Portfolio optimization with two coherent risk measures

Tahsin Deniz Aktürk[1] · **Çağın Ararat**[2]

Abstract

We provide analytical results for a static portfolio optimization problem with two coherent risk measures. The use of two risk measures is motivated by joint decision-making for portfolio selection where the risk perception of the portfolio manager is of primary concern, hence, it appears in the objective function, and the risk perception of an external authority needs to be taken into account as well, which appears in the form of a risk constraint. The problem covers the risk minimization problem with an expected return constraint and the expected return maximization problem with a risk constraint, as special cases. For the general case of an arbitrary joint distribution for the asset returns, under certain conditions, we characterize the optimal portfolio as the optimal Lagrange multiplier associated to an equality-constrained dual problem. Then, we consider the special case of Gaussian returns for which it is possible to identify all cases where an optimal solution exists and to give an explicit formula for the optimal portfolio whenever it exists.

Keywords Portfolio optimization · Coherent risk measure · Mean-risk problem · Markowitz problem

Mathematics Subject Classification 90C11 · 90C20 · 90C90 · 91B30 · 91G10

1 Introduction

The mean-variance portfolio selection problem introduced in the seminal work Markowitz [12] is one of the most well-studied optimization problems. In the basic static version of the problem, one considers multiple correlated assets with known expected returns and covariances, and looks for an allocation of these assets. Considering the trade-off between the linear expected return and the quadratic variance, the problem can be formulated as a biobjective optimization problem whose efficient solutions form the so-called *efficient frontier* on the

✉ Çağın Ararat
cararat@bilkent.edu.tr

Tahsin Deniz Aktürk
akturk@chicagobooth.edu

1 Booth School of Business, The University of Chicago, Chicago, IL, USA

2 Department of Industrial Engineering, Bilkent University, Ankara, Turkey

mean-variance (or mean-standard deviation) plot of all portolios. Merton [13] provides an analytical derivation of the efficient frontier for the general case of $n \geq 2$ assets.

The biobjective mean-variance problem can also be studied in terms of a parametric family of scalar (single-objective) problems. Among the popular *scalarizations* are the ones where one minimizes variance over the set of all portfolios at a given expected return level, which is used as the parameter of the scalar problem. Analogously, one can impose a constraint on the variance using an upper bound parameter and maximizes expected return. Quite naturally, both approaches can be used to verify the analytical results in Merton [13].

Started with Artzner et al. [1], the theory of *coherent risk measures* provides an axiomatic approach to come up with functionals possessing desirable properties for risk measurement purposes. Such properties include monotonicity and translativity (see Sect. 2.2 for precise definitions), which are not satisfied by variance or standard deviation. Canonical examples of coherent risk measures include (negative) expected value and average value-at-risk [16]. In addition, value-at-risk is also known to be a coherent risk measure when considered on a space of Gaussian random variables. Each of these three risk measures is also law-invariant in the sense that two random variables with the same distribution have the same risk.

With a coherent risk measure, one can formulate the corresponding mean-risk portfolio optimization problem by replacing variance with the risk measure. For average value-at-risk, this problem is considered in Rockafellar, Uryasev [15] in the form of risk minimization subject to an expected return constraint. When jointly Gaussian asset returns are assumed, the risk objective function reduces to the sum of a linear function and the square-root of a quadratic form. The special structure of this case is exploited in Landsman [9], Owadally [14], where analytical results are obtained. A more general objective function in which a differentiable function of variance is added to a linear function is considered in Landsman, Makov [10], which also provides closed-form solutions. It should be noted that all of these works assume linear constraints.

In this paper, we consider a "risk-risk problem" where a coherent risk measure is minimized subject to a constraint on a second coherent risk measure. The purpose of using two risk measures is to take into account two risk perceptions when choosing a portfolio. The principle risk measure to be minimized may reflect the risk perception of the portfolio manager while the secondary risk measure in the constraint reflects that of an external authority. Similar to the mean-variance case, the single-objective problem we consider can be seen as a scalarization of a biobjective problem where the objectives are the risk measures of the two bodies who are supposed to choose a portfolio jointly. One advantage of our framework is that it includes both versions of the mean-risk problem as special cases: the one that minimizes risk as well as the one that maximizes expected return.

We first study the risk-risk problem in a general setting where the underlying asset returns are in some L^p space with $p \in [1, +\infty]$ and they have an arbitrary joint distribution with possible correlations. Assuming that the two risk measures are continuous from below so that the suprema in the dual representations are attained at some dual probability measures, we derive a simple dual problem with a linear objective and a linear equality constraint in addition to domain constraints for the dual variables. As the main result of Sect. 3, under certain constraint qualifications, we show that an optimal solution for the portfolio optimization problem can be obtained as the Lagrange multiplier of the equality constraint of the dual problem at optimality.

At the technical level, the risk-risk problem is a finite-dimensional convex optimization problem. We use the standard Slater's condition (see Assumption 3.1 below) as a constraint qualification to guarantee the existence of optimal Lagrange multipliers for the constraints. Then, we work on the Lagrange dual problem and refine it further by introducing additional

dual variables through the dual representations of the two coherent risk measures. This refinement yields a finalized equality-constrained dual problem which is infinite-dimensional due to the dual densities related to the risk measures. To guarantee the existence of an optimal Lagrange multiplier attached to the equality constraint, we need a second constraint qualification. However, the usual Slater's condition with interiority assumptions for the domain constraints is not suitable for this setting due to the fact that many sets in L^q (with $\frac{1}{p} + \frac{1}{q} = 1$), even the positive cone L^q_+, may fail to have an empty interior. For this reason, we use the notion of quasi relative interior and the related mild constraint qualification in Borwein, Lewis [3] (see Assumption 3.2 below), which still guarantees the existence of an optimal Lagrange multiplier.

In Sect. 4, we study the special case where the asset returns are jointly Gaussian and the risk measures are law-invariant. By exploiting the properties of Gaussian distribution and those of the risk measures, the portfolio optimization problem reduces to a problem only with the square-root of a quadratic function and a linear function in the objective and in the constraints. In particular, unlike the above-mentioned works for the Gaussian case, we have a nonlinear constraint that imposes an upper bound on the sum of the square-root of a quadratic form and a linear function.

In the Gaussian case, we observe that the problem can be solved with the help of the hyperbola appearing in the analysis of the mean-variance problem as in Merton [13]. Indeed, as an associated problem, we consider the minimization of a linear function subject to a linear constraint *over this hyperbola*, which is simply a two-dimensional problem and has a clear geometric interpretation. Using this problem, we provide a complete analysis of the main problem. In particular, we identify all cases in which an optimal solution exists, a unique optimal solution exists, the infimum is finite but not attained, and the problem is unbounded (Sect. 4). We provide closed-form expressions for an optimal portfolio, whenever it exists.

2 Mathematical setup

2.1 Portfolios

We are concerned with various risk-averse versions of the portfolio selection problem on a domain of finitely many risky assets with possibly correlated returns in a one-period market model. To introduce the setup of the problem, we let $n \geq 2$ be an integer denoting the number of assets in the market and write $\mathcal{N} = \{1, \ldots, n\}$ for the set of these assets. As usual, we denote by $\mathbb{R}^n$ the n-dimensional real Euclidean space and by $\mathbb{R}^n_+$ the cone of all vectors $x = (x_1, \ldots, x_n)^\mathsf{T} \in \mathbb{R}^n$ with $x_i \geq 0$ for each $i \in \{1, \ldots, n\}$. For $x, z \in \mathbb{R}^n$, we define their scalar product by $x^\mathsf{T} z := \sum_{i=1}^n x_i z_i$. Let us fix a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. We denote by L^0_n the space of all n-dimensional random vectors distinguished up to almost sure equality. For each $p \in [1, +\infty)$, we define $L^p_n = \left\{ X \in L^0_n \mid \mathbb{E}\left[|X|^p \right] \right\} < +\infty$, and for $p = +\infty$, we define $L^\infty_n = \left\{ X \in L^0_n \mid \exists c > 0 \colon \mathbb{P}\{|X| \leq c\} = 1 \right\}$, where $|\cdot|$ is an arbitrary norm on $\mathbb{R}^n$. We write $L^p = L^p_1$ for each $p \in \{0\} \cup [1, +\infty]$.

Let us fix $p \in [1, +\infty]$ and consider a possibly correlated random vector $X = (X_1, \ldots, X_n)^\mathsf{T} \in L^p_n$. For each $i \in \mathcal{N}$, the random variable X_i denotes the return of the ith asset for a fixed period as a multiple of the initial price of that asset. In our context, a portfolio is defined as a vector in $\mathbb{R}^n$ each of whose components denotes the weight of the corresponding asset in the portfolio based on the asset prices at the beginning of the period.

Hence, the set of all portfolios is the set

$$\mathcal{W} := \left\{ w \in \mathbb{R}^n \mid \sum_{i=1}^n w_i = 1 \right\} = \left\{ w \in \mathbb{R}^n \mid \mathbf{1}^\mathsf{T} w = 1 \right\}. \tag{2.1}$$

When shortselling is not allowed, we will restrict ourselves to the portfolios in the subset

$$\mathcal{W}_+ := \mathcal{W} \cap \mathbb{R}_+^n, \tag{2.2}$$

which is the $(n-1)$-dimensional unit simplex. Note that, for a portfolio $w \in \mathcal{W}$, we have $w^\mathsf{T} X \in L^p$, which denotes the return of the porfolio.

2.2 Risk measures

We provide a quick review of the theory of risk measures on L^p with $p \in [1, +\infty]$. The reader is referred to Kaina, Rüschendorf [8] (for $p \in [1, +\infty)$) and to Föllmer, Schied [5] (for $p = +\infty$) for a detailed account of the convex-analytic properties of risk measures on L^p.

For $Y_1, Y_2 \in L^p$, we write $Y_1 \leq Y_2$ if $Y_1(\omega) \leq Y_2(\omega)$ for $\mathbb{P}$-almost every $\omega \in \Omega$ and $Y_1 \sim Y_2$ if Y_1 and Y_2 are identically distributed. A functional $\rho \colon L^p \to \mathbb{R} \cup \{+\infty\}$ is said to be a coherent risk measure if it satisfies the following properties.

  (i) Monotonicity: $Y_1 \leq Y_2$ implies $\rho(Y_1) \geq \rho(Y_2)$ for every $Y_1, Y_2 \in L^p$.
 (ii) Translativity: It holds $\rho(Y + y) = \rho(Y) - y$ for every $Y \in L^p$ and $y \in \mathbb{R}$.
(iii) Subadditivity: It holds $\rho(Y_1 + Y_2) \leq \rho(Y_1) + \rho(Y_2)$ for every $Y_1, Y_2 \in L^p$.
 (iv) Positive homogeneity: It holds $\rho(\lambda Y) = \lambda \rho(Y)$ for every $Y \in L^p$ and $\lambda \geq 0$.

Clearly, positive homogeneity implies the following property.

  (v) Normalization: It holds $\rho(0) = 0$.

Moreover, it is easy to check that, under positive homogeneity, subadditivity is equivalent to the following property.

 (vi) Convexity: It holds $\rho(\lambda Y_1 + (1 - \lambda)Y_2) \leq \lambda \rho(Y_1) + (1 - \lambda)\rho(Y_2)$ for every $Y_1, Y_2 \in \mathcal{Y}$ and $\lambda \in [0, 1]$.

Let ρ be a coherent risk measure. In Sect. 3, we need the following additional property:

(vii) Finiteness: $\rho(Y) < +\infty$ for every $Y \in L^p$.

Let $\mathcal{M}_1(\mathbb{P})$ denote the set of all probability measures on $(\Omega, \mathcal{F})$ that are absolutely continuous with respect to $\mathbb{P}$. Let $q \in [1, +\infty]$ such that $\frac{1}{p} + \frac{1}{q} = 1$ and define

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

Note that $\mathcal{M}_1^q(\mathbb{P}) = \mathcal{M}_1(\mathbb{P})$.

If $p \in [1, +\infty)$, then finiteness is equivalent to having a dual representation of the form

$$\rho(Y) = \max_{\mathbb{Q} \in \mathcal{Q}} \mathbb{E}^{\mathbb{Q}}[-Y], \quad Y \in L^p, \tag{2.3}$$

for some convex set $\mathcal{Q} \subseteq \mathcal{M}_1^q(\mathbb{P})$ of probability measures such that the corresponding set

$$\mathcal{D}(\mathcal{Q}) := \left\{ \frac{d\mathbb{Q}}{d\mathbb{P}} \mid \mathbb{Q} \in \mathcal{Q} \right\}$$

of Radon-Nikodym derivatives is $\sigma(L^q, L^p)$-compact; see ([8], Theorem 2.11). Moreover, finiteness also implies that ρ satisfies the following property [8, Theorem 3.1]:

(viii) Continuity from below: If $Y, Y_1, Y_2, \ldots \in L^p$ such that $Y_1 \leq Y_2 \leq \ldots$ and $\lim_{k\to\infty} Y_k = Y$ $\mathbb{P}$-almost surely, then $\lim_{k\to\infty} \rho(Y_k) = \rho(Y)$.

If $p = +\infty$, then monotonicity and translativity ensure that $\rho(Y) < +\infty$ for every $Y \in L^\infty$ without an additional assumption. Nevertheless, if one further assumes continuity from below, then a representation of the form (2.3) holds for some convex set $\mathcal{Q} \subseteq \mathcal{M}_1(\mathbb{P})$ such that $\mathcal{D}(\mathcal{Q})$ is $\sigma(L^1, L^\infty)$-compact; see ([8], Theorem 3.6).

Finally, we formulate the following additional property that will be needed in Sect. 4.

(ix) Law-invariance: $Y_1 \sim Y_2$ implies $\rho(Y_1) = \rho(Y_2)$ for every $Y_1, Y_2 \in L^p$.

Let us recall three commonly used risk measures: negative expected value, value-at-risk and average value-at-risk.

Example 2.1 (Negative expected value) Let $p = 1$ and take $\rho(Y) = \mathbb{E}[-Y]$ for every $Y \in L^1$. It is easy to check that ρ satisfies properties (i)–(viii) above. In the dual representation (2.3), we simply have $\mathcal{Q} = \{\mathbb{P}\}$ so that $\mathcal{D}(\mathcal{Q}) = \{1\} \subseteq L^\infty$.

Example 2.2 (Value-at-risk) Let $p = 1$. Let $\theta \in (0, 1)$ be a probability level. The *value-at-risk* at level θ for a random variable $Y \in L^1$ is defined as

$$V@R_\theta(Y) := \sup \{r \in \mathbb{R} \mid \mathbb{P}\{Y + r \leq 0\} > \theta\}.$$

It is well-known that $V@R_\theta$ is a law-invariant positively homogeneous risk measure which fails to be convex. However, if X is a Gaussian random vector and $\mathcal{Y}$ is the Gaussian subspace of L^2 spanned by $X_1, \ldots, X_n$ and the constant random variable 1, it holds (see Proposition 4.1 below)

$$V@R_\theta(Y) = \Phi^{-1}(1-\theta)\sqrt{\operatorname{Var}(Y)} - \mathbb{E}[Y]$$

for every $Y \in \mathcal{Y}$, where Φ is the cumulative distribution function of a standard Gaussian random variable Z. In particular, $V@R_\theta(Z) = \Phi^{-1}(1-\theta)$. As $Y \mapsto \sqrt{\operatorname{Var}(Y)}$ is a convex function on L^2, $V@R_\theta$ is a law-invariant coherent risk measure on $\mathcal{Y}$.

