Set-valued shortfall and divergence risk measures

Çağin Ararat* Andreas H. Hamel† Birgit Rudloff‡

May 19, 2014

Abstract

This paper is concerned with the utility-based risk of a financial position in a multi-asset market with frictions. Risk is quantified by set-valued risk measures, and market frictions are modeled by conical/convex random solvency regions representing proportional transaction costs or illiquidity effects, and convex random sets representing trading constraints. First, with a general set-valued risk measure, the effect of having trading opportunities on the risk measure is considered, and a corresponding dual representation theorem is given. Then, assuming individual utility functions for the assets, utility-based shortfall and divergence risk measures are defined, which form two classes of set-valued convex risk measures. Minimal penalty functions are computed in terms of the vector versions of the well-known divergence functionals (generalized relative entropy). As special cases, set-valued versions of the entropic risk measure and the average value at risk are obtained. The general results on the effect of market frictions are applied to the utility-based framework and conditions concerning applicability are presented.

Keywords and phrases: optimized certainty equivalent, shortfall risk, divergence, relative entropy, entropic risk measure, average value at risk, set-valued risk measure, transaction cost, solvency cone, infimal convolution, Lagrangian duality, set optimization

Mathematics Subject Classification (2010): 91B30, 46N10, 46A20, 26E25, 90C46.

1 Introduction

Risk measures for multivariate random variables have recently gained interest in the financial mathematics community. They become important in markets with frictions such as proportional transaction costs or with nonlinear illiquidity effects, but also in other situations where it might be advantageous to work with a random vector instead of a univariate random variable. These multivariate risk measures are functions which assign to a random vector a subset of a finite dimensional Euclidean space and thus, are set-valued functions. In the present paper, the random vectors will model the financial positions in a multi-asset market and their components are in terms of physical units rather than values with respect to a specific numéraire. The pioneering work on set-valued risk measures is provided in [26] for the coherent case. It has been extended to the convex case in [19] and to random market models in [20]. These extensions were possible by an application of the duality theory and, in particular, the Moreau-Fenchel biconjugation theorem for set-valued functions developed in [18]. Extensions to the dynamic framework have been studied in

*Princeton University, Department of Operations Research and Financial Engineering, Princeton, NJ 08540, USA, cararat@princeton.edu.
†Yeshiva University, Department of Mathematical Sciences, New York, NY 10033, USA, hamel@yu.edu
‡Princeton University, Department of Operations Research and Financial Engineering & Bendheim Center for Finance, Princeton, NJ 08540, USA, brudloff@princeton.edu.
Scalar risk measures for multivariate random variables, which can be interpreted as scalarizations of set-valued risk measures (see [14, Section 2.4]) have been studied in [25, 7, 35].

Set-valued generalizations of some well-known coherent risk measures of the scalar framework are also studied such as the set-valued version of the average value at risk in [23, 13, 22] or the set of superhedging portfolios in markets with transaction costs in [20, 27, 12]. Other examples of coherent risk measures for multivariate claims can be found in [3, 9]. To the best of our knowledge, apart from superhedging with certain trading constraints in markets with frictions, which leads to set-valued convex risk measures (see [22]), no other examples have been studied in the convex case yet.

The aim of this paper is to introduce two classes of set-valued convex risk measures, which are the set-valued analogues of the utility- or loss-based shortfall and divergence risk measures. Set-valued shortfall risk measures are defined as extensions of the scalar ones by generalizing the acceptability rule for random variables. Since the values of these risk measures are results of set-valued minimization problems, we study their corresponding Lagrangian dual problems using the new duality constructions in [21]. In analogy with the scalar case, it turns out that the dual objective functions give rise to a different class of convex risk measures, called the divergence risk measures. We compute the minimal penalty functions of these risk measures. As special cases, we obtain set-valued versions of the well-known entropic risk measure and recover the definition of set-valued average value at risk given in [23].

In Section 2, we give some preliminaries on set-valued risk measures. We generalize the notion of market-extension of a risk measure (see [23]) by including trading constraints modeled by random convex sets, and considering issues of liquidation into a certain collection of the assets. In contrast to [23], we allow for a convex (and not necessarily conical) market model to include temporary illiquidity effects in which the bid-ask prices depend on the magnitude of the trade and thus are given by the shape of the limit order book, see for instance [11, 30]. A dual representation result for this new type of market-extension are given in Proposition 2.17. A remarkable property is that, up to taking closures, the market-extension of a regulator risk measure can be seen as the set-valued infimal convolution of the regulator risk measure and set-valued indicator functions of the solvency cones/regions. This mathematical connection is discussed in Section 6.1 of the appendix, Section 6, together with the proofs of Proposition 2.17 and its corollaries.

In Section 3, we review the scalar theory of shortfall and divergence risk measures with slight generalizations on the choice of the loss and divergence functions. A technical remark about the assumptions on the scalar loss and divergence functions in comparison with [5] and [15] is given in Section 6.3 of the appendix. The main part of the paper is Section 4, where set-valued shortfall and divergence risk measures are studied. To make the similarities and differences with the scalar theory more apparent, Section 3 and Section 4 are organized in a parallel way. In Section 5, we define set-valued entropic risk measures as examples of shortfall risk measures and set-valued average value at risk as examples of divergence risk measures. Most of the proofs are collected in Section 6, the appendix.

2 Set-valued risk measures

2.1 Regulator risk measures

Let \( m \) be a strictly positive integer and \( |\cdot| \) an arbitrary fixed norm on \( \mathbb{R}^m \). By \( \mathbb{R}_+^m \) and \( \mathbb{R}_{++}^m \) we denote the set of elements of \( \mathbb{R}^m \) with nonnegative and strictly positive components, respectively. If \( m = 1 \), we abbreviate these symbols to \( \mathbb{R}, \mathbb{R}_+ \) and \( \mathbb{R}_{++} \). Throughout, we consider a probability
space \((\Omega, \mathcal{F}, \mathbb{P})\). We denote by \(L_m^0 := L_m^0(\Omega, \mathcal{F}, \mathbb{P})\) the linear space of random variables taking values in \(\mathbb{R}^m\), where two elements are identified if they are equal \(\mathbb{P}\)-almost surely; and we define

\[
L_m^1 = \{X \in L_m^0 \mid \mathbb{E}[|X|] < +\infty\}, \\
L_m^\infty = \{X \in L_m^0 \mid \text{ess sup} |X| < +\infty\}, \\
(L_m^p)^+ = \{X \in L_m^0 \mid \mathbb{P}\{X \in \mathbb{R}^m_+\} = 1\}, \quad p \in \{1, +\infty\}.
\]

If \(m = 1\), we omit \(m\) from the notation, for instance, we write \(L^\infty\) for \((L^\infty_1)^+\). We model a financial position as an element \(X \in L_m^\infty\), where \(X_\omega\) represents the number of (physical) units in the \(i\)th asset for \(i \in \{1, \ldots, m\}\) when the state of the world \(\omega \in \Omega\) occurs.

The linear space \(\mathbb{R}^m\) is called the space of eligible portfolios. This means that every \(z \in \mathbb{R}^m\) is a potential deposit to be used at initial time in order to compensate for the risk of a financial position. We denote by \(\mathcal{P}(\mathbb{R}^m)\) the power set of \(\mathbb{R}^m\), that is, the set of all subsets of \(\mathbb{R}^m\) including the empty set \(\emptyset\). On \(\mathcal{P}(\mathbb{R}^m)\), the Minkowski addition and multiplication with scalars are defined by \(A + B = \{a + b \mid a \in A, b \in B\}\) and \(sA = \{sa \mid a \in A\}\) for \(A, B \subseteq \mathbb{R}^m\) and \(s \in \mathbb{R}\) with the conventions \(A + \emptyset = \emptyset + B = \emptyset + \emptyset = \emptyset\), \(s\emptyset = \emptyset\) \((s \neq 0)\), and \(0\emptyset = \{0\} \subseteq \mathbb{R}^m\). We also use the shorthand notations \(A - B = A + (-1)B\) and \(z + A = \{z\} + A\). These operations can be defined on the power set \(\mathcal{P}(L_m^p)\) of \(L_m^p\), \(p \in \{0, 1, +\infty\}\), in a similar way.

Risk measures of interest will be functions defined on \(L_m^\infty\) and mapping into \(\mathcal{P}(\mathbb{R}^m)\). To such a function \(R: L_m^\infty \to \mathcal{P}(\mathbb{R}^m)\) we associate the set

\[
\text{graph } R := \{(X, z) \in L_m^\infty \times \mathbb{R}^m \mid z \in R(X)\}
\]

which is usually called the graph of \(R\).

**Definition 2.1.** ([20, Definition 2.3]) A function \(R : L_m^\infty \to \mathcal{P}(\mathbb{R}^m)\) is said to be a (regulator) risk measure if it satisfies the following properties:

(i) Finiteness at 0: \(R(0) \notin \{\emptyset, \mathbb{R}^m\}\).

(ii) Monotonicity: \(Z - X \in (L_m^\infty)^+\) implies \(R(Z) \supseteq R(X)\) for every \(X, Z \in L_m^\infty\).

(iii) Translativity: \(R(X + z) = R(X) - z\) for every \(X \in L_m^\infty\), \(z \in \mathbb{R}^m\).

If \(R\) is a risk measure, finiteness at 0 guarantees finiteness everywhere in the sense that \(R(X) \notin \{\emptyset, \mathbb{R}^m\}\) for every \(X \in L_m^\infty\). This is immediate from translativity and the fact that we are working on \(L_m^\infty\). Besides, using monotonicity and translativity, one easily checks that the values of \(R\) are elements of the set

\[
\mathcal{P}(\mathbb{R}^m, \mathbb{R}^m_+) := \{A \subseteq \mathbb{R}^m \mid A = A + \mathbb{R}^m_+\}. \quad (2.1)
\]

It turns out that \(\mathcal{P}(\mathbb{R}^m, \mathbb{R}^m_+)\) is a convenient image space\(^1\) to study set-valued functions, see [13]. In particular, it is an order complete lattice when equipped with the usual superset relation \(\supseteq\). We have the following infimum and supremum formulae for every nonempty subset \(A\) of \(\mathcal{P}(\mathbb{R}^m, \mathbb{R}^m_+)\):

\[
\inf_{(\mathcal{P}(\mathbb{R}^m, \mathbb{R}^m_+), \supseteq)} A = \bigcup_{A \in A} A, \quad \sup_{(\mathcal{P}(\mathbb{R}^m, \mathbb{R}^m_+), \supseteq)} A = \bigcap_{A \in A} A. \quad (2.2)
\]

We also use the conventions \(\inf_{(\mathcal{P}(\mathbb{R}^m, \mathbb{R}^m_+), \supseteq)} \emptyset = \emptyset\) and \(\sup_{(\mathcal{P}(\mathbb{R}^m, \mathbb{R}^m_+), \supseteq)} \emptyset = \mathbb{R}^m\). Hence, from now on, we limit our treatment to the functions mapping into \(\mathcal{P}(\mathbb{R}^m, \mathbb{R}^m_+)\).

\(^1\)The phrase “image space” for a set-valued function is meant the set (subset of a power set) where the function maps into. This set is not a linear space in general. In particular, \(\mathcal{P}(\mathbb{R}^m, \mathbb{R}^m_+)\) is a conlinear space in the sense of [17].
Definition 2.2. ([20] Definition 2.12) A function \( R : L_\infty^0 \to \mathcal{P}(\mathbb{R}^m, \mathbb{R}^m_+) \) is said to be convex if, for every \( X, Y \in L_\infty^0 \), \( \lambda \in (0, 1) \), we have
\[
R(\lambda X + (1 - \lambda)Y) \supseteq \lambda R(X) + (1 - \lambda)R(Y),
\]
and weak*-closed if graph \( R \) is closed with respect to the product topology of the weak* topology \( \sigma(L_\infty^0, L_1^0) \) and the usual topology on \( \mathbb{R}^m \).

The convexity of a function \( R : L_\infty^0 \to \mathcal{P}(\mathbb{R}^m, \mathbb{R}^m_+) \) is equivalent to the convexity of graph \( R \). Besides, if \( R \) is weak*-closed and convex, then its values are elements of
\[
G(\mathbb{R}^m, \mathbb{R}^m_+) := \{ A \subseteq \mathbb{R}^m : A = \text{cl co}(A + \mathbb{R}^m_+) \},
\]
where \( \text{cl} \) and \( \text{co} \) denote the closure and convex hull operators, respectively. The converse does not hold in general. The set \( (G(\mathbb{R}^m, \mathbb{R}^m_+), \supseteq) \) is also an order complete lattice with the same supremum formula as for \( (\mathcal{P}(\mathbb{R}^m, \mathbb{R}^m_+), \supseteq) \) and the following infimum formula for every nonempty subset \( A \) of \( G(\mathbb{R}^m, \mathbb{R}^m_+) \):
\[
\inf_{(G(\mathbb{R}^m, \mathbb{R}^m_+), \supseteq)} A = \text{cl co} \bigcup_{A \in A} A,
\]
which is motivated by the fact that the union of closed or convex sets is not necessarily closed or convex, respectively. We make the analogous conventions for the infimum and supremum of \( \emptyset \).

Definition 2.3. ([20] Definition 2.1) A set \( A \subseteq L_\infty^0 \) is said to be an acceptance set if \( A + (L_\infty^0)_+ \subseteq A \) and \( \emptyset \neq A \neq L_\infty^0 \).

To a function \( R : L_\infty^0 \to \mathcal{P}(\mathbb{R}^m, \mathbb{R}^m_+) \) we associate the set
\[
A_R := \{ X \in L_\infty^0 : 0 \in R(X) \},
\]
and to a set \( A \subseteq L_\infty^0 \) we associate the function \( R_A : L_\infty^0 \to \mathcal{P}(\mathbb{R}^m) \) given by
\[
R_A(X) := \{ z \in \mathbb{R}^m : X + z \in A \}.
\]
The next proposition establishes a one-to-one correspondence between risk measures and acceptance sets, which is analogous to the one for scalar risk measures.

Proposition 2.4. ([20] Propositions 2.4, 2.5, 2.13) Let \( A \subseteq L_\infty^0 \) be an \((-\), convex, weak*-closed, respectively) acceptance set. Then \( R_A \) is a \((-\), convex, weak*-closed) risk measure and \( A = A_{R_A} \).

Conversely, let \( R \) be a \((-\), convex, weak*-closed) risk measure. Then \( A_R \) is an \((-\), convex, weak*-closed) acceptance set and \( R = R_{A_R} \).

Finally, we state the result about the dual representation of convex weak*-closed risk measures in Proposition 2.6 below. To that end, let \( Q = (Q_1, \ldots, Q_m)^T \) be an \( m \)-dimensional vector probability measure in the sense that \( Q_i \) is a probability measure on \( (\Omega, \mathcal{F}) \) for each \( i \in \{1, \ldots, m\} \). Define \( \mathbb{E}^Q[X] = (\mathbb{E}^{Q_1}[X_1], \ldots, \mathbb{E}^{Q_m}[X_m])^T \) for every \( X \in L_0^0 \) such that the components exist in \( \mathbb{R} \). Denote by \( \mathcal{M}_m(\mathbb{P}) \) the set of all \( m \)-dimensional vector probability measures on \( (\Omega, \mathcal{F}) \) whose components are absolutely continuous with respect to \( \mathbb{P} \). For \( Q \in \mathcal{M}_m(\mathbb{P}) \), we set
\[
\frac{dQ}{d\mathbb{P}} = \left( \frac{dQ_1}{d\mathbb{P}}, \ldots, \frac{dQ_m}{d\mathbb{P}} \right)^T,
\]
where, for each \( i \in \{1, \ldots, m\} \), \( \frac{dQ_i}{d\mathbb{P}} \) denotes the Radon-Nikodym derivative of \( Q_i \) with respect to \( \mathbb{P} \). For \( w \in \mathbb{R}^m_+ \setminus \{0\} \), we define
\[
G(w) := \left\{ z \in \mathbb{R}^m : w^T z \geq 0 \right\}.
\]
In this case, closed risk measure if, and only if, there exists a penalty function 

\[ -\alpha : \mathcal{M}_m(\mathbb{P}) \times (\mathbb{R}_+^m \setminus \{0\}) \rightarrow \mathcal{G}(\mathbb{R}_+^m, \mathbb{R}_+^m) \]  

is said to be a penalty function if it satisfies:

(i) \(\bigcap_{(Q, w) \in \mathcal{M}_m(\mathbb{P}) \times (\mathbb{R}_+^m \setminus \{0\})} \{ -\alpha(Q, w) \neq \emptyset, \text{ and there exists } (Q, w) \in \mathcal{M}_m(\mathbb{P}) \times (\mathbb{R}_+^m \setminus \{0\}) \text{ such that } -\alpha(Q, w) \neq \mathbb{R}_+^m \}.\)

(ii) \(-\alpha(Q, w) = \text{cl}(-\alpha(Q, w) + G(w)) \text{ for every } (Q, w) \in \mathcal{M}_m(\mathbb{P}) \times (\mathbb{R}_+^m \setminus \{0\}).\)

Proposition 2.6. \((\text{[20, Theorem 4.2]}\)) A function \(R : L_\infty^m \rightarrow \mathcal{G}(\mathbb{R}_+^m, \mathbb{R}_+^m)\) is a convex weak*-closed risk measure if, and only if, there exists a penalty function \(-\alpha_R : \mathcal{M}_m(\mathbb{P}) \times (\mathbb{R}_+^m \setminus \{0\}) \rightarrow \mathcal{G}(\mathbb{R}_+^m, \mathbb{R}_+^m)\) such that for every \(X \in L_\infty^m,\)

\[ R(X) = \bigcap_{(Q, w) \in \mathcal{M}_m(\mathbb{P}) \times (\mathbb{R}_+^m \setminus \{0\})} \left[ -\alpha_R(Q, w) + \left( \mathbb{E}^Q [-X] + G(w) \right) \right]. \tag{2.7} \]

In this case, \(2.7\) holds when \(-\alpha_R\) is replaced with the minimal penalty function \(-\alpha_{R_{\text{min}}} : \mathcal{M}_m(\mathbb{P}) \times (\mathbb{R}_+^m \setminus \{0\}) \rightarrow \mathcal{G}(\mathbb{R}_+^m, \mathbb{R}_+^m)\) of \(R\) defined by

\[ -\alpha_{R_{\text{min}}}(Q, w) = \text{cl} \bigcup_{X \in L_\infty^m} \left[ R(X) + \left( \mathbb{E}^Q [X] + G(w) \right) \right], \tag{2.8} \]

and, for every penalty function \(-\alpha_R\) such that \(2.7\) holds, we have \(-\alpha_{R_{\text{min}}}(Q, w) \subseteq -\alpha_R(Q, w)\) for each \((Q, w) \in \mathcal{M}_m(\mathbb{P}) \times (\mathbb{R}_+^m \setminus \{0\}).\)

As in the scalar case, the minimal penalty function of a closed convex risk measure basically coincides with its Fenchel conjugate. In the set-valued case, the transformation from the set-valued conjugate with dual variables \(L_1^m \times (\mathbb{R}_+^m \setminus \{0\})\) to a penalty function with dual variables in \(\mathcal{M}_m(\mathbb{P}) \times (\mathbb{R}_+^m \setminus \{0\})\) requires some extra care; the procedure is described in detail in \([19, 20]\).

2.2 Market-extensions of risk measures

In this section, we consider the problem of minimizing the value of the regulator risk measure over the set of financial positions that can be reached with a given position by trading in the so-called convex market model. The result of the risk minimization, as a function of the given position, is called the market-extension of the regulator risk measure. Market-extensions were introduced in \([20]\) and \([23]\) for the special case of a conical market model. Here, we consider this notion for an arbitrary risk measure, allow a more general convex market model and include the possibility of trading constraints and liquidation issues.

