Journal of Computational and Applied Mathematics 236 (2012) 3373–3393

Contents lists available at SciVerse ScienceDirect

Journal of Computational and Applied Mathematics

journal homepage: www.elsevier.com/locate/cam

Robust portfolio selection involving options under a "marginal + joint" ellipsoidal uncertainty set[☆]

Ai-fan Ling[a,*], Cheng-xian Xu[b]

[a] The Research Center of Financial Development and Risk Prevention, School of Finance, Jiangxi University of Finance & Economics, Nanchang, 330013, PR China
[b] Hangzhou Institute of Service Engineering, Hangzhou Normal University, Hangzhou 310036, PR China

ARTICLE INFO

Article history:
Received 16 September 2010
Received in revised form 18 February 2011

MSC:
91G10
90B25

Keywords:
Robust optimization
Robust portfolio
Probability constraint
Linear matrix inequality

ABSTRACT

In typical robust portfolio selection problems, one mainly finds portfolios with the worst-case return under a given uncertainty set, in which asset returns can be realized. A too large uncertainty set will lead to a too conservative robust portfolio. However, if the given uncertainty set is not large enough, the realized returns of resulting portfolios will be outside of the uncertainty set when an extreme event such as market crash or a large shock of asset returns occurs. The goal of this paper is to propose robust portfolio selection models under so-called "marginal + joint" ellipsoidal uncertainty set and to test the performance of the proposed models. A robust portfolio selection model under a "marginal + joint" ellipsoidal uncertainty set is proposed at first. The model has the advantages of models under the separable uncertainty set and the joint ellipsoidal uncertainty set, and relaxes the requirements on the uncertainty set. Then, one more robust portfolio selection model with option protection is presented by combining options into the proposed robust portfolio selection model. Convex programming approximations with second-order cone and linear matrix inequalities constraints to both models are derived. The proposed robust portfolio selection model with options can hedge risks and generates robust portfolios with well wealth growth rate when an extreme event occurs. Tests on real data of the Chinese stock market and simulated options confirm the property of both the models. Test results show that (1) under the "marginal + joint" uncertainty set, the wealth growth rate and diversification of robust portfolios generated from the first proposed robust portfolio model (without options) are better and greater than those generated from Goldfarb and Iyengar's model, and (2) the robust portfolio selection model with options outperforms the robust portfolio selection model without options when some extreme event occurs.

1. Introduction

Optimal portfolio selection has become a major research area in financial economics. The mean-variance (MV) portfolio selection model, proposed by Markowitz [1,2], provides a fundamental basis for portfolio selection in both theoretical and practical applications. The MV efficient frontier could be derived by solving the MV model and the optimal portfolio can be found when the expected returns and covariance matrix of risk assets and investor's attitude to risk are exactly known. However, Michaud pointed out in [3] that although Markowitz efficiency is a convenient and useful theoretical framework for portfolio selection, in practice it is an error prone procedure that often results in error-maximized and

☆ This work is supported by National Natural Science Foundations of China (No. 71001045, 10971162), Natural Science Foundation of Jiangxi Province of China (No. 20114BAB211008) and Jiangxi University of Finance and Economics Support Program Funds for Outstanding Youths.
* Correspondence to: School of Finance, Jiangxi University of Finance & Economics, Nanchang, 330013, PR China. Tel.: +86 0791 83816792; fax: +86 0791 83816792.
E-mail addresses: aifanling@yahoo.com.cn (A.-f. Ling), mxxu@mail.xjtu.edu.cn (C.-x. Xu).

0377-0427/$ – see front matter
doi:10.1016/j.cam.2012.03.023

investment- irrelevant portfolios. This shows the fact that optimal solutions obtained from the MV model are often sensitive to perturbations in parameters of the problem. Because the estimations of the market parameters are subject to statistical errors, the optimal solutions obtained from the MV model might not be or may be far away from an optimal portfolio of risk assets.

