1 - \lambda)y_2^* \in D_{cg}^+$. If $x \in A_{y_1^*}^m$ or $x \in A_{y_2^*}^m$, then $\min \{I_{A_{y_1^*}^m}(x), I_{A_{y_2^*}^m}(x)\} = 0 \leq I_{A_{\lambda y_1^* + (1-\lambda)y_2^*}^m}(x)$ by the definition of indicator function. On the other hand, suppose that $x \notin A_{y_1^*}^m$ and $x \notin A_{y_2^*}^m$. Then, $\langle y_1^*, -g(x) \rangle > \alpha_f(y_1^*, m)$ and

$\langle y_2^*, -g(x) \rangle > \alpha_f(y_2^*, m)$. Hence,

$$\begin{aligned} \langle \lambda y_1^* + (1 - \lambda)y_2^*, -g(x) \rangle &> \lambda \alpha_f(y_1^*, m) + (1 - \lambda) \alpha_f(y_2^*, m) \\ &= \lambda \sup_{y \in S_m^f} \langle y_1^*, -y \rangle + (1 - \lambda) \sup_{y \in S_m^f} \langle y_2^*, -y \rangle \\ &\geq \sup_{y \in S_m^f} \langle \lambda y_1^* + (1 - \lambda)y_2^*, -y \rangle = \alpha_f(\lambda y_1^* + (1 - \lambda)y_2^*, m). \end{aligned}$$

Therefore, $x \notin A_{\lambda y_1^* + (1 - \lambda)y_2^*}^m$ so that $\min \{I_{A_{y_1^*}^m}(x), I_{A_{y_2^*}^m}(x)\} \leq +\infty = I_{A_{\lambda y_1^* + (1 - \lambda)y_2^*}^m}(x)$. It follows that $y^* \mapsto I_{A_{y^*}^m}(x)$ is quasiconvex, hence so is $y^* \mapsto K_{x^*}^m(x, y^*)$.

Next, we prove (ii). Let $y^* \in \bar{D}^+$. We claim that $\tilde{A}_{y^*}^m$ is a convex set. Indeed, let $x_1, x_2 \in \tilde{A}_{y^*}^m$ and $\lambda \in [0, 1]$. Since $-h_{y^*}^g$ is quasiconvex, we have

$$-h_{y^*}^g(\lambda x_1 + (1 - \lambda)x_2) \leq \max \{ -h_{y^*}^g(x_1), -h_{y^*}^g(x_2) \} < \alpha_f(y^*, m),$$

which implies that $\lambda x_1 + (1 - \lambda)x_2 \in \tilde{A}_{y^*}^m$. Hence, the claim follows. It follows that $I_{\tilde{A}_{y^*}^m}$ is a convex function and $x \mapsto \tilde{K}_{x^*}^m(x, y^*)$ is a concave function.

Let us fix $x \in \mathcal{X}$. We show that $y^* \mapsto I_{\tilde{A}_{y^*}^m}(x)$ is quasiconvex. Let $y_1^*, y_2^* \in \bar{D}^+$ and $\lambda \in [0, 1]$. Since \bar{D}^+ is convex $\lambda y_1^* + (1 - \lambda)y_2^* \in \bar{D}^+$. If $x \in \tilde{A}_{y_1^*}^m$ or $x \in \tilde{A}_{y_2^*}^m$, then we have

$$\min \{I_{\tilde{A}_{y_1^*}^m}(x), I_{\tilde{A}_{y_2^*}^m}(x)\} = 0 \leq I_{\tilde{A}_{\lambda y_1^* + (1 - \lambda)y_2^*}^m}(x).$$

Suppose that $x \notin \tilde{A}_{y_1^*}^m$ and $x \notin \tilde{A}_{y_2^*}^m$. Hence, $\langle y_1^*, -g(x) \rangle \geq \alpha_f(y_1^*, m)$, $\langle y_2^*, -g(x) \rangle \geq \alpha_f(y_2^*, m)$, and

$$\begin{aligned} \langle \lambda y_1^* + (1 - \lambda)y_2^*, -g(x) \rangle &\geq \lambda \alpha_f(y_1^*, m) + (1 - \lambda) \alpha_f(y_2^*, m) \\ &= \lambda \sup_{y \in S_m^f} \langle y_1^*, -y \rangle + (1 - \lambda) \sup_{y \in S_m^f} \langle y_2^*, -y \rangle \\ &\geq \sup_{y \in S_m^f} \langle \lambda y_1^* + (1 - \lambda)y_2^*, -y \rangle = \alpha_f(\lambda y_1^* + (1 - \lambda)y_2^*, m), \end{aligned}$$

which implies $x \notin \tilde{A}_{\lambda y_1^* + (1 - \lambda)y_2^*}^m$. Hence,

$$\min \{I_{\tilde{A}_{y_1^*}^m}(x), I_{\tilde{A}_{y_2^*}^m}(x)\} \leq +\infty = I_{\tilde{A}_{\lambda y_1^* + (1 - \lambda)y_2^*}^m}(x),$$

which completes the proof of quasiconvexity. It follows that $y^* \mapsto \tilde{K}_{x^*}^m(x, y^*)$ is quasiconvex.

Finally, to prove lower semicontinuity, let us define the set

$$E_x^m := \{y^* \in \bar{D}^+ \mid \langle y^*, -g(x) \rangle < \alpha_f(y^*, m)\}.$$

Note that

$$E_x^m = \left\{ y^* \in \bar{D}^+ \mid \langle y^*, -g(x) \rangle < \sup_{y \in S_m^f} \langle y^*, -y \rangle \right\} = \left\{ y^* \in \bar{D}^+ \mid 0 < \sup_{y \in S_m^f} \langle y^*, -y + g(x) \rangle \right\}.$$

Since the supremum of a family of affine functions is lower semicontinuous, it follows that E_x^m is open. On the other hand, for each $y^* \in \bar{D}^+$, it is clear that $y^* \in E_x^m$ if and only if $x \in \tilde{A}_{y^*}^m$, that is, $I_{\tilde{A}_{y^*}^m}(x) = I_{E_x^m}(y^*)$. Hence, we indeed have

$$\tilde{K}_{x^*}^m(x, y^*) = \langle x^*, -x \rangle - I_{\tilde{A}_{y^*}^m}(x) = \langle x^*, -x \rangle - I_{E_x^m}(y^*). \quad (7.4)$$

Since E_x^m is open, the function $I_{E_x^m}$ is upper semicontinuous. By (7.4), $y^* \mapsto \tilde{K}_{x^*}^m(x, y^*)$ is lower semicontinuous. \square

Now, we relate the functions defined in (7.2) to our problem.

Proposition 7.5. *Suppose that f is decreasing, lower semicontinuous and quasiconvex, and that g is D -naturally quasiconcave and D -lower demicontinuous. Then, for each $(x^*, m) \in C^+ \times \mathbb{R}$,*

$$\alpha_{f \circ g}(x^*, m) = \sup_{x \in \mathcal{X}} \inf_{y^* \in \bar{D}^+} K_{x^*}^m(x, y^*).$$

Proof. Let $(x^*, m) \in C^+ \times \mathbb{R}$. Since f is decreasing, lower semicontinuous and quasiconvex, by Remarks 2.11 and 7.2, we have

$$\begin{aligned} \alpha_{f \circ g}(x^*, m) &= \sup_{x \in S_m^{f \circ g}} \langle x^*, -x \rangle = \sup \{ \langle x^*, -x \rangle \mid g(x) \in S_m^f, x \in \mathcal{X} \} \\ &= \sup_{x \in \mathcal{X}} \{ \langle x^*, -x \rangle \mid \forall y^* \in D^+ \setminus \{0\}: \langle y^*, -g(x) \rangle \leq \alpha_f(y^*, m) \} \\ &= \sup_{x \in \mathcal{X}} \{ \langle x^*, -x \rangle \mid \forall y^* \in \bar{D}^+: \langle y^*, -g(x) \rangle \leq \alpha_f(y^*, m) \} = \sup_{x \in B^m} \langle x^*, -x \rangle, \end{aligned}$$

where $B^m := \bigcap_{y^* \in \bar{D}^+} A_{y^*}^m$. Hence,

$$\sup_{x \in B^m} \langle x^*, -x \rangle = \sup_{x \in \mathcal{X}} (\langle x^*, -x \rangle - I_{B^m}(x)) = \sup_{x \in \mathcal{X}} \inf_{y^* \in \bar{D}^+} (\langle x^*, -x \rangle - I_{A_{y^*}^m}(x)) = \sup_{x \in \mathcal{X}} \inf_{y^* \in \bar{D}^+} K_{x^*}^m(x, y^*).$$