Let \(d\) be a strictly positive integer denoting the number of assets in the market. We consider a convex market model in finite discrete time which is typically, as in \([30]\), used to describe a \(d\)-asset market with transaction costs. Let \(T > 0\) be an integer and \((\mathcal{F}_t)_{t=0}^T\) a filtration of \((\Omega, \mathcal{F}, \mathbb{P})\) augmented by the \(\mathbb{P}\)-null sets of \(\mathcal{F}\). The number \(T\) denotes the time horizon, and \((\mathcal{F}_t)_{t=0}^T\) represents the evolution of information over time. We assume that we have no information at time 0, that is, every \(\mathcal{F}_0\)-measurable function is deterministic \(\mathbb{P}\)-almost surely, and we have full information at time \(T\), that is, \(\mathcal{F}_T = \mathcal{F}\). For \(p \in \{0, 1, +\infty\}\), denote by \(L_p^d(\mathcal{F}_t)\) the linear subspace of all \(\mathcal{F}_t\)-measurable random variables in \(L_p^d\). Let \(D : \Omega \rightarrow \mathcal{P}(\mathbb{R}^d)\) be an \(\mathcal{F}_T\)-measurable function, that is, its graph \(\{(\omega, y) \in \Omega \times \mathbb{R}^d \mid y \in D(\omega)\}\) is \(\mathcal{F}_T \otimes \mathcal{B}(\mathbb{R}^d)\)-measurable, where \(\mathcal{B}(\mathbb{R}^d)\) denotes the Borel \(\sigma\)-algebra on \(\mathbb{R}^d\). We define \(L_p^d(\mathcal{F}_t, D) := \{ Y \in L_p^d(\mathcal{F}_t) \mid \mathbb{P}\{ \omega \in \Omega \mid Y(\omega) \in D(\omega)\} = 1\}\) for \(p \in \{0, 1, +\infty\}\).

For each \(t \in \{0, \ldots, T\}\), let \(C_t : \Omega \rightarrow \mathcal{G}(\mathbb{R}^d, \mathbb{R}_+^d)\) be an \(\mathcal{F}_t\)-measurable function such that \(\mathbb{R}_+^d \subseteq C_t(\omega)\) and \(\mathbb{R}_+^d \cap C_t(\omega) = \{0\}\) for each \(t \in \{0, \ldots, T\}\) and \(\omega \in \Omega\). The set \(C_t\) is called the
(random) solvency region at time \( t \), see [11, 30]. It models the bid and ask prices as a function of the magnitude of a trade, for instance, as in [10, 11, 33], and thus directly relates to the shape of the order book. More precisely, \( C_t(\omega) \) is the set of all portfolios which can be exchanged into ones with nonnegative components at time \( t \) when the outcome is \( \omega \in \Omega \). Convex solvency regions allow for the modeling of temporary illiquidity effects in the sense that they cover nonlinear illiquidity effects, but they assume that agents have no market power and thus their trades do not affect the costs of subsequent trades.

**Example 2.7.** An important special case is the conical market model introduced in [28]. Suppose that \( C_t(\omega) \) is a (closed convex) cone for each \( t \in \{0, \ldots, T\} \) and \( \omega \in \Omega \). In this case, the transaction costs are proportional to the size of the orders.

From a financial point of view, it is possible to have additional constraints on the trading opportunities at each time. For instance, trading may be allowed only up to a (possibly state- and time-dependent) threshold level for the assets (Example 2.8), or it may be the case that a certain linear combination of the trading units should not exceed a threshold level (Example 2.9). We model such constraints via convex random sets. Given \( t \in \{0, \ldots, T\} \), let \( D_t : \Omega \to P(\mathbb{R}^d) \) be an \( \mathcal{F}_t \)-measurable function such that \( D_t(\omega) \) is a closed convex set and \( 0 \in C_t(\omega) \cap D_t(\omega) \) for every \( \omega \in \Omega \). Note that \( D_t \) does not necessarily map into \( G(\mathbb{R}^d, \mathbb{R}^d_+) \), and this is why we prefer to work with \( C_t(\omega) \cap D_t(\omega) \) instead of replacing \( C_t \) by \( C_t \cap D_t \).

**Example 2.8.** For each \( t \in \{0, \ldots, T\} \), suppose that

\[
D_t = \bigcap_{i=1}^d (-\infty, (\bar{Y}_t)_i],
\]

for some \( \bar{Y}_t = ((\bar{Y}_t)_1, \ldots, (\bar{Y}_t)_d)^T \in L^0_d(\mathcal{F}_t, \mathbb{R}_d^d) \). In this case, trading in asset \( i \in \{1, \ldots, d\} \) at time \( t \in \{0, \ldots, T\} \) may not exceed the level \( (\bar{Y}_t)_i \).

**Example 2.9.** For each \( t \in \{0, \ldots, T\} \), suppose that

\[
D_t = \{ y \in \mathbb{R}^d \mid A_t^T y \leq B_t \},
\]

for some \( A_t \in L^0_d(\mathcal{F}_t, \mathbb{R}_d^d \setminus \{0\}) \) and \( B_t \in L^0_1(\mathcal{F}_t, \mathbb{R}_+). \) In this case, trading in each asset is unlimited but the linear combination of the trading units with the coefficient vector \( A_t \) cannot exceed the level \( B_t \).

Let

\[
K := -\sum_{t=0}^T L^\infty_{\mathcal{F}_t} (\mathcal{F}_t, C_t \cap D_t), \tag{2.9}
\]

which is the set of all financial positions that can be obtained by trading in the market starting with the zero position. Hence, an investor with a financial position \( Y \in L^\infty_{\mathcal{F}_t} \) can ideally reach any element of the set \( Y + K \) by trading in the market. However, it may be the case that the risk of the resulting position is evaluated only through a (small) selection of the \( d \) assets, in other words, trading has to be done in such a way that the only possibly nonzero components of the resulting position can be in some selected subset of the \( d \) assets. To model such a restriction, we let \( m \in \{1, \ldots, d\} \) and fix a risk measure \( R : L^\infty^m \to \mathcal{P}(\mathbb{R}^m, \mathbb{R}^m_+) \) which will be used for risk evaluation after liquidating the resulting position into \( L^\infty^m \). We make this idea precise by introducing the notion of liquidation function in Definition 2.10 below.
Let us introduce the linear operator \( B : \mathbb{R}^m \to \mathbb{R}^d \) defined by
\[
Bz = (z_1, \ldots, z_m, 0, \ldots, 0)^T.
\]
We will use the composition of \( B \) with random variables in \( L^0_m \). Given \( X \in L^0_m \), \( BX \) denotes the element in \( L^0_d \) defined by \( (BX)(\omega) = B(X(\omega)) \) for \( \omega \in \Omega \). The adjoint \( B^* : \mathbb{R}^d \to \mathbb{R}^m \) of \( B \) is given by
\[
B^*y = (y_1, \ldots, y_m)^T.
\]
Similarly, \( B^* \) can be composed with random variables in \( L^0_d \). With a slight abuse of notation, we will also use \( B^* \) in the context of vector probability measures. Given \( \mathbb{Q} \in \mathcal{M}_d(\mathbb{P}) \), we define \( B^*\mathbb{Q} = (\mathbb{Q}_1, \ldots, \mathbb{Q}_m) \in \mathcal{M}_m(\mathbb{P}) \). Similar to (2.1), we define
\[
\mathcal{P}(L^\infty_m, (L^\infty_m)_+) = \{ A \subseteq L^\infty_m \mid A + (L^\infty_m)_+ = A \}.
\]

**Definition 2.10.** The function \( \Lambda_m : L^\infty_d \to \mathcal{P}(L^\infty_m, (L^\infty_m)_+) \) defined by
\[
\Lambda_m(Y) = \{ X \in L^\infty_m \mid BX \in Y + \mathcal{K} \}
\]
is called the liquidation function associated with \( \mathcal{K} \).

Hence, given \( Y \in L^\infty_d \), \( \Lambda_m(Y) \) is the set of all possible resulting positions in \( Y + \mathcal{K} \) that are already liquidated into the first \( m \) assets. We look for the minimized value of the regulator risk measure \( R \) over the set \( \Lambda_m(Y) \) as the following definition suggests.

**Definition 2.11.** The function \( R^\text{mar} : L^\infty_d \to \mathcal{P}(\mathbb{R}^m, \mathbb{R}^m_+) \) defined by
\[
R^\text{mar}(Y) := (R \circ \Lambda_m)(Y) := \inf_{(\mathcal{P}(\mathbb{R}^m, \mathbb{R}^m_+), \mathcal{P})} \{ R(X) \mid X \in \Lambda_m(Y) \} = \bigcup_{X \in \Lambda_m(Y)} R(X)
\]
is called the market-extension of \( R \).

**Remark 2.12.** In the case of the conical market model described in Example 2.7, when \( \mathcal{D}_t = \mathbb{R}^d \) for each \( t \in \{0, \ldots, T\} \) and no liquidation at \( t = T \) is considered \( (m = d) \), Definition 2.11 recovers the notion of market-extension given in [23, Definition 2.8, Remark 2.9].

**Remark 2.13.** The function \( R^\text{mar} \) can be regarded as a risk measure according to a more general definition than Definition 2.1 as in [20, Definition 2.3]. Indeed, by identifying \( \mathbb{R}^m \) with the subspace \( M := \mathbb{R}^m \times \{ 0 \in \mathbb{R}^{d-m} \} \) of \( \mathbb{R}^d \), we can clearly identify \( R^\text{mar} \) with the function \( Y \mapsto R^\text{mar,}\mathcal{M}(Y) = B(R^\text{mar}(Y)) \) on \( L^\infty_d \) whose image space is \( \mathcal{P}(M, M_+) = \{ A \subseteq M \mid A = A + M_+ \} \), where \( M_+ = \mathbb{R}^m_+ \times \{ 0 \in \mathbb{R}^{d-m} \} \). Using the vocabulary in [20], it is easy to check that \( R^\text{mar,}\mathcal{M} \) is \( M \)-translative and \( (L^\infty_d)_+ \)-monotone but it may fail to have finite values. Indeed, while it is clear that the values of \( R^\text{mar,}\mathcal{M} \) (or equivalently of \( R^\text{mar} \)) are nonempty sets, it is possible that \( R^\text{mar,}\mathcal{M}(Y) = M \), that is, \( R^\text{mar}(Y) = \mathbb{R}^m \) for some \( Y \in L^\infty_d \). Moreover, \( R^\text{mar,}\mathcal{M} \) and \( R^\text{mar} \) are convex if \( R \) is so.

Similar to (2.5), given a function \( F : L^\infty_d \to \mathcal{P}(\mathbb{R}^m, \mathbb{R}^m_+) \), we associate to it the set
\[
A_F := \{ Y \in L^\infty_d \mid 0 \in F(Y) \}.
\]

In particular, in view of Remark 2.13, \( A_{R^\text{mar}} \) can be regarded as the acceptance set of \( R^\text{mar} \). Proposition 2.14 below shows that \( A_{R^\text{mar}} \) can be written in terms of the acceptance set \( A_R \subseteq L^\infty_m \) of the regulator risk measure \( R \) and the set \( \mathcal{K} \) of freely available portfolios.
Proposition 2.14. We have $A_{R^{\text{mar}}} = BA_R - K := \{BX - K \mid X \in A_R, K \in \mathcal{K}\}$.

Proof. For $A \subseteq L_+^\infty$, define $R^m_A : L_+^\infty \to \mathcal{P}(\mathbb{R}^m)$ by

$$R^m_A(Y) = \{z \in \mathbb{R}^m \mid Y + Bz \in A\}. \quad (2.13)$$

Let $Y \in L_+^\infty$. We have

$$R^m_{BA_R - K}(Y) = \{z \in \mathbb{R}^m \mid Y + Bz \in BA_R - K\} = \bigcup_{K \in \mathcal{K}} \{z \in \mathbb{R}^m \mid Y + K + Bz \in BA_R\} = \bigcup_{X \in L_+^\infty, BX \in Y + K} \{z \in \mathbb{R}^m \mid X + z \in A_R\} = \bigcup_{X \in \Lambda_m(Y)} R(X) = R^{\text{mar}}(Y).$$

By a trivial modification of [20, Proposition 2.4] in view of Remark 2.13, which basically tells that acceptance sets and risk measures are in one-to-one relationship, the result follows.

Note that $R^{\text{mar}}$ may fail to be weak*-closed even if $R$ is so. To recover weak*-closedness, we define closed versions of market-extensions.

Definition 2.15. Let $F : L_+^\infty \to \mathcal{P}(\mathbb{R}^m, \mathbb{R}_+^m)$ be a function. The function $\text{cl} \ F := R^{m}_{\text{cl}A_F}$ is called the closed hull of $F$ where the closure of $A_F$ is taken with respect to $\sigma(L_+^\infty, L_+^d)$ and the notation is from (2.12), (2.13). The closed hull of the market-extension of a risk measure $R : L_+^m \to \mathcal{P}(\mathbb{R}^m, \mathbb{R}_+^m)$ is called the closed market-extension of $R$.

Remark 2.16. In analogy with the scalar case, the closed hull of a function $F : L_+^\infty \to \mathcal{P}(\mathbb{R}^m, \mathbb{R}_+^m)$ is the pointwise greatest closed function minorizing it, that is, if $F' : L_+^\infty \to \mathcal{P}(\mathbb{R}^m, \mathbb{R}_+^m)$ is a weak*-closed function such that $F(Y) \subseteq F'(Y)$ for all $Y \in L_+^\infty$, then we have $(\text{cl} F)(Y) \subseteq F'(Y)$ for every $Y \in L_+^\infty$. This is by [18, Corollary 1(ii)].

By Proposition 2.4, the closed hull of a function $F : L_+^\infty \to \mathcal{P}(\mathbb{R}^m, \mathbb{R}_+^m)$ is always weak*-closed. Translativity, monotonicity and convexity are preserved under taking the closed hull (with respect to their corresponding definitions in [20]). Hence, in view of Remark 2.13, the closed market-extension of a convex risk measure $R : L_+^m \to \mathcal{P}(\mathbb{R}^m, \mathbb{R}_+^m)$ can be regarded as a weak*-closed convex risk measure if it has finite values. Proposition 2.17 below gives a dual representation of the closed market-extension in terms of the minimal penalty function of the regulator risk measure under the assumptions of convexity, weak*-closedness, and finiteness of values. The special case of no trading constraints in a convex (conical) market model is given in Corollary 2.18 (Corollary 2.19).

The set of dual variables to be used in the results below is given by

$$\mathcal{W}_{m,d} := \mathcal{M}_m(\mathbb{P}) \times ((\mathbb{R}_+^m \setminus \{0\}) \times \mathbb{R}_+^{d-m}).$$

We will also make use of the homogeneous halfspaces $G(w) := \{y \in \mathbb{R}_+^d \mid w^T y \geq 0\}$ for $w \in \mathbb{R}_+^d \setminus \{0\}$.

Proposition 2.17. Suppose that $R : L_+^\infty \to G(\mathbb{R}^m, \mathbb{R}_+^m)$ is a convex weak*-closed regulator risk measure with minimal penalty function $-\alpha^\text{mar}_R : \mathcal{M}_m(\mathbb{P}) \times (\mathbb{R}_+^m \setminus \{0\}) \to G(\mathbb{R}^m, \mathbb{R}_+^m)$. Assume that $\text{cl} R^{\text{mar}}$ is finite at 0. Then the closed market-extension $\text{cl} R^{\text{mar}} : L_+^\infty \to G(\mathbb{R}^m, \mathbb{R}_+^m)$ can be
regarded as a weak$^*$-closed convex risk measure in view of Remark \ref{rmk:L^ndual}, and it has the following dual representation: For every $Y \in L^\infty_d$,
\[
(\text{cl } R^\text{mar})(Y) = \bigcap_{(Q,w) \in \mathcal{W}_{m,d}} \left[ -\alpha^\text{min}_{\text{cl } R^\text{mar}}(Q, w) + B^* \left( \left( \mathbb{E}^Q[-Y] + G(w) \right) \cap B(\mathbb{R}^m) \right) \right],
\]
where $-\alpha^\text{min}_{\text{cl } R^\text{mar}} : \mathcal{W}_{m,d} \rightarrow \mathcal{G}(\mathbb{R}^m, \mathbb{R}^m_+)$ is defined by
\[
-\alpha^\text{min}_{\text{cl } R^\text{mar}}(Q, w) = -\alpha^R(\mathbb{E}^Q Q, B^* w) + \sum_{t=0}^{T} \bigcup_{U^t \in L^\infty_d(F_t, \mathcal{C}_t \cap \mathcal{D}_t)} B^* \left( \left( \mathbb{E}^Q[U^t] + G(w) \right) \cap B(\mathbb{R}^m) \right).
\]

Recall that the recession cone of a nonempty convex set $C \subseteq \mathbb{R}^d$ is the convex cone $0^+C := \{ y \in \mathbb{R}^d : y + C \subseteq C \}$ and the positive dual cone of a nonempty convex cone $K \subseteq \mathbb{R}^d$ is the convex cone $K^+ := \{ y \in \mathbb{R}^d : \forall k \in K : y^\top k \geq 0 \}$; see \cite[Section 8, p. 61]{2011} and \cite[Section 1.1, p. 7]{2016}, for instance.

**Corollary 2.18.** Under the assumptions of Proposition \ref{prop:regulator}, suppose that $\mathcal{D}_t \equiv \mathbb{R}^d$ for each $t \in \{0, \ldots, T\}$. Then $-\alpha^\text{min}_{\text{cl } R^\text{mar}}$ given by (2.14) is concentrated on the set
\[
\mathcal{W}_{m,d}^{\text{convex}} := \left\{ (Q, w) \in \mathcal{W}_{m,d} : \forall t \in \{0, \ldots, T\} : \text{diag}(w) \mathbb{E} \left[ \frac{dQ}{d\mathbb{P}} \bigg| F_t \right] \in L^1_d(F_t, (0^+C_t)^+) \right\},
\]
where, for each $t \in \{0, \ldots, T\}$, $(0^+C_t)^+ : \Omega \rightarrow \mathcal{G}(\mathbb{R}^d, \mathbb{R}^d_+)$ is the measurable function defined by $(0^+C_t)^+(\omega) := (0^+C_t(\omega))^+$.

In other words, we have $-\alpha^\text{min}_{\text{cl } R^\text{mar}}(Q, w) = \mathbb{R}^m$ for $(Q, w) \in \mathcal{W}_{m,d} \setminus \mathcal{W}_{m,d}^{\text{convex}}$ within the setting of the previous result.

**Corollary 2.19.** Under the assumptions of Proposition \ref{prop:regulator} suppose that $\mathcal{D}_t \equiv \mathbb{R}^d$ for each $t \in \{0, \ldots, T\}$ and that the market model is conical as in Example \ref{ex:market}. Consider the set
\[
\mathcal{W}_{m,d}^{\text{cone}} := \left\{ (Q, w) \in \mathcal{W}_{m,d} : \forall t \in \{0, \ldots, T\} : \text{diag}(w) \mathbb{E} \left[ \frac{dQ}{d\mathbb{P}} \bigg| F_t \right] \in L^1_d(F_t, C_t^+) \right\},
\]
where, for each $t \in \{0, \ldots, T\}$, $C_t^+ : \Omega \rightarrow \mathcal{G}(\mathbb{R}^d, \mathbb{R}^d_+)$ is the measurable function defined by $C_t^+(\omega) := (C_t(\omega))^+$. Then (2.14) reduces to
\[
-\alpha^\text{min}_{\text{cl } R^\text{mar}}(Q, w) = \begin{cases} -\alpha^R(\mathbb{E}^Q Q, B^* w) & \text{if } (Q, w) \in \mathcal{W}_{m,d}^{\text{cone}} \\ \mathbb{R}^m & \text{else} \end{cases},
\]
for each $(Q, w) \in \mathcal{W}_{m,d}$, hence for every $Y \in L^\infty_d$,
\[
(\text{cl } R^\text{mar})(Y) = \bigcap_{(Q,w) \in \mathcal{W}_{m,d}^{\text{cone}}} \left[ -\alpha^R(\mathbb{E}^Q Q, B^* w) + B^* \left( \left( \mathbb{E}^Q[-Y] + G(w) \right) \cap B(\mathbb{R}^m) \right) \right].
\]

The proofs of Proposition \ref{prop:regulator}, Corollary \ref{cor:convex}, Corollary \ref{cor:cone}, above are given in Section \ref{sec:proofs}. They rely on the observation that, roughly speaking, the market-extension of a regulator risk measure is the (set-valued) infimal convolution of the regulator risk measure and the (set-valued) indicator functions of the convex sets $L^\infty_d(F_t, \mathcal{C}_t \cap \mathcal{D}_t)$, $t \in \{0, \ldots, T\}$.\footnote{This is up to taking closures of sets and extending the regulator risk measure from $L^\infty_m$ to $L^\infty_d$ in a trivial way as it will be made precise in Section \ref{sec:proofs}.} This technical observation is discussed in Section \ref{sec:proofs} where the definitions of these notions are also given.
3 Scalar shortfall and divergence risk measures

In this section, we summarize the theory of scalar (utility/loss-based) shortfall and divergence risk measures on $L^\infty = L^\infty_1$ taking values in $\mathbb{R}$. Shortfall risk measures are introduced in [15]. Divergence risk measures are introduced in [4] and analyzed further in [5] with the name optimized certainty equivalent for their negatives. The dual relationship between these risk measures is pointed out in [34] and [5]. In terms of the assumptions on the loss function to be used, we slightly generalize the results of these papers; see Section 6.3 for a comparison.