Emerging modeling tools based on the robust optimization technique are proposed by Soyster [4] and developed by Ben-Tal and Nemirovski [5,6], El Ghaoui and Lebret [7] to overcome the estimation errors of parameters. Many researchers focus on portfolio selection models with either uncertainty parameters or uncertainty distribution that is called robust portfolio selection. For example, Goldfarb and Iyengar [8] proposed a robust factor model for the asset returns, and assumed that the vector of random asset returns $\mathbf{r} \in \mathbb{R}^n$ follows the following factor model

$$\mathbf{r} = \boldsymbol{\mu} + V^T\mathbf{f} + \boldsymbol{\epsilon},$$

where $\boldsymbol{\mu} \in \mathbb{R}^n$ is the mean return vector, $\mathbf{f} \in \mathbb{R}^m$ the random return vector of $m(<n)$ factors that drive the market, $V \in \mathbb{R}^{m\times n}$ the factor loading matrix and $\boldsymbol{\epsilon}$ the vector of residual returns, with $\mathbf{f}$ and $\boldsymbol{\epsilon}$ being independent and ϵ_i, ϵ_j independent for any $i \neq j$. Furthermore, under the assumptions that $\mathbf{f}$ follows a joint normal distribution with mean $\mathbf{0}$ and positive definite covariance matrix F and that ϵ_i follows a normal distribution with mean 0 and variance $d_i(i = 1, \ldots, n)$, they showed that when the parameters $\boldsymbol{\mu}$ and V are selected from the following uncertainty sets

$$\mathcal{S}_m = \{\boldsymbol{\mu} : \boldsymbol{\mu} = \boldsymbol{\mu}_0 + \boldsymbol{\xi}, |\xi_i| \le \gamma_i, i = 1, \ldots, n\},$$
$$\mathcal{S}_v = \{V : V = V_0 + V, \|\mathbf{v}_i\|_Q \le \rho_i, i = 1, \ldots, n\},$$

the resulting robust portfolio selection problem can be reformulated as a second-order cone program (SOCP) which can be solved efficiently by interior point algorithms [9], where constants γ_i and ρ_i depend on confidence levels given by users, and $\mathbf{v}_i$ is the i-th column of the matrix V and $\|\mathbf{a}\|_Q$ denotes the weighted norm of the vector $\mathbf{a}$ with respect to a symmetric positive definite matrix Q.

Differing from Goldfarb and Iyengar's assumptions, El Ghaoui et al. [10] proposed a worst-case Value-at-Risk (VaR) robust portfolio selection model under the assumption that only the first order and second order moments of asset returns are known. Recently, Zhu and Fukushima [11] presented a worst-case conditional VaR portfolio selection model in which mixture distributions with uncertainty parameters are considered. Zhu et al. [12] and Huang et al. [13] extended Zhu and Fukushima's result to robust portfolio selection with downside risk constraints and uncertainty exit time constraints. More robust portfolio selection models with either parameter uncertainties or distribution uncertainties can be found in the recent survey paper provided by Fabozzi et al. [14].

The uncertainty sets $\mathcal{S}_m$ and $\mathcal{S}_v$ considered by Goldfarb and Iyengar [8] are in fact two *separable* uncertainty sets. Halldórsson and Tütüncü [15], Tütüncü and Koenig [16] considered alternative two separable uncertainty sets with box-type structure for the mean and covariance matrix. It is pointed out by Tütüncü and Koenig [16] that all separable uncertainty sets have two common properties: (a) the actual confidence level of an uncertainty set is unknown, and can be much higher than the desired one; and (b) the uncertainty sets are fully or partially box-type. Lu [17,18] extends Goldfarb and Iyengar [8] results to a robust factor model with a joint ellipsoidal uncertainty set to overcome the shortcoming of separable uncertainty sets and proposes a robust mean-variance portfolio selection model.

Robust portfolio selections mentioned above are relatively insensitive to the distributional input parameters and usually outperform the classical MV portfolio selection (e.g. see [19]). Moreover, the realized portfolio return will be greater than or equal to the calculated worst-case portfolio return when asset returns are realized within the considered uncertainty sets. However, the good performance of robust portfolio selection may fail when an extreme event occurs, for example a market crash, for which the asset returns can fall outside the underlying uncertainty sets. The generated robust portfolios will be unprotected or have a weak guarantee in these cases. A straightforward way to avoid this problem is to enlarge uncertainty sets to cover almost all extreme events. But, robust portfolios obtained by this way will be too conservative to perform well under normal market conditions. Hence, how to protect the worst-case portfolio return in these cases becomes a challenging problem for robust portfolio selection researches. Facing this challenge, Lutgens et al. [20] proposed a robust portfolio selection model involving options, and the resulting model is a second-order cone program which may have, in the worst-case, an exponential number of conic constraints. Zymler et al. [21] generalize Lutgens and Sturm's method and present an expectation maximization robust portfolio selection model under an ellipsoidal uncertainty set with derivative insurance guarantee. By combining options into robust portfolio selection model, the realized robust portfolio with options is protected and can outperform the robust portfolio without options under extreme event or normal market conditions.