Therefore, the result follows. \square

Proposition 7.6. *Let $(x^*, m) \in C^+ \times \mathbb{R}$. Then, we have*

$$\inf_{y^* \in D^+ \setminus \{0\}} \alpha_{-h_{y^*}^g}(x^*, \alpha_f(y^*, m)) = \inf_{y^* \in \bar{D}^+} \alpha_{-h_{y^*}^g}(x^*, \alpha_f(y^*, m)) = \inf_{y^* \in \bar{D}^+} \sup_{x \in \mathcal{X}} K_{x^*}^m(x, y^*). \quad (7.5)$$

Proof. Let $\bar{y}^* \in \bar{D}^+$. Clearly, we have

$$\sup_{x \in \mathcal{X}} K_{x^*}^m(x, \bar{y}^*) = \sup_{x \in \mathcal{X}} (\langle x^*, -x \rangle - I_{A_{\bar{y}^*}^m}(x)) = \sup_{x \in A_{\bar{y}^*}^m} \langle x^*, -x \rangle.$$

Hence,

$$\begin{aligned} \inf_{\bar{y}^* \in \bar{D}^+ \setminus \{0\}} \alpha_{-h_{\bar{y}^*}^g}(x^*, \alpha_f(\bar{y}^*, m)) &= \inf_{\bar{y}^* \in \bar{D}^+ \setminus \{0\}} \sup_{x \in \mathcal{X}} \{\langle x^*, -x \rangle \mid -h_{\bar{y}^*}^g(x) \leq \alpha_f(\bar{y}^*, m)\} \\ &= \inf_{\bar{y}^* \in \bar{D}^+ \setminus \{0\}} \sup_{x \in \mathcal{X}} \{\langle x^*, -x \rangle \mid \langle \bar{y}^*, -g(x) \rangle \leq \alpha_f(\bar{y}^*, m)\} \\ &= \inf_{\bar{y}^* \in \bar{D}^+ \setminus \{0\}} \sup_{x \in A_{\bar{y}^*}^m} \langle x^*, -x \rangle = \inf_{\bar{y}^* \in \bar{D}^+ \setminus \{0\}} \sup_{x \in \mathcal{X}} K_{x^*}^m(x, \bar{y}^*), \end{aligned}$$

which completes the proof of the second equality in (7.5). On the other hand, given $y^* \in D^+ \setminus \{0\}$, we may write $y^* = \lambda \bar{y}^*$ for some $\lambda > 0$ and $\bar{y}^* \in \bar{D}^+$, and we have

$$\alpha_{-h_{y^*}^g}(x^*, \alpha_f(y^*, m)) = \sup_{x \in A_{y^*}^m} \langle x^*, -x \rangle = \sup_{x \in A_{\bar{y}^*}^m} \langle x^*, -x \rangle = \alpha_{-h_{\bar{y}^*}^g}(x^*, \alpha_f(\bar{y}^*, m))$$

by Remark 7.2. Hence, the first equality in (7.5) follows as well. \square

From this point on, we work under Assumption 4.5, that is, we assume that \bar{D}^+ is compact while D_π^+ is not necessarily compact. In particular, Proposition 7.6 can be applied to both. With the tools developed above, we are ready to prove the main theorem of the paper. For the completeness of this paper, we give the statements of the well-known minimax equality in Sion [30] and the minimax inequality in Liu [24].

Theorem 7.7 (Sion (1958) [30]). *Let \mathcal{U}, \mathcal{V} be nonempty convex sets of two topological vector spaces, and consider a function $f: \mathcal{U} \times \mathcal{V} \rightarrow \mathbb{R}$. Suppose that f is quasiconcave and upper semicontinuous in its first argument, and quasiconvex and lower semicontinuous in its second argument. Moreover, suppose that one of \mathcal{U}, \mathcal{V} is a compact set. Then, we have*

$$\inf_{u \in \mathcal{U}} \sup_{v \in \mathcal{V}} f(u, v) = \sup_{v \in \mathcal{V}} \inf_{u \in \mathcal{U}} f(u, v).$$

Since the lower semicontinuity of $K_{x^*}^m$ is missing, it seems that we are not able to use Theorem 7.7 in our setting. Instead, the following minimax inequality will be applicable in our proof.

Theorem 7.8 (Liu (1978) [24]). *In the setting of Theorem 7.7, consider two functions $f, \tilde{f}: \mathcal{U} \times \mathcal{V} \rightarrow \overline{\mathbb{R}}$ satisfying the following conditions:*

- (i) f is upper semicontinuous in its first argument and quasiconvex in its second argument.
- (ii) \tilde{f} is quasiconcave in its first argument and lower semicontinuous in its second argument.
- (iii) $\tilde{f}(u, v) \leq f(u, v)$ for all $u \in \mathcal{U}$ and $v \in \mathcal{V}$.
- (iv) \mathcal{U} is compact.

Then, we have

$$\inf_{u \in \mathcal{U}} \sup_{v \in \mathcal{V}} \tilde{f}(u, v) \leq \sup_{v \in \mathcal{V}} \inf_{u \in \mathcal{U}} f(u, v).$$

With the help of Theorem 7.8, we are ready to complete the proof of Theorem 4.6.

Proof of Theorem 4.6. Let $x^* \in C^+ \setminus \{0\}$ and $m \in \mathbb{R}$. For each $y^* \in \bar{D}^+$, since $\tilde{A}_{y^*}^m \subseteq A_{y^*}^m$, we have $I_{\tilde{A}_{y^*}^m}(x) \geq I_{A_{y^*}^m}(x)$ and hence

$$\tilde{K}_{x^*}^m(x, y^*) \leq K_{x^*}^m(x, y^*), \quad x \in \mathcal{X}. \quad (7.6)$$

By Proposition 7.4, $K_{x^*}^m$ is upper semicontinuous and concave in its first variable, and quasiconvex in its second variable; $\tilde{K}_{x^*}^m$ is concave in its first variable, and quasiconvex and lower semicontinuous in its second variable. These properties, together with (7.6), and the convexity and compactness of \bar{D}^+ are sufficient to apply the minimax inequality of Liu [24] (see also Cheng and Lin [11, Thm.

3.1] and Greco and Moschen [18, Cor. 11]) to the functions $K_{x^*}^m, \tilde{K}_{x^*}^m$. Consequently, we obtain

$$\inf_{y^* \in \bar{D}^+} \sup_{x \in \mathcal{X}} \tilde{K}_{x^*}^m(x, y^*) \leq \sup_{x \in \mathcal{X}} \inf_{y^* \in \bar{D}^+} K_{x^*}^m(x, y^*). \quad (7.7)$$

By Proposition 7.3, we have

$$\sup_{x \in \mathcal{X}} \tilde{K}_{x^*}^m(x, y^*) = \sup_{x \in \mathcal{X}} K_{x^*}^m(x, y^*).$$

Hence, (7.7) yields

$$\inf_{y^* \in \bar{D}^+} \sup_{x \in \mathcal{X}} K_{x^*}^m(x, y^*) \leq \sup_{x \in \mathcal{X}} \inf_{y^* \in \bar{D}^+} K_{x^*}^m(x, y^*).$$

However, the reverse inequality already holds by weak duality. Therefore, we get

$$\inf_{y^* \in \bar{D}^+} \sup_{x \in \mathcal{X}} K_{x^*}^m(x, y^*) = \sup_{x \in \mathcal{X}} \inf_{y^* \in \bar{D}^+} K_{x^*}^m(x, y^*).$$

Finally, by Propositions 7.5 and 7.6, we have

$$\begin{aligned} \alpha_{f \circ g}(x^*, m) &= \sup_{x \in \mathcal{X}} \inf_{y^* \in \bar{D}^+} K_{x^*}^m(x, y^*) = \inf_{y^* \in \bar{D}^+} \sup_{x \in \mathcal{X}} K_{x^*}^m(x, y^*) \\ &= \inf_{y^* \in D^+ \setminus \{0\}} \alpha_{-h_{y^*}^g}(x^*, \alpha_f(y^*, m)) = \inf_{y^* \in D^+} \alpha_{-h_{y^*}^g}(x^*, \alpha_f(y^*, m)). \end{aligned}$$