Example 2.3 (Average value-at-risk) Let $\theta \in (0, 1)$ be a probability level. The *average value-at-risk* at level θ for $Y \in L^1$ is defined as

$$AV@R_\theta(Y) := \frac{1}{\theta} \int_0^\theta V@R_u(Y) du.$$

It is well-known that $AV@R_\theta$ is a law-invariant coherent risk measure on L^1. In the dual representation in (2.3), we may take $\mathcal{Q} = \{\mathbb{Q} \in \mathcal{M}_1(\mathbb{P}) \mid \mathbb{P}\{\frac{d\mathbb{Q}}{d\mathbb{P}} \leq \frac{1}{\theta}\} = 1\}$ so that

$$\mathcal{D}(\mathcal{Q}) = \left\{V \in L^\infty \mid \mathbb{P}\left\{0 \leq V \leq \frac{1}{\theta}\right\} = 1\right\}.$$

On the other hand, for every $Y \in \mathcal{Y}$, where $\mathcal{Y}$ is the Gaussian subspace $\mathcal{Y}$ of L^2 in Example 2.2, we have

$$AV@R_\theta(Y) = AV@R_\theta(Z)\sqrt{\operatorname{Var}(Y)} - \mathbb{E}[Y],$$

where Z is a standard Gaussian random variable with

$$AV@R_\theta(Z) = \int_0^\theta \Phi^{-1}(1-u) du.$$

2.3 The portfolio optimization problem

In this section, we formulate the continuous portfolio optimization problem of our interest.

To model risk-aversion, let $\rho_1, \rho_2\colon L^p \to \mathbb{R}$ be two arbitrary coherent risk measures. The aim of the portfolio manager is to choose a portfolio $w \in \mathcal{W}$ that minimizes the type 1 risk $\rho_1(w^\mathsf{T} X)$ while controlling the type 2 risk $\rho_2(w^\mathsf{T} X)$ within a fixed threshold level $r \in \mathbb{R}$, that is, while satisfying

$$\rho_2(w^\mathsf{T} X) \le r,$$

which we refer to as the *risk constraint*. The use of two risk measures makes sense in cases where the portfolio manager has the right to choose the portfolio using ρ_1 as the suitable risk measure for her risk perception but an external regulatory authority with a different risk perception reflected by ρ_2 imposes the risk constraint as an obligation for the portfolio manager. It also makes sense when the portfolio manager wishes to work with two risk measures, the principle one (ρ_1) having a higher seniority than the other (ρ_2). In particular, this framework covers as special cases the problem of maximizing expected return subject to a risk constraint if we take $\rho_1(Y) = \mathbb{E}[-Y]$ for each $Y \in L^p$, as well as the problem of minimizing (the type 1) risk while maintaining a high-enough expected return if we take $\rho_2(Y) = \mathbb{E}[-Y]$ for each $Y \in L^p$.

With these risk considerations, we formulate the continuous portfolio optimization problem with shortselling as

$$\begin{aligned}
\text{minimize} \quad & \rho_1(w^\mathsf{T} X) && (\mathscr{P}(r)) \\
\text{subject to} \quad & \rho_2(w^\mathsf{T} X) \le r \\
& w \in \mathcal{W}.
\end{aligned}$$

In this paper, we provide analytical results for $(\mathscr{P}(r))$ in two cases:

- General case: For a random return vector X with an arbitrary distribution and assuming that ρ_1, ρ_2 are continuous from below, we characterize an optimal solution for $(\mathscr{P}(r))$ as a Lagrange multiplier of an associated dual problem in Sect. 3.
- Gaussian case: For a Gaussian random return vector X and assuming that ρ_1, ρ_2 are law-invariant, we provide a complete analysis of the problem with explicit formulae for an optimal solution and identify the cases where it exists and where it is unique in Sect. 4.

3 The portfolio optimization problem under an arbitrary joint distribution

In this section, we assume that $X \in L_n^p$ for a fixed $p \in [1, +\infty]$ and ρ_1, ρ_2 are arbitrary coherent risk measures on L^p that are finite and continuous from below. In particular, ρ_1, ρ_2 are continuous on L^p; see ([8], Corollary 2.3). Recalling (2.3), ρ_1, ρ_2 admit dual representations of the form

$$\rho_1(Y) = \max_{\mathbb{Q}_1 \in \mathcal{Q}_1} \mathbb{E}^{\mathbb{Q}_1}[-Y], \quad \rho_2(Y) = \max_{\mathbb{Q}_2 \in \mathcal{Q}_2} \mathbb{E}^{\mathbb{Q}_2}[-Y],$$

for each $Y \in L^p$, where $\mathcal{Q}_1, \mathcal{Q}_2$ are convex subsets of $\mathcal{M}_1^q(\mathbb{P})$ such that the corresponding density sets $\mathcal{D}(\mathcal{Q}_1), \mathcal{D}(\mathcal{Q}_2)$ are convex $\sigma(L^q, L^p)$-compact subsets of L^q. For each $j \in$

$\{1, 2\}$, let us define the continuous convex function $g_j \colon \mathbb{R}^n \to \mathbb{R}$ by

$$g_j(w) = \rho_j(w^\mathsf{T} X) = \max_{V \in \mathcal{D}(\mathcal{Q}_j)} \mathbb{E}\left[-V w^\mathsf{T} X\right]$$

for each $w \in \mathbb{R}^n$.

As a preparation for the statement and the proof of the main result, we recall a few notions and facts from convex analysis. Let $\mathcal{X}$ be an Hausdorff locally convex topological linear space with topological dual $\mathcal{Y}$ and bilinear duality mapping $\langle \cdot, \cdot \rangle \colon \mathcal{Y} \times \mathcal{X} \to \mathbb{R}$. For the purposes of this paper, we are interested in three special cases:

(i) $\mathcal{X} = \mathbb{R}^n$ with the usual topology, which yields $\mathcal{Y} = \mathbb{R}^n$ together with $\langle y, x \rangle = y^\mathsf{T} x$ for every $x \in \mathbb{R}^n$, $y \in \mathbb{R}^n$.
(ii) $\mathcal{X} = L^q$ with $q \in [1, +\infty)$ with the weak topology $\sigma(L^q, L^p)$, which yields $\mathcal{Y} = L^p$ together with $\langle Y, U \rangle = \mathbb{E}[UY]$ for every $U \in L^q$, $Y \in L^p$.
(iii) $\mathcal{X} = L^\infty$ with the weak topology $\sigma(L^\infty, L^1)$, which yields $\mathcal{Y} = L^1$ together with $\langle Y, U \rangle = \mathbb{E}[UY]$ for every $U \in L^\infty$, $Y \in L^1$.

Let $A \subseteq \mathcal{X}$ be a set. $\mathrm{cone}(A) := \{\lambda x \mid \lambda \geq 0, x \in A\}$ is called the *conic hull* of A. If A is convex, then $\mathrm{cone}(A)$ is a convex cone. For $x \in A$, the convex cone

$$\mathcal{N}_A(x) := \Big\{y \in \mathcal{Y} \mid \forall x' \in A \colon \langle y, x \rangle \geq \langle y, x' \rangle\Big\}$$

is called the *normal cone* of A at x. The function $I_A \colon \mathcal{X} \to \mathbb{R} \cup \{+\infty\}$ defined by $I_A(x) = 0$ for $x \in A$ and $I_A(x) = +\infty$ for $x \in \mathcal{X} \setminus A$ is called the *indicator function* of A. Note that A is convex if and only if I_A is convex. Let $g \colon \mathcal{X} \to \mathbb{R} \cup \{+\infty\}$ be a function. For a point $x \in \mathcal{X}$, the set $\partial g(x) := \Big\{y \in \mathcal{Y} \mid \forall x' \in \mathcal{X} \colon g(x') \geq g(x) + \langle y, x' - x \rangle\Big\}$ is called the *subdifferential* of g at x. If A is a nonempty convex set, then it is well-known that ([18], Section 2.4) $\partial I_A(x) = \mathcal{N}_A(x)$ for every $x \in A$, and $\partial I_A(x) = \emptyset$ for every $x \in \mathcal{X} \setminus A$. The function $g^* \colon \mathcal{Y} \to \mathbb{R} \cup \{\pm\infty\}$ defined by $g^*(y) := \sup_{x \in \mathcal{X}} (\langle y, x \rangle - g(x))$ for each $y \in \mathcal{Y}$ is called the *conjugate function* of g. Note that $y \in \partial g(x)$ holds if and only if $x \in \partial g^*(y)$ holds for every $x \in \mathcal{X}$, $y \in \mathcal{Y}$ such that g is lower semicontinuous at x.

To formulate a second constraint qualification, we also need the following. For $A \subseteq \mathcal{X}$, the set

$$\mathrm{qri}(A) := \{x \in A \mid \mathcal{N}_A(x) \text{ is a subspace of } \mathcal{Y}\}$$

is called the *quasi relative interior* of A ([3], Proposition 2.8). When $\mathcal{X} = \mathbb{R}^n$, $\mathrm{qri}(A)$ coincides with the relative interior of A. In this case, $\mathrm{qri}(A) \neq \emptyset$ whenever A is nonempty, closed and convex. When $\mathcal{X} = L^q$ ($q \in [1, +\infty)$) is considered with the topology $\sigma(L^q, L^p)$ and A is nonempty, closed and convex, one has $\mathrm{qri}(A) \neq \emptyset$ again thanks to [3, Theorem 2.19]. In particular, if $A = L^q_+ := \{U \in L^q \mid \mathbb{P}\{U \geq 0\} = 1\}$, then $\mathrm{qri}(A) = \{U \in L^q \mid \mathbb{P}\{U > 0\} = 1\}$ by [3, Example 3.11]. (For $q < +\infty$, considering the strong and weak topologies on L^q yield the same quasi relative interior for a convex set by [3, Proposition 2.6].

To be able to study a dual problem with zero duality gap, we work under the following constraint qualification for $(\mathscr{P}(r))$.

Assumption 3.1 (Slater's condition) There exists $w \in \mathcal{W}$ such that $\rho_2(w^\mathsf{T} X) < r$.

The main theorem of this section is Theorem 3.3 below. In its proof, by constructing a Lagrange dual problem for $(\mathscr{P}(r))$ and exploiting the dual representations of ρ_1, ρ_2, we

obtain the following finalized dual problem $(\mathscr{D}(r))$ with an equality constraint.

$$\text{maximize } -r\mathbb{E}[M] - \lambda \qquad\qquad (\mathscr{D}(r))$$
$$\text{subject to } \mathbb{E}[UX] + \mathbb{E}[MX] - \lambda\mathbf{1} = 0$$
$$U \in \mathcal{D}(\mathcal{Q}_1), \; M \in \text{cone}(\mathcal{D}(\mathcal{Q}_2)), \; \lambda \in \mathbb{R}.$$

The theorem states that an optimal solution for $(\mathscr{P}(r))$ can be calculated as the Lagrange multiplier of the equality constraint of $(\mathscr{D}(r))$ at optimality.

In addition to Assumption 3.1 for $(\mathscr{P}(r))$, we use in Theorem 3.3 the following constraint qualification for $(\mathscr{D}(r))$ based on quasi relative interior.

Assumption 3.2 [3, Corollary 4.8] There exist $U \in \text{qri}(\mathcal{D}(\mathcal{Q}_1))$, $M \in \text{qri}(\text{cone}(\mathcal{D}(\mathcal{Q}_2)))$, $\lambda \in \mathbb{R}$ such that

$$\mathbb{E}[UX] + \mathbb{E}[MX] - \lambda\mathbf{1} = 0.$$

Note that Assumption 3.2 simply states that one can find $U \in \text{qri}(\mathcal{D}(\mathcal{Q}_1))$ and $M \in \text{qri}(\text{cone}(\mathcal{D}(\mathcal{Q}_2)))$ such that $\mathbb{E}[UX] + \mathbb{E}[MX]$ is a constant vector in $\mathbb{R}^n$. The comparison of this assumption with the usual interior-based constraint qualifications is discussed in Remark 3.8 after the examples. This remark is followed by Remark 3.9, where we comment on the usefulness of Theorem 3.3 for computations.

Theorem 3.3 *Under Assumptions 3.1 and 3.2, suppose that there exists an optimal solution $(U^*, M^*, \lambda^*) \in L^q \times L^q \times \mathbb{R}$ of $\mathscr{D}(r)$. Then, there exists an optimal Lagrange multiplier $w^* \in \mathbb{R}^n$ associated to the equality constraint of $\mathscr{D}(r)$. Moreover, every $w^* \in \mathbb{R}^n$ that is the Lagrange multiplier of the equality constraint of $\mathscr{D}(r)$ at optimality is an optimal solution for $(\mathscr{P}(r))$.*

We use the following lemma for the proof of Theorem 3.3, which should be known. For completeness, we present its short proof.

Lemma 3.4 *Let $w \in \mathbb{R}^n$, $j \in \{1, 2\}$, and define the attainment set*

$$\mathscr{V}_j(w) := \text{argmax}_{V \in \mathcal{D}(\mathcal{Q}_j)} \mathbb{E}\left[-VX^\mathsf{T}w\right]. \qquad (3.1)$$

Then, one has

$$\partial g_j(w) = \left\{\mathbb{E}[-VX] \mid V \in \mathscr{V}_j(w)\right\}.$$

Proof Since ρ_j is continuous from below, $\mathscr{V}_j(w) \neq \emptyset$. Note that the linear operator $A \colon \mathbb{R}^n \to L^p$ defined by $Aw' := X^\mathsf{T}w'$ for $w' \in \mathbb{R}^n$ has the adjoint operator $A^* \colon L^q \to \mathbb{R}^n$ given by $A^*V = \mathbb{E}[VX]$ for $V \in L^q$. (We consider the $\sigma(L^\infty, L^1)$ topology on L^∞ when $p = +\infty$ so that the dual space of L^∞ is L^1.) Let $w \in \mathbb{R}^n$. Since we have $g_j = \rho_j \circ A$ and ρ_j is continuous on L^p, by subdifferential calculus rules, e.g., by [18, Theorem 2.8.3(iii)],

$$\partial g_j(w) = \left\{A^*V \mid V \in \partial\rho_j(Aw)\right\} = \left\{\mathbb{E}[VX] \mid V \in \partial\rho_j(X^\mathsf{T}w)\right\}.$$

On the other hand, since ρ_j is continuous, for each $Y \in L^p$ and $V \in L^q$, we have $V \in \partial\rho_j(Y)$ if and only if $Y \in \partial\rho_j^*(V)$. On the other hand, for each $V \in L^q$,

$$\rho_j^*(V) = \sup_{Y' \in L^p}\left(\mathbb{E}[VY'] - \rho_j(Y')\right) = \sup_{Y' \in L^p}\left(\mathbb{E}[VY'] + \inf_{V' \in \mathcal{D}(\mathcal{Q}_j)} \mathbb{E}[V'Y']\right)$$
$$= \inf_{V' \in \mathcal{D}(\mathcal{Q}_j)} \sup_{Y' \in L^p} \mathbb{E}[(V + V')Y'] = \inf_{V' \in \mathcal{D}(\mathcal{Q}_j)} I_{\{-V'\}}(V) = I_{-\mathcal{D}(\mathcal{Q}_j)}(V),$$

where we use the minimax theorem [17, Corollary 3.3] for the third equality thanks to the fact that $\mathcal{D}(\mathcal{Q}_j)$ is a convex $\sigma(L^q, L^p)$-compact set. As a result,

$$\partial \rho_j^*(V) = \mathcal{N}_{-\mathcal{D}(\mathcal{Q}_j)}(V) = \left\{ Y \in L^p \mid \forall V' \in \mathcal{D}(\mathcal{Q}_j) \colon \mathbb{E}[VY] \geq \mathbb{E}\left[-V'Y\right] \right\}$$

if $V \in -\mathcal{D}(\mathcal{Q}_j)$ and $\partial \rho_j^*(V) = \emptyset$ if $V \in L^q \setminus -\mathcal{D}(\mathcal{Q}_j)$. Consequently,

$$\begin{aligned}
\partial g_j(w) &= \left\{ \mathbb{E}[VX] \mid V \in \partial \rho_j(X^\mathsf{T} w) \right\} \\
&= \left\{ \mathbb{E}[VX] \mid V \in -\mathcal{D}(\mathcal{Q}_j),\ X^\mathsf{T} w \in \partial \rho_j^*(V) \right\} \\
&= \left\{ \mathbb{E}[VX] \mid V \in -\mathcal{D}(\mathcal{Q}_j), \forall V' \in \mathcal{D}(\mathcal{Q}_j) \colon \mathbb{E}\left[VX^\mathsf{T}w\right] \geq \mathbb{E}\left[-V'X^\mathsf{T}w\right] \right\} \\
&= \left\{ \mathbb{E}[-VX] \mid V \in \mathcal{D}(\mathcal{Q}_j), \forall V' \in \mathcal{D}(\mathcal{Q}_j) \colon \mathbb{E}\left[-VX^\mathsf{T}w\right] \geq \mathbb{E}\left[-V'X^\mathsf{T}w\right] \right\} \\
&= \left\{ \mathbb{E}[-VX] \mid V \in \mathcal{V}_j(w) \right\}
\end{aligned}$$

so that the result follows. $\qquad\square$

Proof of Theorem 3.3 Let us denote by p the optimal value of $(\mathscr{P}(r))$. Thanks to Assumption 3.1, by strong duality for convex optimization, for instance, by [18, Theorem 2.9.3], p is equal to the optimal value of the corresponding Lagrange dual problem, that is,

$$p = \sup_{\nu \geq 0, \lambda \in \mathbb{R}} d(\nu, \lambda), \tag{3.2}$$

where, for each $\nu \geq 0, \lambda \in \mathbb{R}$,

$$d(\nu, \lambda) := \inf_{w \in \mathbb{R}^n} \left(\rho_1(w^\mathsf{T} X) + \nu \left(\rho_2(w^\mathsf{T} X) - r \right) + \lambda \left(\mathbf{1}^\mathsf{T} w - 1 \right) \right).$$