In this paper we will also consider the robust portfolio selection problem with asset returns driven by the factor model. It is assumed that the random parameters $\boldsymbol{\mu}$ and V in the model vary in a joint ellipsoidal uncertainty set instead of separable uncertainty sets. European style options are introduced into the robust portfolio selection model to protect affection of extreme events. The main contributions of this paper are as follows. A robust portfolio selection model is proposed based on a factor model of random asset returns. The model maximizes the expectation of portfolio return subject to a worst-case probability loss. A so-called "marginal + joint" ellipsoidal uncertainty set is introduced for parameters in the model. The model is an extension of [8] robust VaR portfolio selection model with separable uncertainly sets $\mathcal{S}_m \times \mathcal{S}_v$ and [17,18] results. Then options are introduced into the robust portfolio selection model to protect the worst-case portfolio return in extreme events. It is shown that the resulting robust portfolio selection model can provide strong performance guarantees for most of possible realizations of asset returns, and that the resulting model can be reformulated as a tractable convex programming with second-order cone and linear matrix inequality constraints.

The rest of the paper is organized as follows. The robust portfolio selection model without options under a joint ellipsoidal uncertainty set is presented in Section 2. While the robust portfolio selection model with options guarantee is given in Section 3. The definition of the joint ellipsoidal uncertainty set is recalled in Section 4. In Section 5, a tight approximation to the robust portfolio selection model with options is derived and the"marginal + joint" uncertainty set is described. The proposed robust portfolio selection models are reformulated as convex programming with linear matrix inequality constraints in Section 6. Test results are reported in Section 7.

2. Robust portfolio selection without options

Assume that there exist n risky assets (stocks) available in a market. The random returns rate vector of the n risky assets over the investment horizon $[0, T]$ is denoted by $\mathbf{r} = (r_1, \ldots, r_n)^T \in \mathbb{R}^n$ with the interpretation that asset i returns $(1 + r_i)$ dollars for every dollar invested in it. The returns on the assets in different market periods are assumed to be independent. The single period return $\mathbf{r}$ is assumed to be a random vector and satisfy the following factor model

$$\mathbf{r} = \boldsymbol{\mu} + V^T\mathbf{f} + \boldsymbol{\epsilon}, \tag{2.1}$$

where $\boldsymbol{\mu} \in \mathbb{R}^n$ is the mean vector of asset returns, $\mathbf{f} \in \mathbb{R}^m$ is the random return vector of $m(<n)$ factors that drive the market, $V \in \mathbb{R}^{m\times n}$ is the factor loading matrix and $\boldsymbol{\epsilon}$ is the vector of residual returns. The market factor vector $\mathbf{f}$ and residual vector $\boldsymbol{\epsilon}$ satisfy $\mathbf{f} \sim \mathcal{N}(0, F)$ and $\epsilon_i \sim \mathcal{N}(0, d_i)(i = 1, \ldots, n)$, respectively, where $\boldsymbol{\xi} \sim \mathcal{N}(\boldsymbol{\mu}, \Sigma)$ means that $\boldsymbol{\xi}$ follows a multivariate normal distribution with mean vector $\boldsymbol{\mu}$ and covariance matrix Σ. The covariance matrix F is positive definite and will be denoted by $F \succ 0$. Let $\mathbf{d} = (d_1, \ldots, d_{nn})^T$, $d_i > 0$, $i = 1, \ldots, n$, then $D = \mathrm{diag}(\mathbf{d}) \succ 0$ means D is a diagonal matrix with d_i, $i = 1, 2, \ldots, n$ being its diagonal elements. As mentioned in the Introduction, separable uncertainty sets will lead to relatively conservative portfolios. Therefore, differing from the separable uncertainty sets $\mathcal{S}_m \times \mathcal{S}_v$ used by Goldfarb and Iyengar [8], it is assumed in this paper that parameters $\boldsymbol{\mu}$ and V in (2.1) vary in a joint ellipsoidal uncertainty set $\mathcal{S}$ (see definition in the next section) and the robust portfolio selection model is as follows