Finally, by Remark 4.4 and Proposition 7.6 applied to D_π^+ , we have

$$\alpha_{f \circ g}(x^*, m) = \inf_{y_\pi^* \in D_\pi^+} \alpha_{-h_{y_\pi^*}^g}(x^*, \alpha_f(y_\pi^*, m)),$$

which completes the proof. \square

Proof of Corollary 4.8. Let $x^* \in C^+ \setminus \{0\}$ and $s \in \mathbb{R}$. Following the definition of left inverse and using Theorem 4.6, we have

$$\begin{aligned} \alpha_{f \circ g}^{-l}(x^*, s) &= \inf \{m \in \mathbb{R} \mid \alpha_{f \circ g}(x^*, m) \geq s\} \\ &= \inf \left\{ m \in \mathbb{R} \mid \inf_{y^* \in D^+ \setminus \{0\}} \alpha_{-h_{y^*}^g}(x^*, \alpha_f(y^*, m)) \geq s \right\} \\ &= \inf \left\{ m \in \mathbb{R} \mid \forall y^* \in D^+ \setminus \{0\} : \alpha_{-h_{y^*}^g}(x^*, \alpha_f(y^*, m)) \geq s \right\}. \end{aligned}$$

We claim that the following minimax equality holds:

$$\begin{aligned} & \inf \left\{ m \in \mathbb{R} \mid \forall y^* \in D^+ \setminus \{0\} : \alpha_{-h_{y^*}^g}(x^*, \alpha_f(y^*, m)) \geq s \right\} \\ &= \sup_{y^* \in D^+ \setminus \{0\}} \inf \left\{ m \in \mathbb{R} \mid \alpha_{-h_{y^*}^g}(x^*, \alpha_f(y^*, m)) \geq s \right\}. \end{aligned} \quad (7.8)$$

The \geq part of this inequality holds as a weak duality property. Next, we show the \leq part. To get a contradiction, suppose that there exists $\bar{m} \in \mathbb{R}$ such that

$$\begin{aligned} & \inf \left\{ m \in \mathbb{R} \mid \forall y^* \in D^+ \setminus \{0\} : \alpha_{-h_{y^*}^g}(x^*, \alpha_f(y^*, m)) \geq s \right\} \\ &> \bar{m} > \sup_{y^* \in D^+ \setminus \{0\}} \inf \left\{ m \in \mathbb{R} \mid \alpha_{-h_{y^*}^g}(x^*, \alpha_f(y^*, m)) \geq s \right\}. \end{aligned} \quad (7.9)$$

The first inequality in (7.9) implies the existence of $\bar{y}^* \in D^+ \setminus \{0\}$ satisfying

$$\alpha_{-h_{\bar{y}^*}^g}(x^*, \alpha_f(\bar{y}^*, \bar{m})) < s. \quad (7.10)$$

On the other hand, the second inequality in (7.9) implies that $\bar{m} > \inf\{m \in \mathbb{R} \mid \alpha_{-h_{\bar{y}^*}^g}(x^*, \alpha_f(\bar{y}^*, m)) \geq s\}$. Hence, there exists $m_{\bar{y}^*} < \bar{m}$ such that

$$\alpha_{-h_{\bar{y}^*}^g}(x^*, \alpha_f(\bar{y}^*, m_{\bar{y}^*})) \geq s. \quad (7.11)$$

Since α_f is increasing in the second argument by Remark 2.10, we have $\alpha_f(\bar{y}^*, \bar{m}) \geq \alpha_f(\bar{y}^*, m_{\bar{y}^*})$.

Hence, by (7.11), the monotonicity of $\alpha_{-h_{\bar{y}^*}^g}$, and (7.10), we obtain

$$s \leq \alpha_{-h_{\bar{y}^*}^g}(x^*, \alpha_f(\bar{y}^*, m_{\bar{y}^*})) \leq \alpha_{-h_{\bar{y}^*}^g}(x^*, \alpha_f(\bar{y}^*, \bar{m})) < s,$$

which is a contradiction. Hence, (7.8) follows so that

$$\alpha_{f \circ g}^{-l}(x^*, s) = \sup_{y^* \in D^+ \setminus \{0\}} \inf \left\{ m \in \mathbb{R} \mid \alpha_{-h_{y^*}^g}(x^*, \alpha_f(y^*, m)) \geq s \right\}. \quad (7.12)$$

Let $y^* \in D^+ \setminus \{0\}$. We claim that

$$\inf \left\{ m \in \mathbb{R} \mid \alpha_{-h_{y^*}^g}(x^*, \alpha_f(y^*, m)) \geq s \right\} = \inf \left\{ m \in \mathbb{R} \mid \alpha_f(y^*, m) \geq \alpha_{-h_{y^*}^g}^{-l}(x^*, s) \right\}. \quad (7.13)$$

For each $m \in \mathbb{R}$, by the definition of left inverse,

$$\alpha_{-h_{y^*}^g}(x^*, \alpha_f(y^*, m)) \geq s \quad \Rightarrow \quad \alpha_f(y^*, m) \geq \alpha_{-h_{y^*}^g}^{-l}(x^*, s).$$

Hence, the \geq part of (7.13) follows. Next, we prove that \leq part. To get a contradiction, suppose that

$$\inf \left\{ m \in \mathbb{R} \mid \alpha_{-h_{y^*}^g}(x^*, \alpha_f(y^*, m)) \geq s \right\} > \tilde{m} > \inf \left\{ m \in \mathbb{R} \mid \alpha_f(y^*, m) \geq \alpha_{-h_{y^*}^g}^{-l}(x^*, s) \right\}$$

for some $\tilde{m} \in \mathbb{R}$. By the first inequality, we have $\alpha_{-h_{y^*}^g}(x^*, \alpha_f(y^*, \tilde{m})) < s$; and by the second inequality together with the monotonicity of α_f , we have $\alpha_f(y^*, \tilde{m}) \geq \alpha_{-h_{y^*}^g}^{-l}(x^*, s)$. Hence, by the monotonicity of $\alpha_{-h_{y^*}^g}$,

$$s \leq \alpha_{-h_{y^*}^g}(x^*, \alpha_{-h_{y^*}^g}^{-l}(x^*, s)) \leq \alpha_{-h_{y^*}^g}(x^*, \alpha_f(y^*, \tilde{m})) < s,$$

a contradiction. Therefore, (7.13) follows.

Combining (7.12) and (7.13) gives

$$\alpha_{f \circ g}^{-l}(x^*, s) = \sup_{y^* \in D^+ \setminus \{0\}} \inf \left\{ m \in \mathbb{R} \mid \alpha_f(y^*, m) \geq \alpha_{-h_{y^*}^g}^{-l}(x^*, s) \right\} = \sup_{y^* \in D^+ \setminus \{0\}} \alpha_f^{-l}(y^*, \alpha_{-h_{y^*}^g}^{-l}(x^*, s)),$$

which proves (4.5). Combining this with Proposition 4.1 and the monotonicity of g , we get (4.6). \square

Finally, we outline the proofs of the results in Section 4.3. Recall that we work with a monotone convex set $\mathcal{K} \subseteq \mathcal{X}$ with $C \subseteq \mathcal{K}$, and we consider two functions $f: \mathcal{Y} \rightarrow \overline{\mathbb{R}}$ and $g: \mathcal{K} \rightarrow \mathcal{Y}$. Let $x^* \in C^+$ and $m \in \mathbb{R}$. Similar to the constructions for the case $\mathcal{K} = \mathcal{Y}$ above, we define the sets

$$\mathbb{A}_{y^*}^m := \{x \in \mathcal{K} \mid \langle y^*, -g(x) \rangle \leq \alpha_f(y^*, m)\}, \quad \tilde{\mathbb{A}}_{y^*}^m := \{x \in \mathcal{K} \mid \langle y^*, -g(x) \rangle < \alpha_f(y^*, m)\}$$

for each $y^* \in D^+$, and the functions $\mathbb{K}_{x^*}^m, \tilde{\mathbb{K}}_{x^*}^m : \mathcal{K} \times \bar{D}^+ \rightarrow \bar{\mathbb{R}}$ by

$$\mathbb{K}_{x^*}^m(x, y^*) := \langle x^*, -x \rangle - I_{\mathbb{A}_{y^*}^m}(x), \quad \tilde{\mathbb{K}}_{x^*}^m(x, y^*) := \langle x^*, -x \rangle - I_{\tilde{\mathbb{A}}_{y^*}^m}(x).$$

After giving these definitions, by using similar arguments, we can adapt Propositions 7.1, 7.3, 7.4, 7.5 and 7.6, and Remark 7.2 for the following corollary.