The proofs of the results of this section are given in Section 6.2 and most of them inherit the convex duality arguments in [5] rather than the analytic arguments in [15].

**Definition 3.1.** A convex, lower semicontinuous function $f : \mathbb{R} \rightarrow \mathbb{R} \cup \{+\infty\}$ is said to be a loss function if it satisfies the following properties:

(i) $f$ is nondecreasing.
(ii) $0 \in \text{dom } f$.
(iii) $f$ is not identically constant on $\text{dom } f$.

Throughout this section, let $\ell : \mathbb{R} \rightarrow \mathbb{R} \cup \{+\infty\}$ be a loss function. **Definition 3.1** above guarantees that $\text{int } \ell(\mathbb{R}) \neq \emptyset$, where int denotes the interior operator. Assume $x^0 \in \text{int } \ell(\mathbb{R})$.

**Definition 3.2.** The function $\rho_\ell : L^\infty \rightarrow \mathbb{R} \cup \{+\infty\}$ defined by

$$\rho_\ell(X) = \inf \{s \in \mathbb{R} | \mathbb{E}[\ell(-X - s)] \leq x^0\}$$

is called the $\ell$-shortfall risk measure (on $L^\infty$ with threshold level $x^0$).

**Proposition 3.3.** The function $\rho_\ell$ is a convex weak$^*$-lower semicontinuous risk measure in the sense of [16, Definitions 4.1, 4.4]. In particular, $\rho_\ell$ takes values in $\mathbb{R}$.

According to **Definition 3.2** given $X \in L^\infty$, $\rho_\ell(X)$ can be seen as the value of a convex minimization problem. The next proposition computes $\rho_\ell(X)$ as the value of the corresponding Lagrangian dual problem. Its proof in Section 6.2 is an easy application of strong duality.

**Proposition 3.4.** For every $X \in L^\infty$,

$$\rho_\ell(X) = \sup_{\lambda > 0} \left( \inf_{s \in \mathbb{R}} (s + \lambda \mathbb{E}[\ell(-X - s)]) - \lambda x^0 \right). \quad (3.1)$$

Note that $X \mapsto \inf_{s \in \mathbb{R}} (s + \lambda \mathbb{E}[\ell(-X - s)])$ on $L^\infty$ is a monotone and transitive function for each $\lambda \in \mathbb{R}_+$. Our aim is to determine the values of $\lambda$ for which this function is a convex weak$^*$-lower semicontinuous risk measure. To that end, we are interested in the properties of the Legendre-Fenchel conjugate $g : \mathbb{R} \rightarrow \mathbb{R} \cup \{+\infty\}$ of the loss function $\ell$ given by

$$g(y) := \ell^*(-y) = \sup_{x \in \mathbb{R}} (xy - \ell(x)).$$

In the following, we will adopt the convention $(+\infty) \cdot 0 = 0$ as usual in convex analysis, see [32]. We will also use $1/ -\infty = 0$ as well as $1/0 = +\infty$.

**Definition 3.5.** A proper, convex, lower semicontinuous function $\varphi : \mathbb{R} \rightarrow \mathbb{R} \cup \{+\infty\}$ is said to be a divergence function if it satisfies the following properties:

(i) $\text{dom } \varphi \subseteq \mathbb{R}_+$.
(ii) $\varphi$ attains its infimum over $\mathbb{R}$.
(iii) $\varphi$ is not of the form $y \mapsto +\infty \cdot 1_{\{y < 0\}} + (ay + b) \cdot 1_{\{y \geq 0\}}$ with $a \in \mathbb{R}_+ \cup \{+\infty\}$ and $b \in \mathbb{R}$.
Proposition 3.6. Legendre-Fenchel conjugation furnishes a bijection between loss and divergence functions.

Remark 3.7. Let \( \lambda > 0 \). Given a loss function \( f \), \( \lambda f \) is also a loss function. Given a divergence function \( \varphi \), the function \( y \mapsto \varphi(\frac{y}{\lambda}) \) on \( \mathbb{R} \) is also a divergence function. The functions \( f \) and \( \varphi \) are conjugates of each other if, and only if, \( \lambda f \) and \( \varphi \lambda \) are.

Definition 3.8. Let \( \varphi \) be a divergence function, \( \lambda > 0 \) and \( Q \in \mathcal{M}_1(\mathbb{P}) \). The quantity

\[
I_{\varphi,\lambda}(Q \mid \mathbb{P}) := \mathbb{E} \left[ \varphi(\frac{dQ}{d\mathbb{P}}) \right] = \lambda \mathbb{E} \left[ \varphi \left( \frac{1}{\lambda} \frac{dQ}{d\mathbb{P}} \right) \right]
\]

is called the \((\varphi, \lambda)\)-divergence of \( Q \) with respect to \( \mathbb{P} \).

Note that \( g = \ell^* \) is a divergence function, and \( \text{dom}\ g \) is an interval (possibly a singleton) with some endpoints \( \alpha \in \mathbb{R}_+, \beta \in \mathbb{R}_+ \cup \{+\infty\} \) with \( \alpha \leq \beta \). We also have that \( \text{dom}\ g \neq \{0\} \) since otherwise \( g \) would be of the form \( y \mapsto +\infty \cdot 1_{\{y<0\}} + (ay+b) \cdot 1_{\{y\geq0\}} \) for \( a = +\infty \) and some \( b \in \mathbb{R} \). Finally, for each \( \lambda > 0 \), \( y \mapsto g_\lambda(y) := \lambda g(y/\lambda) \) on \( \mathbb{R} \) is a divergence function by Remark 3.7, and the corresponding \((g, \lambda)\)-divergences are defined according to Definition 3.8.

Theorem 3.9. For every \( \lambda > 0 \) and \( X \in L^\infty \),

\[
\inf_{s \in \mathbb{R}} \left( s + \mathbb{E} \left[ \ell(-X - s) \right] \right) = \sup_{Q \in \mathcal{M}_1(\mathbb{P})} \left( \mathbb{E}^Q[-X] - I_{g,\lambda}(Q \mid \mathbb{P}) \right).
\]

Moreover, the quantity in (3.2) is finite if

\[
\lambda \in \mathbb{R}_+ \cup \{0\} \iff y \not\in \text{dom}\ g \}
\]

and it is equal to \(-\infty\) if \( \lambda \not\in \mathbb{R}_+ \).

For every \( \lambda > 0 \), define a function \( \delta_{g,\lambda} : L^\infty \to \mathbb{R} \cup \{-\infty\} \) by

\[
\delta_{g,\lambda}(X) := \sup_{Q \in \mathcal{M}_1(\mathbb{P})} \left( \mathbb{E}^Q[-X] - I_{g,\lambda}(Q \mid \mathbb{P}) \right) - \lambda x^0.
\]

Corollary 3.10. For every \( \lambda > 0 \) and \( X \in L^\infty \),

\[
\rho_\ell(X) = \sup_{\lambda \in 1/\text{dom}\ g} \delta_{g,\lambda}(X).
\]

Proof. By Theorem 3.9 \( \delta_{g,\lambda} \) maps into \( \mathbb{R} \) for \( \lambda \in 1/\text{dom}\ g \), and \( \delta_{g,\lambda} \equiv -\infty \) for \( \lambda \not\in 1/\text{dom}\ g \). The result is an immediate consequence of Proposition 3.4. \( \square \)

Definition 3.11. Let \( \lambda \in 1/\text{dom}\ g \). The function \( \delta_{g,\lambda} : L^\infty \to \mathbb{R} \) defined by (3.4) is called the \((g, \lambda)\)-divergence risk measure on \( L^\infty \) with threshold level \( x^0 \).

By [16] Theorem 4.33, a convex weak*-lower semicontinuous risk measure \( \rho : L^\infty \to \mathbb{R} \) has the dual representation

\[
\rho(X) = \sup_{Q \in \mathcal{M}_1(\mathbb{P})} \left( \mathbb{E}^Q[-X] - \alpha^\text{min}_\rho(Q) \right),
\]

where \( \alpha^\text{min}_\rho : \mathcal{M}_1(\mathbb{P}) \to \mathbb{R} \cup \{+\infty\} \) is the minimal penalty function of \( \rho \) defined by

\[
\alpha^\text{min}_\rho(Q) = \sup_{X \in L^\infty} \left( \mathbb{E}^Q[-X] - \rho(X) \right).
\]

In Propositions 3.12 and 3.13 below, we compute the minimal penalty functions of \( \rho_\ell \) and \( \delta_{g,\lambda} \) for \( \lambda \in 1/\text{dom}\ g \).
Proposition 3.12. Let $\lambda \in 1/\text{dom } g$. The function $\delta_{g,\lambda} : L^\infty \to \mathbb{R}$ is a convex weak*-lower semicontinuous risk measure with minimal penalty function $\alpha^{\min}_{\delta_{g,\lambda}} : \mathcal{M}_1(\mathbb{P}) \to \mathbb{R} \cup \{+\infty\}$ given by

$$\alpha^{\min}_{\delta_{g,\lambda}}(Q) = I_{g,\lambda}(Q | \mathbb{P}) + \lambda x^0.$$ 

Recall that Corollary 3.10 summarizes the duality between the $\ell$-shortfall risk measure $\rho_{\ell}$ and $(g, \lambda)$-divergence risk measures $\delta_{g,\lambda}$ with $\lambda \in 1/\text{dom } g$. Proposition 3.13 tells that there is an analogous duality between the corresponding minimal penalty functions.

Proposition 3.13. The minimal penalty function $\alpha^{\min}_{\rho_{\ell}} : \mathcal{M}_1(\mathbb{P}) \to \mathbb{R} \cup \{+\infty\}$ of $\rho_{\ell}$ is given by

$$\alpha^{\min}_{\rho_{\ell}}(Q) = \inf_{\lambda > 0} (\lambda x^0 + I_{g,\lambda}(Q | \mathbb{P})) = \inf_{\lambda \in 1/\text{dom } g} \alpha^{\min}_{\delta_{g,\lambda}}(Q).$$

4 Set-valued shortfall and divergence risk measures

4.1 Regulator versions

In this section, we introduce set-valued shortfall and divergence risk measures, the central objects of this paper. The proofs are given in Section 6.4.

Throughout Section 4 let $m > 0$ be an integer denoting the number of eligible assets (see Section 2). Corresponding to each such asset $i \in \{1, \ldots, m\}$, we assume that there is a (scalar) loss function $\ell_i : \mathbb{R} \to \mathbb{R} \cup \{+\infty\}$ in the sense of Definition 3.1 we denote by $g_i$ its conjugate function which is a divergence function in the sense of Definition 3.5. We define the vector loss function $\ell : \mathbb{R}^m \to \mathbb{R}^m \cup \{+\infty\}$ by

$$\ell(x) = \begin{cases} (\ell_1(x_1), \ldots, \ell_m(x_m))^T & \text{if } x \in \bigcap_{i=1}^m \text{dom } \ell_i, \\ +\infty & \text{else} \end{cases}$$

and similarly the vector divergence function $g : \mathbb{R}^m \to \mathbb{R}^m \cup \{+\infty\}$ by

$$g(x) = \begin{cases} (g_1(x_1), \ldots, g_m(x_m))^T & \text{if } x \in \bigcap_{i=1}^m \text{dom } g_i, \\ +\infty & \text{else} \end{cases}$$

Here, the element $+\infty$ is added to $\mathbb{R}^m$ as a top element with respect to the partial order generated by $\mathbb{R}^m_+$, that is,

$$\forall z \in \mathbb{R}^m \cup \{+\infty\} : z \leq \mathbb{R}^m_+ + \infty.$$

The addition is extended from $\mathbb{R}^m$ to $\mathbb{R}^m \cup \{+\infty\}$ by $z + (+\infty) = (+\infty) + z = +\infty$ for every $z \in \mathbb{R}^m \cup \{+\infty\}$. Since there hardly is any cause for confusion, we use the same symbol $+\infty$ for every $m \in \{1, 2, \ldots\}$.

Remark 4.1. Of course, one could consider more general loss functions $\ell_i$ that depend on the vector $x \in \mathbb{R}^m$, and not only on the component $x_i$, $i \in \{1, \ldots, m\}$, or even vector loss functions $\ell : \mathbb{R}^d \to \mathbb{R}^m$ with $d > m$. But we want to point out that (1) the interconnectedness of the components of a portfolio $x = X(\omega)$ at time $t$ are already modeled by the prevailing exchange rates $C_t(\omega)$ and trading constraints $D_t(\omega)$ and thus will be included in the market-extension of a risk measure, (2) the dimension reduction, which is motivated by allowing only $m$ of the $d$ assets to be used as eligible assets for risk compensation, is modeled by forcing liquidation into $L^\infty_m$ in Definition 2.11 of the market-extension. This includes the case, where a large number of assets $d$ are denoted in a few ($m < d$) currencies, the currencies are used as eligible assets, and the loss...
functions are just defined for each of the $m$ currencies (but not for each asset individually). Finally, (3) from a risk preference point of view, we assume that there is a complete risk preference for each of the $m$ eligible assets which has a von Neumann-Morgenstern representation: This might be disputable, but it is already (much) more general than the assumption that there is a complete risk preference for multivariate positions (as in [7]) which even has a von Neumann representation generated by a real-valued loss function defined on $\mathbb{R}^m$ (or $\mathbb{R}^d$) as in [5].

Using $\text{dom } \ell := \{ x \in \mathbb{R}^m | \ell(x) \in \mathbb{R}^m \} = \times_{i=1}^m \text{dom } \ell_i \subseteq \mathbb{R}^m$ we note

$$\text{int } (\text{dom } \ell) = \times_{i=1}^m \text{int } \ell_i(\mathbb{R}).$$

The expected loss operator is extended by $E[\ell(X)] = +\infty$ whenever $\mathbb{P}\{X \in \text{dom } \ell\} < 1$. Let $x^0 = (x_1^0, \ldots, x_m^0)^T \in \text{int } (\text{dom } \ell)$. Let $C \in \mathcal{G}(\mathbb{R}^m, \mathbb{R}_+^m)$ be such that $0 \in \mathbb{R}^m$ is a boundary point of $C$.

**Definition 4.2.** The function $R_\ell: L_m^\infty \to \mathcal{P}(\mathbb{R}^m, \mathbb{R}_+^m)$ defined by

$$R_\ell(X) = \{ z \in \mathbb{R}^m | E[\ell(-X - z)] \in x^0 - C \}$$

is called the $\ell$-shortfall risk measure (on $L_m^\infty$ with threshold level $x^0$ and threshold set $C$).

The set $C$ determines a rule with respect to which expected loss vectors are compared to the threshold level: The relation $\leq_C$ defined by $x \leq_C y :\iff y - x \in C$ is reflexive (since $0 \in C$), transitive if $C + C \subseteq C$ and antisymmetric if $C \cap (-C) = \{0\}$ (C is “pointed.”). In particular, if $C$ is a pointed convex cone, then $\leq$ is a partial order which is compatible with the linear structure on $\mathbb{R}^m$. The case $C = \mathbb{R}_+^m$ corresponds to the componentwise ordering of the expected loss vectors.

**Proposition 4.3.** The function $R_\ell: L_m^\infty \to \mathcal{P}(\mathbb{R}^m, \mathbb{R}_+^m)$ is a convex weak*-closed risk measure. In particular, it maps into $\mathcal{G}(\mathbb{R}^m, \mathbb{R}_+^m)$.

Proposition 4.3 above implies that $R_\ell$ has the dual representation given by Proposition 2.6. In Proposition 4.12 at the end of this section, we will compute its minimal penalty function $-\alpha_{R_\ell}^{\text{min}}$; this will be the set-valued analogue of Proposition 3.13.

If $C = \mathbb{R}_+^m$, then $R_\ell$ becomes a trivial generalization of the scalar shortfall risk measures $\rho_{\ell_1}, \ldots, \rho_{\ell_m}$ in the sense that

$$R_\ell(X) = (\rho_{\ell_1}(X_1), \ldots, \rho_{\ell_m}(X_m))^T + \mathbb{R}_+^m,$$

for every $X \in L_m^\infty$. In general, such an explicit representation of $R_\ell$ may not exist. However, given $X \in L_m^\infty$, one may write

$$R_\ell(X) = \inf_{(G(\mathbb{R}^m, \mathbb{R}_+^m), \gtrless)} \{ z + \mathbb{R}_+^m | 0 \in E[\ell(-X - z)] - x^0 + C, z \in \mathbb{R}^m \},$$

that is, $R_\ell(X)$ is the (optimal) value of the set minimization problem

$$\text{minimize} \quad \Phi(z) \quad \text{subject to} \quad 0 \in \Psi(z), z \in \mathbb{R}^m,$$

where $\Phi: \mathbb{R}^m \to \mathcal{G}(\mathbb{R}^m, \mathbb{R}_+^m)$ and $\Psi: \mathbb{R}^m \to \mathcal{G}(\mathbb{R}^m, \mathbb{R}_+^m)$ are the (set-valued) objective function and constraint function, respectively, defined by

$$\Phi(z) = z + \mathbb{R}_+^m \quad \text{and} \quad \Psi(z) = E[\ell(-X - z)] - x^0 + C.$$
Here, it is understood that \( \Psi(z) = 0 \) whenever \( \mathbb{E}[\ell(-X - z)] = +\infty \). A Lagrange duality theory for problems of the form (4.2) is developed in the recent work \cite{21}. Using this theory, we will compute the Lagrangian dual problem for \( R_{\ell}(X) \) in Proposition 4.4 below. It turns out that, after a change of variables provided by Proposition 4.7, the dual objective function has a simple form, which will give rise to set-valued versions of divergence risk measures.

**Proposition 4.4.** For every \( X \in L_m^\infty \),

\[
R_{\ell}(X) = \bigcap_{\lambda, v \in \mathbb{R}_m^n \setminus \{0\}} \left\{ \eta \in \mathbb{R}^m \mid v^T \eta \geq \inf_{z \in \mathbb{R}^m} \left( v^T z + \lambda^T \mathbb{E}[\ell(-X - z)] \right) + \inf_{x \in -x_0 + C} \lambda^T x \right\},
\]  

(4.3)

where \( \lambda^T(+\infty) = +\infty \) whenever \( \lambda \in \mathbb{R}_m^n \setminus \{0\} \).

Let us introduce the vector-valued versions of some notions used in Section 3.

**Definition 4.5.** Let \( r \in \mathbb{R}_m^m \setminus \{0\} \) and \( Q \in \mathcal{M}_m(\mathbb{P}) \). For each \( i \in \{1, \ldots, m\} \), let

\[
I_{g_i, r_i}(Q_i \mid \mathbb{P}) := \begin{cases} r_i \mathbb{E}\left[g_i\left(\frac{1}{r_i} \frac{dQ_i}{d\mathbb{P}}\right)\right] & \text{if } r_i > 0 \\ +\infty & \text{else} \end{cases}.
\]

The element \( I_{g, r}(Q \mid \mathbb{P}) \in \mathbb{R}^m \cup \{+\infty\} \) defined by

\[
I_{g, r}(Q \mid \mathbb{P}) := (I_{g_1, r_1}(Q_1 \mid \mathbb{P}), \ldots, I_{g_m, r_m}(Q_m \mid \mathbb{P}))^T
\]

whenever the right hand side is in \( \mathbb{R}^m \) and \( +\infty \) otherwise is called the vector \((g, r)\)-divergence of \( Q \) with respect to \( \mathbb{P} \).

If \( r_i > 0 \), then \( I_{g_i, r_i}(Q_i \mid \mathbb{P}) \) is the (scalar) \((g_i, r_i)\)-divergence of \( Q_i \) with respect to \( \mathbb{P} \); see Definition 3.8. Recall that \( \mathbb{E}\left[g_i\left(\frac{1}{r_i} \frac{dQ_i}{d\mathbb{P}}\right)\right] = +\infty \) whenever \( \mathbb{P}\left\{\frac{1}{r_i} \frac{dQ_i}{d\mathbb{P}} \in \text{dom } g_i\right\} < 1 \).