$$\begin{array}{lll}(\mathrm{R1}): & \max\limits_{\mathbf{w}} & \min\limits_{(\boldsymbol{\mu},V)\in\mathcal{S}} \mathbb{E}[\mathbf{w}^T\mathbf{r}] \\ & \text{s.t.} & \max\limits_{(\boldsymbol{\mu},V)\in\mathcal{S}} \mathbb{P}\{\mathbf{w}^T\mathbf{r} \le \rho\} \le \alpha, \\ & & \mathbf{w} \in \mathcal{W} = \{\mathbf{w} : \mathbf{w} \ge 0 \quad \mathbf{e}_n^T\mathbf{w} = 1\},\end{array}$$

where $\mathbf{w}$ is the vector of decision variables, ρ is the minimum return that the investor can accept for a given confidence level $\alpha \in (0, 1/2)$, $\mathbf{e}_n \in \mathbb{R}^n$ is the all-one vector, and $\mathbb{E}[\cdot]$ and $\mathbb{P}\{\cdot\}$ denote the expectation and probability operators, respectively. The probability constraint in problem (R1) is also called *chance constraint* or Roy's safety-first rule in literature.

It follows from (2.1) that $\mathbf{f} \sim \mathcal{N}(0, F)$ and $\epsilon_i \sim \mathcal{N}(0, d_i)(i = 1, \ldots, n)$ imply

$$\mathbf{r} \sim N(\boldsymbol{\mu}, V^TFV + D), \tag{2.2}$$

and

$$\mathbf{w}^T\mathbf{r} \sim N(\mathbf{w}^T\boldsymbol{\mu}, \mathbf{w}^T(V^TFV + D)\mathbf{w}). \tag{2.3}$$

Let $\xi \sim \mathcal{N}(0, 1)$, then by the symmetry of normal distribution, for any given $\boldsymbol{\mu}, V$, the chance constraint in problem (R1) can be rewritten as

$$\begin{aligned}\mathbb{P}\{\mathbf{w}^T\mathbf{r} \le \rho\} &= \mathbb{P}\{\mathbf{w}^T\mathbf{r} \ge 2\mathbf{w}^T\boldsymbol{\mu} - \rho\} \\ &= \mathbb{P}\left\{\mathbf{w}^T\boldsymbol{\mu} + \xi\sqrt{\mathbf{w}^T(V^TFV + D)\mathbf{w}} \ge 2\mathbf{w}^T\boldsymbol{\mu} - \rho\right\} \\ &= \mathbb{P}\left\{\xi \ge \frac{\mathbf{w}^T\boldsymbol{\mu} - \rho}{\sqrt{\mathbf{w}^T(V^TFV + D)\mathbf{w}}}\right\} \le \alpha \\ &\Leftrightarrow \mathbb{P}\left\{\xi \le \frac{\mathbf{w}^T\boldsymbol{\mu} - \rho}{\sqrt{\mathbf{w}^T(V^TFV + D)\mathbf{w}}}\right\} \ge 1 - \alpha \\ &\Leftrightarrow \frac{\mathbf{w}^T\boldsymbol{\mu} - \rho}{\sqrt{\mathbf{w}^T(V^TFV + D)\mathbf{w}}} \ge \mathcal{F}_\xi^{-1}(1 - \alpha),\end{aligned} \tag{2.4}$$

where $\mathcal{F}_\xi(\cdot)$ is the cumulative probability function of the random variable ξ and satisfies $\mathcal{F}_\xi^{-1}(1-\alpha) > 0$ since $\alpha \in (0, 1/2)$. Based on (2.4), problem (R1) can be written as the following equivalent form

$$\begin{array}{lll}(\mathrm{R1}') & \max\limits_{\mathbf{w}} & \min\limits_{(\boldsymbol{\mu},V)\in\mathcal{S}} \mathbf{w}^T\boldsymbol{\mu} \\ & \text{s.t.} & \max\limits_{(\boldsymbol{\mu},V)\in\mathcal{S}} \left\{\mathcal{F}_\xi^{-1}(1 - \alpha)\sqrt{\mathbf{w}^T(V^TFV + D)\mathbf{w}} - \mathbf{w}^T\boldsymbol{\mu}\right\} \le -\rho \\ & & \mathbf{w} \in \mathcal{W}.\end{array}$$

From the last inequality in (2.4), the inequality constraint in problem (R1′) implies that

$$\min_{(\boldsymbol{\mu},V)\in\mathcal{S}} \frac{\mathbf{w}^T\boldsymbol{\mu}-\rho}{\sqrt{\mathbf{w}^T(V^TFV+D)\mathbf{w}}} \geq \mathcal{F}_{\xi}^{-1}(1-\alpha).$$

It follows that the worst-case Sharp ratio [22] is required not less than a constant. Comparing with the mean-variance model as considered in [17], problem (R1′) can in fact obtain the portfolio with greater Sharp ratio, i.e. the obtained portfolio based on (R1′) has greater return per unit risk, see the further discussion in Section 7 for detail.