Proof of Corollary 4.12. The proof follows the same reasoning as the proof of Theorem 4.6. \square

Proof of Proposition 4.13. The proof of (4.9) follows the same arguments as the proof of Proposition 4.1. Here, we use Corollary 2.15 instead of Theorem 2.14. The proof of (4.10) follows by the same arguments as in Theorem 4.2. \square

Proof of Proposition 4.14. The proof of Proposition 4.11 is valid for this result. \square

7.2 Proofs for Section 6

Proof of Lemma 6.2. To prove that Λ is lower demicontinuous, by Remark 3.11, we need to prove that $\Lambda^U(M) = \{X \in L^p(\mathbb{R}^n) \mid \Lambda(X) + L^p(\mathbb{R}_+) \subseteq M\}$ is closed for every closed halfspace $M = \{Y \in L^p(\mathbb{R}) \mid \mathbb{E}[Y^*Y] \geq 0\}$, where $Y^* \in L^q(\mathbb{R})$.

We first claim that if $\Lambda(X) + L_+^p(\mathbb{R}) \subseteq M = \{Y \in L^p(\mathbb{R}) \mid \mathbb{E}[Y^*Y] \geq 0\}$ for some $X \in L^p(\mathbb{R}^n)$, then $Y^* \in L^q(\mathbb{R}_+)$. To see this, note that $\mathbb{E}[Y^*(\Lambda(X) + d)] \geq 0$ if and only if $\mathbb{E}[Y^*d] \geq -\mathbb{E}[Y^*\Lambda(X)]$ for every $d \in L_+^p(\mathbb{R})$. Assume that $\mathbb{E}[Y^*d] < 0$ for some $d \in L_+^p(\mathbb{R})$. Since $L_+^p(\mathbb{R})$ is a cone, for every $\lambda > 0$, we have $\lambda d \in L_+^p(\mathbb{R})$. Also, $\lambda \mathbb{E}[Y^*d] \rightarrow -\infty$ as $\lambda \rightarrow 0$. However, $\lambda \mathbb{E}[Y^*d]$ is bounded by $-\mathbb{E}[Y^*\Lambda(X)]$, hence we get a contradiction. Therefore, $\mathbb{E}[Y^*d] \geq 0$ for all $d \in L_+^p(\mathbb{R})$, which implies that $Y^* \in L^q(\mathbb{R}_+)$. This completes the proof of the claim.

In view of the claim, let us take $M = \{Y \in L^p(\mathbb{R}) \mid \mathbb{E}[Y^*Y] \geq 0\}$ for some $Y^* \in L^q(\mathbb{R}_+)$. We aim to show that $\{X \in L^p(\mathbb{R}^n) \mid \Lambda(X) + L^p(\mathbb{R}_+) \subseteq M\}$ is closed. Note that

$$\{X \in L^p(\mathbb{R}^n) \mid \Lambda(X) + L^p(\mathbb{R}_+) \subseteq M\} = \{X \in L^p(\mathbb{R}^n) \mid \mathbb{E}[Y^*\Lambda(X)] \geq 0\}.$$

Let us first consider case (i), where $\tilde{\Lambda}$ is concave and bounded from above. Thanks to concavity, the set $\{X \in L^p(\mathbb{R}^n) \mid \mathbb{E}[Y^*\tilde{\Lambda}(X)] \geq 0\}$ is convex.

Suppose that $p < +\infty$. Take a sequence $(X^k)_{k \in \mathbb{N}}$ in $\{X \in L^p(\mathbb{R}^n) \mid \mathbb{E}[Y^* \Lambda(X)] \geq 0\}$ that converges to some $\tilde{X} \in L^p(\mathbb{R}^n)$ strongly. Hence, there exists a subsequence $(X^{k_\ell})_{\ell \in \mathbb{N}}$ that converges to \tilde{X} almost surely. By the continuity of $\tilde{\Lambda}$, and then reverse Fatou's lemma, we get

$$\begin{aligned} \mathbb{E}[Y^* \Lambda(\tilde{X})] &= \mathbb{E}[Y^* \tilde{\Lambda} \circ \tilde{X}] = \mathbb{E} \left[Y^* \lim_{\ell \rightarrow \infty} \tilde{\Lambda} \circ X^{k_\ell} \right] \\ &\geq \limsup_{\ell \rightarrow \infty} \mathbb{E}[Y^* \tilde{\Lambda} \circ X^{k_\ell}] = \limsup_{\ell \rightarrow \infty} \mathbb{E}[Y^* \Lambda(X^{k_\ell})] \geq 0. \end{aligned} \quad (7.14)$$

Hence, $\tilde{X} \in \{X \in L^p(\mathbb{R}^n) \mid \mathbb{E}[Y^* \Lambda(X)] \geq 0\}$ and this set is closed. Note that we can use reverse Fatou's lemma in the above calculation since $\tilde{\Lambda}$ is bounded from above so that $(Y^* \Lambda(X^{k_\ell}))_{\ell \in \mathbb{N}}$ is bounded from above.

Suppose that $p = +\infty$. To prove weak* closedness, let $r > 0$. By Krein-Šmulian theorem, it is enough to prove that $\{X \in L^\infty(\mathbb{R}^n) \mid \mathbb{E}[Y^* \Lambda(X)] \geq 0, \|X\|_\infty \leq r\}$ is closed in $L^1(\mathbb{R}^n)$. Let $(X^k)_{k \in \mathbb{N}}$ be a sequence in this set that converges to some $\tilde{X} \in L^1(\mathbb{R}^n)$ strongly in $L^1(\mathbb{R}^n)$. Hence, we may find a subsequence $(X^{k_\ell})_{\ell \in \mathbb{N}}$ that converges to \tilde{X} almost surely. Repeating the argument in (7.14), we see that $\mathbb{E}[Y^* \Lambda(\tilde{X})] \geq 0$. On the other hand, we have $\|X^{k_\ell}\| \leq r$ for all $\ell \in \mathbb{N}$ with probability one. Hence, $\|\tilde{X}\| \leq r$ with probability one so that $\|\tilde{X}\|_\infty \leq r$. It follows that $\tilde{X} \in \{X \in L^\infty(\mathbb{R}^n) \mid \mathbb{E}[Y^* \Lambda(X)] \geq 0, \|X\|_\infty \leq r\}$, proving the closedness of this set in $L^1(\mathbb{R}^n)$.

Next we consider case (ii), where $\tilde{\Lambda}$ and hence Λ are linear. In particular, there exists $a \in \mathbb{R}^n$ such that $\tilde{\Lambda}(x) = a^\top x$ for every $x \in \mathbb{R}^n$. Suppose that $p < +\infty$. Let us take a net $(X^k)_{k \in I}$ in $\{X \in L^p(\mathbb{R}^n) \mid \mathbb{E}[Y^* \Lambda(X)] \geq 0\}$ that converges to some $\tilde{X} \in L^p(\mathbb{R}^n)$ weakly, where I is an arbitrary index set. By linearity and weak convergence, we have

$$\mathbb{E}[Y^* \Lambda(\tilde{X})] = \mathbb{E}[Y^* \tilde{\Lambda} \circ \tilde{X}] = \mathbb{E}[(Y^* a)^\top \tilde{X}] = \lim_{k \in I} \mathbb{E}[(Y^* a)^\top X^k] \geq 0,$$

so that $\tilde{X} \in \{X \in L^p(\mathbb{R}^n) \mid \mathbb{E}[Y^* \Lambda(X)] \geq 0\}$, and this set is weakly closed, hence it is also strongly closed. The case $p = +\infty$ can be treated by Krein-Šmulian theorem as above.

For (iii), let us first observe that $(L^p(\mathbb{R}_+^n))^\# = L^p(\mathbb{R}_{++}^n)$ and $(L^p(\mathbb{R}_+))^\# = L^p(\mathbb{R}_{++})$. Now take $X, \bar{X} \in L^p(\mathbb{R}^n)$ with $X \leq_{L^p(\mathbb{R}_{++}^n)} \bar{X}$. Hence, for almost every $\omega \in \Omega$, we have $X(\omega) \leq_{\mathbb{R}_{++}^n} \bar{X}(\omega)$. Since $\tilde{\Lambda}$ is regularly increasing, we have $\Lambda(X)(\omega) = \tilde{\Lambda}(X(\omega)) < \tilde{\Lambda}(\bar{X}(\omega)) = \Lambda(\bar{X})(\omega)$ for almost every $\omega \in \Omega$. Therefore, $\Lambda(X) \leq_{L^p(\mathbb{R}_{++})} \Lambda(\bar{X})$. So Λ is regularly increasing. \square

Proof of Proposition 6.3 . Let $Y^* \in L^q(\mathbb{R}_+) \setminus \{0\}$. Since we have D -concavity, finding the penalty function is a concave maximization problem. Moreover, since the strict sublevel set is nonempty, Slater's condition holds. Hence, we can use strong duality and obtain

$$\begin{aligned}
\alpha_{(-h_{Y^*}^\Lambda)}(X^*, m) &= \sup_{X \in L^p(\mathbb{R}^n)} \left\{ \mathbb{E} \left[-(X^*)^\top X \right] \mid \mathbb{E} [-Y^* \Lambda(X)] \leq m \right\} \\
&= \inf_{\lambda > 0} \sup_{X \in L^p(\mathbb{R}^n)} \left(\mathbb{E} \left[-(X^*)^\top X \right] - \lambda \mathbb{E} [-Y^* \Lambda(X)] + \lambda m \right) \\
&= \inf_{\lambda > 0} \sup_{X \in L^p(\mathbb{R}^n)} \left(\mathbb{E} \left[-(X^*)^\top X + \lambda Y^* \Lambda(X) \right] + \lambda m \right) \\
&= \inf_{\lambda > 0} \left(\mathbb{E} \left[\sup_{x \in \mathbb{R}^n} \left(-(X^*)^\top x + \lambda Y^* \tilde{\Lambda}(x) \right) \right] + \lambda m \right),
\end{aligned}$$

where the second equality is by strong duality (we can ignore the case $\lambda = 0$ as it produces an objective value of $+\infty$) and the fourth equality is by Rockafellar and Wets [29, Thm. 14.60].