**Definition 4.6.** For every \( r \in \mathbb{R}_m^m \setminus \{0\} \), define a function \( \delta_{g, r} : L_m^\infty \to \mathbb{R}^m \cup \{-\infty\} \) by

\[
\delta_{g, r}(X) = (\delta_{g_1, r_1}(X_1), \ldots, \delta_{g_m, r_m}(X_m))^T
\]

whenever the right hand side is in \( \mathbb{R}^m \) and \( \delta_{g, r}(X) = -\infty \) otherwise.

Of course, if \( r_i > 0 \), then \( \delta_{g_i, r_i} \) given by

\[
\delta_{g_i, r_i}(X_i) := \sup_{Q_i \in \mathcal{M}_i(\mathbb{P})} \left( \mathbb{E}[Q_i(-X_i)] - I_{g_i, r_i}(Q_i \mid \mathbb{P}) \right) - r_i x_i^0 = \inf_{z_i \in \mathbb{R}} \left(x_i + r_i \mathbb{E}[\ell_i(-X_i - z_i)]\right) - r_i x_i^0
\]

(4.4)

is the (scalar) \((g_i, r_i)\)-divergence risk measure according to Definition 3.11, and we have \( \delta_{g_i, 0} \equiv -\infty \). The next proposition provides a more useful version of the representation in Proposition 4.4 in terms of the vector-valued functions \((\delta_{g, r})_{r \in \mathbb{R}_m^m} \) as a result of a “change of variables”.

**Proposition 4.7.** For every \( X \in L_m^\infty \),

\[
R_{\ell}(X) = \bigcap_{r, w \in \mathbb{R}_m^m \setminus \{0\}} \left\{ z \in \mathbb{R}^m \mid w^T z \geq w^T \delta_{g, r}(X) + \inf_{x \in C} w^T \text{diag}(r) x \right\},
\]

where \( w^T(-\infty) = -\infty \) whenever \( w \in \mathbb{R}_m^m \setminus \{0\} \).
Given $r \in \mathbb{R}_+^m \setminus \{0\}$, define a set-valued function $D_{g,r} : L^\infty_m \to \mathcal{G}(\mathbb{R}^m, \mathbb{R}_+^m)$ by

$$D_{g,r}(X) = \bigcap_{w \in \mathbb{R}_+^m \setminus \{0\}} \left\{ z \in \mathbb{R}^m \mid w^T z \geq w^T \delta_{g,r}(X) + \inf_{x \in C} w^T \text{diag}(r)x \right\}. \quad (4.5)$$

**Corollary 4.8.** For every $X \in L^\infty_m$,

$$R_\ell(X) = \bigcap_{r \in 1/\text{dom } g} D_{g,r}(X), \quad (4.6)$$

where

$$1/\text{dom } g := \prod_{i=1}^m 1/\text{dom } g_i = \prod_{i=1}^m \left\{ \frac{1}{z_i} \mid 0 \neq z_i \in \text{dom } g_i \right\}. \quad (4.7)$$

**Proof.** By Theorem 3.9, we have $\delta_{g,r}(X) \in \mathbb{R}^m$ for all $X \in L^\infty_m$ if and only if $r \in 1/\text{dom } g$. So $r \notin 1/\text{dom } g$ implies $D_{g,r} \equiv \mathbb{R}^m$. The result directly follows from Proposition 4.7. \qed

**Definition 4.9.** For $r \in 1/\text{dom } g$, $D_{g,r}$ defined by (4.5) is called the $(g,r)$-divergence risk measure on $L^\infty_m$ with threshold level $x^0$ and threshold set $C$. 

Corollary 4.8 states that the $\ell$-shortfall risk measure is the supremum of all $(g,r)$-divergence risk measures with $r \in 1/\text{dom } g$; recall (2.2) for the infimum and supremum formulae for the complete lattice $(\mathcal{P}(\mathbb{R}^m, \mathbb{R}_+^m), \supseteq)$. In general, there is no single $r \in 1/\text{dom } g$ which yields this supremum, i.e. the supremum is not attained in a single argument. Instead, one could look for a set $\Gamma \subseteq 1/\text{dom } g$ such that (4.6) holds with $1/\text{dom } g$ replaced by $\Gamma$ and each $D_{g,r}(X)$ for $r \in \Gamma$ is a maximal element of the set $\{D_{g,r}(X) \mid r \in 1/\text{dom } g\}$ with respect to $\supseteq$. This corresponds to the solution concept for set optimization problems introduced in [24] (see also [21, Definition 3.3]) and will be discussed for the entropic risk measure in Section 5.1.

Since Proposition 4.7 provides a divergence risk measure $D_{g,r}$ in terms of the vector $\delta_{g,r}$ of scalar divergence risk measures, in Proposition 4.10 below, we are able to give a formula that relates the corresponding set-valued and scalar minimal penalty functions.

**Proposition 4.10.** For every $r \in 1/\text{dom } g$ and $X \in L^\infty_m$,

$$D_{g,r}(X) = \delta_{g,r}(X) + \text{diag}(r)C = \inf_{(\mathcal{G}(\mathbb{R}^m, \mathbb{R}_+^m), \supseteq)} \left\{ -z + \text{diag}(r)(\mathbb{E}[\ell(-X + z)] - x^0 + C) \mid z \in \mathbb{R}^m \right\},$$

where $\text{diag}(r)C := \{\text{diag}(r)x \mid x \in C\}$. Moreover, $D_{g,r}$ is a convex weak*-closed risk measure with minimal penalty function $-\alpha_{D_{g,r}}^\text{min}$ given by

$$-\alpha_{D_{g,r}}^\text{min}(Q, w) = \left\{ z \in \mathbb{R}^m \mid w^T z \geq -w^T I_{g,r}(Q \mid \mathbb{P}) + \inf_{x \in -x^0 + C} w^T \text{diag}(r)x \right\}. \quad (4.8)$$

for each $(Q, w) \in \mathcal{M}_m(\mathbb{P}) \times (\mathbb{R}_+^m \setminus \{0\})$.

**Remark 4.11.** Let us give a financial interpretation for $D_{g,r}(X)$ using the representation given in Proposition 4.10 for fixed $r \in 1/\text{dom } g$ and $X \in L^\infty_m$. Suppose that there is an investor with the random portfolio $X$ who wants to choose a deterministic portfolio $z \in \mathbb{R}^m$ to be received at initial time. Hence she will hold $X - z$ at terminal time which corresponds to the expected loss vector $\mathbb{E}[\ell(-X + z)]$. The deterministic portfolios are compared with respect to the componentwise ordering cone $\mathbb{R}_+^m$ and the expected loss vectors with respect to the set $D = -x^0 + C$ as discussed.
Proposition 4.12. If $\alpha$ satisfies the conditions of a convex weak $\mathcal{G}(\mathbb{R}_{+}^{m}, \mathbb{R}_{+}^{2m})$.

Consider $r_i$ as the “relative weight” of $\mathbb{E}[\ell(-X_i + z_i)]$ with respect to $-z_i$ for each $i \in \{1, \ldots, m\}$, we construct the “partially scalarized” problem

$$\minimize \quad -z + r_i + \text{diag}(r) (\mathbb{E}[\ell(-X + z)] - x^0 + C) \quad \text{subject to} \quad z \in \mathbb{R}^m.$$  

Proposition 4.10 shows that the value of this (still set-valued) problem is $D_{g,r}(X)$.

We end this section by formulating the relationship between the minimal penalty functions of shortfall and divergence risk measures.

Proposition 4.12. If $Q \in \mathcal{M}_m(\mathbb{P})$ and $w \in \mathbb{R}_{+}^{m}$, then

$$-\alpha_{R}^{\min}(Q, w) = \left\{ z \in \mathbb{R}^m \mid w^T z \geq \sup_{r \in \mathbb{R}_{+}^{m} \setminus \{0\}} \left( -w^T I_{g_r}(Q \mid \mathbb{P}) + \inf_{x \in -x^0 + C} w^T \text{diag}(r)x \right) \right\}$$

$$= \bigcap_{r \in 1/\text{dom } g} -\alpha_{D_{g,r}}^{\min}(Q, w). \quad (4.9)$$

4.2 Finiteness of the market-extensions

From Remark 2.13 recall that the closed market-extension $\text{cl } R_{mar}$ of a convex risk measure $R$ satisfies the conditions of a convex weak$^*$-closed risk measure except possibly that $\text{cl } R_{mar}(Y) \neq \mathbb{R}^m$ for some $Y \in L_{\mathbb{P}}^\infty$. In this section, we will present sufficient conditions that guarantee this finite-valuedness property for the shortfall and divergence risk measures of Section 4.1. Once this property is established, the dual representation for these closed market-extensions is provided by Proposition 2.13. For simplicity, we assume $m = d$, in which case it is enough to check that the closed market-extension is finite at 0, see Definition 2.1 et seq. We also assume that $\mathcal{D}_t := \mathbb{R}^d$ for each $t \in \{0, \ldots, T\}$ (no explicit trading constraints), and that the market model is conical in the sense of Example 2.7.

Assumption 4.13. Suppose that the solvency cones of the market model share a common supporting halfspace in the sense that there exists $\bar{w} \in \mathbb{R}_+^{d} \setminus \{0\}$ such that for $\mathbb{P}$-almost every $\omega \in \Omega$ and every $t \in \{0, \ldots, T\}$,

$$\bar{w} \in (\mathcal{C}_t(\omega))^+, \quad \inf_{y \in C_t(\omega)} \bar{w}^T y > -\infty.$$  

Remark 4.14. Assume 4.13 states the existence of a halfspace $H = \{z \in \mathbb{R}^d \mid \bar{w}^T z \geq 0\}$ for some $\bar{w} \in \mathbb{R}_+^{d}$ which satisfies $H \supseteq C_t(\omega)$ for $\mathbb{P}$-almost every $\omega \in \Omega$ and all $t \in \{0, \ldots, T\}$. In particular, when the solvency cones are constructed from bid-ask prices (see 23), this is equivalent to the ask prices having a uniform (in time and outcome) lower bound, or equivalently the bid prices having a uniform (in time and outcome) upper bound. That is, $\bar{w}_j \leq \pi_{ij}(\omega, t)\bar{w}_i$ for all
Proposition 4.15. Suppose that Assumption 4.13 holds. Let $r \in 1/\text{dom } g$. If
\[
\inf_{x \in C} w^T \text{diag}(r)x > -\infty,
\] (4.10)
then the closed market-extension $\text{cl } D_{g,r}^{\text{mar}}$ of $D_{g,r}$ has finite values and thus is a convex weak*-closed risk measure.

The proof of Proposition 4.15 is given in Section 6.4.

Proposition 4.16. Suppose that Assumption 4.13 holds. If there exists $r \in 1/\text{dom } g$ such that (4.10) holds, then the closed market-extension $\text{cl } R_{g,r}^{\text{mar}}$ of $R_{g,r}$ has finite values and thus is a convex weak*-closed risk measure.

Proof. From (4.6), for every $Y \in L^\infty_d$, we know that $R_{g,r}(Y) \subseteq D_{g,r}(Y)$; hence, by Remark 2.16, we have $(\text{cl } R_{g,r}^{\text{mar}})(Y) \subseteq (\text{cl } D_{g,r}^{\text{mar}})(Y)$. The result follows from Proposition 4.15.

5 Examples

5.1 Set-valued entropic risk measures

In this section, we assume that the vector loss function $\ell$ of Section 4 is the vector exponential loss function with constant risk aversion vector $\beta \in \mathbb{R}^m_+$, that is, for each $i \in \{1, \ldots, m\}$, we assume
\[
\forall x \in \mathbb{R}: \ell_i(x) = e^{\beta_i x} - 1, \beta_i \neq 0,
\]
which satisfies the conditions in Definition 3.1. The corresponding vector divergence function $g$ is given by
\[
\forall y \in \mathbb{R}: g_i(y) = \frac{y}{\beta_i} \log y - \frac{y}{\beta_i} + 1, \beta_i \neq 0,
\]
for each $i \in \{1, \ldots, m\}$. Here and elsewhere, we make the convention $\log y = -\infty$ for every $y \leq 0$.

Let $x^0 \in \text{int } \ell(\text{dom } \ell) = \mathbb{R}^m_+ \cap (-1, +\infty)$ and $C \in \mathcal{G}(\mathbb{R}^m, \mathbb{R}^m_+)$ with $0 \in \text{bd } C$. We call the corresponding $\ell$-shortfall risk measure $R_{\text{ent}} := R_{\ell}$ the entropic risk measure (with threshold level $x^0$ and threshold set $C$). The next proposition shows that $R_{\text{ent}}$ is in the simple form of “a vector-valued function plus a fixed set”, which is, in general, not the case for an arbitrary loss function.

Proposition 5.1. For every $X \in L^\infty_m$,
\[
R_{\text{ent}}(X) = \rho_{\text{ent}}(X) + \tilde{C},
\]
where
\[
\rho_{\text{ent}}(X) := \left( \frac{1}{\beta_1} \log \mathbb{E} \left[ e^{-\beta_1 X_1} \right], \ldots, \frac{1}{\beta_m} \log \mathbb{E} \left[ e^{-\beta_m X_m} \right] \right)^T,
\]
\[
\tilde{C} := \left\{ z \in \mathbb{R}^m \mid z_i = -\frac{1}{\beta_i} \log (1 + \beta_i (x^0_i - c_i)), i = 1, \ldots, m \right\}, \quad c \in C \cap \left( \bigcap_{j=1}^m (-\infty, x^0_j + \frac{1}{\beta_j}) \right).
\]
Proof. Using the definitions, we have

\[ R^{\text{ent}}(X) = \{ z \in \mathbb{R}^m \mid \mathbb{E}[\ell(-X - z)] \in x^0 - C \} \]

\[ = \left\{ z \in \mathbb{R}^m \mid \exists c \in C \forall i \in \{1, \ldots, m\} : \frac{\mathbb{E}\left[e^{\beta_i(-X_i - z_i)}\right] - 1}{\beta_i} = x_i^0 - c_i \right\} \]

\[ = \left\{ z \in \mathbb{R}^m \mid \exists c \in C \forall i \in \{1, \ldots, m\} : e^{-\beta_i z_i} \mathbb{E}\left[e^{-\beta_i X_i}\right] = 1 + \beta_i(x_i^0 - c_i) \right\} \]

\[ = \left\{ z \in \mathbb{R}^m \mid \exists c \in C \forall i \in \{1, \ldots, m\} : z_i = \frac{1}{\beta_i} \log \frac{\mathbb{E}\left[e^{-\beta_i X_i}\right]}{1 + \beta_i(x_i^0 - c_i)}, \quad 1 + \beta_i(x_i^0 - c_i) \in \mathbb{R}_{++} \right\} \]

\[ = \rho^{\text{ent}}(X) + \tilde{C}. \]

\[ \square \]

Note that the set \(1/{\text{dom}} g\) defined in \((4.7)\) becomes \(\mathbb{R}^m_{++}\). For \(r \in \mathbb{R}^m_{++}\), let \(D^{\text{ent}}_r := D_{g,r}\) be the \((g, r)\)-divergence risk measure defined by Definition \((4.9)\).

**Proposition 5.2.** For every \(r \in \mathbb{R}^m_{++}\) and \(X \in \mathbb{L}_m^\infty\),

\[ D^{\text{ent}}_r(X) = \rho^{\text{ent}}(X) + \bigcap_{w \in \mathbb{R}^d_+ \setminus \{0\}} \left\{ z \in \mathbb{R}^m \mid w^T z \geq \sum_{i=1}^m w_i \left(1 - r_i + \log r_i\right) + \inf_{x \in \mathbb{R}^m \cap C} w^T \text{diag}(r)x \right\}, \]

where \(\rho^{\text{ent}}(X)\) is defined by \((5.1)\).

Proof. Recalling \((4.5)\), it holds

\[ D^{\text{ent}}_r(X) = \bigcap_{w \in \mathbb{R}^m_{++} \setminus \{0\}} \left\{ z \in \mathbb{R}^m \mid w^T z \geq w^T \delta_{g,r}(X) + \inf_{x \in C} w^T \text{diag}(r)x \right\}, \]

where, for each \(i \in \{1, \ldots, m\},\)

\[ \delta_{g,r_i}(X_i) = \inf_{z_i \in \mathbb{R}} \left( z_i + r_i \mathbb{E}[\ell_i(-X_i - z_i)] \right) - r_i x_i^0 \]

\[ = \frac{1}{\beta_i} \log \mathbb{E}\left[e^{-\beta_i X_i}\right] + \frac{1}{\beta_i} \left(1 - r_i + \log r_i\right) - r_i x_i^0 \in \mathbb{R}. \]

The result follows. \(\square\)

Recall from \((4.6)\) that \(R^{\text{ent}}(\cdot)\) is the supremum of all \(D^{\text{ent}}_r(\cdot)\) with \(r \in \mathbb{R}_m^{\text{ent}}_{++}\), that is,

\[ \forall X \in \mathbb{L}_m^\infty : R^{\text{ent}}(X) = \sup_{r \in \mathbb{R}^d_{++}} D^{\text{ent}}_r(X) = \bigcap_{r \in \mathbb{R}_m^{\text{ent}}_{++}} D^{\text{ent}}_r(X). \quad (5.2) \]

If \(m = 1\), then the only choice for \(C\) is \(\mathbb{R}_+\). In this case,

\[ \forall X \in \mathbb{L}_\infty^\infty : R^{\text{ent}}(X) = D^{\text{ent}}_1(X) = \rho^{\text{ent}}(X) - x^0 + \mathbb{R}_+. \]

In other words, the supremum in \((5.2)\) is attained at \(r = 1\). In general, when \(m \geq 2\), we may not be able to find some \(r \in \mathbb{R}_m^{\text{ent}}_{++}\) for which \(R^{\text{ent}}(X) = D^{\text{ent}}_r(X)\). Instead, we will compute a solution to this set maximization problem in the sense of [21, Definition 3.3], that is, we will find a set \(\Gamma \subseteq \mathbb{R}^m_{++}\) such that
Lemma 5.3. Let $r$ be the only vector $r$ can rewrite $D^\text{ent}(X)$ is a maximal element of the collection $\{D^\text{ent}(X) \mid r \in \mathbb{R}^m_+\}$ in the following sense:

$$\forall r \in \mathbb{R}^m_+ : \quad D^\text{ent}(X) \subseteq D^\text{ent}(X) \Rightarrow r = \bar{r}.$$ 

Moreover, the set $\Gamma$ will be independent of the choice of $X$. Indeed, if $r \in \mathbb{R}^m_+$ and $X \in L^\infty_m$, we can rewrite $D^\text{ent}(X)$ as

$$D^\text{ent}(X) = \rho^\text{ent}(X) + \bigcap_{w \in \mathbb{R}^m_+ \setminus \{0\}} \{z \in \mathbb{R}^d \mid w^T z \geq -(f_w(r) + h_w(r))\},$$

where, for $w \in \mathbb{R}^m_+ \setminus \{0\}$, $r \in \mathbb{R}^m_+$,

$$f_w(r) := \sum_{i=1}^m \left( \frac{w_i}{\beta_i} (-1 + r_i - \log r_i) + w_i x_i^0 \right),$$

$$h_w(r) := - \inf_{x \in C} w^T \text{diag}(r)x = \sup_{x \in -C} w^T \text{diag}(r)x.$$ 

For convenience, we sometimes use the notation $[x_i]_{i=1}^m$ for $x = (x_1, \ldots, x_m)^T \in \mathbb{R}^m$.

**Lemma 5.3.** Let $w \in \mathbb{R}^m_+ \setminus \{0\}$. The function $f_w + h_w$ on $\mathbb{R}^m_+$ is either identically $+\infty$ or else it attains its infimum at a unique point $r^w \in \mathbb{R}^m_+$ which is determined by the following property: $r^w$ is the only vector $r \in \mathbb{R}^m_+$ for which $-x^0 + C$ is supported at the point

$$\left[ \frac{1}{\beta_i} \left( 1 - \frac{1}{r_i} \right) \right]_{i=1}^m$$

by the hyperplane with normal direction $\text{diag}(r)w$.

