Tight Bounds for Some Risk Measures, with Applications to Robust Portfolio Selection

Li CHEN * Simai HE † Shuzhong ZHANG ‡

July 17, 2010

Abstract

In this paper we develop tight bounds on the expected values of several risk measures that are of interest to us. This work is motivated by the robust optimization models arising from portfolio selection problems. Indeed the whole paper is centered around robust portfolio models and solutions. The basic setting is to find a portfolio which maximizes (respectively minimizes) the expected utility (respectively disutility) values, in the midst of infinitely many possible ambiguous distributions of the investment returns fitting the given mean and variance estimations. First, we show that the single-stage portfolio selection problem within this framework, whenever the disutility function is in the form of *Lower Partial Moments* (LPM), or *Conditional Value at Risk* (CVaR), or *Value-at-Risk* (VaR), can be solved analytically. The results lead to the solutions for single-stage robust portfolio selection models. Furthermore, the results also lead to a multi-stage *Adjustable Robust Optimization* (ARO) solution when the disutility function is the second order LPM. Exploring beyond the confines of convex optimization, we also consider the so-called S-shaped value function, which plays a key role in the prospect theory of Kahneman and Tversky. The non-robust version of the problem is shown to be NP-hard in general. However, we present an efficient procedure for solving the robust counterpart of the same portfolio selection problem. In this particular case, the consideration of the robustness actually helps to reduce the computational complexity. Finally, we consider the situation whereby we have some additional information about the chance that a quadratic function of the random distribution reaches a certain threshold. That information helps to further reduce the ambiguity in the robust model. We show that the robust optimization problem in that case can be solved by means of Semidefinite Programming (SDP), if no more than two additional chance inequalities are to be incorporated.

Keywords: Portfolio selection, robust optimization, S-shaped function, Chebyshev inequality.

*Department of Systems Engineering and Engineering Management, The Chinese University of Hong Kong, Shatin, Hong Kong. Email: lchen@se.cuhk.edu.hk

†Department of Management Sciences, City University of Hong Kong, Kowloon Tong, Hong Kong. Email: simaihe@cityu.edu.hk

‡Department of Systems Engineering and Engineering Management, The Chinese University of Hong Kong, Shatin, Hong Kong. Email: zhang@se.cuhk.edu.hk. Research supported by Hong Kong RGC Grant CUHK418406.

1 Introduction

The recent financial turmoil certainly has reminded us the role played by risk management. Many investors have learned in a hard way that the stability of the investment return does matter, even at the expenses of some occasional loss of the performance. In Operations Research, the relevant keyword in this context is *robust optimization*, which has been rapidly developed since the pioneering work of Ben-Tal and Nemirovski [5]. On the downside, a possible price one has to pay for the nice *robust* solution is to overcome a computational hurdle: the so-called robust counterparts are typically infinite-dimensional and, unless proper care is chosen in the choice of uncertainty set, the problem may be intractable. Therefore, careful modeling plays a decisive role in robust optimization. This paper aims to present a few more successful cases in this difficult terrain. In particular, we are concerned with robust portfolio selection. Recall that the usual portfolio selection is to maximize (respectively minimize) the expected utility value (respectively disutility value), subject to some physical constraints, under all prospects of the investment. A most famous disutility function is arguably the convex quadratic function, because the minimization of its expected value leads to the mean-variance paradigm of Markowitz [33] and Roy [48]. Since the variance does not differentiate the gain from the loss, Markowitz [34] later proposed to use the *semivariance* instead. To better suit different risk profiles of the investors, Bawa [2] and Fishburn [15] introduced a class of downside risk measure known as the *lower partial moment* (LPM):

$$\mathrm{LPM}_m(r) = \mathsf{E}[(r - X)_+{}^m],$$

where X is the asset return, and r is the return on a benchmark index such as the risk-free rate of return, and m is a parameter, which can take any non-negative value to model the risk attitude of an investor. Specifically, if $m = 0$, then LPM_0 is nothing but the probability of the asset return falling below the benchmark index; if $m = 1$, then LPM_1 is the expected shortfall of the investment, falling below the benchmark index; if $m = 2$, then LPM_2 is an analog of the semi-variance, where, however, the deviation is in reference to the benchmark return instead of the mean.

Nevertheless, not everyone in the field is convinced that the scheme of maximizing (minimizing) the expected value of a concave (convex) utility (disutility) function is all that matters to investors. Kahneman and Tversky [26], for instance, developed an alternative in 1979, known as the *prospect theory*, in which they hypothesized that usually an investor would have a reference point in mind. Judged by the reference point, a loss of a given magnitude matters more than a gain of the same magnitude. Furthermore, the fact that the perception of change in wealth decreases with its distance from the reference point (termed *diminishing sensitivity*) can be modeled by a value function that is concave for gains and convex for losses. Such function is also called the *S-shaped* value function. Tversky and Kahneman [53] established that if the so-called preference homogeneity holds, then the value function of the prospect theory has the following power form:

$$u(x) = \begin{cases} x^\alpha, & x \geq 0, \\ -\lambda(-x)^\beta, & x \leq 0, \end{cases}$$

with the loss aversion implying that $0 < \alpha \leq \beta < 1$ and $\lambda \geq 1$. While lower partial moments or expected S-shaped utility are used to measure the expected loss or utility for some given distribution, Rockafellar and Uryasev [46] proposed to evaluate the risk of an investment by computing the expected loss in the worst $q\%$ of the distribution, known as the *conditional value-at-risk* (CVaR).

Speaking of different disutility functions and their expected value under various assumptions on the underlying distribution, there is an interesting recent finding. Our study [10] showed that if the return of the investment follows the so-called *Q-radial distribution*, then $\text{LPM}_m(r)$ will lead to the same efficient portfolio as the standard mean-variance model, and *vice versa*. The assumptions on the distribution, however, are arguably always subjective. Estimation on the moments of the assets' returns using the historical data, on the other hand, may be considered more objective measurements. For this reason, it is natural to rely on the knowledge of the moments estimations, rather than on the assumption of the entire distribution. Estimating probability bounds using the information of a few estimated moments is a common practice; in the univariate case, e.g. the famous probability inequalities such as the Chebyshev inequality (using the first and second moments) and the Markov inequality (using the first moment) are exactly of this type. For the multivariate case, Zuluaga and Pena [56] derived a special class of generalized Tchebycheff inequalities by transforming them within the framework of conic programming. Beyond the probability bounds, the moment bounds of piece-wise linear utility functions are also very useful, due to the applications in finance and supply chain management. For the simplest two-piece linear utility function $f(X) = \max\{0, X - z\}$, Jensen's inequality can be used to derive an upper bound based on the first and second moments. Scarf [49] used the bound in a min-max newsvendor model wherein X denotes the random demand for a product and z denotes the order quantity. Likewise, Lo [30] used the bound on a call option price where X denotes the stock price and z denotes the strike price. Natarajan and Zhou [38] obtained the explicit and tight upper bound for three-piece linear utility function using the mean and variance information. All these bounds are based on a univariate random variable. Popescu [44] proved that the problem of evaluating the worst case expected utility by optimizing over an n-variate distribution can be in fact reduced to optimization over a univariate distribution with some appropriate mean and variance. Based on this observation, Natarajan, Sim and Uichanco [37] showed that the upper bound for a class of piecewise linear concave utility functions can be computed by solving a single compact second order cone program. They also derived the closed form expression for the worst-case CVaR and its generalizations. Historically, the approach of using the moment cone, via the conic duality theory, to obtain various bounds by Semidefinite Programs (SDP), was due to Nesterov [39], and was formalized by Bertsimas and Popescu [7], and Ben-Tal and Nemirovski [6]. In some cases, such bounds are even explicitly given. For the latter cases, we mention Popescu [42] for a closed-form bound for $\text{LPM}_0(r)$ when up to the third order moments are known, and He, Zhang and Zhang [22] for a closed-form and tight upper bound for $\text{LPM}_0(r)$ when the first, second and fourth moments are known. Since the moment information are sometimes estimated based on the limited historical data, Delage and Ye [11] demonstrated that, disregarding the uncertainty in these estimates can lead to taking poor decisions; furthermore, they proposed to use sample data to help derive confidence regions for the mean and covariance matrix to deal with the ambiguous distributions. In order to reduce the conservativeness of the robust model, Ben-Tal, Bertsimas and Brown [3] proposed a softened framework for robust optimization that relaxes the standard notion of robustness by allowing the decision-maker to vary the protection level in a smooth way across the uncertainty set.

Portfolio selection is the problem of allocating capital across a set of assets in a way that maximizes some measure of performance for a given probability distribution. If the underlying distribution is actually ambiguous then it is natural to consider a *robust* portfolio selection model. In particular, the ambiguous distribution set may be described by the knowledge of its support set and/or its first few moments. There are a number of recent papers along this line, including Natarajan, Pachamanova, and Sim [36], and Delage and Ye [11]. On a different front, we remark that robust optimization in multiple stage setting is in general difficult to model. The so-called adjustable robust optimization (ARO) is one such attempt; see [4, 20]. However, how to incorporate the recourse actions when the time progresses and new information arrives is a hard problem in general.

As we mentioned before, this paper is about robust portfolio selection. In particular, the certainty regarding the ambiguous distribution is its mean and covariance matrix, and possibly some additional knowledge concerning the probability of some projections of the distribution. Throughout the paper, we study various robust portfolio selection models that involve risk measures such as the lower partial moments LPM_m with $m = 0, 1, 2$, the conditional value-at-risk (CVaR), the Value-at-Risk, nondecreasing convex-concave S-shaped value function, and concave utility function with convex, concave or concave-convex derivative. Moreover, we study single-stage decision models as well as multi-stage adjusted (dynamic) robust decision models. Specifically, in Subsection 2.1 we presented a set of tight bounds for the expected value of the lower partial moments and the CVaR, under the condition that the mean and the covariance of the distribution are given. In the rest of Section 2, the results are used to develop various robust portfolio selection models. In Section 3, we give an *adjustable robust optimization* (ARO) solution for the multi-stage portfolio selection problem when the disutility function is the second order lower partial moment. In Section 4, we proceed to consider portfolio selection using the so-called S-shaped value function, which plays an important role in the prospect theory developed by Kahneman and Tversky [26]. Our findings are as follows. First, finding a portfolio that maximizes a given power S-shaped function, when the entire distribution is known (for example to be normally distributed), is in general NP-hard. Surprisingly, we show that finding a *robust* portfolio in this context can be reduced to searching along a single parameter; hence it can be done efficiently. Moreover, we generalize the result to more general value functions, such as monotone convex functions and concave functions with convex, concave or concave-convex derivative. In Section 5, we move on to consider the robust optimization models whereby some additional information regarding the distribution is known. For instance, in addition to the mean and the covariance matrix, we also know that the probability for a quadratic function of the random vector above a certain threshold value is bounded by a known constant. Such robust models can also be solved using SDP. This has some immediate implications. For instance, this implies that if we have some probability estimation of a random variable, then this information can be used to estimate the probability of a *correlated* random variable. In Section 6, we show numerically the relevance of the results that we have developed in the paper.

2 Robust Portfolio Selection Based on the Lower Partial Moments

We shall start our discussion by considering tight bounds on the probabilities and higher order lower partial moments (LPM), using the information about the mean and the covariance of the underlying distribution. Such bounds will naturally lead to robust portfolio optimization models as we shall see later.

Let us introduce some notations first. For any given positive semidefinite matrix $\Gamma \in S^n_{++}$ and vector $\mu \in \Re^n$, we denote $D = \{\pi \mid \ \mathsf{E}_\pi[\xi] = \mu, \mathrm{Cov}_\pi[\xi] = \Gamma \succ 0 \ \}$ to stand for the set of probability distributions with mean μ and covariance Γ. And we also denote $X \sim (\mu, \Gamma)$ to represent the fact that the random vector X belongs to the set whose elements have mean μ and covariance matrix Γ. Since it is a one-to-one correspondence between random variable and probability distribution, we will use the notations $\pi \in D$ and $X \sim (\mu, \Gamma)$ interchangeably in our discussion.

2.1 The Univariate Cases

The case for univariate random variable is a classical one. We shall, however, start with this simple case for completeness. Observe that $\mathrm{LPM}_0(r)$ measures the probability that a random return falls below the target r. The well-known Chebyshev-Cantelli inequality [9] presents its tight moment upper bound exactly.

Lemma 2.1. *It holds that*

$$\sup_{X \sim (\mu,\, \sigma^2)} \mathrm{Prob}\{X \le r\} = \sup_{X \sim (\mu,\, \sigma^2)} \mathrm{LPM}_0(r) = \begin{cases} \frac{1}{1+(r-\mu)^2/\sigma^2}, & \text{if } r < \mu, \\ 1, & \text{if } r \ge \mu. \end{cases}$$

$\mathrm{LPM}_1(r)$ is the expected shortfall of X below the benchmark r. Its upper bound can be derived by Jensen's inequality and has been widely applied: Scarf [49] first introduced this bound to simplify a min-max newsvendor model; Lo [30] used the result to bound European options. Natarajan and Zhou [38] generalized this result to the expected value of any three-piece linear function.

Lemma 2.2. *It holds that*

$$\sup_{X \sim (\mu,\, \sigma^2)} \mathsf{E}[(r - X)_+] = \frac{r - \mu + \sqrt{\sigma^2 + (r-\mu)^2}}{2}.$$

As we will see later, the above measurement of risk is highly related to Conditional Value-at-Risk (CVaR). Given the mean and variance, the worst-case CVaR can also be found explicitly; see [38].

$\mathrm{LPM}_2(r)$ is an analog of the semi-variance, where the reference is to the benchmark return r instead of to the mean. A tight bound on $\mathrm{LPM}_2(r)$ can be established by using the Jensen inequality.

Proposition 2.3. *It holds that*

$$\sup_{X \sim (\mu,\, \sigma^2)} \mathsf{E}[(r - X)_+^2] = [(r - \mu)_+]^2 + \sigma^2.$$

Proof. For any $X \sim (\mu, \sigma^2)$, by Jensen's inequality we have

$$\begin{aligned} \mathsf{E}[(r - X)_+^2] &= \mathsf{E}[(r - X)^2] - \mathsf{E}[(r - X)_-{}^2] \\ &\le \mathsf{E}[(r - X)^2] - ([\mathsf{E}(r - X)]_-)^2 \\ &= \sigma^2 + (r - \mu)^2 - [(r - \mu)_-]^2 \\ &= \sigma^2 + [(r - \mu)_+]^2. \end{aligned}$$

To show the tightness of the bound, consider a sequence of distributions

$$X_n = \begin{cases} \mu + \frac{\sigma}{\sqrt{n-1}}, & \text{with probability } \frac{n-1}{n}, \\ \mu - \sqrt{n-1} \cdot \sigma, & \text{with probability } \frac{1}{n}. \end{cases}$$

It is easy to check that $X_n \sim (\mu, \sigma^2)$, and

$$\mathsf{E}[(r - X_n)_+^2] \longrightarrow \sigma^2 + [(r - \mu)_+]^2,$$

as $n \longrightarrow \infty$. This indicates that the upper bound is indeed tight. □

Remark that $\sup_{X \sim (\mu,\, \sigma^2)} \mathsf{E}[(r - X)_+^m] = +\infty$ for any $m > 2$. This is evident by observing $\mathsf{E}[(r - X_n)_+^m] \to \infty$ where the sequence X_n is taken as the one in the proof of Proposition 2.3. Therefore, we shall only be concerned with the lower partial moments problem LPM_0, LPM_1 and LPM_2 in this paper.

2.2 Portfolio Selection with LPM as Risk Measure

In this section, we will discuss several robust downside risk models. Let us consider the lower partial moments $\text{LPM}_m(r)$, $m = 0, 1, 2$, Conditional Value-at-Risk, and Value-at-Risk as possible risk measures. If the parameters are ambiguous, then the tight upper bounds achieved in last section can be used as the worst-case objective. In fact, for up to two moments, all the models have explicit optimal solutions.

We assume that the first two moments are known, and the entire distribution is otherwise completely free. The portfolio models are as follows. Denote the n financial assets as $S_j,\ j = 1, ..., n$. We assume that the total invested wealth is 1, i.e., $x^{\mathrm{T}} e = 1$, where $x = (x_1, \cdots, x_n)$ is the invested wealth on each asset. We denote by ξ_j the random return rate of the jth financial asset, $j = 1, ..., n$. The benchmark return rate is denoted as $r > 0$. Here the risk measures are the lower partial moments $\text{LPM}_m(r)$. The portfolio selection models (P_m) are formulated as

$$\begin{array}{lll} (P_m) & \min_x & \mathsf{E}[(r - x^{\mathrm{T}} \xi)_+^m], \\ & \text{s.t.} & x^{\mathrm{T}} e = 1\ , \end{array}$$

where 'e' stands for the vector of all-ones with an appropriate dimension, $m = 0$, 1, or 2 represents the degree of the investor's risk aversion. Let us make no assumption on the distribution ξ. Instead, we assume that we have some knowledge regarding the statistical properties of ξ. In this particular case, we assume that we know the first two moments of ξ. In any case, for a given portfolio x, our risk measure is

$$H_m(x;\xi) = \mathsf{E}[(r - x^{\mathrm{T}}\xi)_+^m],$$

the corresponding robust portfolio selection model may be written as

$$\begin{array}{rcl} RP_m & = & \min_x \quad \sup_{\xi\sim(\mu,\Gamma)} \mathsf{E}[(r - \xi^{\mathrm{T}}x)_+^m] \\ & & \text{s.t.} \qquad x^{\mathrm{T}}e = 1 \end{array}$$

where $0 \le m \le 2$. The bounds developed in Section 2.1 become attractive, since these are tight bounds and they can be used to get an explicit expression for the objective function in RP_m. Of course, the bounds in Section 2.1 are good only for univariate distributions, while the ambiguous distribution set D in RP_m involves *multi-dimensional distributions*. For any given ξ, we have all the information about the moments for the distribution $\xi^{\mathrm{T}}x$, once we know the moments of ξ. The opposite is in general not true. Fortunately, if we speak only about the first two moments, then there is actually no loss of any information. To be precise, let us consider two sets

$$\begin{array}{rcl} A & := & \{a^{\mathrm{T}}\xi \mid \mathsf{E}[\xi] = \mu,\ \mathrm{Cov}(\xi) = \Gamma\ \}, \\ B & := & \{\eta \mid \mathsf{E}[\eta] = a^{\mathrm{T}}\mu,\ \mathrm{Var}(\eta) = a^{\mathrm{T}}\Gamma a\ \}. \end{array}$$

Evidently, $A \subseteq B$. The following lemma asserts that the opposite relationship also holds.