Note that for every $x^* \in \mathbb{R}^n$ and $y^* \in \mathbb{R}_+$, we have

$$\sup_{x \in \mathbb{R}^n} (-x^\top x^* + \lambda y^* \tilde{\Lambda}(x)) = \begin{cases} 0 & \text{if } x^* = 0, y^* = 0, \\ \infty & \text{if } x^* \neq 0, y^* = 0, \\ \lambda y^* \tilde{\Phi} \left(\frac{x^*}{\lambda y^*} \right) & \text{if } y^* > 0. \end{cases} \quad (7.15)$$

Therefore, $\alpha_{(-h_{Y^*}^\Lambda)}(X^*, m) = +\infty$ if $Y^* \notin T_{X^*}$, and

$$\alpha_{(-h_{Y^*}^\Lambda)}(X^*, m) = \inf_{\lambda > 0} \left(\mathbb{E} \left[\lambda Y^* \Phi \left(\frac{X^*}{\lambda Y^*} \right) 1_{\{Y^* > 0\}} \right] + \lambda m \right) \quad (7.16)$$

if $Y^* \in T_{X^*}$. Moreover, by Theorem 4.6,

$$\alpha_{\rho \circ \Lambda}(X^*, m) = \inf_{Y^* \in L_+^q(\mathbb{R}) \setminus \{0\}} \alpha_{(-h_{Y^*}^\Lambda)}(X^*, \alpha_\rho(Y^*, m)).$$

By combining this equality with (7.16), it follows that

$$\alpha_{\rho \circ \Lambda}(X^*, m) = \inf_{Y^* \in T_{X^*}} \inf_{\lambda > 0} \left(\mathbb{E} \left[\lambda Y^* \Phi \left(\frac{X^*}{\lambda Y^*} \right) 1_{\{Y^* > 0\}} \right] + \lambda \alpha_\rho(Y^*, m) \right).$$

Then, since T_{X^*} is a cone and α_ρ is positively homogeneous, we get

$$\alpha_{\rho \circ \Lambda}(X^*, m) = \inf_{Y^* \in T_{X^*}} \left(\mathbb{E} \left[Y^* \Phi \left(\frac{X^*}{Y^*} \right) 1_{\{Y^* > 0\}} \right] + \alpha_\rho(Y^*, m) \right),$$

as desired. \square

Proof of Proposition 6.4 . By Corollary 4.8 and Proposition 2.16, we have

$$\begin{aligned} \alpha_{\rho \circ \Lambda}^{-l}(X^*, s) &= \sup_{Y^* \in L^q(\mathbb{R}_+) \setminus \{0\}} \alpha_\rho^{-l} \left(Y^*, \alpha_{-h_{Y^*}^\Lambda}^{-l}(X^*, s) \right) \\ &= \sup_{Y^* \in L^q(\mathbb{R}_+) \setminus \{0\}} \alpha_\rho^{-l} \left(Y^*, \sup_{\gamma \geq 0} (\gamma s - (-h_{Y^*}^\Lambda)^*(-\gamma X^*)) \right) \\ &= \sup_{Y^* \in L^q(\mathbb{R}_+) \setminus \{0\}} \inf \left\{ m \in \mathbb{R} \mid \alpha_\rho(Y^*, m) \geq \sup_{\gamma \geq 0} (\gamma s - (-h_{Y^*}^\Lambda)^*(-\gamma X^*)) \right\} \\ &= \sup_{Y^* \in L^q(\mathbb{R}_+) \setminus \{0\}} \sup_{\gamma \geq 0} \alpha_\rho^{-l} \left(Y^*, \gamma s - (-h_{Y^*}^\Lambda)^*(-\gamma X^*) \right), \end{aligned} \quad (7.17)$$

where the last equality comes from Lemma 2.13. Let us calculate the second argument of α_ρ^{-l} for bounded case $\Phi(0) < +\infty$. For $\gamma = 0$, by using Rockafellar and Wets [29, Thm. 14.60], we have

$$-(-h_{Y^*}^\Lambda)^*(0) = - \sup_{Z \in L^p(\mathbb{R}^n)} \mathbb{E}[Y^* \Lambda(Z)] = -\mathbb{E} \left[\sup_{z \in \mathbb{R}^n} Y^* \Lambda(z) \right] = -\Phi(0) \mathbb{E}[Y^*].$$

Here, the last equality follows by the following simple observation: for every $y^* \in \mathbb{R}_+$,

$$\sup_{z \in \mathbb{R}^n} y^* \Lambda(z) = \begin{cases} 0 & \text{if } y^* = 0, \\ y^* \Phi(0) & \text{else.} \end{cases}$$

For $\gamma > 0$, by Rockafellar and Wets [29, Thm. 14.60], we get

$$(-h_{Y^*}^\Lambda)^*(-\gamma X^*) = \sup_{Z \in L^p(\mathbb{R}^n)} \left(-\mathbb{E} \left[\gamma Z^\top X^* \right] + \mathbb{E} [Y^* \Lambda(Z)] \right) = \mathbb{E} \left[\sup_{z \in \mathbb{R}^n} \left(-\gamma z^\top X^* + Y^* \Lambda(z) \right) \right].$$

Using the calculation in (7.15), it follows that $(-h_{Y^*}^\Lambda)^*(-\gamma X^*) = +\infty$ if $Y^* \notin T_{X^*}$, and

$$(-h_{Y^*}^\Lambda)^*(-\gamma X^*) = \mathbb{E} \left[Y^* \Phi \left(\frac{\gamma X^*}{Y^*} \right) 1_{\{Y^* > 0\}} \right]$$

if $Y^* \in T_{X^*}$. Since α_ρ^{-l} is increasing in the second argument, we can ignore the case $Y^* \notin T_{X^*}$, since the second argument of α_ρ^{-l} will be $-\infty$ in (7.17). By the positive homogeneity of α_ρ , for $\gamma > 0$, we have

$$\alpha_\rho^{-l} \left(Y^*, \gamma s - \mathbb{E} \left[Y^* \Phi \left(\frac{\gamma X^*}{Y^*} \right) 1_{\{Y^* > 0\}} \right] \right) = \alpha_\rho^{-l} \left(\frac{Y^*}{\gamma}, s - \mathbb{E} \left[\frac{Y^*}{\gamma} \Phi \left(\frac{\gamma X^*}{Y^*} \right) 1_{\{Y^* > 0\}} \right] \right).$$

By combining all the findings, we get

$$\begin{aligned} \alpha_{\rho \circ \Lambda}^{-l}(X^*, s) &= \sup_{Y^* \in L^q(\mathbb{R}_+) \setminus \{0\}} \sup_{\gamma \geq 0} \alpha_\rho^{-l} \left(Y^*, \gamma s - (-h_{Y^*}^\Lambda)^*(-\gamma X^*) \right) \\ &= \sup_{Y^* \in L^q(\mathbb{R}_+) \setminus \{0\}} \alpha_\rho^{-l} \left(Y^*, -\Phi(0) \mathbb{E}[Y^*] \right) \vee \sup_{\substack{Y^* \in T_{X^*}, \\ \gamma > 0}} \alpha_\rho^{-l} \left(\frac{Y^*}{\gamma}, s - \mathbb{E} \left[\frac{Y^*}{\gamma} \Phi \left(\frac{\gamma X^*}{Y^*} \right) 1_{\{Y^* > 0\}} \right] \right) \\ &= \sup_{Y^* \in L^q(\mathbb{R}_+) \setminus \{0\}} \alpha_\rho^{-l} \left(Y^*, -\Phi(0) \mathbb{E}[Y^*] \right) \vee \sup_{Y^* \in T_{X^*}} \alpha_\rho^{-l} \left(Y^*, s - \mathbb{E} \left[Y^* \Phi \left(\frac{X^*}{Y^*} \right) 1_{\{Y^* > 0\}} \right] \right), \end{aligned}$$

where the last equation comes from the fact that T_{X^*} is a cone. Now we can pass to the probabilistic setting. For the left side, make the change-of-variables $Y^* = \lambda \frac{d\mathbb{Q}}{d\mathbb{P}}$ where $\lambda > 0$ and $\mathbb{Q} \in \mathcal{M}_1^q(\mathbb{P})$.