**Proof.** First, we extend $f_w$ and $h_w$ from $\mathbb{R}^m_+$ to $\mathbb{R}^m$ with their original definitions so that we have

$$\inf_{r \in \mathbb{R}^m_+} (f_w(r) + h_w(r)) = \inf_{r \in \mathbb{R}^m} (f_w(r) + h_w(r)).$$

Note that $f_w$ is a proper, strictly convex, continuous function and has a unique minimum point. Hence, by [31, Theorem 27.1(d)], $f_w$ has no directions of recession, that is, the recession function $f_w0^+$ of $f_w$ always takes strictly positive values; see [31, p. 66 and p. 69] for definitions. Besides, $h_w$ is a proper, convex, lower semicontinuous function. If $h_w \equiv +\infty$, then the infimum of $f_w + h_w$ is $+\infty$. Suppose that $h_w$ is a proper function. Since $0 \in \text{bd} - C$, $h_w$ always takes nonnegative values. Hence, the infimum of $h_w$ is finite. By [31, Theorem 27.1(a), (i)], this implies that the recession function $h_w0^+$ of $h_w$ always takes nonnegative values. Therefore, $f_w + h_w$ has no directions of recession since $(f_w + h_w)0^+ = f_w0^+ + h_w0^+$ by [31, Theorem 9.3]. Hence, by [31, Theorem 27.1(b), (d)] and the strict convexity of $f_w + h_w$, this function has a unique minimum point $r^w \in \mathbb{R}^m_+$ which is determined by the first order condition

$$0 \in \partial (f_w + h_w)(r^w)$$

$$= \left[ \frac{w_i}{\beta_i} - \frac{w_i}{\beta_i r_i^w} + w_i x_i^0 \right]_{i=1}^m + \left\{ \text{diag}(w) \bar{x} \mid \bar{x} \in -C, \sup_{x \in -C} w^T \text{diag}(r^w)x = w^T \text{diag}(r^w)\bar{x} \right\},$$

that is,

$$\left[ \frac{1}{\beta_i} \left( 1 - \frac{1}{r_i^w} \right) \right]_{i=1}^m \in -x^0 + C, \quad \inf_{x \in -x^0 + C} w^T \text{diag}(r^w)x = \sum_{i=1}^m \frac{w_i r_i^w}{\beta_i} \left( 1 - \frac{1}{r_i^w} \right),$$

which is the claimed property of $r^w$. 

\[ \blacksquare \]
Proposition 5.4. Using the notation in Lemma 5.3, the set
\[ \Gamma := \{ r^w \mid w \in \mathbb{R}_+^m \setminus \{0\}, \ f_w + h_w \text{ is proper} \} \]
is a solution to the maximization problem in (5.2) for every \( X \in L^\infty_m \).

Proof. By Lemma 5.3, it is clear that, for each \( w \in \mathbb{R}_+^m \setminus \{0\} \), we have
\[ \inf_{r \in \mathbb{R}_+^m} (f_w(r) + h_w(r)) = \inf_{r \in \Gamma} (f_w(r) + h_w(r)) = f_w(r^w) + h_w(r^w). \]

Hence, we have
\[
R_{\text{ent}}^w(X) = \bigcap_{r \in \mathbb{R}_+^m} D_{r}^{\text{ent}}(X) \\
= \rho_{\text{ent}}^w(X) + \bigcup_{w \in \mathbb{R}_+^m \setminus \{0\}, \ r \in \mathbb{R}_+^m} \left\{ z \in \mathbb{R}^m \mid w^T z \geq -(f_w(r) + h_w(r)) \right\} \\
= \rho_{\text{ent}}^w(X) + \bigcup_{w \in \mathbb{R}_+^m \setminus \{0\}} \left\{ z \in \mathbb{R}^m \mid w^T z \geq -\inf_{r \in \mathbb{R}_+^d} (f_w(r) + h_w(r)) \right\} \\
= \rho_{\text{ent}}^w(X) + \bigcap_{w \in \mathbb{R}_+^m \setminus \{0\}} \left\{ z \in \mathbb{R}^m \mid w^T z \geq -\inf_{r \in \Gamma} (f_w(r) + h_w(r)) \right\} = \bigcap_{r \in \Gamma} D_{r}^{\text{ent}}(X). \\

Let \( w \in \mathbb{R}_+^m \setminus \{0\} \) such that \( f_w + h_w \) is proper and let \( r \in \mathbb{R}_+^m \). Suppose that \( D_{r}^{\text{ent}}(X) \subseteq D_{r}^{\text{ent}}(X) \). Then,
\[ -(f_w(r) + h_w(r)) = \inf_{z \in D_{r}^{\text{ent}}(X)} w^T z \geq \inf_{z \in D_{r}^{\text{ent}}(X)} w^T z = -(f_w(r^w) + h_w(r^w)), \]
that is, \( f_w(r) + h_w(r) \leq f_w(r^w) + h_w(r^w) \). By Lemma 5.3 this implies that \( r = r^w \). \qed

Finally, we compute the minimal penalty function of \( R_{\text{ent}}^w \) in terms of the vector relative entropies
\[ H(Q \mid P) := \left[ \mathbb{E}^Q_i \left( \frac{dQ_i}{dP} \right) \right]_{i=1}^m \]
of vector probability measures \( Q \in \mathcal{M}_m(\mathbb{P}) \).

Proposition 5.5. For every \( (Q, w) \in \mathcal{M}_m(\mathbb{P}) \times (\mathbb{R}_+^m \setminus \{0\}) \), we have \( -\alpha_{\text{ent}}^w(Q, w) = \mathbb{R}^m \) if \( h_w \equiv +\infty \) and
\[ -\alpha_{\text{ent}}^w(Q, w) = -\text{diag}(\beta)^{-1}H(Q \mid P) - \text{diag}(\beta)^{-1}\log((1 + \text{diag}(\beta)(x^0 - C)) \cap R_+^m) + G^+(w) \]
if \( h_w \) is a proper function. Here, \( \log[A] := \{(\log x_1, \ldots, \log x_m)^T \mid x \in A\} \) for \( A \subseteq \mathbb{R}_+^m \).
Proof. Proposition 4.11 and Lemma 5.3 give

\[ -\alpha_{R_{\ell}}^{\text{min}}(Q, w) = \bigcap_{r \in 1/\text{dom} g} -\alpha_{E_{g,r}}^{\text{min}}(Q, w) \]

\[ = \bigcap_{r \in 1/\text{dom} g} \left\{ z \in \mathbb{R}^m \mid w^T z \geq -w^T I_{g,r}(Q, \mathbb{P}) + \inf_{x \in -x^0 + C} w^T \operatorname{diag}(r) x \right\} \]

\[ = -\left[ \frac{1}{\beta_i} \log \frac{dQ_i}{d\mathbb{P}} \right]_{i=1}^m + \bigcap_{r \in \mathbb{R}^m_{++}} \left\{ z \in \mathbb{R}^m \mid w^T z \geq \sum_{i=1}^m \frac{w_i}{\beta_i} (1 - r_i + \log r_i) \right\} \]

\[ = -\left[ \frac{1}{\beta_i} \log \frac{dQ_i}{d\mathbb{P}} \right]_{i=1}^m + \left\{ z \in \mathbb{R}^m \mid w^T z \geq -(f_w(r) + h_w(r)) \right\} \]

\[ = -\left[ \frac{1}{\beta_i} \log \frac{dQ_i}{d\mathbb{P}} \right]_{i=1}^m + \left\{ z \in \mathbb{R}^m \mid w^T z \geq -(f_w(r^w) + h_w(r^w)) \right\} \]

assuming that \( h_w \) is not identically \(+\infty\) (else \(-\alpha_{R_{\ell}}^{\text{min}}(Q, w) = \mathbb{R}^m \)).

Thus, the minimal penalty function for the entropic risk measure is of the form “- vector relative entropy + a nonhomogeneous halfspace” (except for the trivial case).

5.2 Set-valued average values at risks

In this section, we assume that the vector loss function \( \ell \) of Section 4 is the \textit{(vector) scaled positive part function} with scaling vector \( \alpha \in (0, 1]^m \), that is, for each \( i \in \{1, \ldots, m\} \), we assume

\[ \forall x \in \mathbb{R} : \ell_i(x) = \frac{x^+}{\alpha_i}, \]

which satisfies the conditions in Definition 3.1. The corresponding vector divergence function \( g \) is given by

\[ \forall y \in \mathbb{R} : g_i(y) = \begin{cases} 0 & \text{if } y \in \left[ 0, \frac{1}{\alpha_i} \right], \\ +\infty & \text{else} \end{cases} , \]

for each \( i \in \{1, \ldots, m\} \).

Let \( x^0 \in \text{int} \ell(\text{dom} \ell) = \mathbb{R}^m_{++} \) and \( C \in \mathcal{G}(\mathbb{R}^m, \mathbb{R}^m_{++}) \) with \( 0 \in \text{bd} C \). The corresponding \( \ell \)-shortfall risk measure is given by

\[ R_{\ell}(X) = \left\{ z \in \mathbb{R}^m \mid \mathbb{E} \left[ (z - X)^+ \right] \in \operatorname{diag}(\alpha)(x^0 - C) \right\} , \]

where the positive part function is applied component-wise.

Note that the set \( 1/\text{dom} g \) defined in (4.7) becomes \( \bigcup_{i=1}^m [\alpha_i, +\infty) \). For \( r \in \mathbb{R}^m_{++} \), the \((g, r)\)-divergence risk measure is given by

\[ \forall X \in L^\infty_m : D_{g,r}(X) = \delta_{g,r}(X) + \operatorname{diag}(r)C , \]
where, for each \( i \in \{1, \ldots, m\} \),

\[
\forall X_i \in L_1^\infty : \delta_{g_i,r_i}(X_i) = \inf_{z_i \in \mathbb{R}} \left( z_i + \frac{r_i}{\alpha_i} \mathbb{E} \left[ (z_i - X_i)^+ \right] \right) - r_i x^0_i.
\]

When \( r = (1, \ldots, 1)^T \) and \( C = \mathbb{R}_+^m \), we obtain the set-valued average value at risk in the sense of \cite{23} Definition 2.1 for \( M = \mathbb{R}^m \), which is given by

\[
\text{AV@R}_\alpha(X) := D_{g,1}(X) + x^0 = \left[ \inf_{z_i \in \mathbb{R}} \left( z_i + \frac{1}{\alpha_i} \mathbb{E} \left[ (z_i - X_i)^+ \right] \right) \right]_i^{m} + \mathbb{R}_+^m, \quad X \in L_1^\infty.
\]

Hence, our framework offers the following generalization of set-valued average value at risk to a convex risk measure:

\[
\text{AV@R}_{\alpha,r}(X) := D_{g,r}(X) + \text{diag}(r)x^0 = \left[ \inf_{z_i \in \mathbb{R}} \left( z_i + \frac{r_i}{\alpha_i} \mathbb{E} \left[ (z_i - X_i)^+ \right] \right) \right]_i^{m} + \text{diag}(r)C.
\]

As in the scalar case, this definition even works for \( X \in L_1^m \).

6 Proofs and technical remarks

6.1 Proofs of the results in Section 2.2

In this section, we establish a link between the notions of market-extension and set-valued infimal convolution. Based on this relationship, we give proofs of Proposition 2.17, Corollary 2.18 and Corollary 2.19. We begin by introducing two key concepts from (complete lattice-based) set-valued convex analysis; the reader is referred to \cite{18} for details.

**Definition 6.1.** ([18, Example 1]) Let \( Y \subseteq L_1^\infty \). The function \( \mathcal{T}_{\mathcal{Y}}^m : L_1^\infty \rightarrow \mathcal{G}(\mathbb{R}^m, \mathbb{R}_+^m) \) defined by

\[
\mathcal{T}_{\mathcal{Y}}^m(Y) = \begin{cases} \mathbb{R}_+^m & \text{if } Y \in \mathcal{Y} \\ \emptyset & \text{else} \end{cases}
\]

is called the indicator function of the set \( \mathcal{Y} \).

**Definition 6.2.** ([18, Section 4.4(C)]) Let \( N \) be a strictly positive integer. For each \( n \in \{1, \ldots, N\} \), let \( F^n : L_1^\infty \rightarrow \mathcal{P}(\mathbb{R}^m, \mathbb{R}_+^m) \) be a function. The function \( \Box_{n=1}^N F^n : L_1^\infty \rightarrow \mathcal{G}(\mathbb{R}^m, \mathbb{R}_+^m) \) defined by

\[
(\Box_{n=1}^N F^n)(Y) = \text{cl co} \bigcup_{Y^1, \ldots, Y^N \in L_1^\infty} \left\{ \sum_{n=1}^N F^n(Y^n) \mid Y^1 + \ldots + Y^N = Y \right\}
\]

is called the infimal convolution of \( F^1, \ldots, F^N \).

Recall the linear operator \( B : \mathbb{R}^m \rightarrow \mathbb{R}^d \) defined by (2.10):

\[
Bx = (x_1, \ldots, x_m, 0, \ldots, 0)^T \quad \text{for } x \in \mathbb{R}^m.
\]

Its adjoint \( B^* : \mathbb{R}^d \rightarrow \mathbb{R}^m \) is defined by (2.11):

\[
B^*y = (y_1, \ldots, y_m)^T \quad \text{for } y \in \mathbb{R}^d.
\]

The next proposition basically tells us that the market-extension of a regulator risk measure is the infimal convolution of the regulator risk measure and the indicator function of the negative of the set \( \mathcal{K} \) of freely available portfolios defined by (2.9).
Proposition 6.3. Let \( R : L_d^\infty \to \mathcal{P}(\mathbb{R}^m, \mathbb{R}_+^m) \) be a convex regulator risk measure and let \( \tilde{R} : L_d^\infty \to \mathcal{G}(\mathbb{R}^m, \mathbb{R}_+^m) \) be defined by

\[
\tilde{R}(Y) = \begin{cases} R(B^*Y) & \text{if } Y \in B(L_d^\infty) \\ \emptyset & \text{else} \end{cases}.
\]

Then, for each \( Y \in L_d^\infty \),

\[
\text{cl}(R^\text{mar}(Y)) = (\tilde{R} \square T^m_{\mathcal{K}})(Y) = (\tilde{R} \square T^m_{L_d^\infty}(\mathcal{F}_0, \mathcal{C}_0 \cap \mathcal{D}_0) \square \ldots \square T^m_{L_d^\infty}(\mathcal{F}_T, \mathcal{C}_T \cap \mathcal{D}_T))(Y).
\]

(6.1)

Remark 6.4. It should be noted that, in (6.1) above, we are dealing with the the closure of the set \( R^\text{mar}(Y) \) for a given \( Y \in L_d^\infty \), which is not equal to the value \( \text{cl}(R^\text{mar}(Y)) \) of the closed market-extension \( \text{cl}R^\text{mar} \) at \( Y \), in general.

Proof of Proposition 6.3. For each \( Y \in L_d^\infty \), we have

\[
R^\text{mar}(Y) = \bigcup_{\{X \in L_d^\infty \mid BX \in Y + \mathcal{K}\}} R(X) = \bigcup_{U \in Y + \mathcal{K}} \tilde{R}(U)
\]

\[
= \bigcup_{U, U' \in L_d^\infty} \{\tilde{R}(U) + T^m_{\mathcal{K}}(U') \mid U + U' = Y\}
\]

\[
= \bigcup_{U, U^0, \ldots, U^T \in L_d^\infty} \left\{ \tilde{R}(U) + \sum_{t=0}^T T^m_{L_d^\infty}(\mathcal{F}_t, \mathcal{C}_t \cap \mathcal{D}_t)(U^t) \mid U + U^0 + \ldots + U^T = Y \right\}.
\]

Since each of the functions in the infimal convolution is convex, we can omit the convex hull operator in Definition 6.2 and hence, the result follows.

By Proposition 6.3, the market-extension of a regulator risk measure can be formulated as an infimal convolution up to taking closures. As in the scalar theory, we are able to write the conjugate of an infimal convolution of finitely many convex functions as the sum of the conjugates of these convex functions, which is provided by [18, Lemma 2]. The application of this result is the main step of the proof of Proposition 2.17 below. For completeness, we begin with the definition of conjugate for set-valued functions.

Definition 6.5. ([18, Definition 5]) Let \( F : L_d^\infty \to \mathcal{P}(\mathbb{R}^m, \mathbb{R}_+^m) \) be a function. The function

\[-F^* : L_d^1 \times (\mathbb{R}_+^m \setminus \{0\}) \to \mathbb{R} \]

defined by

\[-F^*(V, v) = \text{cl} \bigcup_{Y \in L_d^\infty} \left( F(Y) + \left\{ z \in \mathbb{R}^m \mid v^T z \geq \mathbb{E} \left[ -V^T Y \right] \right\} \right)\]

is called the (Fenchel) conjugate of \( F \).

Proof of Proposition 2.17. Since \( \text{cl}R^\text{mar} \) has closed values, Proposition 6.3 implies

\[R^\text{mar}(Y) \subseteq (\tilde{R} \square T^m_{L_d^\infty}(\mathcal{F}_0, \mathcal{C}_0 \cap \mathcal{D}_0) \square \ldots \square T^m_{L_d^\infty}(\mathcal{F}_T, \mathcal{C}_T \cap \mathcal{D}_T))(Y) = \text{cl}(R^\text{mar}(Y)) \subseteq (\text{cl}R^\text{mar})(Y),\]

for each \( Y \in L_d^\infty \). By [18] Remark 6 and Lemma 2, all three functions \( R^\text{mar}, \text{cl}(R^\text{mar} \cdot) \), \( \text{cl}R^\text{mar} \) have the same conjugate (on \( L_d^1 \times (\mathbb{R}_+^m \setminus \{0\}) \)) which is given by

\[
- \left( \tilde{R} \square T^m_{L_d^\infty}(\mathcal{F}_0, \mathcal{C}_0 \cap \mathcal{D}_0) \square \ldots \square T^m_{L_d^\infty}(\mathcal{F}_T, \mathcal{C}_T \cap \mathcal{D}_T) \right)^* = -\tilde{R}^* + \sum_{t=0}^T -\left( T^m_{L_d^\infty}(\mathcal{F}_t, \mathcal{C}_t \cap \mathcal{D}_t) \right)^*;
\]

23
note that this is the set-valued version of the rule “the conjugate of the infimal convolution of finitely many convex functions is the sum of their conjugates.” Let \((V, v) \in L^1_d \times (\mathbb{R}_+^m \setminus \{0\})\). Note that the arguments in Remark 2.13 can be repeated for \(\text{cl}(R^{\text{mar}}(\cdot))\) to observe that this function can be regarded as a convex risk measure. Hence, by [20] Proposition 6.7 on the conjugate of a risk measure, for every \((V, v) \in L^1_d \times (\mathbb{R}_+^m \setminus \{0\})\), we have 
\[-(\text{cl}(R^{\text{mar}}(\cdot)))(V, v) = \mathbb{R}^m\] 
unless we have \(V \in -(L^1_d)_+\) and \(v \in \mathbb{E}(-B^*V)\).