Lemma 2.4. *For any $a \neq 0 \in \Re^n$, it holds that $A = B$.*

Proof. We shall need only to show $B \subseteq A$. For any fixed $a \neq 0$, let us take an arbitrary $\eta \in B$. We then construct

$$\xi = \frac{C^{\frac{1}{2}}\beta}{\sqrt{a^{\mathrm{T}}\Gamma a}} + \frac{(\eta - a^{\mathrm{T}}\mu)\Gamma a}{a^{\mathrm{T}}\Gamma a} + \mu,$$

where $C := (a^{\mathrm{T}}\Gamma a)\Gamma - \Gamma a a^{\mathrm{T}}\Gamma$, and $\beta \sim N(0, I_n)$ is independent of η. Now $C \succeq 0$ and $a^{\mathrm{T}}Ca = 0$, and so $a^{\mathrm{T}}C^{\frac{1}{2}} = (0, \cdots, 0)$. Hence $a^{\mathrm{T}}\xi = \eta$, $\mathsf{E}[\xi] = \mu$, $\mathrm{Cov}(\xi) = \Gamma$. In other words, $\eta = a^{\mathrm{T}}\xi \in A$. Therefore $A = B$. □

Remark that the same equivalence result was established in Popescu [44] by a different method.

In light of the equivalence between the two sets, we have

$$\sup_{\xi\sim\,(\mu,\,\Gamma)} \mathsf{E}[(r - \xi^{\mathrm{T}}x)_+^m] = \sup_{\zeta\,\sim\,(x^{\mathrm{T}}\mu,\, x^{\mathrm{T}}\Gamma x)} \mathsf{E}[(r - \zeta)_+^m].$$

Hence, the univariate moments bounds in Section 2.1 can be applied directly. This makes it possible to derive explicit solutions, which we shall present in the theorem below, albeit for the cases $m = 0, 1, 2$ only. Let us first denote

$$\begin{array}{lll} c_0 := e^{\mathrm{T}}\Gamma^{-1}e, & c_1 := e^{\mathrm{T}}\Gamma^{-1}\mu, & c_2 := \mu^{\mathrm{T}}\Gamma^{-1}\mu, \\ b_0 := \frac{c_0}{c_0c_2 - c_1^2}, & b_1 := \frac{c_1}{c_0c_2 - c_1^2}, & b_2 := \frac{c_2}{c_0c_2 - c_1^2}. \end{array}$$

Theorem 2.5. *Suppose* $\Gamma \succ 0$. *Consider the optimization problem*

$$\begin{array}{rcll} v(RP_m) & := & \min_x & \sup_{\zeta \,\sim\, (x^{\mathrm{T}}\mu,\, x^{\mathrm{T}}\Gamma x)} \; \mathsf{E}[(r-\zeta)_+^m] \\ & & \text{s.t.} & x^{\mathrm{T}}e = 1. \end{array}$$

For the cases $m = 0, 1, 2$ *we have the following explicit solutions:*

(a). If $b_1 \geq rb_0$, *then*

$$\begin{aligned} v(RP_0) &= \frac{1}{1 + (\mu - re)^{\mathrm{T}}\Gamma^{-1}(\mu - re)}, \\ x^*_{RP_0} &= \frac{\Gamma^{-1}(\mu - re)}{e^{\mathrm{T}}\Gamma^{-1}\mu - r \cdot e^{\mathrm{T}}\Gamma^{-1}e}. \end{aligned}$$

Else if $b_1 < rb_0$, *then* $v(RP_0) = \frac{1}{1+\frac{1}{b_0}}$.

(b).

$$\begin{aligned} v(RP_1) &= \frac{b_0 r - b_1 + \sqrt{(b_0 r - b_1)^2 + (b_0 b_2 - b_1^2)(b_0 + 1)}}{2 \cdot (b_0 + 1)}, \\ x^*_{RP_1} &= (\Gamma^{-1}\mu \quad \Gamma^{-1}e)\begin{pmatrix} b_0 & -b_1 \\ -b_1 & b_2 \end{pmatrix}\begin{pmatrix} \frac{b_0 \cdot (b_1 + r) + \sqrt{(b_0 r - b_1)^2 + (b_0 b_2 - b_1^2)(b_0 + 1)}}{b_0(b_0 + 1)} \\ 1 \end{pmatrix}. \end{aligned}$$

(c).

$$\begin{aligned} v(RP_2) &= \frac{[(b_0 r - b_1)_+]^2}{b_0(b_0 + 1)} + \frac{1}{c_0}, \\ x^*_{RP_2} &= (\Gamma^{-1}\mu \quad \Gamma^{-1}e)\begin{pmatrix} b_0 & -b_1 \\ -b_1 & b_2 \end{pmatrix}\begin{pmatrix} \frac{(b_0 r - b_1)_+}{b_0(b_0 + 1)} + \frac{b_1}{b_0} \\ 1 \end{pmatrix}. \end{aligned}$$

Proof. Let us denote

$$\begin{aligned} f_0(s,t) &:= \inf_{X \sim (s,t^2)} \mathsf{E}[(r - X)_+^0] = \frac{1}{1 + (r-s)^2/t^2}; \\ f_1(s,t) &:= \inf_{X \sim (s,t^2)} \mathsf{E}[(r - X)_+^1] = \frac{r - s + \sqrt{t^2 + (r-s)^2}}{2}; \\ f_2(s,t) &:= \inf_{X \sim (s,t^2)} \mathsf{E}[(r - X)_+^2] = [(r-s)_+]^2 + t^2. \end{aligned}$$

Now our optimization problem can be expressed by function f_m, i.e.,

$$\begin{aligned} v(RP_m) &= \min_x \{f_m(x^{\mathrm{T}}\mu, \sqrt{x^{\mathrm{T}}\Gamma x}) \mid x^{\mathrm{T}}e = 1\} \\ &= \min_{s \in \Re} \min_x \{f_m(s, \sqrt{x^{\mathrm{T}}\Gamma x}) \mid x^{\mathrm{T}}e = 1,\, x^{\mathrm{T}}\mu = s\} \qquad (2.1) \end{aligned}$$

For any given s and m, the optimal solution x_s^* of the inner optimization problem in (2.1) is a mean-variance efficient solution:

$$\begin{aligned} x_s^* &= \arg\min_x \{f_m(s, \sqrt{x^{\mathrm{T}}\Gamma x}) \mid x^{\mathrm{T}}e = 1,\ x^{\mathrm{T}}\mu = s\} \\ &= \arg\min_x \{x^{\mathrm{T}}\Gamma x \mid x^{\mathrm{T}}e = 1, x^{\mathrm{T}}\mu = s\} \\ &= (\Gamma^{-1}\mu \ \ \Gamma^{-1}e) \begin{pmatrix} b_0 & -b_1 \\ -b_1 & b_2 \end{pmatrix} \begin{pmatrix} s \\ 1 \end{pmatrix}. \end{aligned}$$

Note that the second equality comes from the increasing property of $f_m(s,t)$ in t for all $m = 0, 1, 2$. Since $(x_s^*)^{\mathrm{T}}\Gamma x_s^* = b_0 s^2 - 2b_1 s + b_2$, we have $v(RP_m) = \min_{s\in\Re} f_m(s, \sqrt{b_0 s^2 - 2b_1 s + b_2})$. Finally, by solving the above problem we are led to the solutions:

$$\begin{aligned} s^*_{RP_0} &= \arg\min_{s\in\Re} f_0(s, \sqrt{b_0 s^2 - 2b_1 s + b_2}) \\ &= \arg\max_{s\geq r} \tfrac{(r-s)^2}{b_0 s^2 - 2b_1 s + b_2} \\ &= \begin{cases} \frac{b_2 - b_1 r}{b_1 - b_0 r} & \text{if } b_1 \geq r b_0 \\ +\infty & \text{if } b_1 < r b_0 \end{cases} \end{aligned}$$

$$\begin{aligned} s^*_{RP_1} &= \arg\min_{s\in\Re} f_1(s, \sqrt{b_0 s^2 - 2b_1 s + b_2}) \\ &= \arg\min_{s\in\Re} \tfrac{r - s + \sqrt{b_0 s^2 - 2b_1 s + b_2 + (r-s)^2}}{2} \\ &= \tfrac{b_0\cdot(b_1 + r) + \sqrt{(b_0 r - b_1)^2 + (b_0 b_2 - b_1^2)(b_0 + 1)}}{b_0(b_0+1)}. \end{aligned}$$

$$\begin{aligned} s^*_{RP_2} &= \arg\min_{s\in\Re} f_2(s, \sqrt{b_0 s^2 - 2b_1 s + b_2}) \\ &= \arg\min_{s\in\Re} [(r-s)_+]^2 + b_0 s^2 - 2b_1 s + b_2 \\ &= \tfrac{(b_0 r - b_1)_+}{b_0(b_0+1)} + \tfrac{b_1}{b_0}. \end{aligned}$$

□

Remark 2.6. All the above portfolios are mean-variance efficient, with different locations on the same mean-variance efficient frontier. In particular, $s^*_{RP_0} \geq s^*_{RP_1} \geq s^*_{RP_2}$, which means that for a fixed r, the higher the order of lower partial moments, the more conservative the portfolio.

Remark 2.7. All the explicit results heavily depend on the single constraint $x^{\mathrm{T}}e = 1$, as in the classical mean-variance model. Our approach can be adapted to solve more generally constrained portfolio problems. However, the solutions may no longer be explicit.

Remark 2.8. In case the mean and the covariance matrix are also uncertain, if we assume that they belong to their respective uncertainty sets S_μ and S_Γ, then we can recast the problem as a semidefinite program. In particular, for $m = 1$ or 2, by noticing $f_m(s,t)$ is decreasing in s and increasing in t, we have

$$\begin{aligned} & \min_{x\in X} \max_{\mu\in S_\mu,\ \Gamma\in S_\Gamma} f_m\left(x^{\mathrm{T}}\mu, \sqrt{x^{\mathrm{T}}\Gamma x}\right) \\ = & \min_{x\in X} f_m\left(\min_{\mu\in S_\mu} x^{\mathrm{T}}\mu, \sqrt{\max_{\Gamma\in S_\Gamma} x^{\mathrm{T}}\Gamma x}\right) \\ = & \min_{x,\ y,\ s,\ t} \ \{y \mid y \geq f_m(s,t),\ s \leq \min_{\mu\in S_\mu} x^{\mathrm{T}}\mu,\ t^2 \geq \max_{\Gamma\in S_\Gamma} x^{\mathrm{T}}\Gamma x,\ x \in X, t \geq 0\}. \end{aligned}$$

It is straightforward to recast the first constraint into several linear or second-order cone constraints by adding one or two auxiliary decision variables. The only question left is if the next two constraints are LMI representable, which depends on the particulars of S_μ and S_Γ. As a matter of fact, the minimum mean and maximum variance in these two constraints have been well studied by many researchers, e.g. Lobo *et al.* [31], Nestrov and Nemirivski [40], El Ghaoui and Lebret [12], Goldfarb and Iyengar [16]. In each of the above cases, the constraints in question can be converted into a second-order cone constraint (see Section 3.1 of [16] for a summary).

If the uncertain mean and covariance matrix are coupled and belong to a confidence region

$$S_{\mu,\Gamma} = \{(\mu, \Gamma) \mid (\mu - \mu_0)^{\mathrm{T}}\Gamma_0^{-1}(\mu - \mu_0) \leq \gamma_1, \Gamma + (\mu - \mu_0)(\mu - \mu_0)^{\mathrm{T}} \leq \gamma_2\Gamma_0\}$$

(see Delage and Ye [11]), the robust problem

$$\min_{x \in X} \max_{\xi \sim (\mu,\Gamma) \in S_{\mu,\Gamma}} \mathsf{E}[(r - x^{\mathrm{T}}\xi)_+^m]$$

is equivalent to the following dual problem

$$\begin{array}{lll} \min\limits_{x \in X} & \min_{Q,\ q,\ s,\ t} & s + t \\ & \text{s.t.} & s \geq [(r - x^{\mathrm{T}}\xi)_+^m] - \xi^{\mathrm{T}}Q\xi - \xi^{\mathrm{T}}q, \quad \forall \xi \in \Re^n \\ & & t \geq (\gamma_2\Gamma_0 + \mu_0\mu_0^{\mathrm{T}}) \bullet Q + \mu_0^{\mathrm{T}}q + \sqrt{\gamma_1}\,\|\Gamma_0^{1/2}(q + 2Q\mu_0)\|, \\ & & Q \succ 0. \end{array}$$

(See Lemma 1 of [11]). By the S-lemma, it is easy to verify that when $m = 1$ the first constraint of the dual problem can be cast by LMI's, while for $m = 0$ or 2, this is not the case. Therefore, only for $m = 1$, the robust counterpart can be formulated as an SDP.

Our model in Theorem 2.5(b) is similar to Delage and Ye [11]; we assumed the parameters to be $S = \Re^n$, $\gamma_1 = 0$ and $\gamma_2 = 1$. Because of this, rather than a numerical scheme (Lemma 1 of [11]), we provide an explicit optimal solution for $v(RP_1)$.

2.3 Conditional Value-at-Risk as the Risk Measure

Conditional Value-at-Risk (CVaR), defined as the mean of the tail distribution exceeding VaR, has attracted much attention in recent years. As a measure of risk, CVaR exhibits far better computational properties than VaR. With the help of the expression of $\text{LPM}_1(r)$, for linear loss functions, our robust version CVaR problem can be solved explicitly.

As is common in the CVaR analysis, let $f(x, \xi)$ denote the loss function associated with decision vector $x \in K \subseteq \Re^n$ and random vector $\xi \in \Re^n$. We assume that the cumulative probability distribution function for ξ is $\pi(\cdot)$. We also assume $\mathsf{E}[|f(x, \xi)|] < +\infty$ for each $x \in K$.

Given a decision $x \in K$, the probability of $f(x, \xi)$ not exceeding a threshold α is given by

$$\Psi(x, \alpha) \triangleq \int_{f(x,\ \xi) \leq \alpha} d\pi(\xi).$$

For a given confidence level β (usually greater than 0.9) and a fixed $x \in K$, the value-at-risk is defined as

$$\mathrm{VaR}_\beta(x) \triangleq \min\{\alpha \in \Re \mid \Psi(x,\alpha) \geq \beta\}.$$

The corresponding conditional value-at-risk, denoted by $\mathrm{CVaR}_\beta(x)$, is defined as the expected value of loss that exceeds $\mathrm{VaR}_\beta(x)$; that is,

$$\mathrm{CVaR}_\beta(x) \triangleq \frac{1}{1-\beta}\int_{f(x,\xi)\geq \mathrm{VaR}_\beta(x)} f(x,\xi)d\pi(\xi).$$

Rockafellar and Uryasev [46, 47] demonstrate that the calculation of CVaR can be done by minimizing the following auxiliary function with respect to the variable $\alpha \in \Re$:

$$F_\beta(x,\alpha) \triangleq \alpha + \frac{1}{1-\beta}\int_{\xi\in\Re^n} [f(x,\xi)-\alpha]_+ d\pi(\xi), \tag{2.2}$$

and subsequently

$$\mathrm{CVaR}_\beta(x) = \min_{\alpha\in\Re} F_\beta(x,\alpha).$$

Recall that in Section 2.2, we assumed the distribution $\pi(\cdot)$ is uncertain:

$$\pi \in D \quad = \quad \{\pi \mid \mathsf{E}_\pi[\xi] = \mu, \mathrm{Cov}_\pi(\xi) = \Gamma \succ 0\}.$$

For fixed $x \in K$, the robust optimization counterpart of the portfolio selection problems (with respect to the ambiguity set D) using CVaR or VaR as the risk measure, are formulated by

$$\mathrm{RCVaR}_\beta(x) \triangleq \sup_{\pi\in D} \mathrm{CVaR}_\beta(x), \tag{2.3}$$

and

$$\mathrm{RVaR}_\beta(x) \triangleq \sup_{\pi\in D} \mathrm{VaR}_\beta(x). \tag{2.4}$$

Also with the simple constraint $x^{\mathrm{T}}e = 1$, we find that the robust portfolio problem with CVaR or VaR as risk measure could be solved explicitly.