3. Robust portfolio selection with options

Assume that there exist q derivatives (European options) on n stocks given in Section 2. Let the expiration dates of all options be the end of investment horizon $[0, T]$ and denote the random gross return rate vector of q options by $\mathbf{r}^d = (r_1^d, \ldots, r_q^d)^T \in \mathbb{R}^q$ at the expiration date. Assume that option i on stock j has strike price K_i. If i is a put option with price p_i at time $t = 0$, then its return at time $t = T$ can be expressed as (see also [20,21] for detail)

$$\begin{aligned} r_i^d &= \frac{\max\{0, K_i - S_T^j\}}{p_i} = \frac{\max\{0, K_i - (1+r_j)S_j^0\}}{p_i} \\ &= \max\left\{0, \frac{K_i - (1+r_j)S_0^j}{p_i}\right\} = \max\{0, \widehat{b}_i + a_{ij}(1+r_j)\}, \end{aligned} \tag{3.1}$$

where $\widehat{b}_i = K_i/p_i > 0$ and $a_{ij} = -S_0^j/p_i < 0$, S_0^j is the price of stock j at time $t = 0$, $j = 1, 2, \ldots, n$. It can be seen that r_i^d is a convex piece-wise linear function with respect to r_j. If i is a call option with price c_i at $t = 0$, then its return at $t = T$ is

$$r_i^d = \max\left\{0, \frac{(1+r_j)S_0^j - K_i}{c_i}\right\} = \max\{0, \widehat{b}_i + a_{ij}(1+r_j)\} \tag{3.2}$$

with $\widehat{b}_i = -K_i/c_i < 0$ and $a_{ij} = S_0^j/c_i > 0$.

Then using the component-wise maximization operator 'max' and (3.1) and (3.2), the vector $\mathbf{r}^d$ has the matrix-vector form

$$\mathbf{r}^d = \max\{0, \mathbf{b} + G\mathbf{r}\}, \tag{3.3}$$

where $G = (a_{ij}) \in \mathbb{R}^{q\times n}$ with $a_{ij} = 0$ for $j \neq k$ and $a_{ik} > 0$ or $a_{ik} < 0$ for call or put option i on stock k, $\mathbf{b} = \widehat{\mathbf{b}} + G\mathbf{e}_n$ and $\widehat{\mathbf{b}} = (\widehat{b}_1, \ldots, \widehat{b}_q)^T$.

Let the random return vector for both stocks and options be

$$\widetilde{\mathbf{r}} = \begin{pmatrix} \mathbf{r} \\ \mathbf{r}^d \end{pmatrix} = \begin{pmatrix} \mathbf{r} \\ \max\{0, \mathbf{b} + G\mathbf{r}\} \end{pmatrix} \in \mathbb{R}^{n+q}.$$

A portfolio of an investor holding both stocks and options now can be expressed as a joint vector $\widetilde{\mathbf{w}} = \begin{pmatrix} \mathbf{w} \\ \mathbf{w}^d \end{pmatrix} \in \mathbb{R}^{n+q}$ and satisfies

$$\sum_{i=1}^{n+q} \widetilde{w}_i = \sum_{i=1}^{n} w_i + \sum_{i=1}^{q} w_i^d = 1, \tag{3.4}$$

where $\mathbf{w} = (w_1, \ldots, w_n)^T \in \mathbb{R}^n$ is the weight vector invested to n stocks and $\mathbf{w}^d = (w_1^d, \ldots, w_q^d)^T \in \mathbb{R}^q$ the weight vector invested to q options. Then the robust portfolio selection model with options has the following form

$$\begin{aligned} (\mathrm{R2}): \quad \max_{\widetilde{\mathbf{w}}} \quad & \min_{(\boldsymbol{\mu},V)\in\mathcal{S}} \mathbb{E}[\widetilde{\mathbf{w}}^T\widetilde{\mathbf{r}}] \\ \text{s.t.} \quad & \max_{(\boldsymbol{\mu},V)\in\mathcal{S}} \mathbb{P}\{\widetilde{\mathbf{w}}^T\widetilde{\mathbf{r}} \leq \rho\} \leq \alpha, \\ & \widetilde{\mathbf{w}} \in \widetilde{\mathcal{W}}, \end{aligned}$$