The next step is to pass from \(L^1_d \times (\mathbb{R}_+^m \setminus \{0\})\) to \(\mathcal{W}_{m,d} = \mathcal{M}_d(\mathbb{P}) \times ((\mathbb{R}_+^m \setminus \{0\}) \times \mathbb{R}^{d-m})\) using the “change of variables formula” provided by [20] Lemma 3.4. Application of this lemma yields that for every \(V \in -(L^1_d)_+\) with \(v \in \mathbb{E}(-B^*V)\), there exists \((Q, w) \in \mathcal{W}_{m,d}\) such that for every \(Y \in L^1_f\),
\[
\begin{align*}
\left\{ z \in \mathbb{R}^m \mid v^T z \geq \mathbb{E}\left[(-V)^TY\right] \right\} &= B^*\left(\mathbb{E}^Q[Y] + G(w)\right) \cap B(\mathbb{R}^m),
\end{align*}
\]
and conversely, every \((Q, w) \in \mathcal{W}_{m,d}\) can be obtained by some \(V \in -(L^1_d)_+\) with \(v \in \mathbb{E}(-B^*V)\) such that (6.2) holds for every \(Y \in L^1_f\). (Note that \(B^*(\mathbb{R}^d) = \mathbb{R}^m \times \{0 \in \mathbb{R}^{d-m}\}\). For such corresponding pairs \((V, v)\) and \((Q, w)\), using (6.2), we first observe that
\[
\begin{align*}
-\tilde{R}^* (V, v) &= \text{cl} \bigcup_{Y \in L^1_f} \left( \tilde{R}(Y) + \left\{ z \in \mathbb{R}^m \mid v^T z \geq \mathbb{E}\left[(-V)^TY\right] \right\} \right) \\
&= \text{cl} \bigcup_{Y \in L^1_f} \left( \tilde{R}(Y) + B^*\left(\mathbb{E}^Q[Y] + G(w)\right) \cap B(\mathbb{R}^m) \right) \\
&= \text{cl} \bigcup_{Y \in B(L^1_f)} \left( R(B^*Y) + \mathbb{E}^{B^*Q}[B^*Y] + G(B^*w) \right) \\
&= \text{cl} \bigcup_{X \in L^1_f} \left( R(X) + \mathbb{E}^{B^*Q}[X] + G(B^*w) \right) \\
&= -\alpha^\min_R (B^*Q, B^*w).
\end{align*}
\]

Next, let \(t \in \{0, \ldots, T\}\). For the same such pairs \((V, v)\) and \((Q, w)\), by Definition 6.1 and Definition 6.5, we have
\[
- \left( T^m_{L^1_d(f_t, C_t \cap D_t)} \right)^* (V, v) = \text{cl} \bigcup_{U^t \in L^1_d(f_t, C_t \cap D_t)} \left\{ z \in \mathbb{R}^m \mid v^T z \geq \mathbb{E}\left[(-V)^TU^t\right] \right\}
\]
\[
= \text{cl} \bigcup_{U^t \in L^1_d(f_t, C_t \cap D_t)} B^*\left(\mathbb{E}^Q[U^t] + G(w)\right) \cap B(\mathbb{R}^m).
\]

Finally, note that \(\text{cl} R^{\text{mar}}\) is a convex weak*-closed function that is finite at zero by assumption. Hence, by [18] Theorem 2, which is a biconjugation theorem for set-valued functions, we have
\[
(\text{cl} R^{\text{mar}})(Y) = \bigcap_{V \in -(L^1_d)_+, v = \mathbb{E}(-B^*V)} \left[ (\text{cl} R^{\text{mar}})^*(V, v) + \left\{ z \in \mathbb{R}^m \mid v^T z \geq \mathbb{E}\left[V^TY\right] \right\} \right],
\]
for every \(Y \in L^1_f\), and the above calculations allow for a passage to vector probability measures:
\[
\forall Y \in L^1_f : (\text{cl} R^{\text{mar}})(Y) = \bigcap_{(Q, w) \in \mathcal{W}_{m,d}} \left[ -\alpha^\min_{\text{cl} R^{\text{mar}}}(Q, w) + B^*\left(\mathbb{E}^Q[-Y] + G(w)\right) \cap B(\mathbb{R}^m) \right],
\]
\[\text{As discussed in Remark 2.13, a trivial modification is needed by rewriting the statements using } \mathbb{R}^m \text{ instead of } M = \mathbb{R}^m \times \{0 \in \mathbb{R}^{d-m}\}.\]
for each \((Q, w) \in \mathcal{W}_{m,d}\). □

**Proof of Corollary 2.18.** Let \((Q, w) \in \mathcal{W}_{m,d} \setminus \mathcal{W}_{m,d}^{\text{convex}}\). So there exist \(t \in \{0, \ldots, T\}\) and \(A \in \mathcal{F}_t\) such that \(P(A) > 0\) and \(\text{diag}(w) \mathbb{E} \left[ \frac{dQ}{dP} \mid \mathcal{F}_t \right] (\omega) \notin (0^+ C_t(\omega))^+\) for each \(\omega \in A\). Using the fact that the effective domain of the support function of a nonempty closed convex set in \(\mathbb{R}^d\) is a subset of its recession cone, which is an easy consequence of [31, Corollary 14.2.1] for instance, we see that 
\[
\inf_{y^t \in C_t(\omega)} \left( \text{diag}(w) \mathbb{E} \left[ \frac{dQ}{dP} \mid \mathcal{F}_t \right] (\omega) \right)^T y^t = -\infty \quad \text{for each } \omega \in A.
\]
Note that
\[
\text{cl} \bigcup_{U^t \in L_{m,d}^\infty(\mathcal{F}_t, \mathcal{C}_t)} B^* \left( \left( \mathbb{E}^Q [U^t] + G(w) \right) \cap B(\mathbb{R}^m) \right)
\]
where the last equality is by [32, Theorem 14.60]. Note that the passage to conditional expectations in the third line is necessary for the application of this theorem. Since \(P(A) > 0\), this implies \(\text{cl}\bigcup_{U^t \in L_{m,d}^\infty(\mathcal{F}_t, \mathcal{C}_t)} B^* \left( \left( \mathbb{E}^Q [U^t] + G(w) \right) \cap B(\mathbb{R}^m) \right) = \mathbb{R}^m\). By the computation in the proof of Proposition 2.17, it follows that \(-\alpha_{c_{1,R_{max}}}^\text{min}(Q, w) = \mathbb{R}^m\). □

**Proof of Corollary 2.19.** Let \(t \in \{0, \ldots, T\}\). For each \(\omega \in \Omega\), we have
\[
\inf_{y^t \in C_t(\omega)} \left( \text{diag}(w) \mathbb{E} \left[ \frac{dQ}{dP} \mid \mathcal{F}_t \right] (\omega) \right)^T y^t = \begin{cases} 0 & \text{if } \text{diag}(w) \mathbb{E} \left[ \frac{dQ}{dP} \mid \mathcal{F}_t \right] (\omega) \in (C_t(\omega))^+, \\ -\infty & \text{else} \end{cases}
\]
since \(C_t(\omega)\) is a nonempty closed convex cone. Similar to the calculation in the proof of Corollary 2.18 we have
\[
\text{cl} \bigcup_{U^t \in L_{m,d}^\infty(\mathcal{F}_t, \mathcal{C}_t)} B^* \left( \left( \mathbb{E}^Q [U^t] + G(w) \right) \cap B(\mathbb{R}^m) \right)
\]
from which the result follows immediately. □
6.2 Proofs of the results in Section 3

Proof of Proposition 3.3. Monotonicity, translativity and convexity are trivial. Let \( X \in L^\infty \).

For every \( s \in \mathbb{R} \),

\[
\ell(-\text{ess sup } X - s) \leq \mathbb{E}[\ell(-X - s)] \leq \ell(-\text{ess inf } X - s).
\]

Note that \( \ell \) is strictly increasing on

\[
\ell^{-1}(\text{int} \ell(\mathbb{R})) := \{ x \in \mathbb{R} \mid \ell(x) \in \text{int} \ell(\mathbb{R}) \} = (a, b),
\]

where

\[
a := \inf \left\{ x \in \mathbb{R} \mid \ell(x) > \inf_{y \in \mathbb{R}} \ell(y) \right\} \in \mathbb{R} \cup \{-\infty\},
\]

\[
b := \sup\{ x \in \mathbb{R} \mid \ell(x) < +\infty \} \in \mathbb{R} \cup \{+\infty\}.
\]

Hence, the inverse \( \ell^{-1} \) is well-defined as a function from \( \text{int} \ell(\mathbb{R}) \) to \((a, b)\). We have \( \mathbb{E}[\ell(-X - s)] \leq x^0 \) for each \( s \geq -\text{ess inf } X - \ell^{-1}(x^0) \), and \( \mathbb{E}[\ell(-X - s)] > x^0 \) for each \( s < -\text{ess sup } X - \ell^{-1}(x^0) \). So \( \rho_\ell(X) \in \mathbb{R} \). Besides, \( \mathbb{E}[\ell(-X - \rho_\ell(X))] \leq x^0 \) since the restriction of \( \ell \) on \( \text{dom} \ell \) is a continuous function.

To show weak*-lower semicontinuity, let \((X^n)_{n \in \mathbb{N}}\) be a bounded sequence in \( L^\infty \) converging to some \( X^\infty \in L^\infty \) \( \mathbb{P} \)-almost surely. Then, using Fatou’s lemma together with the fact that the restriction of \( \ell \) on \( \text{dom} \ell \) is nondecreasing and continuous, we have

\[
\mathbb{E}\left[\ell\left(-X^\infty - \liminf_{n \to \infty} \rho_\ell(X^n)\right)\right] = \mathbb{E}\left[\ell\left(\liminf_{n \to \infty} (-X^n - \rho_\ell(X^n))\right)\right] \\
\leq \liminf_{n \to \infty} \mathbb{E}[\ell(-X^n - \rho_\ell(X^n))] \leq x^0.
\]

This implies the so-called Fatou property of \( \rho_\ell \), namely, that \( \rho_\ell(X^\infty) \leq \liminf_{n \to \infty} \rho_\ell(X^n) \). By [16, Theorem 4.33], this is equivalent to the weak*-lower semicontinuity of \( \rho_\ell \). □

Proof of Proposition 3.4. Note that \( m \mapsto \mathbb{E}[\ell(-X - m)] \) is a proper convex function on \( \mathbb{R} \).

Hence, by Definition 3.2, \( \rho_\ell(X) \) is the value of a convex minimization problem. The corresponding Lagrangian dual objective function \( h_X \) is given by

\[
h_X(\lambda) = \inf_{s \in \mathbb{R}} \left( s + \lambda \left( \mathbb{E}[\ell(-X - s)] - x^0 \right) \right),
\]

for every \( \lambda \in \mathbb{R}_+ \). Since \( h_X(0) = -\infty \), the value of the dual problem equals the right hand side of (3.1). Finally, the two sides of (3.1) are equal since the usual Slater’s condition holds: There exists \( \bar{s} \in \mathbb{R} \) such that \( \mathbb{E}[\ell(-X - s)] < x^0 \). This is because we have \( \mathbb{E}[\ell(-X - s)] < x^0 \) for each \( s > -\text{ess inf } X - \ell^{-1}(x^0) \), where \( \ell^{-1} \) is the inverse function on \( \text{int} \ell(\mathbb{R}) \) as in the proof of Proposition 3.3. □

Proof of Proposition 3.6. Let \( f \) be a loss function and \( f^* : \mathbb{R} \to \mathbb{R} \cup \{+\infty\} \) its conjugate function. Note that \( \text{dom} f^* \subseteq \mathbb{R}_+ \) since, for each \( y < 0 \), we have

\[
f^*(y) \geq \sup_{n \in \mathbb{N}} (-ny - f(-n)) \geq \sup_{n \in \mathbb{N}} (-ny) - f(0) = +\infty,
\]

where we use the monotonicity of \( f \) for the second inequality. Clearly, \( f^*(y) \geq -f(0) \) for each \( y \in \mathbb{R} \). Besides, by [31, Theorem 23.3], we have \( \partial f(0) \neq \emptyset \) and, by [31, Theorem 23.5], we have
conclude (3.2), we consider the following cases:

\begin{equation}
L^{*}(g) = -f(0) \text{ for every } y \in \partial f(0). \quad \text{Hence, } f^{*} \text{ attains its infimum. Finally, } f^{*} \text{ is not of the form } y \mapsto \infty, \quad \text{since otherwise we would get } f(x) = (f^{*})^{*}(x) = +\infty \cdot 1_{\{x<0\}} - b \cdot 1_{\{x\leq a\}}, \quad x \in \mathbb{R}, \quad \text{so that } \ell \text{ would be identically constant on } \text{dom } f. \quad \text{Hence, } f^{*} \text{ is a divergence function.}
\end{equation}

Conversely, let \( \phi \) be a divergence function and \( \varphi^{*} : \mathbb{R} \to \mathbb{R} \) its conjugate function. Let \( x^{1}, x^{2} \in \mathbb{R} \) with \( x^{1} \leq x^{2} \). Since \( \text{dom } \phi \subseteq \mathbb{R}_{+} \), we have \( x^{1}y - \varphi(y) \geq x^{2}y - \varphi(y) \) for each \( y \in \text{dom } \phi \) so that \( \varphi^{*}(x^{1}) \geq \varphi^{*}(x^{2}) \). Hence, \( \varphi^{*} \) is nondecreasing. Clearly, \( \varphi^{*}(0) = -\inf_{y \in \mathbb{R}} \varphi(y) \in \mathbb{R} \) so that \( 0 \in \text{dom } \varphi^{*} \). Finally, \( \varphi^{*} \) is not identically constant on \( \text{dom } \varphi^{*} \) as otherwise \( \varphi = (\varphi^{*})^{*} \) would fail to satisfy property (iii) in Definition 3.5. Hence, \( \varphi^{*} \) is a loss function.

**Proof of Theorem 3.9.** Note that the right hand side of (3.2) can be rewritten as a maximization problem on \( L^{1} \):

\begin{equation}
\sup_{Q \in M_{1}(\mathbb{P})} \left( \mathbb{E}^Q [-X] - I_{g, \lambda}(Q | \mathbb{P}) \right) = \sup_{V \in (L^{1})_{+}} \left\{ \mathbb{E} [-XV] - \lambda \mathbb{E} \left[ g \left( \frac{1}{\lambda} V \right) \right] | \mathbb{E} [V] = 1 \right\}.
\end{equation}

The value of the corresponding Lagrangian dual problem is computed as

\begin{align*}
q_{X} & := \inf_{s \in \mathbb{R}} \sup_{V \in (L^{1})_{+}} \left( \mathbb{E} [-XV] - \lambda \mathbb{E} \left[ g \left( \frac{1}{\lambda} V \right) \right] + s(1 - \mathbb{E} [V]) \right) \\
& = \inf_{s \in \mathbb{R}} \left( s + \sup_{V \in (L^{1})_{+}} \mathbb{E} \left[ (-X - s)V - \lambda g \left( \frac{1}{\lambda} V \right) \right] \right) \\
& = \inf_{s \in \mathbb{R}} \left( s + \mathbb{E} \left[ \sup_{z \in \mathbb{R}_{+}} \left( (-X - s)z - \lambda g \left( \frac{1}{\lambda} z \right) \right) \right] \right) \\
& = \inf_{s \in \mathbb{R}} \left( s + \mathbb{E} [g_{\lambda}^{*}(-X - s)] \right),
\end{align*}

where the third equality is due to Theorem 14.60, and \( g_{\lambda}^{*} \) is the conjugate of the divergence function \( g_{\lambda} \); see Remark 3.7. Hence, \( g_{\lambda}^{*} = \lambda \ell \) and \( q_{X} \) equals the left hand side of (3.2). Finally, to conclude (3.2), we consider the following cases:

\begin{enumerate}[(i)]
\item Suppose that \( 1 \in \text{int dom } g_{\lambda} \), that is, \( \lambda \alpha < 1 < \lambda \beta \). (Recall that \( \alpha \) and \( \beta \) are the endpoints of \( \text{dom } g \), see Definition 3.8 et seq.) Then the following constraint qualification holds, for instance, with \( \bar{V} \equiv 1 \):

\begin{equation}
\exists V \in L_{+}^{1}: \mathbb{E} [V] = 1, \quad V \in \text{int dom } g_{\lambda}, \quad \mathbb{P}-\text{almost surely}. \tag{6.3}
\end{equation}

By Corollary 4.8], (6.3) suffices to conclude (3.2). Note that we have

\begin{equation}
-\mathbb{E} [X] - \lambda g \left( \frac{1}{\lambda} \right) \leq \sup_{V \in L_{+}^{1}} \left\{ \mathbb{E} [-XV] - \lambda \mathbb{E} \left[ g \left( \frac{1}{\lambda} V \right) \right] | \mathbb{E} [V] = 1 \right\} \leq -\text{ess inf}_{x} X - \lambda \inf_{x} g(x),
\end{equation}

so that both sides of (3.2) are in \( \mathbb{R} \).

\item Suppose that \( \lambda \alpha = 1 \) or \( \lambda \beta = 1 \). In this case, the only \( V \in L_{+}^{1} \) with \( \mathbb{E} [V] = 1 \) and \( V \in \text{dom } g_{\lambda} \) \( \mathbb{P} \)-almost surely is \( V \equiv 1 \), and hence, the right hand side of (3.2) gives \( -\mathbb{E} [X] - \lambda g \left( \frac{1}{\lambda} \right) \in \mathbb{R} \). Note that (6.3) fails to hold here. If \( \alpha = \beta \), that is, if \( \text{dom } g = \{ \alpha \} \), then \( \ell(x) = x/\lambda - g \left( \frac{1}{\lambda} \right), x \in \mathbb{R} \).
\(\mathbb{R}\), and the left hand side of (3.2) gives \(-\mathbb{E}[X] - \lambda g(1/\lambda)\), showing (3.2). Suppose \(\alpha \neq \beta\) so that \(\text{int dom } g_\lambda \neq \emptyset\). If \(\lambda \beta = 1\), then, using (3.2) for the previous case, we have

\[
\inf_{s \in \mathbb{R}} (s + \lambda \mathbb{E} [\ell(-X - s)]) = \lim_{\varepsilon \downarrow 0} \inf_{s \in \mathbb{R}} (s + (\lambda + \varepsilon) \mathbb{E} [\ell(-X - s)])
\]

\[
= \lim_{\varepsilon \downarrow 0} \sup_{Q \in \mathcal{M}_1(\mathbb{P})} \left( \mathbb{E}^Q [\mathbb{E}^Q [-X] - (\lambda + \varepsilon) \mathbb{E} \left[ g \left( \frac{1}{\lambda + \varepsilon} \frac{dQ}{d\mathbb{P}} \right) \right] \right)
\]

\[
= \sup_{Q \in \mathcal{M}_1(\mathbb{P})} \left( \mathbb{E}^Q [\mathbb{E}^Q [-X] - \lambda \mathbb{E} \left[ g \left( \frac{1}{\lambda} \frac{dQ}{d\mathbb{P}} \right) \right] \right).
\]

Here, the first and the last equalities follow by the fact that the proper, concave, upper semicontinuous function

\[
\mathbb{R} \ni \gamma \rightarrow \inf_{s \in \mathbb{R}} (s + \gamma \mathbb{E} \ell(-X - s)) \in \mathbb{R} \cup \{-\infty\}
\]

is right-continuous at \(\gamma = \lambda\). This argument applies for \(\lambda \alpha = 1\) analogously if we switch the direction of the limit from \(\varepsilon \downarrow 0\) to \(\varepsilon \uparrow 0\) and use the fact that the function in (6.4) is left-continuous at \(\beta = \lambda\). Hence, we obtain (3.2).