Theorem 2.9. *Suppose the loss function is $f(x,\xi) = -x^{\mathrm{T}}\xi$ and the random vector ξ has mean μ and covariance matrix $\Gamma \succ 0$. Let $\beta \in (0.5, 1]$. Consider*

$$\begin{array}{rcll} v(\mathrm{RC}_\beta) & = & \min_x & \sup_{\pi\in D} \mathrm{CVaR}_\beta(x) \\ & & \text{s.t.} & x^{\mathrm{T}}e = 1, \end{array}$$

and

$$\begin{array}{rcll} v(\mathrm{RV}_\beta) & = & \min_x & \sup_{\pi\in D} \mathrm{VaR}_\beta(x) \\ & & \text{s.t.} & x^{\mathrm{T}}e = 1. \end{array}$$

Then the solution of worst-case CVaR *is:*

$$v(\mathrm{RC}_\beta) = \begin{cases} \frac{\sqrt{b_0 b_2 - b_1^2}\sqrt{\frac{\beta b_0}{1-\beta}-1}}{b_0} - \frac{b_1}{b_0}, & \textit{if } \frac{\beta}{1-\beta}\cdot b_0 \geq 1, \\ -\infty, & \textit{if } \frac{\beta}{1-\beta}\cdot b_0 < 1, \end{cases}$$

and when $\frac{\beta}{1-\beta}\cdot b_0 \geq 1$*, with optimal solution*

$$x^*_{RC_\beta} = (\Gamma^{-1}\mu \quad \Gamma^{-1}e) \begin{pmatrix} b_0 & -b_1 \\ -b_1 & b_2 \end{pmatrix} \begin{pmatrix} \frac{\sqrt{b_0 b_2 - b_1^2}}{b_0\sqrt{\frac{\beta b_0}{1-\beta}-1}} + \frac{b_1}{b_0} \\ 1 \end{pmatrix}.$$

The solution of worst-case VaR *is:*

$$RV_\beta \quad = \quad \begin{cases} \frac{\sqrt{b_0 b_2 - b_1^2}\sqrt{\frac{b_0}{4\beta(1-\beta)}-b_0-1}}{b_0} - \frac{b_1}{b_0}, & \textit{if } \frac{b_0}{4\beta(1-\beta)} \geq 1 + b_0, \\ -\infty, & \textit{if } \frac{b_0}{4\beta(1-\beta)} \leq 1 + b_0. \end{cases}$$

When $\frac{b_0}{4\beta(1-\beta)} \geq 1 + b_0$*, the optimal solution is*

$$x^*_{RV_\beta} = (\Gamma^{-1}\mu \quad \Gamma^{-1}e) \begin{pmatrix} b_0 & -b_1 \\ -b_1 & b_2 \end{pmatrix} \begin{pmatrix} \frac{\sqrt{b_0 b_2 - b_1^2}}{b_0\sqrt{\frac{b_0}{4\beta(1-\beta)}-b_0-1}} + \frac{b_1}{b_0} \\ 1 \end{pmatrix}.$$

The proof of the theorem can be found in Appendix A.

Remark 2.10. It is interesting to note that when $\frac{\beta}{1-\beta}\cdot b_0 < 1$ and $\frac{b_0}{4\beta(1-\beta)} \leq 1 + b_0$ then the expected downfall measured by CVaR and VaR are unbounded. In all other cases, the optimal portfolios are mean-variance efficient.

Remark 2.11. In Natarajan, Sim and Uichanco [37], a closed form for the worst-case Conditional Value-at-Risk and its coherent generalization are derived. In Zhu and Fukushima [55], the worst-case CVaR under mixture distribution uncertainty, box uncertainty and ellipsoidal uncertainty is minimized. In El Ghaoui, Oks and Oustry [13], the robust optimization model using worst-case VaR model is investigated. All those results have been applied to robust portfolio optimization and the corresponding problems have been transformed as conic programs (see Problem 2.8 in [37], Section 3.2 in [55], and Theorem 1 in [13]).

3 Multi-Stage Portfolio Selection with Robust LPM_2

It is natural to extend the analysis to the multiple-stage setting. In the spirit of *adjustable robust optimization* (ARO) of Ben-Tal *et al.* [4], we present in this section a tractable ARO model using the $\mathrm{LPM}_2(r)$ risk measure.

Let us consider an K-stage portfolio selection model, and let us denote y_k to be the k-th stage recourse portfolio decision, $k = 1, ..., K$. The objective of the investor is to minimize the terminal LPM_2 risk, i.e., $\mathsf{E}[(r - {y_K}^{\mathrm{T}}\xi_K)_+^2]$. Let ξ_k be the rate of return vector for the k-th stage. In this section, we denote the rate of return of an asset to be the ratio of its final value to its initial value. This convention helps simplify the formulas. With regard to the random vectors ξ_k, we only know their first and second moment estimation: (μ_k, Γ_k). As ξ_k unfolds, we will need to make the $(k+1)$-th stage recourse decision y_{k+1}, before the exact status of ξ_{k+1} is revealed, where $k = 1, 2, ..., K$. When we select y_1, we do not know the actual rate of return ξ_1, and when we select the portfolio y_k $(k \geq 2)$, $\xi_1, ..., \xi_{k-1}$ are known but ξ_k is unknown, and so y_k is an *adjustable* variable depending on the uncertain data $\xi_1, \cdots, \xi_{k-1}$. Our model is thus an *adjustable robust optimization* model. For general problems, the ARO formulations often lead to intractable optimization problems; see [4]. Below we shall elaborate on the following particular ARO formulation of the multi-stage robust portfolio selection model.

Mathematically, the problem is formulated as:

$$
\begin{array}{llll}
V_K(\omega_{K-1}) & = & \min_{y_K} & \sup_{\xi_K \sim (\mu_K,\, \Gamma_K)} \mathsf{E}[(r - y_K^{\mathrm{T}}\xi_K)_+^2] \\
 & & \text{s.t.} & y_K^{\mathrm{T}} e = \omega_{K-1}; \\
V_{K-1}(\omega_{K-2}) & = & \min_{y_{K-1}} & \sup_{\xi_{K-1} \sim (\mu_{K-1},\, \Gamma_{K-1})} \mathsf{E}[V_K(y_{K-1}^{\mathrm{T}}\xi_{K-1})] \\
 & & \text{s.t.} & y_{K-1}^{\mathrm{T}} e = \omega_{K-2} \\
 & \vdots & & \\
V_1(1) & = & \min_{y_1} & \sup_{\xi_1 \sim (\mu_1,\, \Gamma_1)} \mathsf{E}[V_2(y_1^{\mathrm{T}}\xi_1)] \\
 & & \text{s.t.} & {y_1}^{\mathrm{T}} e = 1.
\end{array}
$$

Here we denote $V_{k+1}(\omega_k)$ as the optimal objective value at stage $k+1$, given the wealth $\omega_k = y_k^{\mathrm{T}}\xi_k$ as the result of investment at stage k, $k = 1, ..., K-1$.

We shall present the explicit solutions and values for each stage in the theorem below.

Theorem 3.1. *Suppose that the first and second moments of ξ_k are (μ_k, Γ_k), $k = 1, ..., K$, and furthermore, for $k = 1, ..., K-1$, denote*

$$
\begin{array}{lll}
c_0^k := e^{\mathrm{T}}\tilde{\Gamma}_k^{-1}e, & c_1^k := e^{\mathrm{T}}\tilde{\Gamma}_k^{-1}\mu_k, & c_2^k := \mu_k^{\mathrm{T}}\tilde{\Gamma}_k^{-1}\mu_k, \\
b_0^k := \frac{c_0^k}{c_0^k c_2^k - {c_1^k}^2}, & b_1^k := \frac{c_1^k}{c_0^k c_2^k - {c_1^k}^2}, & b_2^k := \frac{c_2^k}{c_0^k c_2^k - {c_1^k}^2}, \\
q_k = \frac{b_1^{k+1}}{\sqrt{b_0^{k+1}(b_0^{k+1}+1)}}, & r_k = \frac{b_0^{k+1} r_{k+1}}{b_1^{k+1}}, & \tilde{\Gamma}_k = \Gamma_k + \frac{\Gamma_k + \mu_k^{\mathrm{T}}\mu_k}{c_0^{k+1} \cdot q_k^2},
\end{array}
$$

where we assume $r_K = r$, $q_K = 1$. For $k = 1, ..., K$, it holds that

$$
\begin{array}{lll}
V_k(y_{k-1}^{\mathrm{T}}\xi_{k-1}) & = & (q_{K-1} \cdot \cdots \cdot q_{k-1})^2 \cdot \left[(r_{k-1} - y_{k-1}^{\mathrm{T}}\xi_{k-1})_+^2\right] + (q_{K-1} \cdots q_k)^2 \cdot \frac{(y_{k-1}^{\mathrm{T}}\xi_{k-1})^2}{c_0^k}; \\
y_k^* & = & (\tilde{\Gamma}_k^{-1}\mu_k \;\; \tilde{\Gamma}_k^{-1}e) \begin{pmatrix} b_0^k & -b_1^k \\ -b_1^k & b_2^k \end{pmatrix} \begin{pmatrix} \frac{(b_0^k r_k - b_1^k(y_{k-1}^{\mathrm{T}}\xi_{k-1}))_+}{b_0^k(b_0^k+1)} + \frac{b_1^k(y_{k-1}^{\mathrm{T}}\xi_{k-1})}{b_0^k} \\ 1 \end{pmatrix}.
\end{array}
$$

Proof. Recall that

$$\sup_{\eta\sim(u,\sigma)} \mathsf{E}[(r-\eta)_+^2] = (r-u)_+^2 + \sigma^2, \forall r \in \Re$$

and so

$$\begin{array}{rll} & \min_x & \sup_{\xi\sim(\mu,\Gamma)} \mathsf{E}[(r-\xi^{\mathrm{T}}x)_+^2] \\ & \text{s.t.} & x^{\mathrm{T}}e = \omega \\ = & \min_x & (r - x^{\mathrm{T}}\mu)_+^2 + x^{\mathrm{T}}\Gamma x \\ & \text{s.t.} & x^{\mathrm{T}}e = \omega \\ = & \min_z & (r-\omega\cdot z^{\mathrm{T}}\mu)_+^2 + \omega^2\cdot z^{\mathrm{T}}\Gamma z \\ & \text{s.t.} & z^{\mathrm{T}}e = 1 \\ = \quad \omega^2\cdot & \min_z & (r/\omega - z^{\mathrm{T}}\mu)_+^2 + z^{\mathrm{T}}\Gamma z \\ & \text{s.t.} & z^{\mathrm{T}}e = 1. \end{array}$$

The above problem can be explicitly solved, with the optimal value and solution being

$$\begin{aligned} v^* &= \frac{[(b_0 r - b_1\cdot\omega)_+]^2}{b_0(b_0+1)} + \frac{\omega^2}{c_0}, \\ x^* &= (\Gamma^{-1}\mu \;\; \Gamma^{-1}e)\begin{pmatrix} b_0 & -b_1 \\ -b_1 & b_2 \end{pmatrix}\begin{pmatrix} \frac{(b_0 r - b_1\cdot\omega)_+}{b_0(b_0+1)} + \frac{b_1\cdot\omega}{b_0} \\ 1 \end{pmatrix}, \end{aligned}$$

where we use the same notations as in Theorem 2.5; in particular,

$$\begin{array}{lll} c_0 := e^{\mathrm{T}}\Gamma^{-1}e, & c_1 := e^{\mathrm{T}}\Gamma^{-1}\mu, & c_2 := \mu^{\mathrm{T}}\Gamma^{-1}\mu, \\ b_0 := \frac{c_0}{c_0c_2 - c_1^2}, & b_1 := \frac{c_1}{c_0c_2-c_1^2}, & b_2 := \frac{c_2}{c_0c_2-c_1^2}. \end{array}$$

With these relations, we can prove the theorem inductively (in a reversed fashion). For the final stage,

$$\begin{aligned} V_K(y_{K-1}^{\mathrm{T}}\xi_{K-1}) &= \min_{y_K} \; \sup_{\xi_K\sim(\mu_K,\,\Gamma_K)} \mathsf{E}[(r - y_K^{\mathrm{T}}\xi_K)_+^2] \\ &\quad\;\text{s.t.} \;\; y_K^{\mathrm{T}}e = y_{K-1}^{\mathrm{T}}\xi_{K-1} \\ &= \min_{y_K} \; [(r - y_K^{\mathrm{T}}\mu_K)_+^2] + y_K^{\mathrm{T}}\Gamma_K y_K \\ &\quad\;\text{s.t.} \;\; y_K^{\mathrm{T}}e = y_{K-1}^{\mathrm{T}}\xi_{K-1} \\ &= \frac{[(b_0^K r_K - b_1^K\cdot(y_{K-1}^{\mathrm{T}}\xi_{K-1}))_+]^2}{b_0^K(b_0^K+1)} + \frac{(y_{K-1}^{\mathrm{T}}\xi_{K-1})^2}{c_0^K} \\ &= q_{K-1}^2\cdot[(r_{K-1} - y_{K-1}^{\mathrm{T}}\xi_{K-1})_+]^2 + \frac{(y_{K-1}^{\mathrm{T}}\xi_{K-1})^2}{c_0^K} \end{aligned}$$

and

$$y_K^* = (\Gamma_K^{-1}\mu_K \;\; \Gamma_K^{-1}e)\begin{pmatrix} b_0^K & -b_1^K \\ -b_1^K & b_2^K \end{pmatrix}\begin{pmatrix} \frac{(b_0^K r_K - b_1^K\cdot(y_{K-1}^{\mathrm{T}}\xi_{K-1}))_+}{b_0^K(b_0^K+1)} + \frac{b_1^K\cdot(y_{K-1}^{\mathrm{T}}\xi_{K-1})}{b_0^K} \\ 1 \end{pmatrix}.$$

Assume the formula for the optimal value and solution hold true for the $(k+1)$th stage, then for the kth stage, we have

$$
\begin{aligned}
V_k(y_{k-1}^{\mathrm{T}}\xi_{k-1}) &= \min_{y_k}\ \sup_{\xi_k\sim(\mu_k,\,\Gamma_k)} \mathsf{E}[V_{k+1}(y_k^{\mathrm{T}}\xi_k)] \\
&\quad\ \text{s.t.}\quad y_k^{\mathrm{T}}e = y_{k-1}^{\mathrm{T}}\xi_{k-1} \\
&= \min_{y_k}\ \sup_{\xi_k\sim(\mu_k,\,\Gamma_k)} \mathsf{E}\left\{(q_{K-1}\cdots q_k)^2\cdot[(r_k-y_k^{\mathrm{T}}\xi_k)_+]^2+(q_{K-1}\cdots q_{k+1})^2\frac{(y_k^{\mathrm{T}}\xi_k)^2}{c_0^{k+1}}\right\} \\
&\quad\ \text{s.t.}\quad y_k^{\mathrm{T}}e = y_{k-1}^{\mathrm{T}}\xi_{k-1} \\
&= \min_{y_k}\ (q_{K-1}\cdots q_k)^2\cdot\{[(r_k-y_k^{\mathrm{T}}\mu_k)_+^2]+y_k^{\mathrm{T}}\tilde{\Gamma}_k y_k\} \\
&\quad\ \text{s.t.}\quad y_k^{\mathrm{T}}e = y_{k-1}^{\mathrm{T}}\xi_{k-1} \\
&= (q_{K-1}\cdots q_k)^2\cdot\left\{\frac{[(b_0^k r_k-b_1^k\cdot(y_{k-1}^{\mathrm{T}}\xi_{k-1}))_+]^2}{b_0^k(b_0^k+1)}+\frac{(y_{k-1}^{\mathrm{T}}\xi_{k-1})^2}{c_0^k}\right\} \\
&= (q_{K-1}\cdots q_k\cdot q_{k-1})^2\cdot[(r_{k-1}-y_{k-1}^{\mathrm{T}}\xi_{k-1})_+]^2+(q_{K-1}\cdots q_k)^2\cdot\frac{(y_{k-1}^{\mathrm{T}}\xi_{k-1})^2}{c_0^k},
\end{aligned}
$$

and

$$
y_k^* = (\tilde{\Gamma}_k^{-1}\mu_k\ \ \tilde{\Gamma}_k^{-1}e)\begin{pmatrix} b_0^k & -b_1^k \\ -b_1^k & b_2^k \end{pmatrix}\begin{pmatrix} \frac{(b_0^k r_k-b_1^k\cdot(y_{k-1}^{\mathrm{T}}\xi_{k-1}))_+}{b_0^k(b_0^k+1)}+\frac{b_1^k\cdot(y_{k-1}^{\mathrm{T}}\xi_{k-1})}{b_0^k} \\ 1 \end{pmatrix}.
$$

The theorem is thus proven by induction. □

It is crucial that the robust form of the $\mathrm{LPM}_2(r)$ holds the similar structure as its non-robust one. Therefore it is not clear how to extend the result to other risk measures. In case the objective is simply the variance, Li and Ng [29] derived an interesting analytical solution to the multi-stage portfolio selection problem.