where the set $\widetilde{\mathcal{W}} = \{\widetilde{\mathbf{w}} : \widetilde{\mathbf{w}} \geq 0 \text{ satisfies (3.4)}\}$. It is clear that when $G = 0$ and $\mathbf{b} = \mathbf{0}$, problem (R2) reduces to problem (R1). Thus, the robust portfolio selection problem (R2) is a generalization of problem (R1). An advantage of model (R2) is that options introduced into the model can hedge the risk from either a large shock of asset returns or the market crash even though the uncertainty set is not large enough.

It is worthwhile mentioning that the problem (R2) has at least three points that are different from [21]: (1) Zymler et al. [21] assume that the random return vector has the form $\mathbf{r} = \boldsymbol{\mu} + \boldsymbol{\epsilon}$ which is clearly simpler than factor model (2.1). (2) The uncertainty set considered by Zymler et al. [21] is separable. Moreover, the uncertainty set for the covariance matrix is a singleton, this means that the covariance matrix is estimated exactly. (3) Both models use the different method to convert the resulting problem to the solvable SOCP. Additionally, problem (R2) is also different from the problem considered by Zymler et al. [23], in which they extended El Ghaoui et al. [10] worst-case VaR portfolio selection to a robust portfolio selection with non-linear VaR and derivative insurance guarantee. But, in Zymler, Kuhn and Rustem's model, they assume that returns of assets have uncertain distribution, but do not consider the existence of a parameters uncertainty set, that is, parameters are assumed to be known exactly. This means their model is sharply dependent on the estimate of parameters. In our model, we assume that returns of assets have the certainty distribution, but their values belong to an uncertainty set and need not be known exactly. Hence, our model can be viewed to be complementary to Zymler, Kuhn and Rustem's model.

Clearly, it is extremely difficult to directly solve Problem (R2) since the vector $\widetilde{\mathbf{r}}$ is no longer normally distributed and we can not exactly express problem (R2) into a form of (R1′) using known information of the random return vector $\mathbf{r}$.

4. The joint uncertainty set

Before further discussion of problem (R2), we introduce the definition of joint uncertainty set in this section. Interested readers can refer to [8,17] for more detail.

Suppose the market data consists of asset returns, $\{\mathbf{r}^t : t = 1, \dots, p\}$ and factor returns $\{\mathbf{f}^t : t = 1, \dots, p\}$ for $p (p > m + 1)$ periods. Then the linear model (2.1) implies

$$r_i^t = \mu_i + \sum_{j=1}^{m} V_{ji} f_j^t + \epsilon_i^t, \quad i = 1, \dots, n, t = 1, \dots, p.$$

Similar to typical linear regression analysis, we assume that $\{\epsilon_i^t, i = 1, \dots, n\}$ are all normally distributed with $\epsilon_i^t \sim N(0, \sigma_i^2)$ and series independent.

Let $B = (\mathbf{f}^1, \mathbf{f}^2, \dots, \mathbf{f}^p) \in \mathbb{R}^{m \times p}$ be the matrix of factor returns, $\mathbf{y}_i = (r_i^1, r_i^2, \dots, r_i^p)^T$ be the return series of asset i, $A = (\mathbf{e}_p, B^T)$, and $\mathbf{x}_i = (\mu_i, V_{1i}, V_{2i}, \dots, V_{mi})^T$ where $\mathbf{e}_p \in \mathbb{R}^p$ is the all-one vector. Then we have

$$\mathbf{y}_i = A\mathbf{x}_i + \boldsymbol{\epsilon}_i, \quad i = 1, 2, \dots, n,$$

where $\boldsymbol{\epsilon}_i = (\epsilon_i^1, \epsilon_i^2, \dots, \epsilon_i^p)^T$. For any given $\omega \in (0, 1)$, the joint ellipsoidal uncertainty set of parameters $(\boldsymbol{\mu}, V)$ with ω-confidence level is defined as follows [17].