Let us fix $\nu \geq 0, \lambda \in \mathbb{R}$. Using the dual representations of ρ_1, ρ_2,

$$d(\nu, \lambda) = \inf_{w \in \mathbb{R}^n} \left(\max_{U \in \mathcal{D}(\mathcal{Q}_1)} \mathbb{E}\left[-Uw^\mathsf{T}X\right] + \nu \max_{V \in \mathcal{D}(\mathcal{Q}_2)} \mathbb{E}\left[-Vw^\mathsf{T}X\right] + \lambda \mathbf{1}^\mathsf{T} w \right) - r\nu - \lambda.$$

Let $f(w, U, V) := \mathbb{E}\left[-Uw^\mathsf{T}X\right] + \nu \mathbb{E}\left[-Vw^\mathsf{T}X\right] + \lambda \mathbf{1}^\mathsf{T} w$ for each $w \in \mathbb{R}^n$, $U \in \mathcal{D}(\mathcal{Q}_1)$, $V \in \mathcal{D}(\mathcal{Q}_2)$. Note that $w \mapsto f(w, U, V)$ is convex (affine) and continuous, $(U, V) \mapsto f(w, U, V)$ is concave (affine) and $\sigma(L^q, L^p)$-continuous (continuous), and $\mathcal{D}(\mathcal{Q}_1) \times \mathcal{D}(\mathcal{Q}_2)$ is convex and $\sigma(L^q, L^p)$-compact. Hence, the classical minimax theorem [17, Corollary 3.3] ensures that

$$\begin{aligned}
d(\nu, \lambda) &= \sup_{(U,V) \in \mathcal{D}(\mathcal{Q}_1) \times \mathcal{D}(\mathcal{Q}_2)} \inf_{w \in \mathbb{R}^n} \left(\mathbb{E}\left[-Uw^\mathsf{T}X\right] + \nu \mathbb{E}\left[-Vw^\mathsf{T}X\right] + \lambda \mathbf{1}^\mathsf{T} w \right) - r\nu - \lambda \\
&= \sup_{(U,V) \in \mathcal{D}(\mathcal{Q}_1) \times \mathcal{D}(\mathcal{Q}_2)} \inf_{w \in \mathbb{R}^n} \left(\mathbb{E}[-UX] + \nu \mathbb{E}[-VX] + \lambda \mathbf{1} \right)^\mathsf{T} w - r\nu - \lambda.
\end{aligned}$$

Clearly, for every $(U, V) \in \mathcal{D}(\mathcal{Q}_1) \times \mathcal{D}(\mathcal{Q}_2)$,

$$\begin{aligned}
\inf_{w \in \mathbb{R}^n} &\left(\mathbb{E}[-UX] + \nu \mathbb{E}[-VX] + \lambda \mathbf{1} \right)^\mathsf{T} w \\
&= \begin{cases} 0 & \text{if } \mathbb{E}[-UX] + \nu \mathbb{E}[-VX] + \lambda \mathbf{1} = 0, \\ -\infty & \text{else.} \end{cases}
\end{aligned} \tag{3.3}$$

It follows that

$$d(\nu, \lambda) = \begin{cases} -r\nu - \lambda & \text{if } \exists (U, V) \in \mathcal{D}(\mathcal{Q}_1) \times \mathcal{D}(\mathcal{Q}_2) \colon \mathbb{E}[-UX] + \nu \mathbb{E}[-VX] + \lambda \mathbf{1} = 0, \\ -\infty & \text{else.} \end{cases}$$

So the Lagrange dual problem in (3.2) takes the more explicit form

$$\text{maximize} \quad -r\nu - \lambda \qquad\qquad (\tilde{\mathscr{D}}(r))$$
$$\text{subject to} \ \ \mathbb{E}[UX] + \nu \mathbb{E}[VX] - \lambda\mathbf{1} = 0$$
$$U \in \mathcal{D}(\mathcal{Q}_1), \ V \in \mathcal{D}(\mathcal{Q}_2), \ \nu \geq 0, \ \lambda \in \mathbb{R}.$$

To avoid the multiplication of the variables ν, V, we make the following change of variables: Note that if $M \in \text{cone}(\mathcal{D}(\mathcal{Q}_2))$, then there exist $\nu \geq 0$ and $V \in \mathcal{D}(\mathcal{Q}_2)$ such that $M = \nu V$: we simply take $\nu = \mathbb{E}[M]$, and $V = \frac{M}{\nu}$ if $\nu > 0$ and an arbitrary $V \in \mathcal{D}(\mathcal{Q}_2)$ if $\nu = 0$. Conversely, if $\nu \geq 0$ and $V \in \mathcal{D}(\mathcal{Q}_2)$, then $M = \nu V \in \text{cone}(\mathcal{D}(\mathcal{Q}_2))$. These observations allow us to reformulate $(\tilde{\mathscr{D}}(r))$ as $(\mathscr{D}(r))$. Note that both problems have p as their optimal value.

Let $(U^*, M^*, \lambda^*) \in L^q \times L^q \times \mathbb{R}$ be an optimal solution for $\mathscr{D}(r)$. Thanks to Assumption 3.2 and [3, Corollary 4.8], there is strong duality with the corresponding Lagrange dual problem that relaxes the equality constraint, that is, we have

$$p = \inf_{w \in \mathbb{R}^n} \sup_{U \in \mathcal{D}(\mathcal{Q}_1), M \in \text{cone}(\mathcal{D}(\mathcal{Q}_2)), \lambda \in \mathbb{R}} \left(-r\mathbb{E}[M] - \lambda - w^\mathsf{T} \left(\mathbb{E}[UX] + \mathbb{E}[MX] - \lambda\mathbf{1} \right) \right)$$
$$= \inf_{w \in \mathbb{R}^n} \sup_{U \in \mathcal{D}(\mathcal{Q}_1), M \in \text{cone}(\mathcal{D}(\mathcal{Q}_2)), \lambda \in \mathbb{R}} \left(-r\mathbb{E}[M] - \lambda + \mathbb{E}\left[-Uw^\mathsf{T}X\right] + \mathbb{E}\left[-Mw^\mathsf{T}X\right] + \lambda w^\mathsf{T}\mathbf{1} \right).$$

Moreover, [3, Corollary 4.8], also ensures that there exists an optimal Lagrange multiplier $w^* \in \mathbb{R}^n$. By the first-order condition with respect to $U = U^*$, we have

$$0 \in -(w^*)^\mathsf{T}X - \mathcal{N}_{\mathcal{D}(\mathcal{Q}_1)}(U^*),$$

which means that

$$\mathbb{E}\left[-U^*(w^*)^\mathsf{T}X\right] \geq \mathbb{E}\left[-U'(w^*)^\mathsf{T}X\right]$$

for every $U' \in \mathcal{D}(\mathcal{Q}_1)$, that is,

$$\rho_1((w^*)^\mathsf{T}X) = \mathbb{E}\left[-U^*(w^*)^\mathsf{T}X\right].$$

We conclude that $U^* \in \mathscr{V}_1(w^*)$, where $\mathscr{V}_1(w^*)$ is defined by (3.1). In particular, by Lemma 3.4,

$$\mathbb{E}\left[-U^*X\right] \in \partial g_1(w^*). \qquad\qquad (3.4)$$

Similarly, the first-order condition with respect to $M = M^*$ yields

$$\mathbb{E}\left[-M^*\left((w^*)^\mathsf{T}X + r\right)\right] \geq \mathbb{E}\left[-M'\left((w^*)^\mathsf{T}X + r\right)\right]$$

for every $M' \in \text{cone}(\mathcal{D}_2)$, that is,

$$\mathbb{E}\left[-M^*\left((w^*)^\mathsf{T}X + r\right)\right] = \max_{M' \in \text{cone}(\mathcal{D}(\mathcal{Q}_2))} \mathbb{E}\left[-M'\left((w^*)^\mathsf{T}X + r\right)\right]. \qquad (3.5)$$

Since $\text{cone}(\mathcal{D}_2)$ is a cone, the quantity $\sup_{M' \in \text{cone}(\mathcal{D}(\mathcal{Q}_2))} \mathbb{E}\left[-M'((w^*)^\mathsf{T}X + r)\right]$ can either take value 0 or $+\infty$. Since $\mathbb{E}\left[-M^*((w^*)^\mathsf{T}X + r)\right]$ is a finite number, both sides of (3.5)

must be equal to zero. Moreover,

$$
\begin{aligned}
0 &= \max_{M' \in \text{cone}(\mathcal{D}(\mathcal{Q}_2))} \mathbb{E}\left[-M'\left((w^*)^\mathsf{T} X + r\right)\right] \\
&= \left(\sup_{\lambda' \geq 0} \lambda'\right)\left(\max_{V' \in \mathcal{D}(\mathcal{Q}_2)} \mathbb{E}\left[-V'\left((w^*)^\mathsf{T} X + r\right)\right]\right) \\
&= +\infty \cdot \rho_2((w^*)^\mathsf{T} X + r) = +\infty \cdot \left(\rho_2((w^*)^\mathsf{T} X) - r\right).
\end{aligned}
$$

Hence, we have $\rho_2((w^*)^\mathsf{T} X) = r$.

Let $\nu^* = \mathbb{E}\left[M^*\right]$. Suppose first that $\nu^* > 0$. Let $V^* := \frac{M^*}{\nu^*} \in \mathcal{D}(\mathcal{Q}_2)$. Then,

$$
\mathbb{E}\left[-M^*((w^*)^\mathsf{T} X + r)\right] = \nu^* \mathbb{E}\left[-V^*\left((w^*)^\mathsf{T} X + r\right)\right] = 0
$$

so that $\mathbb{E}\left[-V^*(w^*)^\mathsf{T} X\right] = r$. Hence,

$$
\mathbb{E}\left[-V^*(w^*)^\mathsf{T} X\right] = r = \rho_2((w^*)^\mathsf{T} X) = \max_{V' \in \mathcal{D}(\mathcal{Q}_2)} \mathbb{E}\left[-V'(w^*)^\mathsf{T} X\right],
$$

that is, $V^* \in \mathcal{V}_2(w^*)$. In particular,

$$
\mathbb{E}\left[-V^* X\right] \in \partial g_2(w^*).
$$

Next, suppose that $\nu^* = 0$, that is, $M^* = 0$ $\mathbb{P}$-almost surely. Let us pick some $V^* \in \mathcal{V}_2(w^*)$ arbitrarily. (We know that $\mathcal{V}_2(w^*) \neq \emptyset$ since ρ_2 is assumed to be continuous from below.) In both cases, we may write $M^* = \nu^* V^*$ and

$$
\mathbb{E}\left[-M^* X\right] = \nu^* \mathbb{E}\left[-V^* X\right] \in \nu^* \partial g_2(w^*). \tag{3.6}
$$

By the feasibility of (U^*, M^*, λ^*) for $(\mathscr{D}(r))$,

$$
\mathbb{E}\left[-U^* X\right] + \mathbb{E}\left[-M^* X\right] + \lambda^* \mathbf{1} = \mathbb{E}\left[-U^* X\right] + \nu^* \mathbb{E}\left[-V^* X\right] + \lambda^* \mathbf{1} = 0. \tag{3.7}
$$

Hence, by (3.4), (3.6), (3.7), we conclude that

$$
0 \in \partial g_1(w^*) + \nu^* \partial g_2(w^*) + \lambda^* \mathbf{1}.
$$

Finally, by the first-order condition with respect to $\lambda = \lambda^*$, we get

$$
\mathbf{1}^\mathsf{T} w^* = 1,
$$

that is, $w^* \in \mathcal{W}$. Therefore, we establish the Karush-Kuhn-Tucker conditions for $(\mathscr{P}(r))$ at $w = w^*$. By [18, Theorem 2.9.3], we conclude that w^* is an optimal solution for $(\mathscr{P}(r))$. □

Remark 3.5 Let us comment on the roles of the two constraint qualifications and the assumption about the existence of an optimal solution for $(\mathscr{D}(r))$. In the proof of Theorem 3.3, note that Assumption 3.1 already guarantees the existence of an optimal solution $(\bar{\nu}, \bar{\lambda})$ for the Lagrange dual problem in (3.2). Nevertheless, the reformulated problem $(\tilde{\mathscr{D}}(r))$ has two additional variables, U and V, which, together with ν, λ, combine into U, M, λ in the finalized dual problem $(\mathscr{D}(r))$. As a result, the existence of an optimal solution for $(\mathscr{D}(r))$ is not guaranteed *a priori*. Once such an optimal solution is assumed, Assumption 3.2 automatically yields the existence of an optimal Lagrange multiplier for the equality constraint in $(\mathscr{D}(r))$, which is shown to give an optimal solution for the original problem $(\mathscr{P}(r))$.

In the following examples, we consider a few special choices of the risk measures. We work with $p = 1$ in all examples.

Example 3.6 Let ρ_1 be the average value-at-risk at a probability level $\theta \in (0, 1)$ (Example 2.3) and ρ_2 the negative expected value (Example 2.1). In this case:

$$\mathcal{D}(\mathcal{Q}_1) = \left\{ U \in L^\infty \mid \mathbb{P}\left\{ 0 \leq U \leq \frac{1}{\theta} \right\} = 1 \right\},$$

$$\mathrm{qri}(\mathcal{D}(\mathcal{Q}_1)) = \left\{ U \in L^\infty \mid \mathbb{P}\left\{ 0 < U < \frac{1}{\theta} \right\} = 1 \right\},$$

$$\mathrm{cone}(\mathcal{D}(\mathcal{Q}_2)) = \mathrm{cone}(1) = \mathbb{R}_+, \qquad \mathrm{qri}(\mathrm{cone}(\mathcal{D}(\mathcal{Q}_2))) = (0, +\infty),$$

where the quasi relative interiors can be calculated by following a similar procedure as in [3, Example 3.11]. It is easy to observe that Assumption 3.2 is equivalent to the existence of a probability measure $\mathbb{Q}$ on $(\Omega, \mathcal{F})$ that is equivalent to $\mathbb{P}$ such that $\frac{d\mathbb{Q}}{d\mathbb{P}} \leq \frac{1}{\theta}$ $\mathbb{P}$-almost surely and $\mathbb{E}^{\mathbb{Q}}[-X]$ is in the conic convex hull of the set $\{\mathbb{E}[X], \mathbf{1}, -\mathbf{1}\}$. In particular, if $\mathbb{E}[X_1] = \ldots = \mathbb{E}[X_n]$, then this condition is satisfied by $\mathbb{Q} = \mathbb{P}$. Moreover, the dual problem $(\mathscr{D}(r))$ becomes

$$
\begin{aligned}
\text{maximize} \quad & -rm - \lambda \\
\text{subject to} \quad & \mathbb{E}[UX] + m\mathbb{E}[X] - \lambda\mathbf{1} = 0 \\
& \mathbb{E}[U] = 1 \\
& 0 \leq U \leq \frac{1}{\theta} \ \mathbb{P}\text{-almost surely} \\
& U \in L^\infty, \ m \geq 0, \ \lambda \in \mathbb{R},
\end{aligned}
$$

which is a linear programming problem in an infinite-dimensional setting. When $(\Omega, \mathcal{F}, \mathbb{P})$ is a finite probability space, it reduces to a finite-dimensional linear programming problem.