Related to our discussions in this section, Hernández-Hernández and Schied [23] proposed a stochastic control approach to the dynamic maximization of robust utility functionals and obtained an explicit PDE characterization of the optimal strategy in an incomplete diffusion market model where the robust utility functional is defined in terms of a logarithmic utility function and a rather general dynamically consistent penalty function $\gamma(\cdot)$. They characterized the value function and the optimal investment strategy via the solution of a quasi-linear Hamilton-Jacobi-Bellman PDE. Iyengar [25], and Nilim and El Ghaoui [41] studied robust dynamic programming algorithms for finite-state and finite-action Markovian decision processes with uncertain transition probabilities. They proved that both finite and infinite planning horizon problems can be extended to natural robust counterpart when the measures have a certain 'rectangularity' property.

4 Robust Portfolio Selection using the S-Shaped Value Function

As an alternative to the expected utility maximization, Kahneman and Tversky ([26], 1979) developed the so-called prospect theory to count for psychological factors in economical decision making. This work eventually won Nobel prize in economics in 2002. One aspect of the prospect theory is to promote a value function that is S-shaped, in place of an overall concave utility function. The rationale behind such consideration is that, typically a decision maker has a reference point in mind, and he/she would exhibit *diminishing sensitivity* in view of the gain and loss from the reference point. In other words, the value function is concave in the domain of the gains, and is convex in the domain of the losses. Moreover, the loss is more acutely felt than the gain near the reference point.

Mathematically, an S-shaped value function is given as:

$$v(x) = \begin{cases} x^\alpha, & x > 0 \\ -\lambda(-x)^\beta, & x \leq 0, \end{cases}$$

with $0 < \alpha \leq \beta < 1$ and $\lambda \geq 1$. Under the new tenet, an investor is interested in solving the following optimization model:

$$\begin{array}{ll} \max & \mathsf{E}[v(x^{\mathrm{T}}\xi - r)] \\ \text{s.t.} & x \in X, \end{array}$$

where ξ is the return of the assets, r is the reference point, and x is the portfolio to be selected, and X (a convex set) represents the constraint that the investor would like to impose on the portfolio. It is perhaps not surprising that portfolio selection based on the S-shaped value function is difficult in general. This fact is formalized in the next theorem.

Theorem 4.1. *It is NP-hard to solve*

$$\begin{array}{ll} \max & \mathsf{E}[v(\xi^{\mathrm{T}}x)] \\ \text{s.t.} & e^{\mathrm{T}}x = 1, \\ & -1 \leq x_i \leq 1,\ i = 1, ..., n, \end{array}$$

where

$$v(x) = \begin{cases} x^\alpha, & x > 0 \\ -\lambda(-x)^\beta, & x \leq 0, \end{cases}$$

and $\xi \sim N(\mu, \Gamma)$.

Proof. Let $\mu = 0$ and $\sigma(x) = \sqrt{x^{\mathrm{T}}\Gamma x}$. Then,

$$\begin{array}{rcl} \mathsf{E}[v(\xi^{\mathrm{T}}x)] & = & \frac{1}{\sqrt{2\pi}\sigma(x)} \int_0^\infty t^\alpha \cdot e^{-\frac{t^2}{2\sigma(x)^2}} dt - \frac{\lambda}{\sqrt{2\pi}\sigma(x)} \int_0^\infty t^\beta \cdot e^{-\frac{t^2}{2\sigma(x)^2}} dt \\ & = & \sigma(x)^\alpha \cdot (\frac{1}{2}\mathsf{E}_{\eta\sim N(0,1)}|\eta|^\alpha) - \sigma(x)^\beta \cdot (\frac{\lambda}{2}\mathsf{E}_{\eta\sim N(0,1)}|\eta|^\beta) \\ & = & \sigma(x)^\alpha c_1 - \sigma(x)^\beta c_2, \end{array}$$

where $c_1 := \frac{1}{2}\mathsf{E}_{\eta\sim N(0,1)}|\eta|^\alpha > 0$ and $c_2 := \frac{\lambda}{2}\mathsf{E}_{\eta\sim N(0,1)}|\eta|^\beta > 0$.

Since $(c_1 t^\alpha - c_2 t^\beta)' = t^{\alpha-1} \cdot (\alpha c_1 - \beta c_2 t^{\beta-\alpha}) > 0$ when t is small (note $\beta > \alpha$), the quantity $\mathsf{E}[v(\xi^{\mathrm{T}}x)]$ is monotonically increasing in $\sigma(x)$. Therefore,

$$\begin{array}{ll} \max & \mathsf{E}[v(\xi^{\mathrm{T}}x)] \\ \text{s.t.} & e^{\mathrm{T}}x = 1, \\ & -1 \leq x_i \leq 1,\ i = 1, ..., n \end{array}$$

is equivalent to

$$\begin{array}{lll} (MC) & \max & x^{\mathrm{T}}\Gamma x \\ & \text{s.t.} & e^{\mathrm{T}}x = 1, \\ & & -1 \le x_i \le 1,\ i = 1, ..., n, \end{array}$$

if Γ is properly scaled (to be sufficiently small). Consider now the problem of finding the max-cut for a weighted graph of $n = 2m+1$ nodes, with m nodes on one side and $m+1$ nodes on the other side of the cut. Let $\Gamma \succeq 0$ be the Laplacian matrix of the graph. The objective function of (MC) is convex, and so its optimal solution is attained at a vertex. Hence, the $(m, m+1)$ max-cut problem can be cast as (MC). Since the max-cut problem is NP-hard, it follows that the portfolio selection problem based on the S-shaped value function is NP-hard in general. □

It is quite unexpected, however, that the robust counter-part of the optimization model turns out to be easy. The robust optimization model in question is:

$$\begin{array}{ll} \max & \inf_{\xi\sim(\mu,\Gamma)} \mathsf{E}[v(x^{\mathrm{T}}\xi - r)] \\ \text{s.t.} & x \in X. \end{array}$$

To see this we need a few intermediate steps. Let us first introduce a function

$$v_R(\mu, \sigma) := \inf_{\eta\sim(\mu,\, \sigma^2)} \mathsf{E}[v(\eta)].$$

Lemma 4.2. *For any fixed μ, the function $v_R(\mu, \sigma)$ is monotonically non-increasing in σ (> 0).*

Proof. Consider $0 < \bar{\sigma} < \sigma$. For any $\epsilon > 0$ we have a distribution $\bar{\eta}$, satisfying $\mathsf{E}[\bar{\eta}] = \mu$, $\mathrm{Var}(\bar{\eta}) = \bar{\sigma}^2$, and $\mathsf{E}[v(\bar{\eta})] \le v_R(\mu, \bar{\sigma}) + \epsilon$. For any fixed $n > 0$, let us consider a new distribution as follows

$$\eta_n = \begin{cases} \bar{\eta}, & \text{with probability } \frac{n^2-1}{n^2}, \\ -n\sqrt{\sigma^2 - \bar{\sigma}^2}, & \text{with probability } \frac{1}{n^2}. \end{cases}$$

Clearly, as $n \to \infty$ we have

$$\begin{array}{lcll} \mathsf{E}[\eta_n] & = & \mu\frac{n^2-1}{n^2} - \frac{n\sqrt{\sigma^2-\bar{\sigma}^2}}{n^2} & \to \mu, \\ \mathrm{Var}(\eta_n) & = & \sigma^2 - (\mu^2 + \bar{\sigma}^2)/n^2 & \to \sigma^2, \\ \mathsf{E}[v(\eta_n)] & = & \frac{n^2-1}{n^2}\mathsf{E}[v(\bar{\eta})] + \frac{v(-n\sqrt{\sigma^2-\bar{\sigma}^2})}{n^2} & \to \mathsf{E}[v(\bar{\eta})]. \end{array}$$

This shows that $v_R(\mu, \sigma) \le v_R(\mu, \bar{\sigma})$. □

Next, we shall exactly compute the value $v_R(\mu, \sigma)$, and its associated optimal solution.

Theorem 4.3. *It holds that*

$$v_R(\mu, \sigma) = \frac{s^2}{s^2+1}(\mu + \sigma/s)^{\alpha} - \frac{\lambda}{s^2+1}(\sigma s - \mu)^{\beta} \tag{4.5}$$

and an optimal solution is given by

$$\eta^* = \begin{cases} \mu - \sigma s, & \textit{with probability } \frac{1}{s^2+1}; \\ \mu + \sigma/s, & \textit{with probability } \frac{s^2}{s^2+1}, \end{cases}$$

where s is a root for the following function

$$g(x) := (2-\alpha)(\sigma/x+\mu)^{\alpha} + \lambda(2-\beta)(\sigma x-\mu)^{\beta} - \alpha(\sigma x-\mu)(\sigma/x+\mu)^{\alpha-1} - \lambda\beta(\sigma/x+\mu)(\sigma x-\mu)^{\beta-1}, \tag{4.6}$$

where

$$x \in \left(\max\{0, \mu/\sigma\}, \frac{\mu+\sqrt{\mu^2+\sigma^2}}{\sigma}\right].$$

We shall delegate the proof of Theorem 4.3 to Appendix B. The existence of the root can be verified by checking the boundary values of $g(\cdot)$, which are $\lim_{x\downarrow\max\{0,\ \mu/\sigma\}} g(x) = -\infty$ and $g\left(\frac{\mu+\sqrt{\mu^2+\sigma^2}}{\sigma}\right) > 0$. Due to the continuity of function $g(\cdot)$, one may search the root by e.g. the golden-section type line-search method (see Chapter 8 of [32]).

By Lemma 2.4, the robust portfolio selection problem can be reduced to a single-parameter searching problem along the mean-variance efficient frontier:

$$\begin{array}{lll} & \max_{x\in X} & \min_{\xi\sim(\mu,\Gamma)} \mathsf{E}[v(x^{\mathrm{T}}\xi - r)] \\ = & \max_{x\in X} & v_R(x^{\mathrm{T}}\mu - r, \sqrt{x^{\mathrm{T}}\Gamma x}) \\ = & \max_{t\in\Re} & \max_{x\in X} v_R(t-r, \sqrt{x^{\mathrm{T}}\Gamma x}) \\ & \text{s.t.} & x^{\mathrm{T}}\mu = t \\ = & \max_{t\in\Re} & v_R(t-r, \sigma^*(t)), \end{array}$$

where r is the reference return level and

$$[\sigma^*(t)]^2 = \min_{x^{\mathrm{T}}\mu=t,\ x\in X} x^{\mathrm{T}}\Gamma x. \tag{4.7}$$

The last step in the derivation is due to the monotonicity as established in Lemma 4.2. Equivalently, we may pose (4.7) as

$$\sigma^*(t) = \min_{x^{\mathrm{T}}\mu=t,\ x\in X} \|\Gamma^{1/2}x\|.$$

If X is convex, for any $t_1, t_2 \in \Re$ and $\lambda \in [0,1]$, we have $\sigma^*(\lambda t_1 + (1-\lambda)t_2) \le \lambda\sigma^*(t_1) + (1-\lambda)\sigma^*(t_2)$, because of the triangle inequality $\|\Gamma^{\frac{1}{2}}(\lambda\cdot x^*_{t_1} + (1-\lambda)\cdot x^*_{t_2})\| \le \lambda\|\Gamma^{\frac{1}{2}}x^*_{t_1}\| + (1-\lambda)\|\Gamma^{\frac{1}{2}}x^*_{t_2}\|$. Therefore, $\sigma^*(t)$ is in general an increasing and convex function if X is convex. On the premise that $\sigma^*(t)$ is easy to compute, the robust portfolio selection model based on the S-shaped value may be considered easy to solve, since it reduces to searching a single parameter value. For all practical purposes, one can always sample the value of $v_R(t-r, \sigma^*(t))$ for various t's, and then select the highest. Theoretically however, the search can be made efficient, in terms of the polynomial-time computational complexity, if $v_R(t-r, \sigma^*(t))$ is *unimodal* or *quasi-concave*.

Proposition 4.4.

a. If $\sigma^*(t+r) - t\cdot[\sigma^*(t+r)]' \ge 0\ \forall t$, then $v_R(t, \sigma^*(t+r))$ is quasi-concave in t.

b. Suppose that (x_0, r_0) is the optimal solution of

$$\min_{x^{\mathrm{T}}\mu=t,\ x\in X} \quad x^{\mathrm{T}}\Gamma x,$$

where (x,t) is the joint decision vector in the above model. Suppose that $r \geq r_0$ and that the feasible region X is an affine subspace (does not contain the origin). Then we have $\sigma^*(t+r) \geq t[\sigma^*(t+r)]'$ for all t.

Concerning the condition required by Proposition 4.4, the following proposition gives one such example.

In fact, we conjecture that $v_R(t, \sigma^*(t+r))$ is always quasi-concave as long as X is a convex set. Proposition 4.4 may possibly need be further strengthened in that case.

We shall point out here that the S-shaped value function is a two-piece *power* function, hence outside the realm of polynomial functions. This property is interesting. As a matter of fact, if the feasible set of the dual problem is *semi-algebraic*, i.e. a set given by polynomial inequalities, then Popescu [43] already showed that computing optimal moment bounds boils down to SDP. Popescu [43] also pointed out that if the objective function is piecewise polynomial, then the same SDP formulation is possible. In fact, almost all the known moment bounds focus on piecewise polynomials; see [7, 8, 36, 38, 42, 43]. Other from these moment bounds on polynomials, Popescu [44] also studied moment bounds for general non-polynomial utility function with the given mean and variance information, namely,

$$\max_{x\in X} \min_{\xi\sim(\mu,\Gamma)} \mathsf{E}[u(x^{\mathrm{T}}\xi)]$$

which is what we study in this section. Popescu proved that if the inner optimization problem is quasi-concave and satisfies a certain monotonicity condition, then the robust portfolio selection problem can be reduced to a parametric quadratic program. In fact, Popescu [44] elaborated on several specific utility classes in Propositions 5, 7, 8 and 10, where $\min_{\xi\sim(\mu,\Gamma)} \mathsf{E}[u(x^{\mathrm{T}}\xi)]$ indeed satisfies the quasi-concavity. However, the S-shaped function that we investigate here is not included in the classes specified in [44]. Instead of assuming the inner optimization problem to be quasi-concave, we *prove* the quasi-concavity for our S-shaped value function, and propose a single-parameter searching algorithm to solve its robust portfolio selection problem. Similar to the algorithm of Popescu in [44], our algorithm also requires to solve a parametric quadratic program (our parameter is the expected return of assets rather than a general weight parameter).

It is worth mentioning that our method, as well as Popescu's [44] algorithms, require that the utility function must have at most two tangent points with any quadratic function from below. Furthermore, Popescu [44] provided necessary and sufficient conditions for this two-point tangency property to hold, and a necessary condition for the one-point tangency property to hold (see Lemma 1 and Proposition 6 of [44]). In the context of Popescu's characterizations, it is also possible to get easy-computable bounds even if the objective function is more general than a two-piece power function. It only requires that the objective function has at most two tangent points with the quadratic function. Its proof is almost identical to that of Theorem 4.3. As for the tangent points requirement, one may regard Lemma 1 and Proposition 6 of [44] as necessary and/or sufficient conditions. Popescu provided several concrete examples for the conditions to hold; however, they are all monotone convex/concave functions with convex, concave or concave-convex derivative.

Note that for any fixed μ, $g_R(\mu, \sigma)$ is monotonically non-increasing in $\sigma\,(> 0)$, the proof being exactly the same as for Lemma 4.2. With this understanding, we solve Problem (GP) as follows:

1. Get $\sigma^*(t)$ by solving $[\sigma^*(t)]^2 = \min_{x^{\mathrm{T}}\mu=t, x\in X} x^{\mathrm{T}}\Gamma x$.
2. Solve $\max_{t\in\Re} g_R(t, \sigma^*(t))$.

However, the quasi-concavity of $g_R(t, \sigma^*(t))$ is not known in general, although we believe it is true.

5 Robust Portfolio Selection with Chance Information

In this paper we are concerned with robust optimization under distributional uncertainties. So far, the informational structure has been the knowledge of the mean and the covariance of the underlying distribution. However, it is in general always possible to obtain more information regarding the distribution, e.g. we may estimate the chance of a projection of the random vector above a certain threshold via some statistical methods. It turns out that in some cases, this additional information can result in a sharpened robust optimization formulation, which can be solved by SDP. We shall present three such cases in this section to showcase the potential of the technique.

5.1 Additional Chance Constraints

Let us consider robust portfolio selection as in Subsection 2.2, with u being a piecewise linear utility function whose expected value is to be maximized. Now, the additional information with regard to ξ is that

$$\mathrm{Prob}\,\{\xi^{\mathrm{T}} A\,\xi + a^{\mathrm{T}}\xi > r\} \le \beta$$

is known to hold, where A is a certain n-dimensional symmetric matrix, and a is an n-dimensional vector. Together with the knowledge about the first two moments of ξ, the robust portfolio selection model becomes

$$\begin{array}{lll} \max_{x\in X} & \min_{\xi} & \mathsf{E}[u(x^{\mathrm{T}}\xi)] \\ & \text{s.t.} & \xi \sim (\mu, \Gamma) \text{ and } \mathrm{Prob}\,\{\xi^{\mathrm{T}} A\,\xi + a^{\mathrm{T}}\xi > r\} \le \beta. \end{array}$$

Assumption 5.1.

a. *There exists* $\bar{\xi} \in \Re^n$ *such that* $\bar{\xi}^{\mathrm{T}} A\,\bar{\xi} + a^{\mathrm{T}}\bar{\xi} < r$.

b. $\Gamma \succ 0$ *and* $\beta \in (0, 1)$.