(iii) Suppose \(\lambda \alpha > 1\) or \(\lambda \beta < 1\). In this case, there is no \(Y \in L^1_1\) with \(\mathbb{E}[Y] = 1\) and \(Y \in \text{dom } g_\lambda\). \(\mathbb{P}\)-almost surely. Hence, the right hand side of (3.2) gives \(-\infty\). On the other hand, we have

\[
\inf_{s \in \mathbb{R}} (s + \lambda \mathbb{E} [\ell(-X - s)]) \leq \inf_{s \in \mathbb{R}} (s + \lambda \ell(\text{ess inf } X - s)) = -\text{ess inf } X - \sup_{s \in \mathbb{R}} (s - \lambda \ell(s))
\]

\[
= -\text{ess inf } X - \lambda g \left( \frac{1}{\lambda} \right) = -\infty.
\]

Hence, (3.2) holds.

\[\square\]

**Proof of Proposition 3.12.** The definition of \(\delta_{g,\lambda}\) given by (3.4) guarantees monotonicity, translativity, convexity and weak*-lower semicontinuity directly. Note that the function \(Q \mapsto I_{g,\lambda} (Q \mid \mathbb{P}) + \lambda x^0\) on \(\mathcal{M}_1(\mathbb{P})\) is a penalty function in the sense that

\[
\inf_{Q \in \mathcal{M}_1(\mathbb{P})} (\lambda x^0 + I_{g,\lambda} (Q \mid \mathbb{P})) = -\delta_{g,\lambda}(0) \in \mathbb{R}.
\]

Finally, this function is indeed the minimal penalty function of \(\delta_{g,\lambda}\) since, using the definition of minimal penalty function, we have

\[
\alpha_{\text{min}}^{\delta_{g,\lambda}}(Q) = \sup_{X \in L^\infty} \left( \mathbb{E}^Q [-X] - \delta_{g,\lambda}(X) \right) = \lambda x^0 + \sup_{X \in L^\infty} \left( \mathbb{E}^Q [-X] + \sup_{s \in \mathbb{R}} (-s - \lambda \mathbb{E} [\ell(-X - s)]) \right)
\]

\[
= \lambda x^0 + \sup_{s \in \mathbb{R}} \left( -s + \mathbb{E} \left[ \sup_{x \in \mathbb{R}} \left( -\frac{dQ}{d\mathbb{P}} \frac{dQ}{d\mathbb{P}} X - \lambda \ell(-X - s) \right) \right] \right)
\]

\[
= \lambda x^0 + \sup_{s \in \mathbb{R}} \left( -s + \mathbb{E} \left[ \sup_{x \in \mathbb{R}} \left( -\frac{dQ}{d\mathbb{P}} s + g_\lambda \left( \frac{dQ}{d\mathbb{P}} \right) \right) \right] \right) = \lambda x^0 + I_{g,\lambda} (Q \mid \mathbb{P}),
\]

for each \(Q \in \mathcal{M}_1(\mathbb{P})\). Note that, in this computation, the fourth equality follows by [32, Theorem 14.60] and the fifth equality follows from Remark 3.7. \[\square\]
Proof of Proposition 3.13. Let $Q \in \mathcal{M}_1(\mathbb{P})$. By the definition of minimal penalty function and Definition 3.2, we have

$$\alpha_{\rho|\cdot|}^\text{min}(Q) = \sup_{X \in L^\infty} \left( E^Q [-X] - \rho_\ell(X) \right) = \sup_{X \in L^\infty} \left( E^Q [-X] - \inf_{s \in \mathbb{R}} (s + I_{(-\infty,x^0]}(E[\ell(-X - s)]) \right)$$

$$= \sup_{s \in \mathbb{R}} \left( -s - \inf_{X \in L^\infty} \left( E^Q [X] + I_{(-\infty,x^0]}(E[\ell(-X - s)]) \right) \right)$$

$$= \sup_{s \in \mathbb{R}} \sup_{X \in L^\infty} \left( E^Q [X] - I_{(-\infty,x^0]}(E[\ell(X)]) \right) = \sup_{X \in L^\infty} \left( E^Q [X] \right) \{ E[\ell(X)] \leq x^0 \}.$$

The value of the corresponding Lagrangian dual problem is

$$q(Q) := \sup_{\lambda > 0} \sup_{X \in L^\infty} \left( E^Q [X] + \lambda \left( x^0 - E[\ell(X)] \right) \right) = \inf_{\lambda > 0} \left( \lambda x^0 + \sup_{X \in L^\infty} E \left[ \frac{dQ}{d\mathbb{P}} X - \lambda \ell(x) \right] \right)$$

$$= \inf_{\lambda > 0} \left( \lambda x^0 + E \left[ \sup_{x \in \mathbb{R}} \left( \frac{dQ}{d\mathbb{P}} x - \lambda \ell(x) \right) \right] \right) = \inf_{\lambda > 0} \left( \lambda x^0 + I_{g,\lambda}(Q \mid \mathbb{P}) \right),$$

where, we use [32, Theorem 14.60] for the third equality and Remark 3.7 for the fourth equality again. Note that Slater’s condition holds, that is, there exists $\bar{X} \in L^\infty$ such that $E[\ell(X)] < x^0$; take, for example, $\bar{X} \equiv \ell^{-1}(x^0) - 1$, where $\ell^{-1}$ is the inverse function on $\text{int} \ell(\mathbb{R})$ as in the proof of Proposition 3.3. Therefore, $\alpha_{\rho|\cdot|}^\text{min}(Q) = q(Q)$. Note that $I_{g,\lambda}(Q \mid \mathbb{P}) = +\infty$ for every $\lambda \notin 1/\text{dom} g$; see case (iii) in the proof of Theorem 3.9. Hence, we also have $q(Q) = \inf_{\lambda \in 1/\text{dom} g} \alpha_{\delta_{g,\lambda}}^\text{min}(Q)$. \qed

### 6.3 A remark about the scalar loss functions

In [15] and [16], Proposition 3.13 ([15 Theorem 10], [16 Theorem 4.115]) is proved with the additional assumption that $\ell$ maps into $\mathbb{R}$. This assumption implies that the $\ell$-shortfall risk measure is continuous from below and the supremum in (3.5) is attained ([16 Proposition 4.113]). Besides, the same assumption implies the so-called superlinear growth condition on $g$, namely, that $\lim_{y \to +\infty} g(y)/y = +\infty$ ([15 Lemma 11]). Their analytic proof for Proposition 3.13 makes use of this property instead of the dual relationship with divergence risk measures. Using this proposition and assuming that $1 \in \text{dom} g$, they prove Theorem 3.9 for $\lambda = 1$ ([16 Theorem 4.122]), in which case $\delta_{g,1}$ is guaranteed to be a risk measure (has finite values). In our treatment, while 1 may not be in $\text{dom} g$, there exists some $\bar{\lambda} \in 1/\text{dom} g$ such that $1 \in \text{dom} g_{\bar{\lambda}}$ and hence $\delta_{g,\bar{\lambda}}$ is a risk measure.

In [5], on the other hand, the divergence function $g$ is of central importance: In addition to the assumptions here, they assume that $g$ attains its infimum at 1 with value 0, which is equivalent to assuming that $\ell(0) = 0$ and $1 \in \partial \ell(0)$. These assumptions make $g$ a natural divergence function in the sense that the function $Q \mapsto E[g(\frac{dQ}{d\mathbb{P}})]$ on $\mathcal{M}_1(\mathbb{P})$ has positive values and takes the value 0 if $Q = \mathbb{P}$; $E[g(\frac{dQ}{d\mathbb{P}})]$ can be interpreted as the distance between some subjective measure $Q \in \mathcal{M}_1(\mathbb{P})$ and the physical measure $\mathbb{P}$. On the other hand, the additional assumptions on the loss function $\ell$ may be considered as restrictive. Here, we take $\ell$ as the central object by dropping these assumptions and use the convex duality approach as in [5]. Note that Theorem 3.9 ([5 Theorem 4.2]) and Proposition 3.12 ([5 Theorem 4.4]) are proven in [5] for the case $\lambda = 1$. Here, we generalize their proof, basically, by considering the cases where the constraint qualification (6.3), which is also used in the proof of [5] Theorem 4.2], fails to hold.

### 6.4 Proofs of the results in Section 4

**Proof of Proposition 4.3** ($L^\infty_m$)-monotonicity, $\mathbb{R}^m$-translativity and convexity are trivial. Let $X \in L^\infty_m$. Using the proof of Proposition 3.3, we can find $z^1 \in \mathbb{R}^d$ with $E[\ell(-X - z^1)] \in x^0 - \mathbb{R}^m_+$.
and $z^2 \in \mathbb{R}^m$ with $\mathbb{E} [\ell(-X - z^2)] \in x^0 + \mathbb{R}^m_{++}$. So $R_\ell(X) \notin \{0, \mathbb{R}^m\}$.

To show weak* closedness, let $(X_n)_{n \in \mathbb{N}}$ be a bounded sequence in $L^\infty_m$ converging to some $X \in L^\infty_m \mathbb{P}$-almost surely. Let $z \in \mathbb{R}^m$ and suppose that there exists $z^n \in R_{\ell}(X^n)$, for each $n \in \mathbb{N}$, such that $(z^n)_{n \in \mathbb{N}}$ converges to $z$. Using the dominated convergence theorem, the closedness of $x_0 - C$, and the fact that the restriction of $\ell$ on dom $\ell$ is continuous, we have

$$\mathbb{E} [\ell(-X - z)] = \mathbb{E} \left[\ell \left( \lim_{n \to \infty} (-X^n - z^n) \right) \right] = \lim_{n \to \infty} \mathbb{E} [\ell(-X^n - z^n)] \in x^0 - C,$$

that is, $z \in R_{\ell}(X)$. This shows the so-called Fatou property of $R_{\ell}$, namely, that

$$\liminf_{n \to \infty} R_{\ell}(X^n) := \left\{ z \in \mathbb{R}^m \mid \forall n \in \mathbb{N}, \exists z^n \in R_{\ell}(X^n): \lim_{n \to \infty} z^n = z \right\} \subseteq R_{\ell}(X).$$

By [19] Theorem 6.2, this is equivalent to the weak*-closedness of $R_{\ell}$. \hfill \square

**Proof of Proposition 4.4.** Consider the minimization problem (4.2). The halfspace-valued functions $S_{(\lambda, v)} \colon \mathbb{R}^m \to \mathcal{G}(\mathbb{R}^m, \mathbb{R}^m_+)$ for $v \in \mathbb{R}^m_+ \setminus \{0\}$, $\lambda \in \mathbb{R}^m$, defined by

$$S_{(\lambda, v)}(x) = \left\{ z \in \mathbb{R}^m \mid v^T z \geq \lambda^T x \right\},$$

will be used as set-valued substitutes for the (continuous) linear functionals in the scalar duality theory as in [13, 21]. Here, we have two types of dual variables: The variable $\lambda \in \mathbb{R}^m$ is the usual vector of Lagrange multipliers which will be used to scalarize the values of the constraint function $\Psi$ whereas the variable $v \in \mathbb{R}^m_+$ is the weight vector which will be used to scalarize the values of the objective function $\Phi$. Fix $X \in L^\infty_m$. The set-valued Lagrangian $L \colon \mathbb{R}^m \times \mathbb{R}^m \times (\mathbb{R}^m_+ \setminus \{0\}) \to \mathcal{G}(\mathbb{R}^m, \mathbb{R}^m_+)$ and the objective function $H \colon \mathbb{R}^m \times (\mathbb{R}^m_+ \setminus \{0\}) \to \mathcal{G}(\mathbb{R}^m, \mathbb{R}^m_+)$ of the dual problem for (4.2) are defined by

$$L(z, \lambda, v) = \mathcal{C} \left( \Phi(z) + \inf_{\mathcal{G}(\mathbb{R}^m, \mathbb{R}^m_+), z} \left\{ S_{(\lambda, v)}(x) \mid x \in \Psi(z) \right\} \right),$$

$$H(\lambda, v) = \inf_{\mathcal{G}(\mathbb{R}^m, \mathbb{R}^m_+), z} \left\{ L(z, \lambda, v) \mid z \in \mathbb{R}^m \right\},$$

see [21] Section 2 and Section 6.1. Using the definition of $\Psi$ and the formula (2.4) for infimum in $\mathcal{G}(\mathbb{R}^m, \mathbb{R}^m_+)$, we obtain for $(z, \lambda, v) \in \mathbb{R}^m \times \mathbb{R}^m \times (\mathbb{R}^m_+ \setminus \{0\})$

$$L(z, \lambda, v) = \mathcal{C} \left( z + \mathbb{R}^m_+ + \mathcal{C} \{ \eta \in \mathbb{R}^m \mid v^T \eta \geq \inf_{x \in (\mathbb{E}[\ell(-X - z)] - x^0 + C) \cap \mathbb{R}^m} S_{(\lambda, v)}(x) \} \right)$$

$$= \mathcal{C} \left( z + \mathbb{R}^m_+ + \left\{ \eta \in \mathbb{R}^m \mid v^T \eta \geq \inf_{x \in (\mathbb{E}[\ell(-X - z)] - x^0 + C) \cap \mathbb{R}^m} \lambda^T x \right\} \right)$$

$$= \left\{ \eta \in \mathbb{R}^m \mid \mathbb{E} [\ell(-X - z)] \in \mathbb{R}^m, \ v^T \eta \geq \inf_{z \in \mathbb{R}^m} \left( v^T z + \lambda^T \mathbb{E} [\ell(-X - z)] + \inf_{x \in -x^0 + C} \lambda^T x \right) \right\}.$$  

This gives for $(\lambda, v) \in \mathbb{R}^m \times (\mathbb{R}^m_+ \setminus \{0\})$

$$H(\lambda, v) = \inf_{z \in \mathbb{R}^m} L(z, \lambda, v) = \mathcal{C} \left( \left\{ \eta \in \mathbb{R}^m \mid \mathbb{E} [\ell(-X - z)] \in \mathbb{R}^m, \ v^T \eta \geq \inf_{z \in \mathbb{R}^m} \left( v^T z + \lambda^T \mathbb{E} [\ell(-X - z)] + \inf_{x \in -x^0 + C} \lambda^T x \right) \right\} \right).$$  

(6.6)
Note that \( L(z, 0, v) = \emptyset \) whenever \( \mathbb{E} [\ell(-X - z)] \not\in \mathbb{R}^m \), but this cannot happen for all \( z \in \mathbb{R}^m \) since \( X \in L_m^\infty \). Hence \( H(0, v) = \mathbb{R}^m \) for every \( v \in \mathbb{R}^m_+ \setminus \{0\} \). Suppose \( \lambda \not\in \mathbb{R}^m_+ \). Since \( C + \mathbb{R}^m_+ \subseteq C \), there exists \( \bar{x} \in C \) such that for every \( n \in \mathbb{N} \), we have \( n \bar{x} \in C \) and \( \lambda^T \bar{x} < 0 \). Hence \( \inf_{x \in C} \lambda^T x = -\infty \) and \( H(\lambda, v) = \mathbb{R}^m \) for every \( v \in \mathbb{R}^m_+ \setminus \{0\} \). The dual optimal value is the supremum over the dual variables: Since the supremum in \( G(\mathbb{R}^m, \mathbb{R}^m_+) \) is the intersection, we have to take it over the \( H(\lambda, v) \) with \( \lambda, v \) running only through \( \mathbb{R}^m_+ \setminus \{0\} \), that is, the dual optimal value is

\[
d_X := \sup_{(\mathbb{R}^m, \mathbb{R}^m_+)} \{ H(\lambda, v) \mid \lambda \in \mathbb{R}^m, v \in \mathbb{R}^m_+ \setminus \{0\} \} = \bigcap_{\lambda \in \mathbb{R}^m, v \in \mathbb{R}^m_+ \setminus \{0\}} H(\lambda, v). \quad (6.7)
\]

Noting that for \( \lambda, v \in \mathbb{R}^m_+ \setminus \{0\} \)

\[
H(\lambda, v) = \left\{ \eta \in \mathbb{R}^m \mid v^T \eta \geq \inf_{z \in \mathbb{R}^m} \left( v^T z + \lambda^T \mathbb{E} [\ell(-X - z)] \right) + \inf_{x \in -\bar{x} + C} \lambda^T x \right\}
\]

since by convention \( \lambda^T (+\infty) = +\infty \) we see that \( d_X \) equals the right hand side of (4.3). Finally, the two sides of (4.3) are equal since the set-valued version of Slater’s condition holds (see proof of Proposition 3.4). This follows as for the scalar version (see proof of Proposition 3.4).

**Proof of Proposition 4.7** With (6.6) in view, we define for \( r, w \in \mathbb{R}^m_+ \setminus \{0\} \)

\[
M(r, w) = \left\{ \eta \in \mathbb{R}^m \mid w^T \eta \geq w^T \delta_{\ell_2}(X) + \inf_{x \in C} w^T \text{diag}(r)x \right\}, \quad (6.8)
\]

and we will show

\[
\bigcap \{ H(\lambda, v) \mid \lambda, v \in \mathbb{R}^m_+ \setminus \{0\} \} = \bigcap \{ M(r, w) \mid r, w \in \mathbb{R}^m_+ \setminus \{0\}, r^T w > 0 \}.
\]

First, if \( r, w \in \mathbb{R}^m_+ \setminus \{0\} \) and \( r^T w > 0 \), we define \( \lambda_i = r_i w_i \) and \( \lambda_i = w_i \) for \( i \in \{1, \ldots, m\} \). Then \( \lambda, v \in \mathbb{R}^m_+ \setminus \{0\} \) as well as \( H(\lambda, v) = M(r, w) \); see Definition 4.6 and (4.4). This means that the intersection on the left hand side runs over at least as many sets as the one on the right hand side, hence “\( \subseteq \)" holds true.

Conversely, if \( \lambda, v \in \mathbb{R}^m_+ \setminus \{0\} \) we define for \( n \in \mathbb{N} \) and \( i \in \{1, \ldots, m\} \)

\[
r_i^n = \frac{\lambda_i}{v_i} \quad \text{and} \quad w_i^n = v_i \quad \text{whenever} \quad v_i > 0,
\]

\[
r_i^n = 1 \quad \text{and} \quad w_i^n = \lambda_i \quad \text{whenever} \quad v_i = 0, \lambda_i = 0,
\]

\[
r_i^n = n \lambda_i \quad \text{and} \quad w_i^n = \frac{1}{n} \quad \text{whenever} \quad v_i = 0, \lambda_i > 0.
\]

Then \( r^n, w^n \in \mathbb{R}^m_+ \setminus \{0\} \) and \( \lambda_i = r_i^n w_i^n \). Moreover, \( (r^n)^T w^n > 0 \) since \( \lambda \not= 0 \). Assume

\[
\eta \in \bigcap \{ M(r, w) \mid r, w \in \mathbb{R}^m_+ \setminus \{0\}, r^T w > 0 \}.
\]

If there is no \( i \in \{1, \ldots, m\} \) satisfying \( v_i = 0 \) and \( \lambda_i > 0 \), then \( v = w^n \) and \( H(\lambda, v) = M(r^n, w^n) \) for every \( n \in \mathbb{N} \), hence \( \eta \in H(\lambda, v) \). Next, assume there is at least one \( j \in \{1, \ldots, m\} \) with \( v_j = 0 \), \( \lambda_j > 0 \) as well as \( g_j(0) = +\infty \). Then

\[
\inf_{z_j \in \mathbb{R}} \mathbb{E} [\ell_j(-X_j - z_j)] = \inf_{y \in \mathbb{R}} \ell_j(y) = -g_j(0) = -\infty. \quad (6.9)
\]

31
since $\ell_j$ is nondecreasing and $X \in L^\infty$. This implies
\[
\sum_{i=1}^m \inf_{z_i \in \mathbb{R}} \{v_i z_i + \lambda_i \mathbb{E} [\ell_i (-X_i - z_i)]\} + \inf_{x \in -x^0 + C} \lambda^T x = -\infty
\]
since $v_j = 0$ and none of the terms in the sum is $+\infty$. Hence $\eta \in H(\lambda, v)$. Finally, assume that there is at least one $j \in \{1, \ldots, m\}$ with $v_j = 0$, $\lambda_j > 0$ as well as $g_j(0) \in \mathbb{R}$. Since $\eta \in M(r^n, w^n)$ for every $n \in \mathbb{N}$, it follows
\[
\sum_{i: v_i > 0} v_i \eta_i + \sum_{i: v_i = \lambda_i = 0} \frac{\eta_i}{n} \geq \sum_{i: v_i > 0} \inf_{z_i \in \mathbb{R}} \{v_i z_i + \lambda_i \mathbb{E} [\ell_i (-X_i - z_i)]\}
+ \sum_{i: v_i = \lambda_i > 0} \frac{1}{n} \inf_{z_i \in \mathbb{R}} \{z_i + n \lambda_i \mathbb{E} [\ell_i (-X_i - z_i)]\} + \inf_{x \in -x^0 + C} \lambda^T x.
\]
(6.10)
Whenever $\lambda_j > 0$ we obtain
\[
- \text{ess sup} X - n \lambda_j g_j \left( \frac{1}{n \lambda_j} \right) \leq \inf_{z_j \in \mathbb{R}} \{z_j + n \lambda_j \mathbb{E} [\ell_j (-X_j - z_j)]\} \leq - \text{ess inf} X - n \lambda_j g_j \left( \frac{1}{n \lambda_j} \right)
\]
for each $n \in \mathbb{N}$. This can be checked with a similar calculation as in (6.5). Since $g_j$ is convex and lower semicontinuous, the restriction of $g_j$ to $\text{cl dom} g_j$ is a continuous function (see Proposition 2.1.6), and we obtain
\[
\lim_{n \to \infty} \frac{1}{n} \inf_{z_j \in \mathbb{R}} \{z_j + n \lambda_j \mathbb{E} [\ell_j (-X_j - z_j)]\} = - \lim_{n \to \infty} \lambda_j g_j \left( \frac{1}{n \lambda_j} \right) = - \lambda_j g_j(0) = \lambda_j \inf_{y \in \mathbb{R}} \ell_j(y)
\]
whenever $v_j = 0$, $\lambda_j > 0$ and $g_j(0) \in \mathbb{R}$. Using the first two equalities in (6.9) and taking the limit in (6.10) as $n \to \infty$, we finally obtain
\[
v^T \eta \geq \sum_{i: v_i > 0} \inf_{z_i \in \mathbb{R}} \{v_i z_i + \lambda_i \mathbb{E} [\ell_i (-X_i - z_i)]\} + \sum_{i: v_i = \lambda_i > 0} \lambda_i \inf_{y \in \mathbb{R}} \ell_i(y) + \inf_{x \in -x^0 + C} \lambda^T x
\]
\[
= \sum_{i: v_i > 0} \inf_{z_i \in \mathbb{R}} \{v_i z_i + \lambda_i \mathbb{E} [\ell_i (-X_i - z_i)]\} + \sum_{i: v_i = \lambda_i > 0} \inf_{z_i \in \mathbb{R}} \{v_i z_i + \lambda_i \mathbb{E} [\ell_i (-X_i - z_i)]\} + \inf_{x \in -x^0 + C} \lambda^T x
\]
\[
= \sum_{i=1}^m \inf_{z_i \in \mathbb{R}} \{v_i z_i + \lambda_i \mathbb{E} [\ell_i (-X_i - z_i)]\} + \inf_{x \in -x^0 + C} \lambda^T x,
\]
that is, $\eta \in H(\lambda, v)$. Note that $M(r, w) = \mathbb{R}^m$ whenever $w^T r = 0$, thus
\[
\bigcap_{\lambda, v \in \mathbb{R}^m \setminus \{0\}} H(\lambda, v) = \bigcap_{r, w \in \mathbb{R}^m \setminus \{0\}} M(r, w).
\]

\[\square\]

**Proof of Proposition 4.10.** Let $X \in L^\infty$. We have $\delta_{g,r}(X) \in \mathbb{R}^m$ so that
\[
D_{g,r}(X) = \delta_{g,r}(X) + \bigcap_{w \in \mathbb{R}^m \setminus \{0\}} \left\{ z \in \mathbb{R}^m \mid w^T z \geq \inf_{x \in C} w^T \text{diag}(r)x \right\}
\]
\[
= \delta_{g,r}(X) + \bigcap_{w \in \mathbb{R}^m \setminus \{0\}} \left\{ z \in \mathbb{R}^m \mid w^T z \geq \inf_{x \in \text{diag}(r)C} w^T x \right\} = \delta_{g,r}(X) + \text{diag}(r)C,
\]
32
where the last equality is due to $C \in \mathcal{G}(\mathbb{R}^m, \mathbb{R}^m_+)$ and the fact that a closed convex set is the intersection of all closed halfspaces containing it; see [21, (5.2)]. This representation also guarantees that $D_{g,r}(X) \notin \{\emptyset, \mathbb{R}^m\}$. The second representation of $D_{g,r}(X)$ as a set-valued infimum is obvious from the definition of $\delta_{g,r}(X)$ given by Definition 4.6.