Example 3.7 We switch the roles of negative expected value and average value-at-risk in Example 3.6 so that

$$\mathcal{D}(\mathcal{Q}_1) = \mathrm{qri}(\mathcal{D}(\mathcal{Q}_1)) = \{1\} \subseteq L^\infty,$$

$$\mathrm{cone}(\mathcal{D}(\mathcal{Q}_2)) = \left\{ M \in L^\infty \mid \mathbb{P}\left\{ 0 \leq M \leq \frac{\mathbb{E}[M]}{\theta} \right\} = 1 \right\},$$

$$\mathrm{qri}(\mathrm{cone}(\mathcal{D}(\mathcal{Q}_2))) = \left\{ M \in L^\infty \mid \mathbb{P}\left\{ 0 < M < \frac{\mathbb{E}[M]}{\theta} \right\} = 1 \right\}.$$

In this case, Assumption 3.2 is equivalent to the existence of a finite measure with density M such that $\theta M < \mathbb{E}[M]$ $\mathbb{P}$-almost surely and $\mathbb{E}[-MX]$ is in the unbounded polyhedral set $\{\mathbb{E}[X] - \lambda\mathbf{1} \mid \lambda \in \mathbb{R}\}$. In particular, if $\mathbb{E}[X_1] = \ldots = \mathbb{E}[X_n]$, then this condition is satisfied by $M \equiv 1$. Moreover, the dual problem $(\mathscr{D}(r))$ becomes

$$
\begin{aligned}
\text{maximize} \quad & -r\mathbb{E}[M] - \lambda \\
\text{subject to} \quad & \mathbb{E}[X] + \mathbb{E}[MX] - \lambda\mathbf{1} = 0 \\
& 0 \leq \theta M \leq \mathbb{E}[M] \ \mathbb{P}\text{-almost surely} \\
& M \in L^\infty, \ \lambda \in \mathbb{R},
\end{aligned}
$$

which reduces to a finite-dimensional linear programming problem when $(\Omega, \mathcal{F}, \mathbb{P})$ is a finite probability space.

Remark 3.8 Note that Assumption 3.2 is a constraint qualification for the dual problem $(\mathscr{D}(r))$. A more standard alternative of it would assume the existence of $U \in \text{int}(\mathcal{D}(\mathcal{Q}_1))$, $M \in \text{int}(\text{cone}(\mathcal{D}(\mathcal{Q}_2)))$, $\lambda \in \mathbb{R}$ such that $\mathbb{E}[UX] + \mathbb{E}[MX] - \lambda \mathbf{1} = 0$, where int denotes topological interior. However, in infinite-dimensional spaces, many convex sets that show up in applications have empty interior. For instance, it is well-known that, for $q \in [1, +\infty)$, we have int $L_+^q = \emptyset$ unless L^q is a finite-dimensional space, that is, the underlying probability space is isomorphic to a finite probability space (see, for instance, [7, Example 10.1.3] and [6, Example 2.12]). Similarly, we have int$(\mathcal{D}(\mathcal{Q}_1)) = \emptyset$ in Example 3.6 and int$(\text{cone}(\mathcal{D}(\mathcal{Q}_2))) = \emptyset$ in Example 3.7 unless L^q is a finite-dimensional space. Hence, the standard alternative is never satisfied for our examples whenever we deviate from the finite-dimensional case. Hence, Assumption 3.2, which uses quasi relative interior instead of usual interior, is much weaker than its interior-based counterpart (see, for instance, [18, Theorem 2.9.6] for a strong duality theorem with an interior-based constraint qualification).

Remark 3.9 Let us comment on the benefit of Theorem 3.3 for computing an optimal solution for the primal portfolio optimization problem $(\mathscr{P}(r))$. When $(\Omega, \mathcal{F}, \mathbb{P})$ is a finite probability space, the dual problem $(\mathscr{D}(r))$ reduces to a finite-dimensional convex optimization problem with the set constraints $U \in \mathcal{D}(\mathcal{Q}_1)$ and $M \in \text{cone}(\mathcal{D}(\mathcal{Q}_2))$. If these constraints can be represented by finitely many (convex) inequalities, then $(\mathscr{D}(r))$ can be solved by commercial software for convex optimization such as CVX, which also return the value of the Lagrange multipliers for the constraints at (approximate) optimality. Hence, without an additional procedure, an optimal solution for $(\mathscr{P}(r))$ is readily computed by the solver as the dual multiplier of the constraint $\mathbb{E}[UX] + \mathbb{E}[MX] - \lambda \mathbf{1} = 0$ at optimality. As noted in Examples 3.6 and 3.7, when the classical coherent risk measures negative expected value and average value-at-risk are used in the problem, the set constraints $U \in \mathcal{D}(\mathcal{Q}_1)$ and $M \in \text{cone}(\mathcal{D}(\mathcal{Q}_2))$ are easily represented by finitely many linear inequalities so that $(\mathscr{D}(r))$ even reduces to a linear programming problem. In this case, many more commerical solvers are available (for instance, CPLEX, Gurobi) and they also return the values of the dual variables associated to the constraints at optimality. Hence, in the cases where $(\mathscr{D}(r))$ reduces to a standard convex/linear optimization problem and it has an optimal solution, thanks to Theorem 3.3, an optimal portfolio for $(\mathscr{P}(r))$ is returned by commercial solvers. It should also be noted that the idea of recovering a primal optimal solution from the dual multipliers of the dual problem is pretty well-known in the nonsmooth/stochastic optimization literature (see, [2,11], for instance). Hence, at a high level, Theorem 3.3 can be considered as a result in the same spirit.

We finish this section by providing an analogous dual problem and an optimality result for the case where shortselling is not allowed, namely, for the problem

$$\text{minimize} \quad \rho_1(w^\mathsf{T} X) \qquad\qquad (\mathscr{P}_+(r))$$
$$\text{subject to} \quad \rho_2(w^\mathsf{T} X) \le r$$
$$w \in \mathcal{W}_+,$$

where $\mathcal{W}_+$ is defined by (2.2). The analysis of $(\mathscr{P}_+(r))$ is very similar to that of $(\mathscr{P}(r))$ and it yields the finalized dual problem

$$\text{maximize} \quad -r\mathbb{E}[M] - \lambda \qquad\qquad (\mathscr{D}_+(r))$$
$$\text{subject to} \quad \mathbb{E}[UX] + \mathbb{E}[MX] - \lambda \mathbf{1} \le 0$$
$$U \in \mathcal{D}(\mathcal{Q}_1), \ M \in \text{cone}(\mathcal{D}(\mathcal{Q}_2)), \ \lambda \in \mathbb{R}.$$

We have the following duality result which works under modified versions of Assumptions 3.1 and 3.2.

Theorem 3.10 *Assume that there exists $w \in \mathcal{W}_+$ such that $\rho_2(w^T X) < r$ and $w_i > 0$ for every $i \in \mathcal{N}$. Assume further that there exist $U \in \text{qri}(\mathcal{D}(\mathcal{Q}_1))$, $M \in \text{qri}(\text{cone}(\mathcal{D}(\mathcal{Q}_2)))$, $\lambda \in \mathbb{R}$ such that $\mathbb{E}[UX_i] + \mathbb{E}[MX_i] - \lambda < 0$ for every $i \in \mathcal{N}$. Suppose that there exists an optimal solution $(U^*, M^*, \lambda^*) \in L^q \times L^q \times \mathbb{R}$ of $(\mathscr{D}_+(r))$. Then, there exists an optimal Lagrange multiplier $w^* \in \mathbb{R}^n_+$ associated to the inequality constraint of $(\mathscr{D}_+(r))$, and w^* is an optimal solution for $(\mathscr{P}_+(r))$.*

Proof The proof goes along the same lines as the proof of Theorem 3.3. The only important change is in (3.3):

$$\inf_{w \in \mathbb{R}^n_+} (\mathbb{E}[-UX] + \nu\mathbb{E}[-VX] + \lambda\mathbf{1})^T w = \begin{cases} 0 & \text{if } \mathbb{E}[-UX] + \nu\mathbb{E}[-VX] + \lambda\mathbf{1} \geq 0, \\ -\infty & \text{else,} \end{cases}$$

which is the reason for having an inequality constraint in $(\mathscr{D}_+(r))$. The rest follows in a standard manner. □

4 The portfolio optimization problem under the multivariate Gaussian distribution

In this section, we study the problem $(\mathscr{P}(r))$ under the special case that X is a Gaussian random vector and ρ_1, ρ_2 are law-invariant. Under these assumptions, it turns out that the analysis of $(\mathscr{P}(r))$ can be performed in terms of the hyperbola appearing in the classical Markowitz problem and an optimal solution for $(\mathscr{P}(r))$ can be calculated with an explicit formula whenever it exists. The aim of this section is to provide an analysis that is peculiar to the Gaussian case; hence, we follow a route that is quite different from the general duality-based approach in Sect. 3.

We assume that $X = (X_1, \ldots, X_n)^T \in L^2_n$ is a Gaussian random vector with mean vector $m = (m_1, \ldots, m_n)^T$ and covariance matrix $C \in \mathbb{R}^{n \times n}$. We further assume that m and $\mathbf{1} := (1, \ldots, 1)^T \in \mathbb{R}^n$ are linearly independent and that C is a nonsingular matrix with inverse C^{-1}. Hence, C is a symmetric positive definite matrix with strictly positive eigenvalues.

Note that, for a portfolio $w \in \mathcal{W}$, its return $w^T X$ is a Gaussian random variable. A simple calculation yields that the corresponding expected value and variance are given by

$$\mu_w := \mathbb{E}[w^T X] = m^T w, \quad \sigma_w^2 := \text{Var}(w^T X) = w^T C w, \tag{4.1}$$

respectively. For every $w \in \mathcal{W}$, we may write

$$w^T X = \mathbb{E}[w^T X] + \sqrt{\text{Var}(w^T X)} Z \tag{4.2}$$

for some standard Gaussian random variable Z (with zero mean and unit variance). Using this, we provide an explicit expression for the values of a generic law-invariant coherent risk measure ρ next.

Proposition 4.1 *Let ρ be a coherent, law-invariant and finite risk measure on L^2. For every $w \in \mathcal{W}$, it holds*

$$\rho(w^T X) = \rho(Z)\sqrt{w^T C w} - m^T w,$$

where Z is an arbitrary standard Gaussian random variable.

Proof Let $w \in \mathcal{W}$. Using (4.2), we obtain

$$\rho(w^\mathsf{T} X) = \rho\left(\sqrt{\mathrm{Var}(w^\mathsf{T} X)} Z + \mathbb{E}\left[w^\mathsf{T} X\right]\right) = \rho(Z)\sqrt{\mathrm{Var}(w^\mathsf{T} X)} - \mathbb{E}\left[w^\mathsf{T} X\right]$$

thanks to the translativity and positive homogeneity of ρ. Finally, the number $\rho(Z)$ is free of the choice of the standard Gaussian random variable Z thanks to the law-invariance of ρ. □

With a slight abuse of notation, we define $\rho_j := \rho_j(Z) \geq 0$ for each $j \in \{1, 2\}$, where Z is a generic standard Gaussian random variable. Thanks to Proposition 4.1, we may rewrite $(\mathscr{P}(r))$ as

$$\begin{aligned}
\text{minimize} \quad & \rho_1 \sqrt{w^\mathsf{T} C w} - m^\mathsf{T} w && (\mathscr{P}(r)) \\
\text{subject to} \quad & \rho_2 \sqrt{w^\mathsf{T} C w} - m^\mathsf{T} w \leq r \\
& \mathbf{1}^\mathsf{T} w = 1 \\
& w \in \mathbb{R}^n.
\end{aligned}$$

In what follows, we provide an analytical solution for $(\mathscr{P}(r))$, whenever it exists, under all possible relationships among the parameters m, C, r, ρ_1, ρ_2. To that end, let us introduce the constants

$$\alpha := m^\mathsf{T} C^{-1} m, \quad \beta := m^\mathsf{T} C^{-1} \mathbf{1} = \mathbf{1}^\mathsf{T} C^{-1} m, \quad \gamma := \mathbf{1}^\mathsf{T} C^{-1} \mathbf{1}, \quad \delta := \alpha \gamma - \beta^2,$$

which also appear in the analysis of the classical Markowitz problem. As a consequence of the positive definiteness of C, it is well-known and easy to check that $\alpha, \gamma, \delta > 0$.

4.1 The Markowitz hyperbola

The analysis of the n-dimensional portfolio optimization problem $(\mathscr{P}(r))$ is based on an associated two-dimensional optimization problem whose decision variables stand for the standard deviation and expected return of a portfolio. Note that every portfolio $w \in \mathcal{W}$ induces a standard deviation-expected return pair $(\sigma_w, \mu_w) \in \mathbb{R}^2$ of $(\mathscr{M}(r))$ through the definitions $\sigma_w = \sqrt{w^\mathsf{T} C w}$, $\mu_w = m^\mathsf{T} r$. The structure of the set $\{(\sigma_w, \mu_w) \mid w \in \mathcal{W}\}$ is very well-known: this set is the convex hull of the right wing of a hyperbola. The precise version of this classical result is recalled in the next lemma.

Lemma 4.2 *Let $\mu \in \mathbb{R}$ and consider the problem of finding the portfolio with minimum variance among all the portfolios with expected return level μ:*

$$\begin{aligned}
\text{minimize} \quad & w^\mathsf{T} C w && (\mathscr{A}(\mu)) \\
\text{subject to} \quad & m^\mathsf{T} w = \mu \\
& \mathbf{1}^\mathsf{T} w = 1 \\
& w \in \mathbb{R}^n.
\end{aligned}$$

The problem $\mathscr{A}(\mu)$ has a unique optimal solution given by

$$w(\mu) := \frac{1}{\delta}\left((\gamma \mu - \beta) C^{-1} m + (\alpha - \beta \mu) C^{-1} \mathbf{1}\right) \tag{4.3}$$

with corresponding expected return $\mu_{w(\mu)} = \mu$ and standard deviation

$$\sigma_{w(\mu)} = \sqrt{\frac{1}{\gamma} + \frac{\gamma}{\delta}\left(\mu - \frac{\beta}{\gamma}\right)^2}.$$

$$\left\{ (\sigma_{w(\mu)}, \mu) \mid \mu \in \mathbb{R} \right\} = \mathcal{H}_+ := \mathcal{H} \cap (\mathbb{R}_+ \times \mathbb{R}),$$

where $\mathcal{H}$ is a hyperbola defined by

$$\mathcal{H} := \left\{ (\sigma, \mu) \in \mathbb{R}^2 \mid \sigma^2 - \frac{\gamma}{\delta}\left(\mu - \frac{\beta}{\gamma} \right)^2 = \frac{1}{\gamma} \right\},$$

whose asymptotes are specified by the equations

$$\mu = \frac{\beta}{\gamma} \pm \sqrt{\frac{\delta}{\gamma}} \sigma.$$

In particular, for every point (σ, μ) on the right wing $\mathcal{H}_+$, there exists a unique portfolio $w \in \mathcal{W}$ such that $(\sigma, \mu) = (\sigma_w, \mu_w)$. Hence, Let $\mathrm{co}\,\mathcal{H}_+$ be the convex hull of $\mathcal{H}_+$, that is,

$$\mathrm{co}\,\mathcal{H}_+ = \left\{ (\sigma, \mu) \in \mathbb{R}_+ \times \mathbb{R} \mid \sigma^2 - \frac{\gamma}{\delta}\left(\mu - \frac{\beta}{\gamma} \right)^2 \geq \frac{1}{\gamma} \right\}.$$

For every $(\sigma, \mu) \in \mathrm{co}\,\mathcal{H}_+$, there exists a portfolio $w \in \mathcal{W}$ such that $(\sigma, \mu) = (\sigma_w, \mu_w)$. In particular,

$$\{(\sigma_w, \mu_w) \mid w \in \mathcal{W}\} = \mathrm{co}\,\mathcal{H}_+. \tag{4.4}$$

Proof These are well-known results from the analysis of the classical Markowitz problem. The reader may refer to the original derivation in Merton [13] as well as many textbooks covering portfolio optimization, for instance, [4, Chapter 3].   □

Note that the point $(\frac{1}{\sqrt{\gamma}}, \frac{\beta}{\gamma})$ is the corner point of the right wing $\mathcal{H}_+$; in particular, for every $(\sigma, \mu) \in \mathrm{co}\,\mathcal{H}_+$, it holds $\sigma \geq \frac{1}{\sqrt{\gamma}}$. For each $\sigma \geq \frac{1}{\sqrt{\gamma}}$, let

$$\mu(\sigma) := \frac{\beta}{\gamma} + \sqrt{\frac{\delta}{\gamma}\sigma^2 - \frac{\delta}{\gamma^2}}.$$

In particular, for every $(\sigma, \mu) \in \mathrm{co}\,\mathcal{H}_+$, it holds $\mu \leq \mu(\sigma)$.