$$\mathcal{S} := S_{\boldsymbol{\mu},V}(\omega) = \left\{ (\boldsymbol{\mu}, V) : \sum_{i=1}^{n} \frac{(\mathbf{x}_i - \overline{\mathbf{x}}_i)^T (A^T A)(\mathbf{x}_i - \overline{\mathbf{x}}_i)}{s_i^2} \le (m+1)c(\omega) \right\} \tag{4.1}$$

for certain $c(\omega)$, where

$$\overline{\mathbf{x}}_i = (A^T A)^{-1} A^T \mathbf{y}_i \tag{4.2}$$

is the least squares estimate of the parameter vector $\mathbf{x}_i$, and

$$s_i^2 = \frac{\| \mathbf{y}_i - A\overline{\mathbf{x}}_i \|^2}{p - m - 1} \tag{4.3}$$

is the unbiased estimate of σ_i^2. It will be shown in Section 6 that the uncertainty set $\mathcal{S}$ can be expressed as a quadratic function constraint with respect to $\mathbf{x}_i (i = 1, \dots, n)$ (see (6.10)) and hence $\mathcal{S}$ is convex. It is shown (see [17]) that $\mathcal{S}$ is an ω-confidence uncertainty set of parameters $(\boldsymbol{\mu}, V)$ for certain $c(\omega)$, if and only if $\mathbb{P}\left\{\sum_{i=1}^{n} \mathcal{Y}^i \le c(\omega)\right\} = \omega$, i.e. $c(\omega)$ is the ω-critical value of $\sum_{i=1}^{n} \mathcal{Y}^i$, where

$$\mathcal{Y}^i = \frac{(\mathbf{x}_i - \overline{\mathbf{x}}_i)^T (A^T A)(\mathbf{x}_i - \overline{\mathbf{x}}_i)}{(m+1)s_i^2}, \quad i = 1, 2, \dots, n$$

are independent and distributed according to the F-distribution with $m+1$ degrees of freedom in the numerator and $p-m-1$ degrees of freedom in the denominator.

5. An approximation to problem (R2)

As mentioned in Section 3, it is extremely difficult to directly solve problem (R2). In this section we will give an approximation to problem (R2) such that explicit expressions for the problem can be obtained. For this purpose, following lemmas are required and its proof is straightforward.

Lemma 5.1. *Let $\mathbf{a} \in \mathbb{R}^n$ be a nonnegative constant vector. Then, for any $\mathbf{x} \in \mathbb{R}^n$, we have*

$$\mathbf{a}^T \max(\mathbf{0}, \mathbf{x}) = \max\{\mathbf{x}^T \mathbf{y} : \mathbf{0} \le \mathbf{y} \le \mathbf{a}\} \tag{5.1}$$

and $\mathbf{a}^T \max(\mathbf{0}, \mathbf{x})$ is a convex function with respect to $\mathbf{x}$.

Proof. The convexity of the function $\mathbf{a}^T\max(\mathbf{0},\mathbf{x})$ is a well known result of convex analysis and thus its proof is omitted. Next we only prove (5.1). For any $\mathbf{x}\in\mathbb{R}^n$, notice that $\max(\mathbf{0},\mathbf{x})\geq\mathbf{0}$ and $\max(\mathbf{0},\mathbf{x})\geq\mathbf{x}$. Hence, the left hand side of (5.1) can be expressed as a linear program, that is,

$$\mathbf{a}^T\max(\mathbf{0},\mathbf{x})=\min_{\mathbf{h}}\{\mathbf{a}^T\mathbf{h}:\mathbf{h}\geq\mathbf{0},\mathbf{h}\geq\mathbf{x}\}.$$

The dual problem of this linear programs is $\max_{\mathbf{y}}\{\mathbf{x}^T\mathbf{y}:\mathbf{0}\leq\mathbf{y}\leq\mathbf{a}\}$. This obtains the equality (5.1). □

Although the random vector $\mathbf{r}$ follows a multivariate normal distribution, we can not obtain the distribution of the random vector $\mathbf{r}^d$ since (3.3) is in fact a nonlinear function in $\mathbf{r}$. However, a lower bound for the objective function of problem (R2) can be obtained based on Lemma 5.1. In fact it follows from Lemma 5.1 and Jensen's inequality [24] that for any $(\boldsymbol{\mu},V)\in\mathcal{S}$, we have