By using the positive homogeneity of α_ρ , we have

$$\alpha_\rho^{-l} \left(Y^*, -\Phi(0) \mathbb{E}[Y^*] \right) = \alpha_\rho^{-l} \left(\lambda \frac{d\mathbb{Q}}{d\mathbb{P}}, -\Phi(0) \mathbb{E} \left[\lambda \frac{d\mathbb{Q}}{d\mathbb{P}} \right] \right) = \alpha_\rho^{-l} \left(\frac{d\mathbb{Q}}{d\mathbb{P}}, -\Phi(0) \mathbb{E} \left[\frac{d\mathbb{Q}}{d\mathbb{P}} \right] \right) = \alpha_\rho^{-l} \left(\frac{d\mathbb{Q}}{d\mathbb{P}}, -\Phi(0) \right),$$

which gives

$$\sup_{Y^* \in L^q(\mathbb{R}_+) \setminus \{0\}} \alpha_\rho^{-l} \left(Y^*, -\Phi(0) \mathbb{E}[Y^*] \right) = \sup_{\mathbb{Q} \in \mathcal{M}_1^q(\mathbb{P})} \alpha_\rho^{-l} \left(\frac{d\mathbb{Q}}{d\mathbb{P}}, -\Phi(0) \right).$$

For the other part, we can make the change-of-variables $X^* = w \cdot \frac{d\mathbb{S}}{d\mathbb{P}}$ and $Y^* = \lambda \frac{d\mathbb{Q}}{d\mathbb{P}}$ as before and get

$$\sup_{Y^* \in T_{X^*}} \alpha_\rho^{-l} \left(Y^*, s - \mathbb{E} \left[Y^* \Phi \left(\frac{X^*}{Y^*} \right) 1_{\{Y^* > 0\}} \right] \right) = \sup_{\substack{\mathbb{Q} \in \mathcal{M}_1^q(\mathbb{P}), \lambda > 0 \\ w_i \mathbb{S}_i \ll \mathbb{Q}}} \alpha_\rho^{-l} \left(\frac{d\mathbb{Q}}{d\mathbb{P}}, \frac{s}{\lambda} - \mathbb{E}_{\mathbb{Q}} \left[\Phi \left(\frac{w \cdot d\mathbb{S}}{\lambda d\mathbb{Q}} \right) \right] \right).$$

Finally, we have

$$\alpha_{\rho \circ \Lambda}^{-l} \left(w \cdot \frac{d\mathbb{S}}{d\mathbb{P}}, s \right) = \sup_{\mathbb{Q} \in \mathcal{M}_1^q(\mathbb{P})} \alpha_\rho^{-l} \left(\frac{d\mathbb{Q}}{d\mathbb{P}}, -\Phi(0) \right) \vee \sup_{\substack{\mathbb{Q} \in \mathcal{M}_1^q(\mathbb{P}), \lambda > 0 \\ w_i \mathbb{S}_i \ll \mathbb{Q}}} \alpha_\rho^{-l} \left(\frac{d\mathbb{Q}}{d\mathbb{P}}, \frac{s}{\lambda} - \mathbb{E}_{\mathbb{Q}} \left[\Phi \left(\frac{w \cdot d\mathbb{S}}{\lambda d\mathbb{Q}} \right) \right] \right).$$

For the unbounded case $\Phi(0) = \infty$, we can omit the first term above by the monotonicity of α_ρ^{-l} . \square

Proof of Proposition 6.5. By Proposition 4.11, we have the following

$$R(X) = \rho \circ \Lambda(X) = \sup_{X^* \in L^q(\mathbb{R}_+^n) \setminus \{0\}} \sup_{Y^* \in L^q(\mathbb{R}_+) \setminus \{0\}} \alpha_\rho^{-l} \left(Y^*, -\mathbb{E} \left[X^\top X^* \right] - (-h_{Y^*}^\Lambda)^*(-X^*) \right).$$

We calculate the second argument of α_ρ^{-l} . By Rockafellar and Wets [29, Thm. 14.60], we get

$$(-h_{Y^*}^\Lambda)^*(-X^*) = \sup_{Z \in L^p(\mathbb{R}^n)} \left(-\mathbb{E} \left[Z^\top X^* \right] + \mathbb{E} [Y^* \Lambda(Z)] \right) = \mathbb{E} \left[\sup_{z \in \mathbb{R}^n} \left(-z^\top X^* + Y^* \tilde{\Lambda}(z) \right) \right].$$

By the calculation in (7.15), we have $(-h_{Y^*}^\Lambda)^*(-X^*) = +\infty$ if $Y \notin T_{X^*}$, and

$$(-h_{Y^*}^\Lambda)^*(-X^*) = \mathbb{E} \left[Y^* \Phi \left(\frac{X^*}{Y^*} \right) 1_{\{Y^* > 0\}} \right]$$

if $Y^* \in T_{X^*}$. Since α_ρ^{-l} is increasing in the second argument, we can ignore the case $Y^* \notin T_{X^*}$ since the second argument will be $-\infty$. Therefore, we have

$$R(X) = \sup_{X^* \in L^q(\mathbb{R}_+^n)} \sup_{Y^* \in T_{X^*}} \alpha_\rho^{-l} \left(Y^*, -\mathbb{E} \left[X^\top X \right] - \mathbb{E} \left[Y^* \Phi \left(\frac{X^*}{Y^*} \right) 1_{\{Y^* > 0\}} \right] \right).$$

We can make the change-of-variables $X^* = w \cdot \frac{d\mathbb{S}}{d\mathbb{P}}$ and $Y^* = \lambda \frac{d\mathbb{Q}}{d\mathbb{P}}$ as before and we get

$$\begin{aligned} R(X) &= \sup_{X^* \in L^q(\mathbb{R}_+^n)} \sup_{Y^* \in T_{X^*}} \alpha_\rho^{-l} \left(Y^*, -\mathbb{E} \left[X^\top X \right] - \mathbb{E} \left[Y^* \Phi \left(\frac{X^*}{Y^*} \right) 1_{\{Y^* > 0\}} \right] \right) \\ &= \sup_{\substack{w \in \mathbb{R}_+^n \setminus \{0\}, \mathbb{S} \in \mathcal{M}_n^+(\mathbb{P}) \\ \mathbb{Q} \in \mathcal{M}_1^q(\mathbb{P}), w_i \mathbb{S}_i \ll \mathbb{Q}}} \alpha_\rho^{-l} \left(\frac{d\mathbb{Q}}{d\mathbb{P}}, -\mathbb{E}_{\mathbb{Q}} \left[\Phi \left(\frac{w \cdot d\mathbb{S}}{d\mathbb{Q}} \right) \right] - w^\top \mathbb{E}_{\mathbb{S}} [X] \right), \end{aligned}$$

after using the positive homogeneity of α_ρ and writing w instead of $\frac{w}{\lambda}$. \square

Proof of Proposition 6.11. Since we have concavity, finding the penalty function is a concave

maximization problem. Thanks to Slater's condition holds, we can use strong duality and obtain

$$\begin{aligned}
\alpha_{(-h_{Y^*}^\Lambda)}(X^*, m) &= \sup_{X \in L^p(\mathbb{R}_+^n)} \left\{ \mathbb{E} \left[-X^\top X^* \right] \mid \mathbb{E} \left[-Y^* \Lambda(X) \right] \leq m \right\} \\
&= \inf_{\lambda \geq 0} \sup_{X \in L^p(\mathbb{R}_+^n)} \left(\mathbb{E} \left[-X^\top X^* + \lambda Y^* \Lambda(X) \right] + \lambda m \right) \\
&= \inf_{\lambda \geq 0} \mathbb{E} \left[\sup_{x \in \mathbb{R}_+^n} \left(-x^\top X^* + \lambda Y^* \tilde{\Lambda}(x) + \lambda m \right) \right],
\end{aligned}$$

where last equality is by Rockafellar and Wets [29, Thm. 14.60]. For $\lambda = 0$, by using the fact that $X^* \in L^q(\mathbb{R}_+^n)$, we reach

$$\sup_{X \in L^p(\mathbb{R}_+^n)} \left(\mathbb{E} \left[-X^\top X^* + \lambda Y^* \Lambda(X) \right] + \lambda m \right) = \sup_{X \in L^p(\mathbb{R}_+^n)} \mathbb{E} \left[-X^\top X^* \right] = 0.$$