4.2 The associated two-dimensional problems

The relation (4.4) in Lemma 4.2 motivates us to introduce a related problem expressed as

$$\begin{aligned} \text{minimize} \quad & \rho_1 \sigma - \mu & (\mathscr{M}(r)) \\ \text{subject to} \quad & \rho_2 \sigma - \mu \leq r \\ & (\sigma, \mu) \in \mathrm{co}\,\mathcal{H}_+. \end{aligned}$$

Indeed, for every feasible solution $w \in \mathbb{R}^n$ of $(\mathscr{P}(r))$, the point (σ_w, μ_w) is a feasible solution of $(\mathscr{M}(r))$ and the corresponding objective function values are equal. Moreover, by the last part of Lemma 4.2, for every feasible solution $(\sigma, \mu) \in \mathbb{R}^2$ of $(\mathscr{M}(r))$, there exists a feasible solution $w \in \mathbb{R}^n$ of $(\mathscr{P}(r))$ such that $(\sigma, \mu) = (\sigma_w, \mu_w)$ and the corresponding objective function values are equal. It follows that for an optimal solution $w \in \mathbb{R}^n$ of $(\mathscr{P}(r))$, supposing that it exists, the induced feasible solution (σ_w, μ_w) of $(\mathscr{M}(r))$ is also optimal for $(\mathscr{M}(r))$. On the other hand, since $\rho_1 \sigma - \mu \geq \rho_1 \sigma - \mu(\sigma)$ and $r \geq \rho_2 \sigma - \mu \geq \rho_2 \sigma - \mu(\sigma)$

for every feasible solution (σ, μ) of $(\mathscr{M}(r))$, an optimal solution of $(\mathscr{M}(r))$, whenever it exists, must be on the upper half of $\mathcal{H}_+$, that is, it must be of the form $(\sigma, \mu(\sigma))$ for some $\sigma \geq \frac{1}{\sqrt{\gamma}}$. By the uniqueness part of Lemma 4.2, such an optimal solution corresponds to a unique portfolio given by the formula in (4.3). Consequently, to figure out the optimal value and the possible optimal solutions of $(\mathscr{P}(r))$, it suffices to carry out the same analysis for $(\mathscr{M}(r))$ and then to recover an optimal solution of $(\mathscr{P}(r))$ using (4.3) whenever there is an optimal solution of $(\mathscr{M}(r))$.

Before providing a joint analysis of $(\mathscr{M}(r))$ and $(\mathscr{P}(r))$, we start by solving an "unconstrained" problem, namely, the problem of minimizing the objective function of $(\mathscr{M}(r))$ over the whole set $\mathrm{co}\,\mathcal{H}_+$, without the additional risk constraint.

Proposition 4.3 *Consider the auxiliary problem*

$$\text{minimize} \quad \rho_1 \sigma - \mu \qquad\qquad (\mathscr{M}_{\mathscr{A}})$$

$$\text{subject to} \quad (\sigma, \mu) \in \mathrm{co}\,\mathcal{H}_+.$$

(i) *Suppose that $\rho_1 < \sqrt{\frac{\delta}{\gamma}}$. Then, $(\mathscr{M}_{\mathscr{A}})$ is an unbounded problem with optimal value $-\infty$.*

(ii) *Suppose that $\rho_1 = \sqrt{\frac{\delta}{\gamma}}$. Then, $(\mathscr{M}_{\mathscr{A}})$ has a finite infimum that is equal to $-\frac{\beta}{\gamma}$ but the infimum is not attained by a feasible point.*

(iii) *Suppose that $\rho_1 > \sqrt{\frac{\delta}{\gamma}}$. Then, the unique optimal solution of $(\mathscr{M}_{\mathscr{A}})$ is (σ^*, μ^*), where*

$$\sigma^* := \frac{\rho_1}{\sqrt{\gamma \rho_1^2 - \delta}}, \quad \mu^* := \frac{\beta}{\gamma} + \frac{\delta}{\gamma \sqrt{\gamma \rho_1^2 - \delta}}. \qquad\qquad (4.5)$$

Moreover, the unique portfolio w^* with $(\sigma_{w^*} = \sigma^*, \mu_{w^*} = \mu^*)$ is given by

$$w^* = w(\mu^*) = \frac{1}{\sqrt{\gamma \rho_1^2 - \delta}} C^{-1} m + \left(\frac{1}{\gamma} - \frac{\beta}{\gamma \sqrt{\gamma \rho_1^2 - \delta}} \right) C^{-1} \mathbf{1}.$$

Proof (i) Suppose that $\rho_1 < \sqrt{\frac{\delta}{\gamma}}$. A standard exercise in calculus yields that

$$\lim_{\frac{1}{\sqrt{\gamma}} \leq \sigma \uparrow +\infty} (\rho_1 \sigma - \mu(\sigma)) = \lim_{\sigma_+(r) \leq \sigma \uparrow +\infty} \left(\rho_1 \sigma - \frac{\beta}{\gamma} - \sqrt{\frac{\delta}{\gamma} \sigma^2 - \frac{\delta}{\gamma^2}} \right) = -\infty.$$

Since the objective function diverges to $-\infty$ on a subset of $\mathrm{co}\,\mathcal{H}_+$, it follows that $(\mathscr{M}_{\mathscr{A}})$ is an unbounded problem with optimal value $-\infty$.

(ii) Suppose that $\rho_1 = \sqrt{\frac{\delta}{\gamma}}$. In this case, the limit evaluated in the previous case yields

$$\lim_{\frac{1}{\sqrt{\gamma}} \leq \sigma \uparrow +\infty} (\rho_1 \sigma - \mu(\sigma)) = -\frac{\beta}{\gamma}.$$

On the other hand, since the hyperbola $\mathcal{H}$ and its asymptote $\left\{ (\sigma, \mu) \in \mathbb{R}^2 \mid \sqrt{\frac{\delta}{\gamma}} \sigma - \mu = -\frac{\beta}{\gamma} \right\}$ do not intersect, there is no feasible solution $(\bar{\sigma}, \bar{\mu})$ of $(\mathscr{M}_{\mathscr{A}})$ such that

$$\rho_1 \bar{\sigma} - \bar{\mu} = \sqrt{\frac{\delta}{\gamma}} \bar{\sigma} - \bar{\mu} = -\frac{\beta}{\gamma}.$$

Hence, the infimum of $(\mathscr{M}_{\mathscr{A}})$ is equal to $-\frac{\beta}{\gamma}$ but it is not attained by a feasible solution.

(iii) Suppose that $\rho_1 > \sqrt{\frac{\delta}{\gamma}}$. Since every point $(\sigma, \mu) \in \mathrm{co}\,\mathcal{H}_+$ has $\rho_1\sigma - \mu \geq \rho_1\sigma - \mu(\sigma)$, if (σ, μ) is an optimal solution of $(\mathcal{M}_{\mathscr{A}})$, then it must satisfy $\mu = \mu(\sigma)$. Moreover, since $\mathrm{co}\,\mathcal{H}_+$ is a convex set, by the well-known first-order condition, a point $(\sigma, \mu(\sigma))$ is an optimal solution of $(\mathcal{M}_{\mathscr{A}})$ if and only if the negative of the gradient of the objective function at (σ, μ), which is $(-\rho_1, 1)$ in this case, is a normal direction of the feasible region $\mathrm{co}\,\mathcal{H}_+$ at (σ, μ), that is,

$$(-\rho_1, 1) \in \left\{ (x, y) \in \mathbb{R}^2 \mid \frac{d\mu(\sigma)}{d\sigma} y + x = 0, x \leq 0 \right\},$$

where the derivative is calculated as

$$\frac{d\mu(\sigma)}{d\sigma} = \sqrt{\frac{\delta}{\gamma}} \frac{\sigma}{\sqrt{\sigma^2 - \frac{1}{\gamma}}}.$$

Hence, $(\sigma, \mu(\sigma))$ is an optimal solution of $(\mathcal{M}_{\mathscr{A}})$ if and only if

$$\sqrt{\frac{\delta}{\gamma}} \frac{\sigma}{\sqrt{\sigma^2 - \frac{1}{\gamma}}} = \rho_1,$$

that is,

$$\sigma = \sigma^* = \frac{\rho_1}{\sqrt{\gamma\rho_1^2 - \gamma}}.$$

Consequently, we also have $\mu(\sigma) = \mu^*$. Hence, (σ^*, μ^*) is the unique optimal solution of $(\mathcal{M}_{\mathscr{A}})$. The corresponding portfolio $w^* = w(\mu^*)$ can be calculated easily using (4.3). □

4.3 Main theorems

In this section, we present complete solutions for $(\mathcal{M}(r))$ and $(\mathscr{P}(r))$. To that end, we provide three main theorems based on the slope of the line

$$\mathcal{L}(r) := \left\{ (\sigma, \mu) \in \mathbb{R}^2 \mid \rho_2\sigma - \mu = r \right\}$$

related to the risk constraint. It turns out that the comparison between the slope ρ_2 of $\mathcal{L}(r)$ and the (positive) slope $\sqrt{\frac{\delta}{\gamma}}$ of the asymptote of $\mathcal{H}$ is critical for the analysis.

Theorem 4.4 *Let $r \in \mathbb{R}$ and suppose that $\rho_2 < \sqrt{\frac{\delta}{\gamma}}$. Then, the hyperbola $\mathcal{H}$ and the line $\mathcal{L}(r)$ intersect at two distinct points $(\sigma_-(r), \mu_-(r))$ and $(\sigma_+(r), \mu_+(r))$ defined by*

$$\sigma_\pm(r) := \frac{-(\gamma r + \beta)\rho_2 \pm \sqrt{\delta(\gamma r^2 + 2\beta r + \alpha - \rho_2^2)}}{\delta - \gamma\rho_2^2}, \tag{4.6}$$

$$\mu_\pm(r) := \frac{-\delta r - \beta\rho_2^2 \pm \rho_2\sqrt{\delta(\gamma r^2 + 2\beta r + \alpha - \rho_2^2)}}{\delta - \gamma\rho_2^2}. \tag{4.7}$$

In particular, $\sigma_+(r) > 0 > \sigma_-(r)$. Moreover, one of the following cases holds for $(\mathcal{M}(r))$.

(i) *Suppose that $\rho_1 < \sqrt{\frac{\delta}{\gamma}}$. Then, $(\mathcal{M}(r))$ and $(\mathcal{P}(r))$ are unbounded problems with common optimal value $-\infty$.*

(ii) *Suppose that $\rho_1 = \sqrt{\frac{\delta}{\gamma}}$. Then, $(\mathcal{M}(r))$ and $(\mathcal{P}(r))$ have a common finite infimum that is equal to $-\frac{\beta}{\gamma}$ but the infimum is not attained by a feasible solution in both problems.*

(iii) *Suppose that $\rho_1 > \sqrt{\frac{\delta}{\gamma}}$. Let*

$$r^* := \rho_2 \sigma^* - \mu^* = \frac{\rho_1 \rho_2 \gamma - \delta}{\gamma \sqrt{\gamma \rho_1^2 - \delta}} - \frac{\beta}{\gamma}, \quad r_0 := \rho_2 \sigma_0 - \mu_0 = \frac{\rho_2}{\sqrt{\gamma}} - \frac{\beta}{\gamma}. \quad (4.8)$$

It holds $r^* \le r_0$. Moreover, the unique optimal solution (σ^*, μ^*) of $(\mathcal{M}_{\mathscr{A}})$ is also the unique optimal solution of $(\mathcal{M}(r))$ and the corresponding portfolio w^* is the unique optimal solution of $(\mathcal{P}(r))$ if and only if $r \ge r^*$. In particular, this is the case when $r \ge r_0$. If $r < r^*$, then $(\sigma_+(r), \mu_+(r))$ is the unique optimal solution of $(\mathcal{M}(r))$ and

$$w_+(r) := w(\mu_+(r)) = \frac{1}{\delta - \gamma \rho_2^2} \left[\left(-\gamma r - \beta + \frac{\gamma \rho_2}{\delta} \sqrt{\delta(\gamma \rho_2^2 + 2\beta r + \alpha - \rho_2^2)} \right) C^{-1} m \right.$$

$$\left. + \left(\beta r + \alpha - \rho_2^2 - \frac{\beta \rho_2}{\delta} \sqrt{\delta(\gamma \rho_2^2 + 2\beta r + \alpha - \rho_2^2)} \right) C^{-1} \mathbf{1} \right]$$

is the unique optimal solution of $(\mathcal{P}(r))$.

Proof By the definitions of $\mathcal{H}$ and $\mathcal{L}(r)$, a point $(\sigma, \mu) \in \mathcal{H} \cap \mathcal{L}(r)$ must satisfy

$$\sigma^2 - \frac{\gamma}{\delta} \left(\mu - \frac{\beta}{\gamma} \right)^2 = \sigma^2 - \frac{\gamma}{\delta} \left(\rho_2 \sigma - r - \frac{\beta}{\gamma} \right)^2 = \frac{1}{\gamma},$$

that is,

$$\left(1 - \frac{\gamma}{\delta} \rho_2^2 \right) \sigma^2 + 2 \frac{\gamma}{\delta} \left(r + \frac{\beta}{\gamma} \right) \rho_2 \sigma - \frac{\gamma}{\delta} \left(r + \frac{\beta}{\gamma} \right)^2 - \frac{1}{\gamma} = 0. \quad (4.9)$$

Note that (4.9) is a quadratic equation in σ whose discriminant is given by

$$\Delta(r) := 4 \frac{\gamma^2}{\delta^2} \left(r + \frac{\beta}{\gamma} \right)^2 \rho_2^2 + 4 \left(1 - \frac{\gamma}{\delta} \rho_2^2 \right) \left[\frac{\gamma}{\delta} \left(r + \frac{\beta}{\gamma} \right)^2 + \frac{1}{\gamma} \right]$$

$$= 4 \frac{\gamma}{\delta} \left(r + \frac{\beta}{\gamma} \right)^2 + 4 \frac{1}{\gamma} - 4 \frac{1}{\delta} \rho_2^2 \quad (4.10)$$

$$= \frac{4}{\delta} \left(\gamma r^2 + 2\beta r + \frac{\beta^2 + \delta}{\gamma} - \rho_2^2 \right)$$

$$= \frac{4}{\delta} \left(\gamma r^2 + 2\beta r + \alpha - \rho_2^2 \right). \quad (4.11)$$

Using (4.11), one can easily check that $r \mapsto \Delta(r)$ is a strictly convex quadratic function on $\mathbb{R}$ whose minimum value is given by

$$\min_{r \in \mathbb{R}} \Delta(r) = \frac{4}{\delta} \left(\frac{\delta}{\gamma} - \rho_2^2 \right). \quad (4.12)$$

Since $\rho_2 < \sqrt{\frac{\delta}{\gamma}}$ by assumption, we see that $\Delta(r) > 0$ for every $r \in \mathbb{R}$ so that the quadratic equation (4.9) has two distinct real solutions $\sigma_-(r), \sigma_+(r)$ given by (4.6). Moreover, by