Following the moments cone approach developed by Popescu [42, 43, 44], we obtain various bounds by SDP via the conic duality theory. From Assumption 5.1(b), the strong duality holds. So, the above problem can be recast as

$$\begin{array}{lll} \max_{x\in X} \max_{Z,z_0,z_1,z_3} & z_0 + z_1^{\mathrm{T}}\mu + Z \bullet (\Gamma + \mu\mu^{\mathrm{T}}) - \beta z_3 \\ \text{s.t.} & z_0 + z_1^{\mathrm{T}}\xi + \xi^{\mathrm{T}} Z\xi \le z_3 \mathbf{1}_{\{\xi^{\mathrm{T}} A\, \xi + a^{\mathrm{T}}\xi > r\}} + u(x^{\mathrm{T}}\xi), \ \forall\, \xi \in \Re^n \\ & z_3 \ge 0. \end{array}$$

Suppose that $u(y) := \min\{c + by, 0\}$ with $b, c \in \Re$. The above problem can be further written as

$$\begin{array}{lll} \max_{x\in X} \max_{Z,z_0,z_1,z_3} & z_0 + z_1^{\mathrm{T}}\mu + Z \bullet (\Gamma + \mu\mu^{\mathrm{T}}) - \beta z_3 \\ \text{s.t.} & z_0 + z_1^{\mathrm{T}}\xi + \xi^{\mathrm{T}} Z\xi \le z_3 + c + b(x^{\mathrm{T}}\xi), \ \forall\, \xi \in \Re^n \\ & z_0 + z_1^{\mathrm{T}}\xi + \xi^{\mathrm{T}} Z\xi \le z_3, \ \forall\, \xi \in \Re^n \\ & z_0 + z_1^{\mathrm{T}}\xi + \xi^{\mathrm{T}} Z\xi \le c + b(x^{\mathrm{T}}\xi), \ \forall\, \xi^{\mathrm{T}} A\, \xi + a^{\mathrm{T}}\xi \le r \\ & z_0 + z_1^{\mathrm{T}}\xi + \xi^{\mathrm{T}} Z\xi \le 0, \ \forall\, \xi^{\mathrm{T}} A\, \xi + a^{\mathrm{T}}\xi \le r \\ & z_3 \ge 0. \end{array}$$

Using the S-lemma under Assumption 5.1(a) (see [6]), the constraints can be written as Linear Matrix Inequalities (LMI) as shown below:

$$\begin{array}{ll} \max_{x\in X} \max_{Z,z_0,z_1,z_3,s,t} & z_0 + z_1^{\mathrm{T}}\mu + Z \bullet (\Gamma + \mu\mu^{\mathrm{T}}) - \beta z_3 \\ \text{s.t.} & \begin{pmatrix} z_0 - z_3 - c & (z_1^{\mathrm{T}} - bx^{\mathrm{T}})/2 \\ (z_1 - bx)/2 & Z \end{pmatrix} \preceq 0 \\ & \begin{pmatrix} z_0 - z_3 & z_1^{\mathrm{T}}/2 \\ z_1/2 & Z \end{pmatrix} \preceq 0 \\ & \begin{pmatrix} z_0 - c + tr & (z_1^{\mathrm{T}} - bx^{\mathrm{T}} - ta^{\mathrm{T}})/2 \\ (z_1 - bx - ta)/2 & Z - tA \end{pmatrix} \preceq 0 \\ & \begin{pmatrix} z_0 + sr & (z_1^{\mathrm{T}} - sa^{\mathrm{T}})/2 \\ (z_1 - sa)/2 & Z - sA \end{pmatrix} \preceq 0 \\ & s, t, z_3 \ge 0. \end{array}$$

Therefore, the robust portfolio selection problem can be solved by SDP (using e.g. SeDuMi of Sturm [51], or via CVX of Grant, Boyd, and Ye [17, 18] which has a friendly interface). It is possible to extend the method to include one more chance information in the robust optimization formulation, based on the extended S-lemma of Sturm and Zhang [52]. The problem in question is:

$$\begin{array}{ll} \max_{x\in X} & \min_{\xi\sim(\mu,\Gamma)} \mathsf{E}[u(x^{\mathrm{T}}\xi)] \\ \text{s.t.} & \text{Prob}\,\{\xi^{\mathrm{T}} A\, \xi + a^{\mathrm{T}}\xi > r_1\} \le \beta_1, \\ & \text{Prob}\,\{d^{\mathrm{T}}\xi > r_2\} \le \beta_2. \end{array}$$

To apply the results in [52], we need the following technical assumptions:

Assumption 5.2.

a. $A \succeq 0$ *and there exists* $\bar\xi \in \Re^n$ *such that* $\bar\xi^{\mathrm{T}} A\, \bar\xi + a^{\mathrm{T}}\bar\xi < r_1$ *and* $d^{\mathrm{T}}\bar\xi < r_2$.

b. $\Gamma \succ 0$ *and* $\beta_1, \beta_2 \in (0,1)$.

Similar as before, its dual formulation is

$$
\begin{array}{rl}
\max_{x\in X} \max_{Z,z_0,z_1,z_3,z_4} & z_0 + z_1^{\mathrm{T}}\mu + Z \bullet (\Gamma + \mu\mu^{\mathrm{T}}) - \beta_1 z_3 - \beta_2 z_4 \\
\text{s.t.} & z_0 + z_1^{\mathrm{T}}\xi + \xi^{\mathrm{T}} Z\xi \le z_3 \mathbf{1}_{\{\xi^{\mathrm{T}} A\,\xi + a^{\mathrm{T}}\xi > r_1\}} + z_4 \mathbf{1}_{\{d^{\mathrm{T}}\xi > r_2\}} + u(x^{\mathrm{T}}\xi),\ \forall\, \xi \in \Re^n \\
& z_3, z_4 \ge 0,
\end{array}
$$

which can be further written as

$$
\begin{array}{rl}
\max_{x\in X} \max_{Z,z_0,z_1,z_3,z_4} & z_0 + z_1^{\mathrm{T}}\mu + Z \bullet (\Gamma + \mu\mu^{\mathrm{T}}) - \beta_1 z_3 - \beta_2 z_4 \\
\text{s.t.} & z_0 + z_1^{\mathrm{T}}\xi + \xi^{\mathrm{T}} Z\xi \le z_3 + z_4 + c + b(x^{\mathrm{T}}\xi),\ \forall\, \xi \in \Re^n \\
& z_0 + z_1^{\mathrm{T}}\xi + \xi^{\mathrm{T}} Z\xi \le z_3 + c + b(x^{\mathrm{T}}\xi),\ \forall\, d^{\mathrm{T}}\xi \le r_2 \\
& z_0 + z_1^{\mathrm{T}}\xi + \xi^{\mathrm{T}} Z\xi \le z_3 + z_4,\ \forall\, \xi \in \Re^n \\
& z_0 + z_1^{\mathrm{T}}\xi + \xi^{\mathrm{T}} Z\xi \le z_3,\ \forall\, d^{\mathrm{T}}\xi \le r_2 \\
& z_0 + z_1^{\mathrm{T}}\xi + \xi^{\mathrm{T}} Z\xi \le c + b(x^{\mathrm{T}}\xi) + z_4,\ \forall\, \xi^{\mathrm{T}} A\,\xi + a^{\mathrm{T}}\xi \le r_1 \\
& z_0 + z_1^{\mathrm{T}}\xi + \xi^{\mathrm{T}} Z\xi \le c + b(x^{\mathrm{T}}\xi),\ \forall\, \xi^{\mathrm{T}} A\,\xi + a^{\mathrm{T}}\xi \le r_1,\ d^{\mathrm{T}}\xi \le r_2 \\
& z_0 + z_1^{\mathrm{T}}\xi + \xi^{\mathrm{T}} Z\xi \le z_4,\ \forall\, \xi^{\mathrm{T}} A\,\xi + a^{\mathrm{T}}\xi \le r_1 \\
& z_0 + z_1^{\mathrm{T}}\xi + \xi^{\mathrm{T}} Z\xi \le 0,\ \forall\, \xi^{\mathrm{T}} A\,\xi + a^{\mathrm{T}}\xi \le r_1,\ d^{\mathrm{T}}\xi \le r_2 \\
& z_3, z_4 \ge 0.
\end{array}
$$

By Corollary 7 in [52], under the Assumption 5.2(a) we have the following equivalent conic formulation:

$$
\begin{array}{ll}
\max & z_0 + z_1^{\mathrm{T}}\mu + Z \bullet (\Gamma + \mu\mu^{\mathrm{T}}) - \beta_1 z_3 - \beta_2 z_4 \\
\text{s.t.} & \begin{pmatrix} z_0 - z_3 - z_4 - c & (z_1^{\mathrm{T}} - bx^{\mathrm{T}})/2 \\ (z_1 - bx)/2 & Z \end{pmatrix} \preceq 0, \\
& \begin{pmatrix} z_0 - z_3 - c + t_1 r_2 & (z_1^{\mathrm{T}} - bx^{\mathrm{T}} - t_1 d^{\mathrm{T}})/2 \\ (z_1 - bx - t_1 d)/2 & Z \end{pmatrix} \preceq 0, \\
& \begin{pmatrix} z_0 - z_3 - z_4 & z_1^{\mathrm{T}}/2 \\ z_1/2 & Z \end{pmatrix} \preceq 0, \\
& \begin{pmatrix} z_0 - z_3 + t_2 r_2 & (z_1^{\mathrm{T}} - t_2 d^{\mathrm{T}})/2 \\ (z_1 - t_2 d)/2 & Z \end{pmatrix} \preceq 0, \\
& \begin{pmatrix} z_0 - z_4 - c + t_3 r_1 & (z_1^{\mathrm{T}} - t_3 a^{\mathrm{T}} - bx^{\mathrm{T}})/2 \\ (z_1 - t_3 a - bx)/2 & Z - t_3 A \end{pmatrix} \preceq 0, \\
& \begin{pmatrix} z_0 - c + t_4 r_1 + 2 s_1 r_2 & \frac{z_1^{\mathrm{T}} - bx^{\mathrm{T}} - t_4 a^{\mathrm{T}}}{2} + r_2 v^{\mathrm{T}} - s_1 d^{\mathrm{T}} \\ \frac{z_1 - bx - t_4 a}{2} + r_2 v - s_1 d & Z - t_4 A - v d^{\mathrm{T}} - d v^{\mathrm{T}} \end{pmatrix} \preceq 0, \\
& \begin{pmatrix} z_0 - z_4 + t_5 r_1 & (z_1^{\mathrm{T}} - t_5 a^{\mathrm{T}})/2 \\ (z_1 - t_5 a)/2 & Z - t_5 A \end{pmatrix} \preceq 0, \\
& \begin{pmatrix} z_0 + t_6 r_1 + 2 s_2 r_2 & \frac{z_1^{\mathrm{T}} - t_6 a^{\mathrm{T}}}{2} + r_2 w^{\mathrm{T}} - s_2 d^{\mathrm{T}} \\ \frac{z_1 - t_6 a}{2} + r_2 w - s_2 d & Z - t_6 A - w d^{\mathrm{T}} - d w^{\mathrm{T}} \end{pmatrix} \preceq 0, \\
& \begin{pmatrix} s_1 \\ v \end{pmatrix} = \begin{pmatrix} r_2 + 1 & r_2 - 1 & 0 \\ -a & -a & 2R^{\mathrm{T}} \end{pmatrix} y_1, \\
& \begin{pmatrix} s_2 \\ w \end{pmatrix} = \begin{pmatrix} r_2 + 1 & r_2 - 1 & 0 \\ -a & -a & 2R^{\mathrm{T}} \end{pmatrix} y_2, \\
& y_1, y_2 \in \mathrm{SOC}(2 + \mathrm{rank(A)}), \\
& z_3, z_4, t_1, \cdots, t_6 \geq 0,
\end{array}
$$

where the decision variables are: $x, Z, z_0, z_1, z_3, z_4, v, w, y_1, y_2, s_1, s_2, t_1, \cdots, t_6$, R is a $\mathrm{rank}(A) \times n$ matrix satisfying $A = R^{\mathrm{T}} R$, and SOC denotes the *second-order cone* with

$$
\mathrm{SOC}(m) := \left\{ x = (x_1, \cdots, x_m)^{\mathrm{T}} \in \Re^m \mid x_m \geq \sqrt{x_1^2 + \cdots + x_{m-1}^2} \right\}.
$$

This model fits perfectly within the domain of SeDuMi [51] of Sturm.

5.2 Extending Lo's Option Bound

An immediate application of the results in Section 5.1 is to extend the option bound due to Lo [30], which yields a closed-form upper bound on the expected payoff of a European call option when the first two moments (i.e. the mean and variance) of the underlying asset price at maturity is known. In practice, it is possible to get more information about the distribution of the underlying asset. For instance, one may estimate the statistics of a correlated asset, in the hope that this information will help to sharpen the bound as given in Lo [30]. In this subsection, we consider such bound when a probability bound of another correlated asset is available. Specifically, we consider two different assets, s_1 and s_2, whose mean vector $\mu \in \Re^2$ and covariance matrix Γ (a 2 by 2 positive semidefinite matrix) are given. Suppose that we have an estimation on the probability of stock s_2 above some reference point r. The problem is to get the tightest possible expected value of a European call option on stock s_1. Mathematically, the problem under consideration is:

$$\begin{array}{ll} \max_{(s_1,s_2)\sim(\mu,\Gamma)} & \mathsf{E}[(s_1-k)_+] \\ \text{s.t.} & \text{Prob}\,\{s_2 > r\} \le \beta. \end{array}$$

To avoid trivial cases, using the Chebyshev-Cantelli bound as shown in Proposition 2.1, we shall consider the parameters to satisfy:

$$\beta \in \begin{cases} \left[0,\ \frac{\Gamma_{22}}{(r-\mu_2)^2+\Gamma_{22}}\right], & \text{and } r \ge \mu_2, \\ \\ \left[\frac{(r-\mu_2)^2}{(r-\mu_2)^2+\Gamma_{22}},\ 1\right], & \text{and } r \le \mu_2. \end{cases}$$

Similar as before, the dual is given by

$$\begin{array}{ll} \min_{Z,z_0,z_1,z_3} & z_0 + z_1^{\mathrm{T}}\mu + Z \bullet (\Gamma + \mu\mu^{\mathrm{T}}) + \beta z_3 \\ \text{s.t.} & z_0 + z_1^{\mathrm{T}} \begin{pmatrix} s_1 \\ s_2 \end{pmatrix} + \begin{pmatrix} s_1 \\ s_2 \end{pmatrix}^{\mathrm{T}} Z \begin{pmatrix} s_1 \\ s_2 \end{pmatrix} + z_3 \mathbf{1}_{\{s_2>r\}} \ge (s_1-k)_+,\ \forall\,(s_1,s_2) \in \Re^2 \\ & z_3 \ge 0, \end{array}$$

which, by the S-lemma, can be cast as an SDP:

$$\begin{array}{ll} \min_{Z,z_0,z_1,z_3,w,t} & z_0 + z_1^{\mathrm{T}}\mu + Z \bullet (\Gamma + \mu\mu^{\mathrm{T}}) + \beta z_3 \\ \text{s.t.} & \begin{pmatrix} z_0 + z_3 + k & 0.5\cdot(z_1^{\mathrm{T}} - (1,0)) \\ 0.5\cdot(z_1 - (1,0)^{\mathrm{T}}) & Z \end{pmatrix} \succeq 0, \\ & \begin{pmatrix} z_0 + z_3 & 0.5\cdot z_1^{\mathrm{T}} \\ 0.5\cdot z_1 & Z \end{pmatrix} \succeq 0, \\ & \begin{pmatrix} z_0 + k - wr & 0.5\cdot(z_1^{\mathrm{T}} + (-1,w)) \\ 0.5\cdot(z_1 + (-1,w)^{\mathrm{T}}) & Z \end{pmatrix} \succeq 0, \\ & \begin{pmatrix} z_0 - tr & 0.5\cdot(z_1^{\mathrm{T}} + (0,t)) \\ 0.5\cdot(z_1 + (0,t)^{\mathrm{T}}) & Z \end{pmatrix} \succeq 0, \\ & z_3, w, t \ge 0. \end{array}$$

The above bound is computable using an SDP solver, and is tighter than the corresponding bound of Lo which uses only the first two moments of the underlying asset s_1. The numerical performance of the method will be presented at Section 6.