On the other hand, by the calculation in (7.15), we have

$$\alpha_{(-h_{Y^*}^\Lambda)}(X^*, m) = 0 \wedge \inf_{\lambda > 0} \left(\lambda m + \mathbb{E} \left[1_{\{Y^* > 0\}} \lambda Y^* \Phi \left(\frac{X^*}{\lambda Y^*} \right) \right] \right),$$

and by Corollary 4.12, we obtain

$$\begin{aligned}
\alpha_{\rho \circ \Lambda}(X^*, m) &= \inf_{Y^* \in L^q(\mathbb{R}_+) \setminus \{0\}} \alpha_{(-h_{Y^*}^\Lambda)}(X^*, \alpha_\rho(Y^*, m)) \\
&= \inf_{Y^* \in L^q(\mathbb{R}_+) \setminus \{0\}} 0 \wedge \inf_{\lambda > 0} \left(\lambda \alpha_\rho(Y^*, m) + \mathbb{E} \left[1_{\{Y^* > 0\}} \lambda Y^* \Phi \left(\frac{X^*}{\lambda Y^*} \right) \right] \right) \\
&= 0 \wedge \inf_{Y^* \in L^q(\mathbb{R}_+) \setminus \{0\}} \left(\alpha_\rho(Y^*, m) + \mathbb{E} \left[1_{\{Y^* > 0\}} Y^* \Phi \left(\frac{X^*}{Y^*} \right) \right] \right),
\end{aligned}$$

where last line follows as α is positively homogeneous in its first component and $L^q(\mathbb{R}_+)$ is a cone.

Next, let us fix some arbitrary $n \in \mathbb{N}$ and take

$$Y_n^* := \left(1 - \frac{1}{n} \right) Y^* 1_{\{Y^* > 0\}} + \frac{1}{n} 1_{\{Y^* = 0\}} \in L^q(\mathbb{R}_{++}).$$

Then, we have

$$\begin{aligned}
& \inf_{\bar{Y}^* \in L^q(\mathbb{R}_{++})} \left(\alpha_\rho(\bar{Y}^*, m) + \mathbb{E} \left[1_{\{\bar{Y}^* > 0\}} \bar{Y}^* \Phi \left(\frac{X^*}{\bar{Y}^*} \right) \right] \right) \\
& \leq \alpha_\rho(Y_n^*, m) + \mathbb{E} \left[1_{\{Y_n^* > 0\}} Y_n^* \Phi \left(\frac{X^*}{Y_n^*} \right) \right] \\
& = \sup_{Y \in S_n^p} -\mathbb{E}[Y Y_n^*] + \mathbb{E} \left[1_{\{Y_n^* > 0\}} \sup_{x \in \mathbb{R}_+^d} \left(-X^{*T} x + Y_n^* \tilde{\Lambda}(x) \right) \right] \\
& \leq \left(1 - \frac{1}{n} \right) \alpha_\rho(Y^* 1_{\{Y^* > 0\}}, m) + \frac{1}{n} \alpha_\rho(1_{\{Y^* = 0\}}, m) + \left(1 - \frac{1}{n} \right) \mathbb{E} \left[1_{\{Y^* > 0\}} Y^* \Phi \left(\frac{X^*}{Y^*} \right) \right] \\
& \quad + \frac{1}{n} \mathbb{E} \left[1_{\{1_{\{Y^* = 0\}} > 0\}} 1_{\{Y^* = 0\}} \Phi \left(\frac{X^*}{1_{\{Y^* = 0\}}} \right) \right],
\end{aligned}$$

where the last equality comes from the fact that supremum of affine functions is convex and indicator function of a convex set is a convex function. These inequalities are valid for every $n \in \mathbb{N}$, hence by sending n to ∞ , we get

$$\begin{aligned}
& \inf_{\bar{Y}^* \in L^q(\mathbb{R}_{++})} \left(\alpha_\rho(\bar{Y}^*, m) + \mathbb{E} \left[1_{\{\bar{Y}^* > 0\}} \bar{Y}^* \Phi \left(\frac{X^*}{\bar{Y}^*} \right) \right] \right) \\
& \leq \alpha_\rho(Y^* 1_{\{Y^* > 0\}}, m) + \mathbb{E} \left[1_{\{Y^* > 0\}} Y^* \Phi \left(\frac{X^*}{Y^*} \right) \right] = \alpha_\rho(Y^*, m) + \mathbb{E} \left[1_{\{Y^* > 0\}} Y^* \Phi \left(\frac{X^*}{Y^*} \right) \right],
\end{aligned}$$

where last equality is trivial since it is the set where $Y^* = 0$ and does not affect the expectation. Since this inequality true for every $Y^* \in L^q(\mathbb{R}_+) \setminus \{0\}$, by taking infimum we will have the following

$$\begin{aligned}
& \inf_{Y^* \in L^q(\mathbb{R}_{++})} \left(\alpha_\rho(Y^*, m) + \mathbb{E} \left[1_{\{Y^* > 0\}} Y^* \Phi \left(\frac{X^*}{Y^*} \right) \right] \right) \\
& \leq \inf_{Y^* \in L^q(\mathbb{R}_+) \setminus \{0\}} \left(\alpha_\rho(Y^*, m) + \mathbb{E} \left[1_{\{Y^* > 0\}} Y^* \Phi \left(\frac{X^*}{Y^*} \right) \right] \right).
\end{aligned}$$

Also since $L^q(\mathbb{R}_{++}) \subseteq L^q(\mathbb{R}_+) \setminus \{0\}$, the reverse inequality holds as well, hence we obtain

$$\begin{aligned}
& \inf_{Y^* \in L^q(\mathbb{R}_{++})} \left(\alpha_\rho(Y^*, m) + \mathbb{E} \left[1_{\{Y^* > 0\}} Y^* \Phi \left(\frac{X^*}{Y^*} \right) \right] \right) \\
& = \inf_{Y^* \in L^q(\mathbb{R}_+) \setminus \{0\}} \left(\alpha_\rho(Y^*, m) + \mathbb{E} \left[1_{\{Y^* > 0\}} Y^* \Phi \left(\frac{X^*}{Y^*} \right) \right] \right), \tag{7.18}
\end{aligned}$$

as desired. \square

Proof of Proposition 6.12. By Proposition 4.14 we have

$$R(X) = \rho \circ \Lambda(X) = \sup_{X^* \in L^q(\mathbb{R}_+^n) \setminus \{0\}} \sup_{Y^* \in L^q(\mathbb{R}_+) \setminus \{0\}} \alpha_\rho^{-l} \left(Y^*, -\mathbb{E} \left[X^\top X^* \right] - (-h_{Y^*}^\Lambda)^*(-X^*) \right).$$

We will calculate the second argument. By using Rockafellar and Wets [29, Thm. 14.60], we get

$$(-h_{Y^*}^\Lambda)^*(-X^*) = \sup_{Z \in L^p(\mathbb{R}_+^n)} \left(-\mathbb{E} \left[Z^\top X^* \right] + \mathbb{E} \left[Y^* \Lambda(Z) \right] \right) = \mathbb{E} \left[\sup_{z \in \mathbb{R}_+^n} \left(-z^\top X^* + Y^* \tilde{\Lambda}(z) \right) \right].$$

By (7.15), we have

$$(-h_{Y^*}^\Lambda)^*(-X^*) = \mathbb{E} \left[1_{\{Y^* > 0\}} Y^* \Phi \left(\frac{X^*}{Y^*} \right) \right].$$