([4.10](#)) and the assumption that $\rho_2 < \sqrt{\frac{\delta}{\gamma}}$, we have

$$\delta \left(\gamma r^2 + 2\beta r + \alpha - \rho_2^2 \right)$$
$$= \frac{\delta^2}{4} \Delta(r) > \gamma \delta \left(r + \frac{\beta}{\gamma} \right)^2 \geq \gamma^2 \left(r + \frac{\beta}{\gamma} \right)^2 \rho_2^2 = (\gamma r + \beta)^2 \rho_2^2, \qquad (4.13)$$

which implies that $\sigma_-(r) < 0$ and $\sigma_+(r) > 0$. The corresponding expected return values $\mu_-(r), \mu_+(r)$ given by ([4.7](#)) are calculated from the defining equation of $\mathcal{L}(r)$ so that

$$\mathcal{H} \cap \mathcal{L}(r) = \{ (\sigma_-(r), \mu_-(r)), (\sigma_+(r), \mu_+(r)) \}.$$

Next, we consider the three possible cases for $(\mathcal{M}(r))$. As a preparation, we first claim that every $(\bar{\sigma}, \bar{\mu}) \in \mathcal{H}_+$ with $\bar{\sigma} \geq \sigma_+(r)$ is also a feasible solution of $(\mathcal{M}(r))$. In other words, we claim that the set

$$\mathcal{S} := \{ (\bar{\sigma}, \bar{\mu}) \in \mathbb{R}_+ \times \mathbb{R} \mid \bar{\mu} = \mu(\bar{\sigma}), \bar{\sigma} \geq \sigma_+(r) \}$$

is a subset of the feasible region of $(\mathcal{M}(r))$, that is, $\rho_2 \bar{\sigma} - \bar{\mu} \leq r$ for every $(\bar{\sigma}, \bar{\mu}) \in \mathcal{S}$. Indeed, since $\rho_2 \leq \sqrt{\frac{\delta}{\gamma}}$, we have

$$\frac{d}{d\sigma} (\rho_2 \sigma - \mu(\sigma)) = \rho_2 - \frac{\sqrt{\frac{\delta}{\gamma}} \sigma}{\sqrt{\sigma^2 - \frac{1}{\gamma}}} \leq \sqrt{\frac{\delta}{\gamma}} \left(1 - \frac{\sigma}{\sqrt{\sigma^2 - \frac{1}{\gamma}}} \right) < 0 \qquad (4.14)$$

for every $\sigma > \frac{1}{\sqrt{\gamma}}$. Since we also have $\mu(\sigma_+(r)) \geq \mu_+(r)$, it follows that every $(\bar{\sigma}, \bar{\mu}) \in \mathcal{S}$ satisfies

$$r = \rho_2 \sigma_+(r) - \mu_+(r) \geq \rho_2 \sigma_+(r) - \mu(\sigma_+(r)) > \rho_2 \bar{\sigma} - \mu(\bar{\sigma}) = \rho_2 \bar{\sigma} - \bar{\mu}$$

so that it is feasible for $(\mathcal{M}(r))$. Hence, the claim follows.

(i) Suppose that $\rho_1 < \sqrt{\frac{\delta}{\gamma}}$. A standard exercise in calculus yields that

$$\lim_{\sigma_+(r) \leq \sigma \uparrow +\infty} (\rho_1 \sigma - \mu(\sigma)) = \lim_{\sigma_+(r) \leq \sigma \uparrow +\infty} \left(\rho_1 \sigma - \frac{\beta}{\gamma} - \sqrt{\frac{\delta}{\gamma} \sigma^2 - \frac{\delta}{\gamma^2}} \right) = -\infty.$$

Since the objective function diverges to $-\infty$ on $\mathcal{S}$, it follows that $(\mathcal{M}(r))$ is an unbounded problem with optimal value $-\infty$.

(ii) Suppose that $\rho_1 = \sqrt{\frac{\delta}{\gamma}}$. In this case, the limit evaluated in the previous case yields

$$\lim_{\sigma_+(r) \leq \sigma \uparrow +\infty} (\rho_1 \sigma - \mu(\sigma)) = -\frac{\beta}{\gamma}.$$

On the other hand, since the hyperbola $\mathcal{H}$ and its asymptote $\left\{ (\sigma, \mu) \in \mathbb{R}^2 \mid \sqrt{\frac{\delta}{\gamma}} \sigma - \mu = -\frac{\beta}{\gamma} \right\}$ do not intersect, there is no feasible solution (σ, μ) of $(\mathcal{M}(r))$ such that

$$\rho_1 \sigma - \mu = \sqrt{\frac{\delta}{\gamma}} \sigma - \mu = -\frac{\beta}{\gamma}.$$

Hence, the infimum in $(\mathcal{M}(r))$ is equal to $-\frac{\beta}{\gamma}$ but it is not attained by a feasible solution.

(iii) Suppose that $\rho_1 > \sqrt{\frac{\delta}{\gamma}}$. Note that the feasible region of $(\mathcal{M}(r))$ is a subset of that of $(\mathcal{M}_{\mathscr{A}})$. Hence, in view of Proposition 4.3, the unique optimal solution (σ^*, μ^*) of $(\mathcal{M}_{\mathscr{A}})$ is also the unique optimal solution of $(\mathcal{M}(r))$ if and only if it is feasible for $(\mathcal{M}(r))$, that is,

$$r^* = \rho_2 \sigma^* - \mu^* \le r,$$

where r^* is defined by (4.8).

Next, we show that $r^* \le r_0$, where r_0 is defined by (4.8). So we show that

$$\frac{\rho_1 \rho_2 \gamma - \delta}{\gamma \sqrt{\gamma \rho_1^2 - \delta}} - \frac{\beta}{\gamma} \le \frac{\rho_2}{\sqrt{\gamma}} - \frac{\beta}{\gamma},$$

which is equivalent to

$$\rho_1 \rho_2 \gamma - \delta \le \rho_2 \sqrt{\gamma} \sqrt{\gamma \rho_1^2 - \delta}. \tag{4.15}$$

If $\gamma \rho_1 \rho_2 - \delta \le 0$, then (4.15) holds trivially. Suppose that $\gamma \rho_1 \rho_2 - \delta > 0$. In this case, (4.15) is equivalent to

$$\gamma^2 \rho_1^2 \rho_2^2 + \delta^2 - 2\gamma \delta \rho_1 \rho_2 = (\gamma \rho_1 \rho_2 - \delta)^2 \le \gamma \rho_2^2 \left(\gamma \rho_1^2 - \delta \right) = \gamma^2 \rho_1^2 \rho_2^2 - \gamma \delta \rho_2^2,$$

which is equivalent to

$$\delta - 2\gamma \rho_1 \rho_2 + \gamma \rho_2^2 \le 0.$$

But the last inequality follows from the supposition and the assumption that $\rho_1 > \sqrt{\frac{\delta}{\gamma}} > \rho_2$ since

$$\delta - 2\gamma \rho_1 \rho_2 + \gamma \rho_2^2 \le \gamma \rho_1 \rho_2 - 2\gamma \rho_1 \rho_2 + \gamma \rho_2^2 = \gamma \rho_2 (\rho_2 - \rho_1) \le 0.$$

Consequently, (4.15) holds when $\gamma \rho_1 \rho_2 - \delta > 0$ as well. Hence, $r^* \le r_0$.

Finally, we consider the case $r < r^*$, that is, (σ^*, μ^*) is not feasible for $(\mathcal{M}(r))$. In this case, we prove that $(\sigma_+(r), \mu_+(r))$ is the unique optimal solution of $(\mathcal{M}(r))$. To that end, note that we have $r < r_0$ in this case so that

$$\rho_2 \sigma_+(r) - \mu_+(r) = r < r_0 = \rho_2 \sigma_0 - \mu_0 \le \rho_2 \sigma_+(r) - \mu_0.$$

This implies $\mu_+(r) > \mu_0$. In particular, $\mu_+(r) = \mu(\sigma_+(r))$. Next, let $(\bar{\sigma}, \bar{\mu})$ be a feasible solution of $(\mathcal{M}(r))$ with $(\bar{\sigma}, \bar{\mu}) \ne (\sigma_+(r), \mu_+(r))$. We first claim that $\bar{\sigma} > \sigma_+(r)$. To get a contradiction, suppose $\bar{\sigma} \le \sigma_+(r)$. By (4.14), $\sigma \mapsto \rho_2 \sigma - \mu(\sigma)$ is a decreasing function. Using this and the fact that $\bar{\mu} \le \mu(\bar{\sigma})$, we obtain

$$r \ge \rho_2 \bar{\sigma} - \bar{\mu} \ge \rho_2 \bar{\sigma} - \mu(\bar{\sigma}) \ge \rho_2 \sigma_+(r) - \mu_+(r) = r,$$

which yields $\rho_2 \bar{\sigma} - \bar{\mu} = r$ and $\bar{\mu} = \mu(\bar{\sigma})$. This implies $(\bar{\sigma}, \bar{\mu}) \in \mathcal{H} \cap \mathcal{L}(r)$ and hence $(\bar{\sigma}, \bar{\mu}) = (\sigma_+(r), \mu_+(r))$, which is a contradiction. Hence, the claim follows. On the other hand, using the assumption $\rho_1 > \sqrt{\frac{\delta}{\gamma}}$, we notice that

$$\frac{d}{d\sigma} (\rho_1 \sigma - \mu(\sigma)) = \rho_1 - \frac{\sqrt{\frac{\delta}{\gamma}} \sigma}{\sqrt{\sigma^2 - \frac{1}{\gamma}}} > 0 \quad \Leftrightarrow \quad \sigma > \frac{\rho_1}{\sqrt{\gamma \rho_1^2 - \delta}} = \sigma^*, \tag{4.16}$$

that is $\sigma \mapsto \rho_1 \sigma - \mu(\sigma)$ is a strictly increasing function for $\sigma > \sigma^*$. Moreover, we have $\bar{\sigma} > \sigma_+(r) > \sigma^*$. Indeed, the first inequality is by the previous claim. The second

inequality holds as otherwise, (σ^*, μ^*) would be feasible for $(\mathscr{M}(r))$ by the preparatory claim preceding the analysis of the three cases, which is excluded by the assumption $r < r^*$. Since we also have $\bar{\mu} \leq \mu(\bar{\sigma})$ and $\mu_+(r) = \mu(\sigma_+(r))$, it follows that

$$\rho_1 \bar{\sigma} - \bar{\mu} \geq \rho_1 \bar{\sigma} - \mu(\bar{\sigma}) > \rho_1 \sigma_+(r) - \mu(\sigma_+(r)) = \rho_1 \sigma_+(r) - \mu_+(r),$$

that is, $(\bar{\sigma}, \bar{\mu})$ is not optimal for $(\mathscr{M}(r))$. Hence, $(\sigma_+(r), \mu_+(r))$ is the unique optimal solution of $(\mathscr{M}(r))$.

$\square$

Theorem 4.5 *Let $r \in \mathbb{R}$ and suppose that $\rho_2 > \sqrt{\frac{\delta}{\gamma}}$. Then, the hyperbola $\mathcal{H}$ and the line $\mathcal{L}(r)$ intersect precisely at two points, $(\sigma_-(r), \mu_-(r))$ and $(\sigma_+(r), \mu_+(r))$ defined by (4.6), if and only if $r \leq r_-$ or $r \geq r_+$, where*

$$r_{\pm} := \frac{-\beta \pm \sqrt{\gamma \rho_2^2 - \delta}}{\gamma}. \tag{4.17}$$

In particular, it holds $\sigma_+(r) \leq \sigma_-(r) < 0$ if $r \leq r_-$, it holds $0 < \sigma_+(r) \leq \sigma_-(r)$ if $r \geq r_+$. Moreover, the points $(\sigma_-(r), \mu_-(r))$ and $(\sigma_+(r), \mu_+(r))$ are identical if and only if $r = r_-$ or $r = r_+$. The hyperbola $\mathcal{H}$ and the line $\mathcal{L}(r)$ do not intersect at all if and only if $r_- < r < r_+$. Consequently, $(\mathscr{M}(r))$ is feasible if and only if $r \geq r_+$.

Suppose that $r \geq r_+$. Then, one of the following cases holds for $(\mathscr{M}(r))$.

(i) *Suppose that $\rho_1 \leq \sqrt{\frac{\delta}{\gamma}}$. Then, $(\sigma_-(r), \mu_-(r))$ is the unique optimal solution of $(\mathscr{M}(r))$ and*

$$w_-(r) := w(\mu_-(r)) = \frac{1}{\gamma \rho_2^2 - \delta} \left[\left(\gamma r + \beta + \frac{\gamma \rho_2}{\delta} \sqrt{\delta(\gamma \rho_2^2 + 2\beta r + \alpha - \rho_2^2)} \right) C^{-1} m \right.$$
$$\left. + \left(-\beta r - \alpha + \rho_2^2 - \frac{\beta \rho_2}{\delta} \sqrt{\delta(\gamma \rho_2^2 + 2\beta r + \alpha - \rho_2^2)} \right) C^{-1} \mathbf{1} \right]$$

is the unique optimal solution of $(\mathscr{P}(r))$.

(ii) *Suppose that $\rho_1 > \sqrt{\frac{\delta}{\gamma}}$. Then, the unique optimal solution (σ^*, μ^*) of $(\mathscr{M}_{\mathscr{A}})$ is also the unique optimal solution of $(\mathscr{M}(r))$ and the corresponding portfolio w^* is the unique optimal solution of $(\mathscr{P}(r))$ if and only if $r \geq r^*$, where r^* is defined by (4.8).*

It holds $r_+ = r^*$ if $\rho_1 = \rho_2$ and $r_+ < r^*$ if $\rho_1 \neq \rho_2$. Suppose that $\rho_1 \neq \rho_2$ and $r_+ \leq r < r^*$. Then, one of the following cases holds for $(\mathscr{M}(r))$.

   (a) *If $\rho_1 < \rho_2$, then $(\sigma_-(r), \mu_-(r))$ is the unique optimal solution of $(\mathscr{M}(r))$ and $w_-(r)$ is the unique optimal solution of $(\mathscr{P}(r))$.*

   (b) *If $\rho_1 > \rho_2$, then $(\sigma_+(r), \mu_+(r))$ is the unique optimal solution of $(\mathscr{M}(r))$ and $w_+(r)$ is the unique optimal solution of $(\mathscr{P}(r))$.*

Proof As in the proof of Theorem 4.4, a point $(\sigma, \mu) \in \mathcal{H} \cap \mathcal{L}(r)$ satisfies (4.9), which is a quadratic equation in σ with discriminant $\Delta(r)$ given by (4.11). However, since $\rho_2 > \sqrt{\frac{\delta}{\gamma}}$, the minimum in (4.12) is strictly negative: $\min_{r \in \mathbb{R}} \Delta(r) < 0$. Moreover, we have $\Delta(r) = 0$ if and only if $r \in \{r_-, r_+\}$, where $r_\pm$ are defined by (4.17); $\Delta(r) < 0$ if and only if $r_- < r < r_+$; $\Delta(r) > 0$ if and only if $r < r_-$ or $r > r_+$. Hence, $\mathcal{H} \cap \mathcal{L}(r)$ is nonempty if and only if $r \leq r_-$ or $r \geq r_+$, and the intersection consists of $(\sigma_-(r), \mu_-(r))$, $(\sigma_+(r), \mu_+(r))$ in this

case. Mimicing the arguments for (4.13), this time with $\rho_2 > \sqrt{\frac{\delta}{\gamma}}$, gives

$$\delta(\gamma r^2 + 2\beta r + \alpha - \rho_2^2) < (\gamma r + \beta)^2 \rho_2^2.$$

It follows that $\sigma_+(r) \leq \sigma_-(r) < 0$ if $r \leq r_- \leq -\frac{\beta}{\gamma}$ and $0 < \sigma_+(r) \leq \sigma_-(r)$ if $r \geq r_+ \geq -\frac{\beta}{\gamma}$. The rest of the claims in the first paragraph of the theorem follows immediately.