5.3 Extending Chebyshev's Probability Bound

The same approach can be used to improve the bound on probability if some other probability bound of a related random variable is known. A straightforward application is to strengthen the

Chebyshev type inequality, using the first two moments. In particular, let us use the same setting as in Subsection 5.2, and find the tightest upper and lower bounds of the probability $\text{Prob}\,\{s_1 \geq k\}$ with the information that $(s_1, s_2) \sim (\mu, \Gamma)$, and that $\text{Prob}\,(s_2 > r) \leq \beta$, where

$$\beta \in \begin{cases} \left[0,\ \frac{\Gamma_{22}}{(r-\mu_2)^2+\Gamma_{22}}\right], & \text{and } r \geq \mu_2, \\ \\ \left[\frac{(r-\mu_2)^2}{(r-\mu_2)^2+\Gamma_{22}},\ 1\right], & \text{and } r \leq \mu_2. \end{cases}$$

Consider the upper bounding problem first, which is now cast as

$$\begin{array}{ll} \max_{(s_1,s_2)\sim(\mu,\Gamma)} & \text{Prob}\,\{s_1 \geq k\} \\ \text{s.t.} & \text{Prob}\,\{s_2 > r\} \leq \beta, \end{array}$$

with its dual problem:

$$\begin{array}{ll} \min_{Z,z_0,z_1,z_3} & z_0 + z_1^{\mathrm{T}}\mu + Z \bullet (\Gamma + \mu\mu^{\mathrm{T}}) + \beta z_3 \\ \text{s.t.} & z_0 + z_1^{\mathrm{T}} \begin{pmatrix} s_1 \\ s_2 \end{pmatrix} + \begin{pmatrix} s_1 \\ s_2 \end{pmatrix}^{\mathrm{T}} Z \begin{pmatrix} s_1 \\ s_2 \end{pmatrix} + z_3 \mathbf{1}_{\{s_2 > r\}} \geq \mathbf{1}_{\{s_1 \geq k\}},\ \forall\,(s_1,s_2) \in \Re^2 \\ & z_3 \geq 0. \end{array}$$

Writing out the first constraint explicitly, we have

$$\begin{array}{ll} \min_{Z,z_0,z_1,z_3} & z_0 + z_1^{\mathrm{T}}\mu + Z \bullet (\Gamma + \mu\mu^{\mathrm{T}}) + \beta z_3 \\ \text{s.t.} & z_0 + z_1^{\mathrm{T}} \begin{pmatrix} s_1 \\ s_2 \end{pmatrix} + \begin{pmatrix} s_1 \\ s_2 \end{pmatrix}^{\mathrm{T}} Z \begin{pmatrix} s_1 \\ s_2 \end{pmatrix} + z_3 \geq 1, \forall\ s_1 \geq k, \\ & z_0 + z_1^{\mathrm{T}} \begin{pmatrix} s_1 \\ s_2 \end{pmatrix} + \begin{pmatrix} s_1 \\ s_2 \end{pmatrix}^{\mathrm{T}} Z \begin{pmatrix} s_1 \\ s_2 \end{pmatrix} + z_3 \geq 0,\ \forall\,(s_1,s_2) \in \Re^2, \\ & z_0 + z_1^{\mathrm{T}} \begin{pmatrix} s_1 \\ s_2 \end{pmatrix} + \begin{pmatrix} s_1 \\ s_2 \end{pmatrix}^{\mathrm{T}} Z \begin{pmatrix} s_1 \\ s_2 \end{pmatrix} \geq 1,\ \forall\ s_1 \geq k, s_2 \leq r, \\ & z_0 + z_1^{\mathrm{T}} \begin{pmatrix} s_1 \\ s_2 \end{pmatrix} + \begin{pmatrix} s_1 \\ s_2 \end{pmatrix}^{\mathrm{T}} Z \begin{pmatrix} s_1 \\ s_2 \end{pmatrix} \geq 0,\ \forall\ s_2 \leq r, \\ & z_3 \geq 0. \end{array}$$

Using the S-lemma and the extended S-lemma ([52]), we have

$$\begin{array}{ll} \min_{Z,z_0,z_1,z_3,w,t,p,q} & z_0 + z_1^{\mathrm{T}}\mu + Z \bullet (\Gamma + \mu\mu^{\mathrm{T}}) + \beta z_3 \\ \text{s.t.} & \begin{pmatrix} z_0 + z_3 + wk - 1 & 0.5 \cdot (z_1^{\mathrm{T}} - (w,0)) \\ 0.5 \cdot (z_1 - (w,0)^{\mathrm{T}}) & Z \end{pmatrix} \succeq 0, \\ & \begin{pmatrix} z_0 + z_3 & 0.5 \cdot z_1^{\mathrm{T}} \\ 0.5 \cdot z_1 & Z \end{pmatrix} \succeq 0, \\ & \begin{pmatrix} z_0 - 1 + p\,k - q\,r & \frac{z_1^{\mathrm{T}} + (-p,q)}{2} \\ \frac{z_1 + (-p,q)^{\mathrm{T}}}{2} & Z \end{pmatrix} \succeq 0, \\ & \begin{pmatrix} z_0 - tr & 0.5 \cdot (z_1^{\mathrm{T}} + (0,t)) \\ 0.5 \cdot (z_1 + (0,t)^{\mathrm{T}}) & Z \end{pmatrix} \succeq 0, \\ & z_3, w, t, p, q \geq 0. \end{array}$$

In addition, the lower bound follows immediately by observing

$$\left\{ \begin{array}{ll} \min_{(s_1,s_2)\sim(\mu,\Gamma)} & \text{Prob}\,\{s_1 \geq k\} \\ \text{s.t.} & \text{Prob}\,\{s_2 > r\} \leq \beta \end{array} \right\} = \left\{ \begin{array}{lll} 1- & \max_{(\hat{s}_1,s_2)\sim(\hat{\mu},\hat{\Gamma})} & \text{Prob}\,\{\hat{s}_1 \geq -k\} \\ & \text{s.t.} & \text{Prob}\,\{s_2 > r\} \leq \beta \end{array} \right\}$$

where $\hat{\mu} = (-\mu_1, \mu_2)^{\mathrm{T}}$ and $\hat{\Gamma} = \begin{pmatrix} \Gamma_{11} & -\Gamma_{12} \\ -\Gamma_{21} & \Gamma_{22} \end{pmatrix}$.

6 Numerical Results

In this section, we shall present the results of some numerical experiments to show the effectiveness of the methods proposed in this paper. Specifically, we shall conduct two sets of numerical experiments. First, we shall test the approach presented in Section 4 to select a robust portfolio maximizing the S-shaped value function. Secondly, we evaluate the performance of the new conic bounds which improve Lo's option bound and the classical Chebyshev type probability bound, with varying correlation coefficients.

In the numerical experiments, we use the historical data of the Hong Kong Stock Market. We selected five constituent stocks of the Hang Seng index from July 10, 2008 to January 12, 2009; they are: CHEUNG KONG (0001.HK), HSBC (0005.HK), CATHAY PAC AIR (0293.HK), HKEX (0388.HK) and CHINA MOBILE (0941.HK). The stock data (closing prices) were collected from Yahoo!Finance, from which the mean and variance are estimated as

$$\begin{aligned} \mu &= [0.003684902, 0.004492878, 0.005115208, 0.003893002, 0.003487849]\,, \\ \Gamma &= \begin{bmatrix} 0.001779092 & 0.001204961 & 0.001436253 & 0.001463577 & 0.001311733 \\ 0.001204961 & 0.001386942 & 0.001360971 & 0.001284149 & 0.001113368 \\ 0.001436253 & 0.001360971 & 0.002252675 & 0.001477246 & 0.001210239 \\ 0.001463577 & 0.001284149 & 0.001477246 & 0.002060033 & 0.001486818 \\ 0.001311733 & 0.001113368 & 0.001210239 & 0.001486818 & 0.001651268 \end{bmatrix}. \end{aligned}$$

6.1 Evaluation of the Robust and Non-Robust Portfolios for the S-Shaped Value Function

In this section, we aim to test the performance of the robust portfolio based on the S-shaped value function. The setup for the tests is as follows. To avoid the intractable case, let us consider $X = \{x \mid e^{\mathrm{T}}x = 1\}$. Assuming ξ to be normally distributed with mean and covariance matrix (μ, Γ), and introducing the notations $s = x^{\mathrm{T}}\mu - r,\ t = \sqrt{x^{\mathrm{T}}\Gamma x}$, we have

$$\mathsf{E}[v(\xi^{\mathrm{T}}x - r)] = \mathsf{E}_{\eta\sim N(0,1)}[v(t\eta + s)].$$

The non-robust (taking ξ to be normally distributed by assumption) optimal portfolio is

$$x_{NR} \in \{x \mid x^{\mathrm{T}}\mu = r + s^*, x^{\mathrm{T}}\Gamma x = (t^*)^2, x^{\mathrm{T}}e = 1\}$$

where the pair (s^*, t^*) is the optimal solution to

$$\max_{s,t} \mathsf{E}_{\eta\sim N(0,1)}[v(t\eta + s)]$$

subject to $t \geq \sqrt{b_0(s+r)^2 - 2b_1(s+r) + b_2}$, where b_0, b_1, b_2 are defined before Theorem 2.5.

Basically, we first solve the inner two-variable maximization problem to get the optimal s^* and t^*, and then we solve the equations

$$x^{\mathrm{T}}e = 1, x^{\mathrm{T}}\Gamma x = (t^*)^2, x^{\mathrm{T}}\mu = r + s^*,$$

to get an optimal portfolio x^*.

On the other hand, the robust portfolio x_R can be found by searching along the classical Markowitz's [33] mean-variance efficient frontier to maximize $v_R(t - r, \sigma^*(t))$ as stipulated in Section 4.

We then compare the (sampled) expected S-shaped values:

$$\mathsf{E}[v(\eta^{\mathrm{T}}\Gamma^{\frac{1}{2}}x_{NR} + \mu^{\mathrm{T}}x_{NR} - r)] \quad \text{vs.} \quad \mathsf{E}\left[v\left(\eta^{\mathrm{T}}\Gamma^{\frac{1}{2}}x_R + \mu^{\mathrm{T}}x_R - r\right)\right]$$

where η is actually *not* normal, but log normal (shifted), i.e. $\eta \sim \mathit{LogN}(0, I)$ while maintaining $\Gamma^{\frac{1}{2}}\eta \sim (\mu, \Gamma)$.

For each pair of parameters $(\alpha, \beta, \lambda, r)$, we compute the optimal portfolios for both the non-robust and robust models and conduct 5 simulation runs in order to ensure the stability of simulation results. In each simulation run, we generate 10^5 realizations of the standard lognormal (normal) random vector. The S-shaped function values are then evaluated according to the non-robust and the robust portfolios respectively. We report the probability of the robust portfolio outperforming the non-robust one, the statistical mean and the statistical standard deviation of these 10^5 simulated S-shaped function values, tabulated in tables. In Tables 1 and 2, we present the simulations results regarding the situation where the actual distribution is *lognormal*, while in Tables 3 and 4 we present the results where the actual distribution is *normal*. The first column of each table indicates the simulation run. The probability of over 10^5 realized values that robust models outperform non-robust ones are shown in column two. The average values of these 10^5 realized S-shaped function values for the robust and the non-robust portfolios are reported in columns three and four respectively. The fifth column indicates the relative difference between the non-robust and robust S-shaped function values in the percentage. Similarly, the standard deviations over 10^5 realized values for robust and non-robust models are reported in columns six and seven respectively. The last column reports the relative difference in terms of percentage. The average of the 5 simulation runs are shown in the last row.

Some observations are in order here. In the case where the actual distribution is *lognormal*, it is clear from the fifth column of Tables 1 and 2 that the non-robust model yields a slightly better expected value (less than 1%) than the robust version. However, as shown in the last column the solution found by the robust model has a reduced variability than that of the non-robust portfolio.

	Rob>NonRob	Mean			Standard Deviation		
Run	Prob	Rob(a)	NonRob(b)	%100(b-a)/b	Rob (c)	NonRob (d)	%100(d-c)/d
1	43.22%	0.5303	0.5346	0.80%	0.1392	0.1425	2.33%
2	43.36%	0.5304	0.5347	0.79%	0.1393	0.1426	2.26%
3	43.53%	0.5309	0.5351	0.79%	0.1402	0.1436	2.36%
4	43.49%	0.5310	0.5352	0.78%	0.1397	0.1431	2.34%
5	43.25%	0.5310	0.5353	0.81%	0.1400	0.1436	2.50%
Avg	43.37%	0.5307	0.5350	0.80%	0.1397	0.1431	2.36%

Table 1: Lognormal Simulation for Non-rob vs Rob Portfolios: $(\alpha, \beta, \lambda, r) = (0.5, 0.8, 1.5, 0)$

	Rob>NonRob	Mean			Standard Deviation		
Run	Prob	Rob(a)	NonRob(b)	%100(b-a)/b	Rob (c)	NonRob (d)	%100(d-c)/d
1	43.36%	0.7682	0.7704	0.28%	0.0801	0.0821	2.51%
2	43.50%	0.7681	0.7702	0.28%	0.0799	0.0820	2.55%
3	43.50%	0.7680	0.7702	0.28%	0.0798	0.0819	2.48%
4	43.15%	0.7679	0.7701	0.29%	0.0797	0.0819	2.61%
5	43.26%	0.7679	0.7702	0.29%	0.0799	0.0820	2.55%
Avg	43.35%	0.7680	0.7702	0.28%	0.0799	0.0820	2.54%

Table 2: Lognormal Simulation for Non-rob vs Rob Portfolios: $(\alpha, \beta, \lambda, r) = (0.2, 0.9, 1.5, 0.005)$

	Rob>NonRob	Mean			Standard Deviation		
Run	Prob	Rob(a)	NonRob(b)	%100(b-a)/b	Rob (c)	NonRob (d)	%100(d-c)/d
1	44.19%	0.0526	0.0585	10.03%	0.2219	0.2311	3.97%
2	44.54%	0.0536	0.0592	9.39%	0.2218	0.2311	3.98%
3	44.43%	0.0527	0.0585	9.84%	0.2214	0.2306	4.00%
4	44.41%	0.0531	0.0589	9.92%	0.2222	0.2314	3.99%
5	44.45%	0.0522	0.0580	9.99%	0.2220	0.2315	4.12%
Avg	44.40%	0.0529	0.0586	9.83%	0.2219	0.2311	4.01%

Table 3: Normal Simulation for Non-rob vs Rob Portfolios: $(\alpha, \beta, \lambda, r) = (0.5, 0.8, 1.5, 0)$

	Rob>NonRob	Mean			Standard Deviation		
Run	Prob	Rob(a)	NonRob(b)	%100(b-a)/b	Rob (c)	NonRob (d)	%100(d-c)/d
1	44.30%	0.2117	0.2213	4.34%	0.3492	0.3573	2.27%
2	44.28%	0.2134	0.2231	4.33%	0.3494	0.3573	2.23%
3	44.21%	0.2138	0.2238	4.45%	0.3493	0.3574	2.29%
4	44.17%	0.2115	0.2216	4.55%	0.3492	0.3574	2.29%
5	44.38%	0.2131	0.2226	4.26%	0.3490	0.3572	2.29%
Avg	44.27%	0.2127	0.2225	4.39%	0.3492	0.3573	2.27%

Table 4: Normal Simulation for Non-rob vs Rob Portfolios: $(\alpha, \beta, \lambda, r) = (0.2, 0.9, 1.5, 0.005)$

The first column shows that there are over 40% of the chance a robust solution may actually outperform the non-robust solution.

However, if the distribution is actually normal, then the robust model performs suboptimally. In Tables 3 and 4, we use the same parameters as in Tables 1 and 2 respectively. Although the robust solutions still have over 40% chance to beat the non-robust ones, from the fifth columns we observe that the quality loss on the mean value is close to 10%. This observation seems to confirm the common belief that the more ambiguous the distribution, the better the robust solution. In all the cases, the average performance loss of the robust solution is no more than 10%, and so it can be considered as a practical alternative when the nonrobust formulation is intractable.

6.2 New Option Bounds vs. Lo's Bounds

Pick HSBC (0005.HK) as the underlying stock and Cheung Kung (0001.HK) as another correlated stock. The mean and variance of the stock prices are

$$\mu = [104.1357143 \quad 88.08095238];$$

$$\Gamma_1 = 417.0165543; \Gamma_2 = 279.8505143.$$

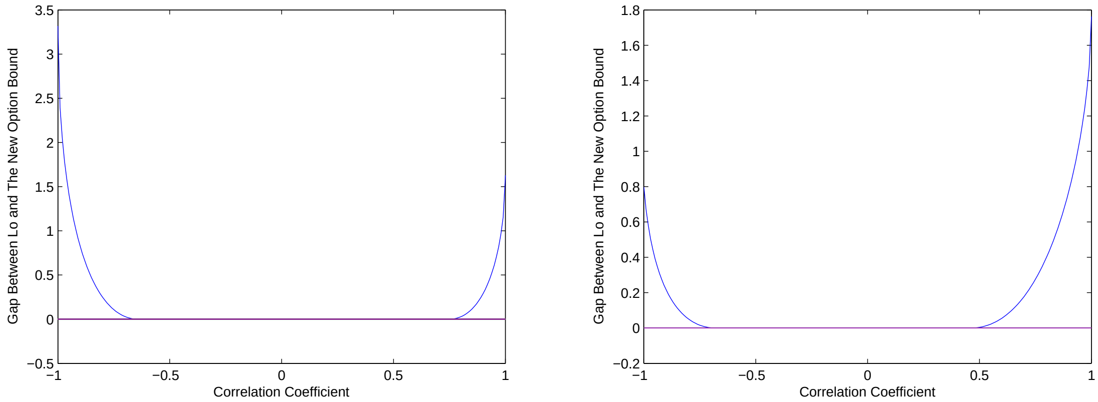

Figure 1: New Option Bound with Strike Price $k = 100$ (lhs) and 150 (rhs)

We choose

$$r = 0.5 \cdot \mu(2);$$
$$\beta = \min\left\{1.02 \cdot \frac{(r-\mu(2))^2}{(\Gamma_2 + (r-\mu(2))^2)}, 1\right\}.$$

With different levels of correlation, the new option bounds and their gap with Lo's bounds do vary. We demonstrate the effect due to the correlation coefficient on the gap with different strike prices in Figure 1. For in-the-money strike price \$100 and out-of-money strike price \$150, when the two stocks are highly correlated, the new option bounds have been significantly improved.