Now, let us complete the proof by using Lemma 2.13 as follows:

$$\begin{aligned} & \sup_{Y^* \in L^q(\mathbb{R}_+) \setminus \{0\}} \alpha_\rho^{-l} \left(Y^*, -\mathbb{E} \left[X^\top X^* \right] - \mathbb{E} \left[1_{\{Y^* > 0\}} Y^* \Phi \left(\frac{X^*}{Y^*} \right) \right] \right) \\ &= \sup_{Y^* \in L^q(\mathbb{R}_+) \setminus \{0\}} \inf \left\{ m \in \mathbb{R} \mid \alpha_\rho(Y^*, m) \geq -\mathbb{E} \left[X^\top X^* \right] - \mathbb{E} \left[1_{\{Y^* > 0\}} Y^* \Phi \left(\frac{X^*}{Y^*} \right) \right] \right\} \\ &= \sup_{Y^* \in L^q(\mathbb{R}_+) \setminus \{0\}} \inf \left\{ m \in \mathbb{R} \mid \alpha_\rho(Y^*, m) + \mathbb{E} \left[1_{\{Y^* > 0\}} Y^* \Phi \left(\frac{X^*}{Y^*} \right) \right] \geq -\mathbb{E} \left[X^\top X^* \right] \right\} \\ &= \inf \left\{ m \in \mathbb{R} \mid \forall Y^* \in L^q(\mathbb{R}_+) \setminus \{0\}: \alpha_\rho(Y^*, m) + \mathbb{E} \left[1_{\{Y^* > 0\}} Y^* \Phi \left(\frac{X^*}{Y^*} \right) \right] \geq -\mathbb{E} \left[X^\top X^* \right] \right\} \\ &= \inf \left\{ m \in \mathbb{R} \mid \inf_{Y^* \in L^q(\mathbb{R}_+) \setminus \{0\}} \left(\alpha_\rho(Y^*, m) + \mathbb{E} \left[1_{\{Y^* > 0\}} Y^* \Phi \left(\frac{X^*}{Y^*} \right) \right] \right) \geq -\mathbb{E} \left[X^\top X^* \right] \right\} \\ &= \inf \left\{ m \in \mathbb{R} \mid \inf_{Y^* \in L^q(\mathbb{R}_{++})} \left(\alpha_\rho(Y^*, m) + \mathbb{E} \left[1_{\{Y^* > 0\}} Y^* \Phi \left(\frac{X^*}{Y^*} \right) \right] \right) \geq -\mathbb{E} \left[X^\top X^* \right] \right\} \\ &= \sup_{Y^* \in L^q(\mathbb{R}_{++})} \inf \left\{ m \in \mathbb{R} \mid \alpha_\rho(Y^*, m) \geq -\mathbb{E} \left[X^\top X^* \right] - \mathbb{E} \left[1_{\{Y^* > 0\}} Y^* \Phi \left(\frac{X^*}{Y^*} \right) \right] \right\} \\ &= \sup_{Y^* \in L^q(\mathbb{R}_{++})} \alpha_\rho^{-l} \left(Y^*, -\mathbb{E} \left[X^\top X^* \right] - \mathbb{E} \left[1_{\{Y^* > 0\}} Y^* \Phi \left(\frac{X^*}{Y^*} \right) \right] \right). \end{aligned}$$

Here, we use (7.18) in the fifth equality and Lemma 2.13 in the sixth equality. \square

Acknowledgments

This work is supported by the 3501 program of TÜBİTAK (Scientific & Technological Research Council of Turkey), Project No. 117F438. The authors thank Özlem Çavuş and Elisa Mastrogia-

como for their valuable feedback on a draft version of the paper.

References

- [1] Charalambos D. Aliprantis and Kim C. Border. *Infinite Dimensional Analysis: A Hitchhiker's Guide*. Springer, third edition edition, 2006.
- [2] Çağın Ararat. Portfolio optimization with two quasiconvex risk measures. *Turkish Journal of Mathematics*, 45:695–717, 2021.
- [3] Çağın Ararat and Birgit Rudloff. Dual representations for systemic risk measures. *Mathematics and Financial Economics*, 14:139–174, 2020.
- [4] Philippe Artzner, Freddy Delbaen, Jean-Marc Eber, and David Heath. Coherent measures of risk. *Mathematical Finance*, 9:203–228, 1999.
- [5] Francesca Biagini, Jean-Pierre Fouque, Marco Frittelli, and Thilo Meyer-Brandis. A unified approach to systemic risk measures via acceptance sets. *Mathematical Finance*, 29:329–367, 2019.
- [6] Radu Ioan Boț, Sorin-Mihai Grad, and Gert Wanka. Generalized Moreau-Rockafellar results for composed convex functions. *Optimization*, 58(7):917–933, 2009.
- [7] James V. Burke, Tim Hoheisel, and Quang V. Nguyen. A study of convex convex-composite functions via infimal convolution with applications. *Mathematics of Operations Research*, to appear, 2021. URL <https://doi.org/10.1287/moor.2020.1099>.
- [8] Simone Cerreia-Vioglio, Fabio Maccheroni, Massimo Marinacci, and Luigi Montrucchio. Risk measures: rationality and diversification. *Mathematical Finance*, 21:743–774, 2011.
- [9] Simone Cerreia-Vioglio, Fabio Maccheroni, Massimo Marinacci, and Luigi Montrucchio. Complete monotone quasiconcave duality. *Mathematics of Operations Research*, 36:321–339, 2011.
- [10] Chen Chen, Garud Iyengar, and Ciamac C. Moallemi. An axiomatic approach to systemic risk. *Management Science*, 59:1373–1388, 2013.
- [11] Cao-Zong Cheng and Bor-Luh Lin. Nonlinear two functions minimax theorems. In Biagio Ricceri and Stephen Simons, editors, *Minimax Theory and Applications*, pages 1–20. Kluwer Academic Publishers, 1998.
- [12] Samuel Drapeau and Michael Kupper. Risk preferences and their robust representation. *Mathematics of Operations Research*, 38(1):28–62, 2013.
- [13] Larry Eisenberg and Thomas H. Noe. Systemic risk in financial systems. *Management Science*, 47:236–249, 2001.
- [14] Zachary Feinstein, Birgit Rudloff, and Stefan Weber. Measures of systemic risk. *SIAM Journal on Financial Mathematics*, 8:672–708, 2017.
- [15] Hans Föllmer and Alexander Schied. *Stochastic Finance: An Introduction in Discrete Time*. De Gruyter, fourth edition edition, 2016.

- [16] Marco Frittelli and Matteo Maggis. Complete duality for quasiconvex dynamic risk measures on modules of the l^p -type. *Statistics and Risk Modeling*, 31(1):103–128, 2014.
- [17] Jochen Glück and Martin R. Weber. Almost interior points in ordered banach spaces and the long-term behaviour of strongly positive operator semigroups. *Studia Mathematica*, 254: 237–263, 2020.
- [18] Gabriele H. Greco and Maria Pia Moschen. A minimax inequality for marginally semicontinuous functions. In Biagio Ricceri and Stephen Simons, editors, *Minimax Theory and Applications*, pages 41–51. Kluwer Academic Publishers, 1998.
- [19] Truong Xuan Duc Ha. Demicontinuity, generalized convexity and loose saddle points of set-valued maps. *Optimization*, 51:293–308, 2002.
- [20] Andreas H. Hamel and Frank Heyde. Duality for set-valued measures of risk. *SIAM Journal on Financial Mathematics*, 1:66–95, 2010.
- [21] Andreas H. Hamel, Frank Heyde, and Birgit Rudloff. Set-valued risk measures for conical market models. *Mathematics and Financial Economics*, 5:1–28, 2011.
- [22] Sigrid Källblad. Risk- and ambiguity-averse portfolio optimization with quasiconvex utility functionals. *Finance and Stochastics*, 21:397–425, 2017.
- [23] Daishi Kuroiwa. Convexity for set-valued maps. *Applied Mathematics Letters*, 9(2):97–101, 1996.
- [24] Fon-Che Liu. A note on the von Neumann-Sion minimax principle. *Bulletin of the Institute of Mathematics. Academia Sinica*, 6(2, part 2):517–523, 1978.
- [25] Elisa Mastrogiacomo and Emanuela Rosazza Gianin. Portfolio optimization with quasiconvex risk measures. *Mathematics of Operations Research*, 40:1042–1059, 2015.
- [26] Jean-Paul Penot and Michel Volle. On quasi-convex duality. *Mathematics of Operations Research*, 15(4):597–625, 1990.
- [27] Michael Reed and Barry Simon. *I: Functional Analysis*. Methods of Modern Mathematical Physics. Academic Press, revised and enlarged edition edition, 1980.
- [28] R. Tyrrell Rockafellar. *Convex Analysis*. Princeton University Press, 1970.
- [29] R. Tyrrell Rockafellar and Roger J.-B. Wets. *Variational Analysis*. Springer, third printing, 2009.
- [30] Maurice Sion. On general minimax theorems. *Pacific Journal of Mathematics*, 8:171–176, 1958.
- [31] Tamaki Tanaka. Generalized quasiconvexities, cone saddle points, and minimax theorem for vector-valued functions. *Journal of Optimization Theory and Applications*, 81:355–377, 1994.
- [32] Constantin Zălinescu. *Convex Analysis in General Vector Spaces*. World Scientific, 2002.