For the rest of the proof, suppose that $r \geq r_+$. We consider the three possible cases for $(\mathcal{M}(r))$ next. As a preparation, we first show that $\mu(\sigma_-(r)) = \mu_-(r)$. To that end, it suffices to show that

$$\frac{\delta r + \beta \rho_2^2 + \rho_2\sqrt{\delta(\gamma r^2 + 2\beta r + \alpha - \rho_2^2)}}{\gamma \rho_2^2 - \delta} = \mu_-(r) \geq \mu_0 = \frac{\beta}{\gamma},$$

which is equivalent to

$$\delta \gamma r + \gamma \beta \rho_2^2 + \gamma \rho_2\sqrt{\delta(\gamma r^2 + 2\beta r + \alpha - \rho_2^2)} \geq \gamma \beta \rho_2^2 - \delta \beta,$$

that is,

$$\delta(\gamma r + \beta) + \gamma \rho_2\sqrt{\delta(\gamma r^2 + 2\beta r + \alpha - \rho_2^2)} \geq 0. \qquad (4.18)$$

On the other hand, since $r \geq r_+$, we have $\gamma r + \beta \geq \sqrt{\gamma \rho_2^2 - \delta} > 0$ from which (4.18) follows. Hence, $\mu(\sigma_-(r)) = \mu_-(r)$.

(i) Suppose that $\rho_1 \leq \sqrt{\frac{\delta}{\gamma}}$. Let $(\bar{\sigma}, \bar{\mu})$ be a feasible solution of $(\mathcal{M}(r))$ with $(\bar{\sigma}, \bar{\mu}) \neq (\sigma_-(r), \mu_-(r))$. We claim that $\bar{\sigma} \leq \sigma_-(r)$. To get a contradiction, suppose that $\bar{\sigma} > \sigma_-(r)$. Similar to (4.16), we notice that $\sigma \mapsto \rho_2\sigma - \mu(\sigma)$ is a strictly increasing function for $\sigma > \frac{\rho_2}{\sqrt{\gamma \rho_2^2 - \delta}}$ thanks to the assumption $\rho_2 > \sqrt{\frac{\delta}{\gamma}}$. On the other hand, since we assume that $r \geq r_+$, we have

$$\sigma_-(r) = \frac{(\gamma r + \beta)\rho_2 + \sqrt{\delta(\gamma r^2 + 2\beta r + \alpha - \rho_2^2)}}{\gamma \rho_2^2 - \delta} \geq \frac{(\gamma r_+ + \beta)\rho_2}{\gamma \rho_2^2 - \delta} = \frac{\rho_2}{\sqrt{\gamma \rho_2^2 - \delta}}.$$

Since we also have $\bar{\mu} \leq \mu(\bar{\sigma})$ and $\mu(\sigma_-(r)) = \mu_-(r)$, it follows that

$$r \geq \rho_2\bar{\sigma} - \bar{\mu} \geq \rho_2\bar{\sigma} - \mu(\bar{\sigma}) > \rho_2\sigma_-(r) - \mu(\sigma_-(r)) = \rho_2\sigma_-(r) - \mu_-(r) = r,$$

which is a contradiction. Hence, $\bar{\sigma} \leq \sigma_-(r)$. Moreover, we further have $\bar{\sigma} < \sigma_-(r)$ as otherwise $\bar{\sigma} = \sigma_-(r)$ would imply

$$r \geq \rho_2\bar{\sigma} - \bar{\mu} = \rho_2\sigma_-(r) - \bar{\mu} \geq \rho_2\sigma_-(r) - \mu(\bar{\sigma}) = \rho_2\sigma_-(r) - \mu_-(r) = r$$

so that $(\bar{\sigma}, \bar{\mu}) = (\sigma_-(r), \mu_-(r))$, which is a contradiction. On the other hand, similar to (4.14), we can argue that $\sigma \mapsto \rho_1\sigma - \mu(\sigma)$ is a strictly decreasing function for $\sigma > \frac{1}{\sqrt{\gamma}}$ thanks to the assumption $\rho_1 \leq \sqrt{\frac{\delta}{\gamma}}$. Consequently, $\bar{\sigma} < \sigma_-(r)$ implies

$$\rho_1\bar{\sigma} - \bar{\mu} \geq \rho_1\bar{\sigma} - \mu(\bar{\sigma}) > \rho_1\sigma_-(r) - \mu(\sigma_-(r)) = \rho_1\sigma_-(r) - \mu_-(r),$$

that is, $(\bar{\sigma}, \bar{\mu})$ is not optimal for $(\mathcal{M}(r))$. Hence, $(\sigma_-(r), \mu_-(r))$ is the unique optimal solution of $(\mathcal{M}(r))$.

(ii) Suppose that $\rho_1 > \sqrt{\frac{\delta}{\gamma}}$. In this case, as in the proof of (iii) of Theorem 4.7, we note that the unique optimal solution (σ^*, μ^*) of $(\mathscr{M}_{\mathscr{A}})$ is also the unique optimal solution of $(\mathscr{M}(r))$ if and only if $r \geq r^*$, where r^* is defined by (4.8). From the definitions, it is clear that $r_+ = r^*$ if $\rho_1 = \rho_2$. Suppose that $\rho_1 \neq \rho_2$. We first claim that

$$r_+ < r^*. \tag{4.19}$$

Indeed, supposing otherwise would yield $\sqrt{(\gamma\rho_1^2 - \delta)(\gamma\rho_2^2 - \delta)} \geq \rho_1\rho_2\gamma - \delta$, which is equivalent to $(\gamma\rho_1^2 - \delta)(\gamma\rho_2^2 - \delta) \geq (\rho_1\rho_2\gamma - \delta)^2$ as we have $\rho_1\rho_2 > \frac{\delta}{\gamma}$ by the assumptions on ρ_1, ρ_2. Further simplification would yield the inequality $0 \geq \gamma\delta(\rho_1 - \rho_2)^2$, which is a contradiction since $\rho_1 \neq \rho_2$ by supposition. Hence, the claim follows.

In view of (4.19), it remains to figure out the optimal solution of $(\mathscr{M}(r))$ under the condition that

$$r_+ \leq r < r^*, \tag{4.20}$$

which we assume for the rest of the proof.

(a) Let us assume that $\rho_1 < \rho_2$. We prove that $(\sigma_-(r), \mu_-(r))$ is the unique optimal solution of $(\mathscr{M}(r))$. To that end, let $(\bar{\sigma}, \bar{\mu})$ be a feasible solution of $(\mathscr{M}(r))$ such that $(\bar{\sigma}, \bar{\mu}) \neq (\sigma_-(r), \mu_-(r))$. Following the same arguments as in the proof of (i), one can check that $\bar{\sigma} < \sigma_-(r)$.

Next, we show $\sigma_-(r) < \sigma^*$. Note that $\sigma_-(r) < \sigma^*$ is equivalent to

$$\frac{(\gamma r + \beta)\rho_2 + \sqrt{\delta(\gamma r^2 + 2\beta r + \alpha - \rho_2^2)}}{\gamma\rho_2^2 - \delta} < \frac{\rho_1}{\sqrt{\gamma\rho_1^2 - \delta}},$$

that is,

$$\sqrt{\delta(\gamma r^2 + 2\beta r + \alpha - \rho_2^2)} < (\gamma r + \beta)\rho_2 - \frac{\rho_1(\gamma\rho_2^2 - \delta)}{\sqrt{\gamma\rho_1^2 - \delta}}. \tag{4.21}$$

However, (4.21) does not hold true when its right hand side is negative, that is, when

$$r \geq r_{++} := \frac{\rho_1(\gamma\rho_2^2 - \delta)}{\gamma\rho_2\sqrt{\gamma\rho_1^2 - \delta}} - \frac{\beta}{\gamma}.$$

On the other hand, note that

$$r^* < r_{++} \quad \Leftrightarrow \quad \frac{\rho_1(\gamma\rho_2^2 - \delta)}{\gamma\rho_2\sqrt{\gamma\rho_1^2 - \delta}} - \frac{\beta}{\gamma} > \frac{\rho_1\rho_2\gamma - \delta}{\gamma\sqrt{\gamma\rho_1^2 - \delta}} - \frac{\beta}{\gamma}$$

$$\Leftrightarrow \quad \gamma\rho_1\rho_2^2 - \rho_1\delta > \gamma\rho_1\rho_2^2 - \delta\rho_2$$

$$\Leftrightarrow \quad \rho_1 < \rho_2. \tag{4.22}$$

Hence, in view of (4.20), we always have $r < r_{++}$ so that (4.21) can be rewritten as

$$\delta(\gamma r^2 + 2\beta r + \alpha - \rho_2^2) < (\gamma r + \beta)^2\rho_2^2$$

$$+ \frac{\rho_1^2(\gamma\rho_2^2 - \delta)^2}{(\gamma\rho_1^2 - \delta)} - 2(\gamma r + \beta)\frac{\rho_1\rho_2(\gamma\rho_2^2 - \delta)}{\sqrt{\gamma\rho_1^2 - \delta}},$$

$$\tag{4.23}$$

which is equivalent to

$$0 < \gamma r^2 + 2\left(\beta - \frac{\gamma\rho_1\rho_2}{\sqrt{\gamma\rho_1^2 - \delta}}\right)r + \alpha + \frac{\rho_1^2(\gamma\rho_2^2 - \delta)}{\gamma\rho_1^2 - \delta} - \frac{2\beta\rho_1\rho_2}{\sqrt{\gamma\rho_1^2 - \delta}}.$$

(4.24)

The discriminant of the quadratic function of r in the right hand side of (4.24) is calculated as $\frac{4\delta^2}{\gamma\rho_1^2 - \delta} > 0$ so that this function has two distict real zeros and the smaller of these zeros is precisely r^*. Since we assume (4.20), (4.24) always holds and we have $\sigma_-(r) < \sigma^*$.

On the other hand, by (4.16), $\sigma \mapsto \rho_1\sigma - \mu(\sigma)$ is a strictly decreasing function for $\sigma < \sigma^*$. Hence, $\bar{\sigma} < \sigma_-(r) < \sigma^*$ implies

$$\rho_1\bar{\sigma} - \bar{\mu} \geq \rho_1\bar{\sigma} - \mu(\bar{\sigma}) > \rho_1\sigma_-(r) - \mu(\sigma_-(r)) = \rho_1\sigma_-(r) - \mu_-(r).$$

(4.25)

We conclude that $(\bar{\sigma}, \bar{\mu})$ is not optimal for $(\mathcal{M}(r))$. Hence, $(\sigma_-(r), \mu_-(r))$ is the unique optimal solution.

(b) Let us assume that $\rho_1 > \rho_2$. We prove that $(\sigma_+(r), \mu_+(r))$ is the unique optimal solution of $(\mathcal{M}(r))$. To that end, let $(\bar{\sigma}, \bar{\mu})$ be a feasible solution of $(\mathcal{M}(r))$ with $(\bar{\sigma}, \bar{\mu}) \neq (\sigma_+(r), \mu_+(r))$. We claim that $\bar{\sigma} \geq \sigma_+(r)$. To get a contradiction, suppose that $\bar{\sigma} < \sigma_+(r)$.

Similar to the proof of (i), we notice that $\sigma \mapsto \rho_2\sigma - \mu(\sigma)$ is a strictly decreasing function for $\sigma < \frac{\rho_2}{\sqrt{\gamma\rho_2^2 - \delta}}$. Next, we show that

$$\frac{(\gamma r + \beta)\rho_2 - \sqrt{\delta(\gamma r^2 + 2\beta r + \alpha - \rho_2^2)}}{\gamma\rho_2^2 - \delta} = \sigma_+(r) \leq \frac{\rho_2}{\sqrt{\gamma\rho_2^2 - \delta}},$$

(4.26)

which is equivalent to

$$(\gamma r + \beta)\rho_2 - \rho_2\sqrt{\gamma\rho_2^2 - \delta} \leq \sqrt{\delta(\gamma r^2 + 2\beta r + \alpha - \rho_2^2)}$$

(4.27)

It is easy to check that the left hand side of (4.27) is positive thanks to the assumption $r \geq r_+$. Hence, (4.27) is equivalent to

$$\rho_2^2\left(\gamma^2 r^2 + 2\beta r + \beta^2\right)$$
$$+\rho_2^2(\gamma\rho_2^2 - \delta) - 2\rho_2^2\sqrt{\gamma\rho_2^2 - \delta}(\gamma r + \beta) \leq \delta(\gamma r^2 + 2\beta r + \alpha - \rho_2^2),$$

that is,

$$(\gamma\rho_2^2 - \delta)\gamma r^2 + 2\left(\beta(\gamma\rho_2^2 - \delta) - \gamma\rho_2^2\sqrt{\gamma\rho_2^2 - \delta}\right)r$$
$$+\rho_2^2(\beta^2 + \gamma\rho_2^2) - \alpha\delta - 2\beta\rho_2^2\sqrt{\gamma\rho_2^2 - \delta} \leq 0.$$

(4.28)

One can check that the discriminant of the quadratic function of r on the left hand side of (4.28) is calculated as $4\delta^2(\gamma\rho_2^2 - \delta) > 0$ so that it has two distinct zeros

which are given as

$$-\frac{\beta}{\gamma} + \frac{\gamma\rho_2^2 \pm \delta}{\gamma\sqrt{\gamma\rho_2^2 - \delta}}.$$

Hence, (4.28) holds if and only if r is between these two zeros. Note that the smaller zero is equal to r_+ and we have $r \geq r_+$ by assumption. Next, we show that

$$r^* = -\frac{\beta}{\gamma} + \frac{\rho_1\rho_2\gamma - \delta}{\gamma\sqrt{\gamma\rho_1^2 - \delta}} < -\frac{\beta}{\gamma} + \frac{\gamma\rho_2^2 + \delta}{\gamma\sqrt{\gamma\rho_2^2 - \delta}}, \tag{4.29}$$

which is equivalent to

$$\sqrt{\gamma\rho_1^2 - \delta}(\gamma\rho_2^2 + \delta) > (\rho_1\rho_2\gamma - \delta)\sqrt{\gamma\rho_2^2 - \delta}. \tag{4.30}$$

By the assumptions $\rho_1 > \sqrt{\frac{\delta}{\gamma}}$, $\rho_2 > \sqrt{\frac{\delta}{\gamma}}$, the right hand side of (4.30) is positive so that (4.30) is equivalent to

$$(\gamma\rho_1^2 - \delta)(\delta^2\rho_2^4 + \delta^2 + 2\gamma\delta\rho_2^2) > (\gamma\rho_2^2 - \delta)(\gamma^2\rho_1^2\rho_2^2 + \delta^2 - 2\gamma\delta\rho_1\rho_2),$$

that is,

$$\delta\rho_1^2 + 3\gamma\rho_1^2\rho_2^2 - \gamma\rho_2^4 - 3\delta\rho_2^2 + 2\gamma\rho_1\rho_2^3 - 2\delta\rho_1\rho_2$$
$$= \delta\rho_1^2 - \gamma\rho_2^4 + 3\rho_2^2(\gamma\rho_1^2 - \delta) + 2\rho_1\rho_2(\gamma\rho_2^2 - \delta) > 0.$$

However, since $\rho_1 > \rho_2$, we have

$$\delta\rho_1^2 - \gamma\rho_2^4 + 3\rho_2^2(\gamma\rho_1^2 - \delta) + 2\rho_1\rho_2(\gamma\rho_2^2 - \delta)$$
$$> \delta\rho_2^2 - \gamma\rho_2^4 + 3\rho_2^2(\gamma\rho_2^2 - \delta) + 2\rho_2^2(\gamma\rho_2^2 - \delta)$$
$$= \rho_2^2(\delta - \gamma\rho_2^2) + 5\rho_2^2(\gamma\rho_2^2 - \delta)$$
$$= 4\rho_2^2(\gamma\rho_2^2 - \delta) > 0$$

so that (4.29) holds. Consequently, the assumption (4.20) guarantees that (4.26) holds.