6.3 New Chebyshev Bounds vs. Original Chebyshev Bounds

Using the same parameters as in Section 6.2, we compare the extended Chebyshev probability probability upper and lower bounds; that is

$$\begin{array}{rlll} 0 & \leq \text{Prob}\{s_1 \geq k\} & \leq \frac{\Gamma_1}{(k-\mu_1)^2+\Gamma_1}, & \text{if } k \geq \mu_1, \\ \frac{(k-\mu_1)^2}{(k-\mu_1)^2+\Gamma_1} & \leq \text{Prob}\{s_1 \geq k\} & \leq 1, & \text{if } k \leq \mu_1. \end{array}$$

We observe from the numerical results that for different reference level k, if the level of correlation is high (either positive or negative), then the new bounds are substantially improved as shown from Figures 2 to 3. In the plots, the horizontal line is the Chebyshev upper (lower) bound, which is used as a benchmark for the new bounds.

References

[1] A.V. Balakrishnan, *Applied functional analysis*, 2nd ed., Springer, New York, 1981.

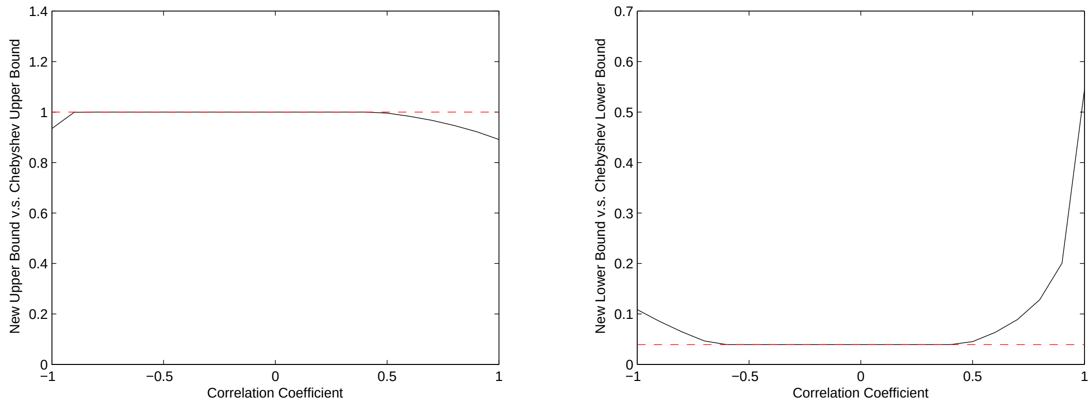

Figure 2: New Chebyshev Upper (lhs) and Lower (rhs) Bounds with $k = 100$

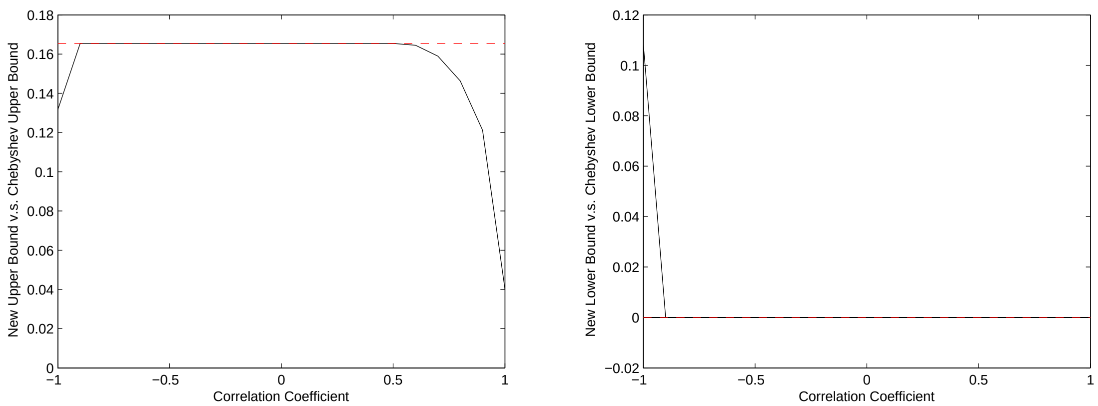

Figure 3: New Chebyshev Upper (lhs) and Lower (rhs) Bounds with $k = 150$

[2] V.S. Bawa, Optimal rules for ordering uncertain prospects, *Journal of Financial Economics*, 2 (1), pp. 95–121, 1975.

[3] A. Ben-Tal and D. Bertsimas and D. B. Brown, A Soft Robust Approach to Optimization under Ambiguity, *Operations Research*, to appear, 2009.

[4] A. Ben-Tal, A. Goryashko, E. Guslitzer and A. Nemirovski, Adjustable Robust Solutions of Uncertain Linear Programs, *Mathematical Programming*, 99 (2), pp. 351–376, 2004.

[5] A. Ben-Tal and A. Nemirovski, Robust convex optimization, *Mathematics of Operations Research*, 23 (4), pp. 769–805, 1998.

[6] A. Ben-Tal and A. Nemirovski, *Lectures on Modern Convex Optimization: Analysis, Algorithm, and Engineering Applications*, MPS/SIAM Ser. on Optimization, SIAM, Philadelphia, 2001.

[7] D. Bertsimas and I. Popescu, On the Relation Between Option and Stock prices: A Convex Optimization Approach, *Operation Research*, 50 (2), pp. 358–374, 2002.

[8] D. Bertsimas and I. Popescu, Optimal Inequality in Probability Theory: A Convex Optimization Approach, *SIAM Journal on Optimization*, 15 (3), pp. 780–804, 2005.

[9] F.P. Cantelli, Intorno ad un teorema fundamentale della teoria del rischio, *Boll. Assoc. Attuar. Ital. (Milan)*, pp. 1–23, 1910.

[10] L. Chen, S. He, and S. Zhang, When all risk-adjusted performance measures are the same: In praise of the Sharpe ratio, *Technical Report SEEM2007-09, Department of Systems Engineering & Engineering Management, The Chinese University of Hong Kong*, 2007.

[11] E. Delage and Y. Ye, Distributionally Robust Optimization under Moment Uncertainty with Application to Data-Driven Problems, *Working Paper, Stanford University*, 2008.

[12] L. El Ghaoui and H. Lebret, Robust Solutions to Least-Square Problems with Uncertain Data, *SIAM Journal on Matrix Analysis and Applications*, 18 (4), pp. 1035-1064, 1997.

[13] L. El Ghaoui, M. Oks and F. Oustry, Worst-Case Value-at-Risk and Robust Portfolio Optimization: A Conic Programming Approach, *Operations Research*, 51 (4), pp. 543-556, 2003.

[14] P. Embrechts, A. McNeil and D. Stranmann, Correlation and Dependence in Risk Management: Properties and Pitfalls, *Risk Management: Value at Risk and Beyond*, eds. M. Dempster and H.K. Moffatt, Cambridge University Press, 2001.

[15] P.C. Fishburn, Mean-Risk Analysis with Risk Associated with Below-Target Returns, *American Economics Review*, 67 (2), pp. 116–126, 1977.

[16] D. Goldfarb and G. Iyengar, Robust Portfolio Selection Problem, *Mathematics of Operations Rsearch*, 28 (1), pp. 1–38, 2003.

[17] M. Grant and S. Boyd, CVX: Matlab software for disciplined convex programming [web page and software], *http://www.stanford.edu/ boyd/cvx*, November 2007.

[18] M. Grant, S. Boyd, and Y. Ye, CVX Users' Guide: for cvx version 1.1 (build 615), *http://www.stanford.edu/ boyd/cvx/cvx_usrguide.pdf*, March 2008.

[19] M. Grötschel, L. Lovász, and A. Schrijver, *Geometric Algorithms and Combinatorial Optimization: Algorithms and Combinatorics*, Springer, Berlin-Heidelberg, 1988.

[20] E. Guslitzer, Uncertainty-Immunized Solutions in Linear Programming, *M.Sc. Thesis*, The Technion, Haifa, Israel, June 2002.

[21] S. He, Z.Q. Luo, J. Nie, and S. Zhang, Semidefnite Relaxation Bounds for Indefinite Homogeneous Quadratic Optimization, *SIAM Journal on Optimization*, 19, pp. 503–523, 2008.

[22] S. He, J. Zhang, and S. Zhang, Bounding Probability of Small Deviation: A Fourth Moment Approach, *Technical Report SEEM2007-09, Department of Systems Engineering & Engineering Management, The Chinese University of Hong Kong*, 2007.

[23] D. Hernández-Hernández and A. Schied, A Control Approach to Robust Utility Maximization with Logarithmic Utility and Time Consistent Penalties, *Stochastic Processes and their Applications*, 117 (8), pp. 980-1000, 2007.

[24] K. Isii, The extrema of probaiblity determined by generalized moments. I. Bounded random variables, *Annals of the Institute of Statistical Mathematics*, 12, pp. 164–168, 1960.

[25] G. Iyengar, Robust Dynamic Programming, *Mathematics of Operations Research*, 30 (2), pp. 257-280, 2005.

[26] D. Kahneman and A. Tversky, Prospect theory: An analysis of decisions under risk, *Econometrica*, 43, pp. 263–291, 1979.

[27] S. Karlin and W.J. Studden, Tchebysheff Systems: With Applications in Analysis and Statistics, Interscience, John Wiley and Sons, Pure Appl. Math. 15, New York, 1966.

[28] J.B. Lasserre, Bounds on measures satisfying moment conditions, *Ann. Appl. Probab.*, 12, pp. 1114–1137, 2002.

[29] D. Li and W. Ng, Optimal Dynamic Portfolio Selection: Multiperiod Mean-Variance Formulation, *Mathematical Finance*, 10 (3), pp. 387-406, 2000.

[30] A. Lo, Semiparametric upper bounds for option prices and expected payoffs, *Journal of Financial Economics*, 19, pp. 373–388, 1987.

[31] M.S. Lobo, L. Vandenberghe, S. Boyd and H. Lebret, Applications of second-order cone programming, *Linear Algebra and its Applications*, 284(1-3), pp. 193-228, 1998.

[32] D.G. Luenberger and Y. Ye, *Linear And Nonlinear Programming*, 3rd edition, Springer-Verlag, New York, 2008.

[33] H.M. Markowitz, Portfolio selection, *Journal of Finance*, 7 (1), pp. 79–91, 1952.

[34] H.M. Markowitz, Porfolio selection: Efficient diversification of investment, *John Willey*, New York, 1959.

[35] R. Mises, The Limits of a Distribution Function if Two Expected Values Are Given, *Ann. Math. Statist.*, 10, pp. 99–104, 1939.

[36] K. Natarajan, D. Pachamanova, and M. Sim, Incorporating Asymmetric Distributional Information in Robust Value-at-Risk Optimization, *Management Science*, 54 (3), pp. 573-585, 2008.

[37] K. Natarajan, M. Sim and J. Uichanco, Tractable Robust Expected Utility and Risk Models for Portfolio Optimization, *Working Paper, National University of Singapore*, March 2008.

[38] K. Natarajan and L. Zhou, A Mean-Variance Bound For A Three-Piece Linear Function, *Probability in Engineering and Informational Sciences*, 21 (4), pp. 611-621, 2007.

[39] Yu. Nesterov, Structure of Non-Negative Polynomial and Optimization Problems, *Preprint DP 9749, Louvain-la-Neuve, Belgium*, 1997.

[40] Yu. Nesterov and A. Nemirovski, Interior-Point Polynomial Algorithms in Convex Programming, *SIAM Studies in Applied Mathematics*, 1994.

[41] A. Nilim and L. El Ghaoui, Robust Solutions to Markov Decision Problems with Uncertain Transition Matrices, *Operations Research*, 53 (5), pp. 780-798, 2005.

[42] I. Popescu, Applications of Optimization in Probability, Finance and Revenue Management, *Ph.D. Dissertation, Applied Mathematics Department and Operations Research Center, MIT, Cambridge, MA*, 1999.

[43] I. Popescu, A Semidefinite Programming Approach to Optimal Moment Bounds for Convex Classes of Distributions, *Mathematics of Operation Research*, 50 (3), pp. 632-657, 2005.

[44] I. Popescu, Robust Mean-Covariance Solutions for Stochastic Optimization, *Operations Research*, 55 (1), pp. 98-112, 2007.

[45] S.I. Resnick, *A Probability Path*, 2nd edition, Birkhäuser, Bonston, 1999.

[46] R.T. Rockafellar and S. Uryasev, Optimization of Conditional Value-at-Risk, *Journal of Risk*, 2, pp. 21–41, 2000.

[47] R.T. Rockafellar and S. Uryasev, Conditional Value-at-Risk for General Loss Distributions, *Journal of Banking and Finance*, 26, pp. 1443–1471, 2002.

[48] A.D. Roy, Safety first and the holding of assets, *Econometrica*, 20 (3), pp. 431–449, 1952.

[49] H. Scarf, A min-max solution of an inventory problem, *Studies in The Mathematical Theory of Inventory and Production. Stanford University Press,California*, pp. 201-209, 1958.

[50] J. Smith, Generalized Chebyshev Inequalities: Theory and Applications in Decision Analysis, *Operations Research*, 43, pp. 807–825, 1995.

[51] J.F. Sturm, Using SeDuMi 1.02, a MATLAB toolbox for optimization over symmetric cones, *Optimization Methods and Software*, 11 (12), pp. 625–653, 1999.

[52] J.F. Sturm and S. Zhang, On Cones of Nonnegative Quadratic Functions, *Mathematics of Operations Research*, 28(2), pp. 246-267, 2003.

[53] A. Tversky and D. Kahneman, Advances in Prospect Theory: Cumulative Representation of Uncertainty, *Journal of Risk and Uncertainty*, 5, pp. 147-176, 1992.

[54] M. Zelen, Bounds on a Distribution Function That Are Functions of Moments to Order Four, *J. Res. Nat. Bur. Stand.*, 53, pp. 377–381, 1954.

[55] S.S. Zhu and M. Fukushima, Worst-Case Conditional Value-at-Risk with Application to Robust Portfolio Management, *Operations Research*, 57(5), pp. 1155–1168, 2009.

[56] L. Zuluaga and J. Peña, A Conic Programming Approach to Generalized Tchebycheff Inequalities, *Mathematics of Operations Research*, 30(2), pp. 369–388, 2005.

A Proof of Theorem 2.9

For the linear loss function $f(x,\xi) = -x^{\mathrm{T}}\xi$, its CVaR auxiliary function is

$$F_\beta(x,\alpha) = \alpha + \frac{1}{1-\beta}\mathsf{E}[-x^{\mathrm{T}}\xi - \alpha]_+.$$

Using Proposition 2.2, Lemma 2.4 and Theorem 2 in [55], we have

$$\begin{aligned}
\mathrm{RC}_\beta(x) : &= \sup_{\pi\in D} \mathrm{CVaR}_\beta(x) \\
&= \min_{\alpha\in\Re} \sup_{\zeta\,\sim\,(x^{\mathrm{T}}\mu,\, x^{\mathrm{T}}\Gamma x)} \alpha + \frac{1}{1-\beta}\mathsf{E}[(-\alpha-\zeta)_+] \\
&= \min_{\alpha\in\Re} \alpha + \frac{1}{1-\beta}\sup_{\zeta\,\sim\,(x^{\mathrm{T}}\mu,\, x^{\mathrm{T}}\Gamma x)} \mathsf{E}\left[(-\alpha-\zeta)_+\right] \\
&= \min_{\alpha\in\Re} \alpha + \frac{1}{1-\beta}\left[\frac{1}{2}\sqrt{x^{\mathrm{T}}\Gamma x + (x^{\mathrm{T}}\mu+\alpha)^2} + \frac{-\alpha - x^{\mathrm{T}}\mu}{2}\right]
\end{aligned}$$

Let

$$h_\beta(\alpha) := \alpha + \frac{1}{2(1-\beta)}\left[-\alpha - x^{\mathrm{T}}\mu + \sqrt{x^{\mathrm{T}}\Gamma x + (x^{\mathrm{T}}\mu+\alpha)^2}\right].$$

By the first order optimality condition, the optimum α_x^* minimizing $h_\beta(\alpha)$ for given x satisfies

$$\alpha_x^* = \frac{2\beta-1}{2\sqrt{\beta(1-\beta)}}\cdot\sqrt{x^{\mathrm{T}}\Gamma x} - x^{\mathrm{T}}\mu \qquad (A.8)$$

$$\mathrm{RC}_\beta(x) = h(\alpha_x^*) = \sqrt{\frac{\beta}{1-\beta}}\cdot\sqrt{x^{\mathrm{T}}\Gamma x} - x^{\mathrm{T}}\mu \qquad (A.9)$$

and

$$\begin{array}{rcll} RC_\beta & = & \min_x & \mathrm{RC}_\beta(x) \\ & & \text{s.t.} & x^{\mathrm{T}} e = 1 \ , \end{array}$$

is equivalent to

$$\begin{array}{lll} \min_{s\in\Re} & \min_x & \sqrt{\frac{\beta}{1-\beta}} \cdot \sqrt{x^{\mathrm{T}}\Gamma x} - s \\ & \text{s.t.} & x^{\mathrm{T}} e = 1, \\ & & x^{\mathrm{T}}\mu = s. \end{array}$$