Let $w \in \mathbb{R}^m_+\setminus\{0\}$ and $M(r, w)$ as in (6.8). For the moment, let us denote by $-\alpha$ the function defined by the right hand side of (4.8). By (4.4), we have

$$M(r, w) = \left\{ z \in \mathbb{R}^m \mid w^T z \geq \sum_{i=1}^{m} \sup_{Q_i \in \mathcal{M}_1(\mathcal{P})} \left\{ w_i E^Q_i [-X_i] - w_i I_{g_i, r_i}(Q_i | \mathcal{P}) \right\} + \inf_{x \in -x^0 + C} w^T \text{diag}(r)x \right\}$$

$$= \bigcap_{Q \in \mathcal{M}_m(\mathcal{P})} \left\{ z \in \mathbb{R}^m \mid w^T z \geq w^T \left( E^Q [-X] - I_{g, r}(Q | \mathcal{P}) \right) + \inf_{x \in -x^0 + C} w^T \text{diag}(r)x \right\}$$

$$= \bigcap_{Q \in \mathcal{M}_m(\mathcal{P})} \left( -\alpha(Q, w) + E^Q [-X] \right).$$

Hence,

$$D_{g,r}(X) = \bigcap_{w \in \mathbb{R}^m_+ \setminus\{0\}} M(r, w) = \bigcap_{(Q,w) \in \mathcal{M}_m(\mathcal{P}) \times (\mathbb{R}^m_+ \setminus\{0\})} \left( -\alpha(Q, w) + E^Q [-X] \right).$$

This representation of $D_{g,r}$ guarantees monotonicity, translativity, convexity and weak*-closedness.

Note that $-\alpha$ is a penalty function in the sense of Definition 2.5 because

$$\bigcap_{(Q,w) \in \mathcal{M}_m(\mathcal{P}) \times (\mathbb{R}^m_+ \setminus\{0\})} -\alpha(Q, w) = D_{g,r}(0) \notin \emptyset \mathbb{R}^d.$$ 

Finally, we show that $-\alpha$ is indeed the minimal penalty function of $D_{g,r}$. By the definition of the minimal penalty function, we obtain for $Q \in \mathcal{M}_m(\mathcal{P})$, $w \in \mathbb{R}^m_+ \setminus\{0\}$

$$-\alpha_{D_{g,r}}(Q, w) = \text{cl} \bigcup_{X \in L^\infty_m} \left( E^Q [X] + G(w) + D_{g,r}(X) \right)$$

$$= \text{cl} \bigcup_{X \in L^\infty_m} \left( E^Q [X] + \delta_{g,r}(X) + \text{diag}(r)C + G(w) \right)$$

$$= \text{cl} \bigcup_{X \in L^\infty_m} \left( E^Q [X] + \delta_{g,r}(X) + \left\{ z \in \mathbb{R}^m \mid w^T z \geq \inf_{x \in C} w^T \text{diag}(r)x \right\} \right)$$

$$= \left\{ z \in \mathbb{R}^m \mid w^T z \geq \inf_{X \in L^\infty_m} \sum_{i=1}^{m} \inf_{X_i \in L^\infty_i} \left( E^Q_i [X_i] + \delta_{g_i, r_i}(X_i) \right) + \inf_{x \in C} \text{diag}(r)x \right\}$$

$$= \left\{ z \in \mathbb{R}^m \mid w^T z \geq \sum_{i=1}^{m} w_i \inf_{X_i \in L^\infty_i} \left( E^Q_i [X_i] + \delta_{g_i, r_i}(X_i) \right) + \inf_{x \in C} \text{diag}(r)x \right\}$$

$$= \left\{ z \in \mathbb{R}^m \mid w^T z \geq \sum_{i=1}^{m} w_i (-I_{g_i, r_i}(Q_i | \mathcal{P}) - r_i x_i^0) + \inf_{x \in C} \text{diag}(r)x \right\}$$

$$= -\alpha(Q, w),$$

where the sixth equality follows from the analogous scalar result established in Proposition 3.12. Hence, $-\alpha_{D_{g,r}} = -\alpha$, that is, (4.8) holds. □
Proof of Proposition 4.12. By the definition of minimal penalty function, we obtain for $Q \in \mathcal{M}_m(\mathbb{P})$, $w \in \mathbb{R}^m_{++}$,

$$-\alpha_{R^e}^{\min}(Q, w) = \operatorname{cl} \bigcup_{X \in L_m^\infty} \left( \mathbb{E}^Q[X] + G(w) + R_e(X) \right)$$

$$= \operatorname{cl} \bigcup_{X \in L_m^\infty} \left( \mathbb{E}^Q[X] + G(w) + \bigcup_{z \in \mathbb{R}_m^m} \{ z + \mathbb{R}_m^m \mid 0 \in \mathbb{E}[\ell(-X - z)] - x^0 + C \} \right)$$

$$= \operatorname{cl} \bigcup_{X \in L_m^\infty} \bigcup_{z \in \mathbb{R}_m^m} \left( \mathbb{E}^Q[X] + G(w) \mid 0 \in \mathbb{E}[\ell(-X - z)] - x^0 + C \right)$$

$$= \operatorname{cl} \bigcup_{X \in L_m^\infty} \left( \mathbb{E}^Q[-X] + G(w) \mid 0 \in \mathbb{E}[\ell(X)] - x^0 + C \right)$$

$$= \inf_{X \in L_m^\infty} \left( \mathbb{E}^Q[-X] + G(w) \mid 0 \in \mathbb{E}[\ell(X)] - x^0 + C \right)$$

where the infimum in the last line is taken in $(G(\mathbb{R}_m^m, \mathbb{R}_+^m), \sup)$. Next, we compute the value of the Lagrangian dual problem for this convex set-valued minimization problem using the approach of [21]. Denote by $L: L_m^\infty \times \mathbb{R}_+^m \times \mathbb{R}_+^m \setminus \{ 0 \} \to G(\mathbb{R}_m^m, \mathbb{R}_+^m)$ the Lagrangian and by $H: L_m^\infty \times \mathbb{R}_+^m \setminus \{ 0 \} \to G(\mathbb{R}_m^m, \mathbb{R}_+^m)$ the Lagrangian dual objective function. For $X \in L_m^\infty$, $\lambda \in \mathbb{R}_+^m$, $v \in \mathbb{R}_+^m \setminus \{ 0 \}$, we have $L(X, \lambda, v) = M^m$ if $v \notin \{ sw \mid s > 0 \}$ and

$$L(X, \lambda, v) = \mathbb{E}^Q[-X] + G(sw) + \left\{ z \in \mathbb{R}_m^m \mid sw^T z \geq \lambda^T \mathbb{E}[\ell(X)] + \inf_{x \in -x^0 + C} \lambda^T x \right\}$$

$$= \left\{ z \in \mathbb{R}_m^m \mid sw^T z \geq sw^T \mathbb{E}^Q[-X] + \lambda^T \mathbb{E}[\ell(X)] + \inf_{x \in -x^0 + C} \frac{1}{s} \lambda^T x \right\}.$$

whenever $v = sw$ for some $s > 0$. Observe $G(sw) = G(w)$ for all $s > 0$. Hence, for $\lambda \in \mathbb{R}_+^m$ and $s > 0$

$$H(\lambda, sw) = \left\{ z \in \mathbb{R}_m^m \mid w^T z \geq \inf_{X \in L_m^\infty} \left( w^T \mathbb{E}^Q[-X] + \frac{1}{s} \lambda^T \mathbb{E}[\ell(X)] \right) + \inf_{x \in -x^0 + C} \frac{1}{s} \lambda^T x \right\} = H(\frac{\lambda}{s}, w).$$

The dual problem now is

$$\sup \left\{ H \left( \frac{\lambda}{s}, w \right) \mid s > 0, \lambda \in \mathbb{R}_+^m \right\} = \sup \left\{ H(\lambda, w) \mid \lambda \in \mathbb{R}_+^m \right\}.$$

Since $w_i > 0$ for all $i \in \{ 1, \ldots, m \}$ by assumption we have

$$H(\lambda, w) = M(r, w)$$

with $r_i = \frac{\lambda}{w_i}$ for $i = 1, \ldots, m$ and

$$M(r, w) := \left\{ z \in \mathbb{R}_m^m \mid w^T z \geq \inf_{X \in L_m^\infty} \left( w^T \mathbb{E}^Q[-X] + w^T \operatorname{diag}(r) \mathbb{E}[\ell(X)] \right) + \inf_{x \in -x^0 + C} w^T \operatorname{diag}(r)x \right\}.$$

Note that

$$\inf_{X \in L_m^\infty} \left( w^T \mathbb{E}^Q[-X] + w^T \operatorname{diag}(r) \mathbb{E}[\ell(X)] \right) = \sum_{i=1}^m w_i \inf_{x_i \in \mathbb{R}} \left( w_i \mathbb{E} \left[ -\frac{dQ_i}{dp} X_i \right] + r_i \mathbb{E} [\ell_i(X_i)] \right)$$

$$= \sum_{i=1}^m w_i \mathbb{E} \left[ \inf_{x_i \in \mathbb{R}} \left( -\frac{dQ_i}{dp} x_i + r_i \ell_i(x_i) \right) \right]$$

$$= w^T I_{g,r}(Q \mid \mathbb{P}).$$

34
Therefore, the value of the dual problem equals the middle term in (4.9). Note that the set-valued version of Slater’s condition holds, that is, there exists $\bar{X} \in L^\infty_m$ such that $(E[\ell(\bar{X})] - x^0 + C) \cap -\mathbb{R}^m_+ \neq \emptyset$. This is immediate from the scalar version as in the proof of Proposition 3.13. Hence, the first equality in (4.9) holds by [21, Theorem 6.6]. Since $I_{g,r}(Q \mid P) \notin \mathbb{R}^m$ for $r \notin 1/\text{dom } g$, we also have the second equality in (4.9). 

**Proof of Proposition 4.15.** Let $i \in \{1, \ldots, m\}$. From Remark 3.7, recall that for every $s \in \mathbb{R}$,

$$r_i \ell_i(s) = \sup_{y \in \mathbb{R}} \left( sy - r_i g_i \left( \frac{y}{r_i} \right) \right) \geq s - r_i g_i \left( \frac{1}{r_i} \right).$$

Hence, given $X \in L^\infty_m$, we have

$$\delta_{g,r}(X) = \inf_{z_i \in \mathbb{R}} (z_i + r_i E[\ell_i(-X_i - z_i)]) - r_i x_i^0 \geq -E[X] - r_i g_i \left( \frac{1}{r_i} \right) - r_i x_i^0. \quad (6.11)$$

So we have

$$\inf_{z \in D_{g,r}^\infty(X)} \bar{w}^T z = \inf_{z \in \mathcal{K}} \inf_{z \in D_{g,r}(X - Z)} \bar{w}^T z = \inf_{Y \in \mathcal{K}} \inf_{z \in \bar{w}^T z \in \delta_{g,r}(X - Y) + \text{diag}(r)C} = \inf_{Y \in \mathcal{K}} \bar{w}^T \delta_{g,r}(X - Y) + \inf_{x \in C} \bar{w}^T \text{diag}(r) x$$

$$\geq -\bar{w}^T E[X] + \inf_{Y \in \mathcal{K}} \bar{w}^T E[Y] - \sum_{i=1}^m \bar{w}_i r_i g_i \left( \frac{1}{r_i} \right) + \inf_{x \in -x^0 + C} \bar{w}^T \text{diag}(r) x$$

$$= -\bar{w}^T E[X] + \sum_{t=0}^T \left( \inf_{Z_t \in L^\infty_m(F_t, F_t)} \bar{w}^T E[Z_t] - \sum_{i=1}^m \bar{w}_i r_i g_i \left( \frac{1}{r_i} \right) \right) + \inf_{x \in -x^0 + C} \bar{w}^T \text{diag}(r) x,$$

where the inequality in the middle follows from (6.11). By the same argument as in the proof of Corollary 2.19, the hypotheses guarantee the finiteness of the quantity in the last line. Hence,

$$D_{g,r}^\infty(X) \subseteq \left\{ \bar{z} \in \mathbb{R}^m \mid \bar{w}^T \bar{z} \geq \inf_{z \in D_{g,r}^\infty(X)} \bar{w}^T z \right\} \subseteq -E[X] + H, \quad X \in L^\infty_m,$$

where

$$H := \left\{ z \in \mathbb{R}^m \mid \bar{w}^T z \geq \inf_{Z \in \mathcal{K}} \bar{w}^T E[Z] - \sum_{i=1}^m \bar{w}_i r_i g_i \left( \frac{1}{r_i} \right) + \inf_{x \in -x^0 + C} \bar{w}^T \text{diag}(r) x \right\} \neq \mathbb{R}^m.$$

Note that $L^\infty_m \ni X \mapsto -E[X] + H \in G(\mathbb{R}^m, \mathbb{R}^m_+)$ is a convex weak*-closed function. Hence, by Remark 2.16, for every $X \in L^\infty_m$,

$$\text{cl } D_{g,r}^\infty(X) \subseteq -E[X] + H \neq \mathbb{R}^m.$$

**Acknowledgments**

Thanks to Zachary Feinstein and Samuel Drapeau for useful comments on the set-valued entropic risk measure and optimized certainty equivalents. Birgit Rudloff’s research was supported by NSF award DMS-1007938.
References

[1] F. Astic and N. Touzi, No arbitrage conditions and liquidity, Journal of Mathematical Economics 43 (2007), 692–708.

[2] I. Ben Tahar and E. Lepinette, Vector-valued coherent risk measure processes, International Journal of Theoretical and Applied Finance 17 (2014), 1450011.

[3] I. Ben Tahar, Tail conditional expectation for vector-valued risks, SFB 649 Discussion Papers SFB649DP2006-029, Humboldt University, Collaborative Research Center 649, 2006.

[4] A. Ben-Tal and M. Teboulle, Expected utility, penalty functions and duality in stochastic nonlinear programming, Management Science 32 (1986), 1445–1466.

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

[6] J. M. Borwein and A. S. Lewis, Partially finite convex programming I, Mathematical Programming 57 (1992), 15–48.

[7] Ch. Burgert and L. Rüschendorf, Consistent risk measures for portfolio vectors, Insurance: Mathematics and Economics 38 (2006), 289–297.

[8] L. Campi and M. P. Owen, Multivariate utility maximization with proportional transaction costs, Finance and Stochastics 15 (2011), 461–499.

[9] I. Cascos and I. Molchanov, Multivariate risk measures: a constructive approach based on selections, ArXiv e-prints, February 2013.

[10] U. Çetin, R. A. Jarrow and P. Protter, Liquidity risk and arbitrage pricing theory, Finance and Stochastics 8 (2004), 311–341.

[11] U. Çetin and L. C. G. Rogers, Modelling liquidity effects in discrete time, Mathematical Finance 17 (2007), 15–29.

[12] Z. Feinstein and B. Rudloff, Time consistency of dynamic risk measures in markets with transaction costs, Quantitative Finance 13 (2013), 1473–1489.

[13] Z. Feinstein and B. Rudloff, Multi-portfolio time consistency for set-valued convex and coherent risk measures, Finance and Stochastics (forthcoming).

[14] Z. Feinstein and B. Rudloff, A comparison of techniques for dynamic risk measures with transaction costs, Set Optimization and Applications in Finance, Springer PROMS Series (forthcoming).

[15] H. Föllmer and A. Schied, Convex measures of risk and trading constraints, Finance and Stochastics 6 (2002), 429–447.

[16] H. Föllmer and A. Schied, Stochastic finance: an introduction in discrete time, Walter de Gruyter Berlin New York, third revised and extended edition 2011.

[17] A. H. Hamel, Variational principles on metric and uniform spaces, Habilitationsschrift, Martin-Luther Universität Halle-Wittenberg, http://sundoc.bibliothek.uni-halle.de/habil-online/05/05H167/habil.pdf (2005)
[18] A. H. Hamel, A duality theory for set-valued functions I: Fenchel conjugation theory, Set-Valued and Variational Analysis 17 (2009), 153–182.

[19] A. H. Hamel and F. Heyde, Duality for set-valued measures of risk, SIAM Journal on Financial Mathematics 1 (2010), 66–95.

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

[21] A. H. Hamel and A. Löhne, Lagrange duality in set optimization, Journal of Optimization Theory and Applications 161 (2014), 368–397.

[22] A. H. Hamel, A. Löhne and B. Rudloff, A Benson type algorithm for linear vector optimization and applications, Journal of Global Optimization, DOI: 10.1007/s10898-013-0098-2 (2013).

[23] A. H. Hamel, B. Rudloff and M. Yankova, Set-valued average value at risk and its computation, Mathematics and Financial Economics 7 (2013), 229–246.

[24] F. Heyde and A. Löhne, Solution concepts in vector optimization: a fresh look at an old story, Optimization 60 (2011), 1421-1440.

[25] E. Jouini and H. Kallal, Martingales and arbitrage in securities markets with transaction costs, Journal of Economic Theory 66 (1995), 178–197.

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

[27] A. Löhne and B. Rudloff, An algorithm for calculating the set of superhedging portfolios in markets with transaction costs, International Journal of Theoretical and Applied Finance 17 (2014), 1450012.

[28] Y. M. Kabanov, Hedging and liquidation under transaction costs in currency markets, Finance and Stochastics 3 (1999), 237–248.

[29] Y. M. Kabanov, M. Safarian, Markets with transaction costs, Springer-Verlag Berlin, 2009.

[30] T. Pennanen and I. Penner, Hedging of claims with physical delivery under convex transaction costs, SIAM Journal on Financial Mathematics 1 (2010), 158–178.

[31] R. T. Rockafellar, Convex analysis, Princeton University Press, 1970.

[32] R. T. Rockafellar, R. J-B Wets, Variational analysis, Grundlehren der mathematischen Wissenschaften 317, Springer-Verlag Berlin Heidelberg, 1998, corrected third printing 2009.

[33] L. C. G. Rogers and S. Singh, The cost of illiquidity and its effects on hedging, Mathematical Finance 20 (2010), 597–615.

[34] A. Schied, Optimal investments for risk- and ambiguity-averse preferences: a duality approach, Finance and Stochastics 11 (2007), 107–129.

[35] S. Weber, W. Anderson, A.-M. Hamm, T. Knispel, M. Liese and T. Salfeld, Liquidity-adjusted risk measures, Mathematics and Financial Economics 7 (2013), 69–91.

[36] C. Zalinescu, Convex analysis in general vector spaces, World Scientific, 2002.