Next, we show that $\mu(\sigma_+(r)) = \mu_+(r)$. To that end, it suffices to show that

$$\frac{\delta r + \beta\rho_2^2 - \rho_2\sqrt{\delta(\gamma r^2 + 2\beta r + \alpha - \rho_2^2)}}{\gamma\rho_2^2 - \delta} = \mu_+(r) \geq \mu_0 = \frac{\beta}{\gamma}, \tag{4.31}$$

which is equivalent to

$$\delta\gamma r + \gamma\beta\rho_2^2 - \gamma\rho_2\sqrt{\delta(\gamma r^2 + 2\beta r + \alpha - \rho_2^2)} \geq \gamma\beta\rho_2^2 - \delta\beta,$$

that is,

$$\sqrt{\delta}(\gamma r + \beta) \geq \gamma\rho_2\sqrt{\gamma r^2 + 2\beta r + \alpha - \rho_2^2}. \tag{4.32}$$

On the other hand, since $r \geq r_+$, we have $\gamma r + \beta \geq \sqrt{\gamma\rho_2^2 - \delta} > 0$ so that (4.32) is equivalent to

$$0 \geq -\delta(\gamma^2 r^2 + \beta^2 + 2\gamma\beta r) + \gamma^2\rho_2^2(\gamma r^2 + 2\beta r + \alpha - \rho_2^2) \tag{4.33}$$
$$= (\gamma\rho_2^2 - \delta)\gamma(\gamma r^2 + 2\beta r) - \delta\beta^2 + \gamma^2\rho_2^2(\alpha - \rho_2^2). \tag{4.34}$$

Note that the discriminant of the quadratic function of r in (4.34) is $4\gamma^3\rho_2^2(\gamma\rho_2^2 - \delta)^2 > 0$ so that it has two distinct zeros given by

$$-\frac{\beta}{\gamma} \pm \frac{\rho_2}{\sqrt{\gamma}}.$$

Hence, the inequality in (4.33) holds if and only if r is between these two zeros. Clearly, we have $-\frac{\beta}{\gamma} - \frac{\rho_2}{\sqrt{\gamma}} < -\frac{\beta}{\gamma} \le r_+ \le r$. On the other hand, note that the larger zero is equal to r_0 and it is easy to check that $r^* \le r_0$ if and only if $\rho_1 \ge \frac{\gamma\rho_2^2 + \delta}{2\gamma\rho_2}$ and we also have $\rho_2 > \frac{\gamma\rho_2^2 + \delta}{2\gamma\rho_2}$ since $\rho_2 > \sqrt{\delta}\gamma$. Consequently, the assumption $\rho_1 > \rho_2$ implies that $r^* \le r_0$ holds so that r is between the two zeros of the quadratic function in (4.34). Hence, the inequality in (4.33) holds, (4.31) holds and we have $\mu(\sigma_+(r)) = \mu_+(r)$.

Hence, $\bar{\sigma} < \sigma_+(r) \le \frac{\rho_2}{\sqrt{\gamma\rho_2^2 - \delta}}$ and $\mu(\sigma_+(r)) = \mu_+(r)$ imply that

$$r \ge \rho_2\bar{\sigma} - \bar{\mu} \ge \rho_2\bar{\sigma} - \mu(\bar{\sigma}) > \rho_2\sigma_+(r) - \mu(\sigma_+(r)) = \rho_2\sigma_+(r) - \mu_+(r) = r,$$

which is a contradiction. Hence, $\bar{\sigma} \ge \sigma_+(r)$. Moreover, we further have $\bar{\sigma} > \sigma_+(r)$ as otherwise $\bar{\sigma} = \sigma_+(r)$ would imply

$$r \ge \rho_2\bar{\sigma} - \bar{\mu} = \rho_2\sigma_+(r) - \bar{\mu} \ge \rho_2\sigma_+(r) - \mu(\bar{\sigma}) = \rho_2\sigma_+(r) - \mu_+(r) = r$$

so that $(\bar{\sigma}, \bar{\mu}) = (\sigma_+(r), \mu_+(r))$, which is a contradiction.

In view of (4.22), we have $r^* > r_{++}$. We show that $\sigma_+(r) > \sigma^*$. Note that $\sigma_+(r) > \sigma^*$ is equivalent to

$$\frac{(\gamma r + \beta)\rho_2 - \sqrt{\delta(\gamma r^2 + 2\beta r + \alpha - \rho_2^2)}}{\gamma\rho_2^2 - \delta} > \frac{\rho_1}{\sqrt{\gamma\rho_1^2 - \delta}},$$

that is,

$$\sqrt{\delta(\gamma r^2 + 2\beta r + \alpha - \rho_2^2)} < (\gamma r + \beta)\rho_2 - \frac{\rho_1(\gamma\rho_2^2 - \delta)}{\sqrt{\gamma\rho_1^2 - \delta}}. \tag{4.35}$$

Since $r^* > r_{++}$, the right hand side of (4.35) is strictly positive so that (4.35) is equivalent to (4.23) as well as to (4.24). Repeating the same analysis of the quadratic function in (4.24), we see that this function has two distinct real zeros and the smaller of these zeros is precisely r^*. Since we assume (4.20), (4.24) always holds and we have $\sigma_+(r) > \sigma^*$.

Consequently, $\bar{\sigma} > \sigma_+(r) > \sigma^*$ and the fact that $\sigma \mapsto \rho_1\sigma - \mu(\sigma)$ is strictly increasing for $\sigma > \sigma^*$ imply

$$\rho_1\bar{\sigma} - \bar{\mu} \ge \rho_1\bar{\sigma} - \mu(\bar{\sigma}) > \rho_1\sigma_+(r) - \mu(\sigma_+(r)) = \rho_1\sigma_+(r) - \mu_+(r),$$

that is, $(\bar{\sigma}, \bar{\mu})$ is not optimal for $(\mathcal{M}(r))$. Hence, $(\sigma_+(r), \mu_+(r))$ is the unique optimal solution of $(\mathcal{M}(r))$.

$\square$

Remark 4.6 Theorem 4.5 can be used to understand if solving a risk-risk problem is significantly different from solving a mean-risk problem in the following way. Suppose that the

risk measure in the constraint satisfies $\rho_2 > \sqrt{\frac{\delta}{\gamma}}$ and we have $r \geq r_+$. When the risk measure in the objective is simply the negative expected value, we have $\rho_1 = 0$. In this case, by (i) of Theorem 4.5, $w_-(r)$ is the unique optimal portfolio. However, the optimal portfolio does not change at all even if we use a nontrivial coherent risk measure in the objective such as value-at-risk or average value-at-risk as long as $\rho_1 \leq \sqrt{\frac{\delta}{\gamma}}$. Hence, the risk-risk problem with $\rho_1 \leq \sqrt{\frac{\delta}{\gamma}}$ is practically the same as the mean-risk problem. On the other hand, by (ii) of Theorem 4.5, for $\rho_1 > \sqrt{\frac{\delta}{\gamma}}$ with $r \geq r^*$, the risk-risk problem becomes significantly different from the mean-risk problem since the unique optimal portfolio is w^*, which depends on the choices of ρ_1 and ρ_2.

Theorem 4.7 *Let $r \in \mathbb{R}$ and suppose that $\rho_2 = \sqrt{\frac{\delta}{\gamma}}$. Then, the hyperbola $\mathcal{H}$ and the line $\mathcal{L}(r)$ intersect precisely at the single point $(\hat{\sigma}(r), \hat{\mu}(r))$ defined by*

$$\hat{\sigma}(r) := \frac{\gamma r^2 + 2\beta r + \alpha}{2\rho_2(\gamma r + \beta)}, \quad \hat{\mu}(r) := \frac{\alpha - \gamma r^2}{2(\gamma r + \beta)} \tag{4.36}$$

if and only if $r \neq -\frac{\beta}{\gamma}$. In particular, $\bar{\sigma}(r) < 0$ if $r < -\frac{\beta}{\gamma}$, it holds $\bar{\sigma}(r) > 0$ if $r > -\frac{\beta}{\gamma}$. The hyperbola $\mathcal{H}$ and the line $\mathcal{L}(r)$ do not intersect at all if and only if $r = -\frac{\beta}{\gamma}$. Consequently, $(\mathcal{M}(r))$ is feasible if and only if $r > -\frac{\beta}{\gamma}$.

Suppose that $r > -\frac{\beta}{\gamma}$. Then, one of the following cases holds for $(\mathcal{M}(r))$.

(i) *Suppose that $\rho_1 < \sqrt{\frac{\delta}{\gamma}}$. Then, $(\mathcal{M}(r))$ and $(\mathcal{P}(r))$ are unbounded problems with common optimal value $-\infty$.*
(ii) *Suppose that $\rho_1 = \sqrt{\frac{\delta}{\gamma}}$. Then, $(\mathcal{M}(r))$ and $(\mathcal{P}(r))$ have a common finite infimum that is equal to $-\frac{\beta}{\gamma}$ but the infimum is not attained by a feasible solution in both problems.*
(iii) *Suppose that $\rho_1 > \sqrt{\frac{\delta}{\gamma}}$. Then, the unique optimal solution (σ^*, μ^*) of $(\mathcal{M}_{\mathscr{A}})$ is also the unique optimal solution of $(\mathcal{M}(r))$ and the corresponding portfolio w^* is the unique optimal solution of $(\mathcal{P}(r))$ if and only if $r \geq r^*$, where r^* is defined by (4.8). If $r < r^*$, then $(\hat{\sigma}(r), \hat{\mu}(r))$ is the unique optimal solution of $(\mathcal{M}(r))$ and*

$$\hat{w}(r) := w(\hat{\mu}) = \frac{1}{2\delta} \left[\left(\frac{\delta}{\gamma r + \beta} - \gamma r - \beta \right) C^{-1} m + \left(\alpha + \frac{\gamma r(\beta r + \alpha)}{\gamma r + \beta} \right) C^{-1} \mathbf{1} \right]$$

is the unique optimal solution of $(\mathcal{P}(r))$.

Proof As in the proof of Theorem 4.4, a point $(\sigma, \mu) \in \mathcal{H} \cap \mathcal{L}(r)$ satisfies (4.9), which reduces to a linear equation in σ as we have $\rho_2 = \sqrt{\frac{\delta}{\gamma}}$. Suppose that $r \neq -\frac{\beta}{\gamma}$. Then, the unique solution of this equation is given by

$$\hat{\sigma}(r) = \frac{r + \frac{\beta}{\gamma}}{2\rho_2} + \frac{\frac{1}{\gamma}}{2\rho_2 \frac{\gamma}{\delta} \left(r + \frac{\beta}{\gamma} \right)} = \frac{(\gamma r + \beta)^2 + \delta}{2\rho_2 \gamma (r\gamma + \beta)} = \frac{\gamma r^2 + 2\beta r + \alpha}{2\rho_2(\gamma r + \beta)}. \tag{4.37}$$

The corresponding mean value of the point is given by

$$\hat{\mu}(r) = \rho_2 \hat{\sigma} - r = \frac{\gamma r^2 + 2\beta r + \alpha - 2(\gamma r + \beta)r}{2(\gamma r + \beta)} = \frac{\alpha - \gamma r^2}{2(\gamma r + \beta)}.$$

Hence, $\mathcal{H} \cap \mathcal{L}(r) = \left\{ (\hat{\sigma}(r), \hat{\mu}(r)) \right\}$ as defined by (4.36). Moreover, from the third expression in (4.37), it is clear that $\hat{\sigma}(r) < 0$ if $r < -\frac{\beta}{\gamma}$ and $\hat{\sigma}(r) > 0$ if $r > -\frac{\beta}{\gamma}$. If $r = -\frac{\beta}{\gamma}$, then (4.9) has no solution so that $\mathcal{H} \cap \mathcal{L}(r) \neq \emptyset$. It follows that $(\mathcal{M}(r))$ is feasible if and only if $r > -\frac{\beta}{\gamma}$.

For the rest of the proof, suppose that $r > -\frac{\beta}{\gamma}$. Note that an analogue of the preparatory claim in the proof of Theorem 4.4 can be shown here with the same arguments: every $(\bar{\sigma}, \bar{\mu}) \in \mathcal{H}_+$ with $\bar{\sigma} \geq \hat{\sigma}(r)$ is also a feasible solution of $(\mathcal{M}(r))$. Similarly, (i) and (ii) here can be shown here by repeating the same arguments as in the proofs of (i) and (ii) of Theorem 4.4. Hence, we consider only the case $\rho_1 > \sqrt{\frac{\delta}{\gamma}}$ here. As in the proof of (iii) of Theorem 4.7, we note that the unique optimal solution (σ^*, μ^*) of $(\mathcal{M}_{\mathscr{A}})$ is also the unique optimal solution of $(\mathcal{M}(r))$ if and only if $r \geq r^*$, where r^* is defined by (4.8). Since $\rho_1 > \sqrt{\frac{\delta}{\gamma}} = \rho_2$, we have $r^* > -\frac{\beta}{\gamma}$. Suppose that $r < r^*$. In this case, $(\hat{\sigma}(r), \hat{\mu}(r))$ is the unique optimal solution of $(\mathcal{M}(r))$. This can be shown using similar arguments as in the proof of (ii)b. of Theorem 4.5. To avoid repetitions, the details are omitted. $\qquad \square$

5 Conclusion

In this paper, we look at the static portfolio optimization problem with two coherent risk measures. We consider the case where the asset returns take arbitrary joint distributions, and characterize the optimal portfolio through the associated dual problem. The dual problem is an infinite-dimensional convex optimization problem with a linear objective function and a linear constraint besides the convex constraints on some dual variables. We detect some special cases where the dual problem reduces to a linear programming problem. In the second part of the paper, under the restriction that asset returns are jointly Gaussian, we identify all parameter configurations under which an optimal portfolio exists and provide an explicit formula for it.

References

1. Artzner, P., Delbaen, F., Eber, J.-M., Heath, D.: Coherent measures of risk. Math. Finance **9**(3), 203–228 (1999)
2. Barahona, F., Anbil, R.: The volume algorithm: producing primal solutions with a subgradient method. Math. Program. **87**(3), 385–399 (2000)
3. Borwein, J.M., Lewis, A.S.: Partially finite convex programming, Part I: quasi relative interiors and duality theory. Math. Program. **57**(1), 15–48 (1992)
4. Capiński, M., Zastawniak, T.: Mathematics for Finance, 2nd edn. Springer, London (2011)
5. Föllmer, H., Schied, A.: Stochastic Finance: An Introduction in Discrete Time. De Gruyter Textbook Series, 2nd edn. De Gruyter, Berlin (2016)
6. Glück, J., Weber, M.R.: Almost interior points in ordered Banach spaces and the long-term behaviour of strongly positive operator semigroups. arXiv:1901.03306 (2019)
7. Hale, J.K., Magalhães, L.T., Oliva, W.: Dynamics in Infinite Dimensions, 2nd edn. Springer, New York (2002)
8. Kaina, M., Rüschendorf, L.: On convex risk measures on L^p-spaces. Math. Methods Oper. Res. **69**, 475–95 (2009)

9. Landsman, Z.: Minimization of the root of a quadratic functional under an affine equality constraint. J. Comput. Appl. Math. **216**, 319–327 (2008)

10. Landsman, Z., Makov, U.: Minimization of a function of a quadratic functional with application to optimal portfolio selection. J. Optim. Theory Appl. **170**, 308–322 (2016)

11. Larsson, T., Patriksson, M., Strömberg, A.: Ergodic, primal convergence in dual subgradient schemes for convex programming. Math. Program. **86**(2), 283–312 (1999)

12. Markowitz, H.: Portfolio selection. J. Finance **7**(1), 77–91 (1952)

13. Merton, R.C.: An analytic derivation of the efficient portfolio frontier. J. Financ. Quant. Anal. **74**, 1851–1872 (1972)

14. Owadally, I.: An improved closed-form solution for the constrained minimization of the root of a quadratic functional. J. Comput. Appl. Math. **236**, 4428–4435 (2011)

15. Rockafellar, R.T., Uryasev, S.: Optimization of conditional value-at-risk. J. Risk **2**(3), 21–41 (2000)

16. Rockafellar, R.T., Uryasev, S.: Conditional value-at-risk for general loss distributions. J. Bank. Finance **26**(7), 1443–1471 (2002)

17. Sion, M.: On general minimax theorems. Pac. J. Math. **8**(1), 171–176 (1958)

18. Zalinescu, C.: Convex Analysis in General Vector Spaces. World Scientific, Singapore (2002)