Since the optimal solution for

$$\begin{array}{ll} \min & x^{\mathrm{T}}\Gamma x \\ \text{s.t.} & x^{\mathrm{T}} e = 1 \\ & x^{\mathrm{T}}\mu = s \end{array}$$

is known to be

$$x^*(s) = (\Gamma^{-1}\mu \ \ \Gamma^{-1}e) \begin{pmatrix} b_0 & -b_1 \\ -b_1 & b_2 \end{pmatrix} \begin{pmatrix} s \\ 1 \end{pmatrix},$$

and

$$x^*(s)^{\mathrm{T}}\Gamma\, x^*(s) = b_0 \cdot s^2 - 2b_1 \cdot s + b_2.$$

Thus

$$\mathrm{RC}_\beta = \min_{s\in\Re} h_\beta(s) := \sqrt{\frac{\beta}{1-\beta}} \cdot \sqrt{b_0 \cdot s^2 - 2b_1 \cdot s + b_2} - s.$$

The first order optimality condition gives

$$h'_\beta(s) = \sqrt{\frac{\beta}{1-\beta}} \cdot \frac{b_0 s - b_1}{\sqrt{b_0 \cdot s^2 - 2b_1 \cdot s + b_2}} - 1 = 0.$$

Solving this equation we get

$$s^*_{RC} = \begin{cases} \frac{\sqrt{b_0 b_2 - b_1^2}}{b_0\sqrt{\frac{\beta b_0}{1-\beta}-1}} + \frac{b_1}{b_0}, & \text{if } \frac{\beta}{1-\beta}\cdot b_0 \geq 1, \\ +\infty, & \text{if } \frac{\beta}{1-\beta}\cdot b_0 < 1; \end{cases}$$

$$x^*_{RC_\beta} = \begin{cases} (\Gamma^{-1}\mu \ \ \Gamma^{-1}e) \begin{pmatrix} b_0 & -b_1 \\ -b_1 & b_2 \end{pmatrix} \begin{pmatrix} \frac{\sqrt{b_0 b_2 - b_1^2}}{b_0\sqrt{\frac{\beta b_0}{1-\beta}-1}} + \frac{b_1}{b_0} \\ 1 \end{pmatrix}, & \text{if } \frac{\beta}{1-\beta}\cdot b_0 \geq 1, \\ (\Gamma^{-1}\mu \ \ \Gamma^{-1}e) \begin{pmatrix} b_0 & -b_1 \\ -b_1 & b_2 \end{pmatrix} \begin{pmatrix} +\infty \\ 1 \end{pmatrix}, & \text{if } \frac{\beta}{1-\beta}\cdot b_0 < 1; \end{cases}$$

$$\mathrm{RC}_\beta = \begin{cases} \frac{\sqrt{b_0 b_2 - b_1^2}\sqrt{\frac{\beta b_0}{1-\beta}-1}}{b_0} - \frac{b_1}{b_0}, & \text{if } \frac{\beta}{1-\beta}\cdot b_0 \geq 1, \\ -\infty, & \text{if } \frac{\beta}{1-\beta}\cdot b_0 < 1. \end{cases}$$

By Theorem 1 in [47], we have

$$\mathrm{RV}_\beta(x) := \sup_{\pi\in D} \mathrm{VaR}_\beta(x) = \alpha^*_x = \frac{2\beta - 1}{2\sqrt{\beta(1-\beta)}} \cdot \sqrt{x^{\mathrm{T}}\Gamma x} - x^{\mathrm{T}}\mu.$$

Following similar analysis, we have

$$s^*_{RV} = \begin{cases} \frac{\sqrt{b_0 b_2 - b_1^2}}{b_0\sqrt{\frac{b_0}{4\beta(1-\beta)}-b_0-1}} + \frac{b_1}{b_0}, & \text{if } \frac{b_0}{4\beta(1-\beta)} \geq 1+b_0; \\ +\infty, & \text{if } \frac{b_0}{4\beta(1-\beta)} \leq 1+b_0; \end{cases}$$

$$x^*_{RV_\beta} = (\Gamma^{-1}\mu \;\; \Gamma^{-1}e) \begin{pmatrix} b_0 & -b_1 \\ -b_1 & b_2 \end{pmatrix} \begin{pmatrix} \frac{\sqrt{b_0 b_2 - b_1^2}}{b_0\sqrt{\frac{b_0}{4\beta(1-\beta)}-b_0-1}} + \frac{b_1}{b_0} \\ 1 \end{pmatrix}, \text{ if } \frac{b_0}{4\beta(1-\beta)} \geq 1+b_0;$$

$$\text{RV}_\beta = \begin{cases} \frac{\sqrt{b_0 b_2 - b_1^2}\sqrt{\frac{b_0}{4\beta(1-\beta)}-b_0-1}}{b_0} - \frac{b_1}{b_0}, & \text{if } \frac{b_0}{4\beta(1-\beta)} \geq 1+b_0, \\ -\infty, & \text{if } \frac{b_0}{4\beta(1-\beta)} \leq 1+b_0. \end{cases}$$

B Proof of Theorem 4.3

The proof is constructive. The main idea is to construct a pair of primal-dual feasible solutions for this infinite-dimensional linear program, and make sure that they satisfy the complementary slackness conditions. (In the case of linear programming, this ensures optimality). The optimality will then follow. In particular, consider our primal problem:

$$v_P = \min_{\eta \sim (\mu,\sigma^2)} \mathsf{E}[v(\eta)],$$

and its dual problem

$$\begin{array}{rcll} v_D & = & \max_{y_0,y_1,y_2} & y_0 + y_1\mu + (\mu^2+\sigma^2)y_2 \\ & & \text{s.t.} & y_0 + y_1 x + y_2 x^2 \leq v(x), \quad \forall x \in \Re. \end{array}$$

Here we shall elaborate the relationship between the primal and dual optimal solutions as follows. For any point $x \in \Re$, the primal optimal solution η^* has nonzero probability on x if and only if the dual solution (y_0^*, y_1^*, y_2^*) satisfies $y_0^* + y_1^* x + y_2^* x^2 = v(x)$. In other words, the touching points of the smooth curves $f(x) := y_0 + y_1 x + y_2 x^2$ and $v(x)$ should be the points where the primal optimal solution (a distribution) places all its masses, according to the complementary slackness condition.

As illustrated by Figure 4, the two functions will have two tangent points at most, and indeed there will be two, in order to satisfy the primal feasibility. Let us denote these points as $-a$ and b, where $a, b \geq 0$. Due to the tangency condition, they must satisfy the following two equations:

$$\begin{cases} f(-a) - v(-a) = 0, \\ f'(-a) - v'(-a) = 0, \end{cases} \tag{B.10}$$

and

$$\begin{cases} f(b) - v(b) = 0, \\ f'(b) - v'(b) = 0. \end{cases} \tag{B.11}$$

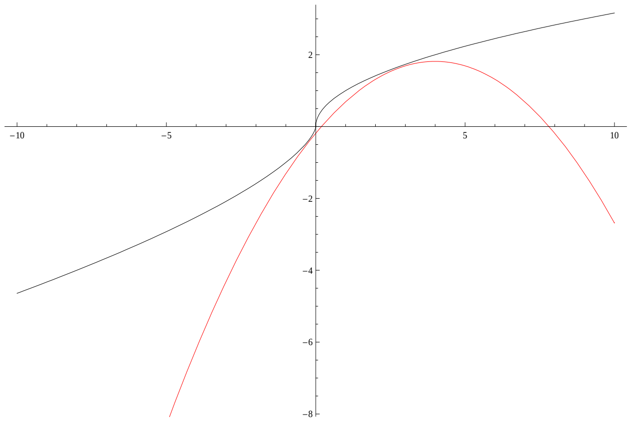

Figure 4: The S-shaped and quadratic functions, in view of the dual problem

Writing them down explicitly leads to

$$(2-\alpha)b^{\alpha} + \lambda(2-\beta)a^{\beta} - \alpha a b^{\alpha-1} - \lambda\beta b a^{\beta-1} = 0. \tag{B.12}$$

On the other hand, the two-point distribution will have to satisfy the first and second moment constraints. This can be achieved using a single variable s with $s \in \left(\max\{0, \mu/\sigma\}, \frac{\mu+\sqrt{\mu^2+\sigma^2}}{\sigma}\right]$, and the distribution is constructed explicitly as:

$$\eta = \begin{cases} \mu - \sigma s, & \text{with probability } \frac{1}{s^2+1}; \\ \mu + \sigma/s, & \text{with probability } \frac{s^2}{s^2+1}. \end{cases} \tag{B.13}$$

Substituting $a = \sigma s - \mu$ and $b = \mu + \sigma/s$ into the above equation (B.12), we deduce that this can be done by setting s to be a root of the following function

$$g(x) := (2-\alpha)(\sigma/x+\mu)^{\alpha} + \lambda(2-\beta)(\sigma x-\mu)^{\beta} - \alpha(\sigma x-\mu)(\sigma/x+\mu)^{\alpha-1} - \lambda\beta(\sigma/x+\mu)(\sigma x-\mu)^{\beta-1},$$

at the interval $x \in \left(\max\{0, \mu/\sigma\}, \frac{\mu+\sqrt{\mu^2+\sigma^2}}{\sigma}\right]$. It is easy to verify that $g(x)$ changes its sign at the interval, hence a root always exists. And we also have $0 < a < b$, as $\max\{0, \mu/\sigma\} \le s \le \frac{\mu+\sqrt{\mu^2+\sigma^2}}{\sigma}$. With one root s and the corresponding a, b values in hand, we claim the solution

$$\begin{aligned} y_2 &:= \frac{(1-\alpha)b^{\alpha}+\lambda(1-\beta)a^{\beta}}{a^2-b^2}; \\ y_1 &:= \alpha b^{\alpha-1} - 2y_2 b; \\ y_0 &:= (1-\alpha)b^{\alpha} + y_2 b^2, \end{aligned}$$

is a dual feasible solution with $y_2 < 0$ $(0 < a < b)$. In fact, the dual feasibility is equivalent to the nonnegativity of the following function $w(x) := v(x) - (y_0 + y_1 x + y_2 x^2)$. Taking the first and the second derivatives, it is easy to observe that $w''(x)$ is an increasing function in $(-\infty, 0]$ and $[0, \infty)$ separately, and it has one unique root $q > 0$. Thus, $w'(x)$ is increasing for $x \in \Re_- \cup (q, +\infty)$ and $w'(x)$ is decreasing for $x \in [0, q]$. Thus, there can be at most three roots for $w'(x)$. Therefore,

$w(x)$ can only be W-shaped. From (B.10) and (B.11), it is clear that $-a$ and b are the two minimal extreme points, so we conclude that $w(x) \geq 0, \forall x \in \Re$. Moreover, the dual feasible solution satisfies the complementary relationship with the primal feasible solution (B.13) that we had identified before. Therefore the strong duality holds by this complementary duality pair. The proof is complete by the construction.

C Proof of Proposition 4.4

a. From the proof of Theorem 4.3, we know that an optimal solution for $v_R(t, \sigma^*(t+r))$ is a two-point distribution, i.e.

$$\eta^* = \begin{cases} t - \sigma^*(t+r) \cdot s, & \text{with probability } \frac{1}{s^2+1}; \\ t + \sigma^*(t+r)/s, & \text{with probability } \frac{s^2}{s^2+1}. \end{cases}$$

For any fixed s, let us denote

$$a(t;s) := \sigma^*(t+r) \cdot s - t,\ b(t;s) := t + \sigma^*(t+r)/s, \text{ and } h(t;s) := \frac{s^2 \cdot b(t;s)^\alpha}{s^2+1} - \frac{\lambda \cdot a(t;s)^\beta}{s^2+1},$$

and so

$$v_R(t, \sigma^*(t+r)) = \inf_{s \geq \max\{0, t/\sigma^*(t+r)\}} h(t;s).$$

By the definition of quasi-concavity (i.e. the upper level sets of $v_R(t, \sigma^*(t+r))$ are always convex), it suffices to show that $h(t;s)$ is quasi-concave in t for any fixed s, because an upper level set of $v_R(t, \sigma^*(t+r))$ is the intersection of the upper level sets of $h(t;s)$, for any $s \geq \max\{0, t/\sigma^*(t+r)\}$. To simplify the notation, let us drop the dependence on s in $a(t;s)$, $b(t;s)$ and $h(t;s)$. For our purpose, it suffices to show that $h(t)$ does not have any local minimum point. Consider the first and second derivatives of $h(t)$:

$$\begin{aligned} h'(t) &= \frac{s^2}{s^2+1} \cdot \alpha b(t)^{\alpha-1} b'(t) - \frac{\lambda\beta \cdot a(t)^{\beta-1} a'(t)}{s^2+1}; \\ h''(t) &= \frac{s^2}{s^2+1} \cdot \alpha[(\alpha-1) b(t)^{\alpha-2} (b'(t))^2 + b(t)^{\alpha-1} b''(t)] \\ &\quad - \frac{\lambda\beta}{s^2+1}[(\beta-1) a(t)^{\beta-2} (a'(t))^2 + a(t)^{\beta-1} a''(t)]. \end{aligned}$$

For any stationary point of $h(t)$ (i.e. t satisfying $h'(t) = 0$), we have

$$h''(t) = \frac{s^2 \cdot \alpha b(t)^{\alpha-1} b'(t)}{s^2+1}[(\alpha-1) b'(t)/b(t) - (\beta-1) a'(t)/a(t) + b''(t)/b(t) - a''(t)/a(t)].$$

We aim to show that whenever $h'(t) = 0$ it follows that $h''(t) \leq 0$, since this would imply the quasi-concavity of h. Since $b'(t) = 1 + [\sigma^*(t+r)]'/s > 0$, and so if $h'(t) = 0$ we have $a'(t) > 0$. It remains to prove

$$(\alpha-1) b(t)'/b(t) - (\beta-1) a(t)'/a(t) + b(t)''/b(t) - a(t)''/a(t) \leq 0. \tag{C.14}$$

Observe that

$$b''(t)/b(t) - a''(t)/a(t) = \frac{-t\cdot[\sigma^*(t+r)]''}{a(t)b(t)}\left(\frac{1}{s}+s\right) \leq 0, \tag{C.15}$$

$$\frac{b'(t)}{b(t)} - \frac{a'(t)}{a(t)} = \frac{s+\frac{1}{s}}{a(t)b(t)}\left[\sigma^*(t+r) - t\cdot[\sigma^*(t+r)]'\right] \geq 0 \tag{C.16}$$

and so

$$\begin{aligned} & (\alpha-1)b'(t)/b(t) - (\beta-1)a'(t)/a(t) + b''(t)/b(t) - a''(t)/a(t) \\ \leq\ & (\beta-1)[b'(t)/b(t) - a'(t)/a(t)] + b''(t)/b(t) - a''(t)/a(t) \leq 0 \end{aligned}$$

where we used (C.15) and (C.16), and the fact that $0 < \alpha \leq \beta \leq 1$. This proves (C.14), and the proposition follows as a consequence.

b. For a fixed t, let us denote $V(t) = (\sigma^*(t))^2$, and suppose that x_t is the optimal portfolio for the expected return $t+r_0$. Since x_0 is the minimizer of $x^{\mathrm{T}}\Gamma\, x$ with $x \in X$, we have $(x-x_0)^{\mathrm{T}}\Gamma x_0 \geq 0$ for all $x \in X$, and so $x^{\mathrm{T}}\Gamma x_0 \geq x_0^{\mathrm{T}}\Gamma x_0 \geq 0$. Because the feasible region is affine linear, we know that for any $s > 1$, $y_s := sx_t - (s-1)x_0 \in X$ and $\mu^{\mathrm{T}} y_s = s(t+r_0) - (s-1)r_0 = st + r_0$. Therefore,

$$\begin{aligned} V(st+r_0) &\leq y_s^{\mathrm{T}}\Gamma y_s \\ &= s^2 x_t^{\mathrm{T}}\Gamma x_t + (s-1)^2 x_0^{\mathrm{T}}\Gamma x_0 - 2s(s-1)x_t^{\mathrm{T}}\Gamma x_0 \\ &\leq s^2 x_t^{\mathrm{T}}\Gamma x_t + (s-1)^2 x_0^{\mathrm{T}}\Gamma x_0 \\ &= s^2 V(t+r_0) + (s-1)^2 V(r_0). \end{aligned}$$

Consequently,

$$\begin{aligned} V(st+r_0)\,|'_{s=1} &= \lim_{s\to 1^+} \frac{V(st+r_0) - V(t+r_0)}{s-1} \\ &\leq \lim_{s\to 1^+} \frac{(s^2-1)V(t+r_0) + (s-1)^2 V(r_0)}{s-1} \\ &= 2V(t+r_0). \end{aligned}$$

Therefore,

$$2[\sigma^*(t+r_0)]^2 = 2V(t+r_0) \geq V(st+r_0)\,|'_{s=1} = 2t\sigma^*(t+r_0)[\sigma^*(t+r_0)]'.$$

Since $0 \notin X$ and $\Gamma \succ 0$, it follows that $\sigma^*(t+r_0) > 0$, and so $\sigma^*(t+r_0) \geq t[\sigma^*(t+r_0)]'$ for any t. Now replacing t with $t+r-r_0$, since $r > r_0$ we also have $\sigma^*(t+r) \geq (t+r-r_0)[\sigma^*(t+r)]' \geq t[\sigma^*(t+r)]'$.