$$\begin{aligned}\mathbb{E}[\widetilde{\mathbf{w}}^T\widetilde{\mathbf{r}}]&=\mathbb{E}[\mathbf{w}^T\mathbf{r}]+\mathbb{E}[(\mathbf{w}^d)^T\mathbf{r}^d]\\&=\mathbb{E}[\mathbf{w}^T\mathbf{r}]+\mathbb{E}[(\mathbf{w}^d)^T\max(0,(\mathbf{b}+G\mathbf{r}))]\\&=\mathbb{E}[\mathbf{w}^T\mathbf{r}]+\mathbb{E}[\max\{(\mathbf{b}+G\mathbf{r})^T\mathbf{y}:\mathbf{0}\leq\mathbf{y}\leq\mathbf{w}^d\}]\\&\geq\mathbb{E}[\mathbf{w}^T\mathbf{r}]+\max_{\mathbf{0}\leq\mathbf{y}\leq\mathbf{w}^d}\mathbb{E}[(\mathbf{b}+G\mathbf{r})^T\mathbf{y}]\\&=\mathbf{w}^T\boldsymbol{\mu}+\max_{\mathbf{0}\leq\mathbf{y}\leq\mathbf{w}^d}\{(\mathbf{b}+G\boldsymbol{\mu})^T\mathbf{y}\}\\&=\max_{\mathbf{0}\leq\mathbf{y}\leq\mathbf{w}^d}\{(\mathbf{w}+G^T\mathbf{y})^T\boldsymbol{\mu}+\mathbf{b}^T\mathbf{y}\},\end{aligned}\tag{5.2}$$

where the last equality is based on the fact that $\mathbf{w}^T\boldsymbol{\mu}$ is independent of $\mathbf{y}$.

Let $\mathcal{S}_\mu$ be the projection of $\mathcal{S}$ on the subspace in which $\boldsymbol{\mu}$ exists. Then S_μ is a convex set since $\mathcal{S}$ is convex. Then from (5.2) and the fact that $(\mathbf{w}+G^T\mathbf{y})^T\boldsymbol{\mu}+\mathbf{b}^T\mathbf{y}$ is independent of V, we have

$$\begin{aligned}\min_{(\boldsymbol{\mu},V)\in\mathcal{S}}\mathbb{E}[\widetilde{\mathbf{w}}^T\widetilde{\mathbf{r}}]&\geq\min_{(\boldsymbol{\mu},V)\in\mathcal{S}}\max_{\mathbf{0}\leq\mathbf{y}\leq\mathbf{w}^d}\{(\mathbf{w}+G^T\mathbf{y})^T\boldsymbol{\mu}+\mathbf{b}^T\mathbf{y}\}\\&=\min_{\boldsymbol{\mu}\in\mathcal{S}_\mu}\max_{\mathbf{0}\leq\mathbf{y}\leq\mathbf{w}^d}\{(\mathbf{w}+G^T\mathbf{y})^T\boldsymbol{\mu}+\mathbf{b}^T\mathbf{y}\}\\&=\max_{\mathbf{0}\leq\mathbf{y}\leq\mathbf{w}^d}\min_{\boldsymbol{\mu}\in\mathcal{S}_\mu}\{(\mathbf{w}+G^T\mathbf{y})^T\boldsymbol{\mu}+\mathbf{b}^T\mathbf{y}\}\\&=\max_{\mathbf{0}\leq\mathbf{y}\leq\mathbf{w}^d}\min_{(\boldsymbol{\mu},V)\in\mathcal{S}}\{(\mathbf{w}+G^T\mathbf{y})^T\boldsymbol{\mu}+\mathbf{b}^T\mathbf{y}\},\end{aligned}\tag{5.3}$$

where the second equality from the bottom is based on the max–min theorem of convex function (e.g. see [25] for detail).

For the probability constraint in problem (R2), from the definition of vector $\widetilde{\mathbf{r}}$ we have

$$\begin{aligned}\mathbb{P}\{\widetilde{\mathbf{w}}^T\widetilde{\mathbf{r}}\leq\rho\}&=\mathbb{P}\{\mathbf{w}^T\mathbf{r}+(\mathbf{w}^d)^T\mathbf{r}^d\leq\rho\}\\&=\mathbb{P}\{\mathbf{w}^T\mathbf{r}+(\mathbf{w}^d)^T\max(0,\mathbf{b}+G\mathbf{r})\leq\rho\}\\&=\mathbb{P}\left\{\mathbf{w}^T\mathbf{r}+\max_{\mathbf{0}\leq\mathbf{y}\leq\mathbf{w}^d}(\mathbf{b}+G\mathbf{r})^T\mathbf{y}\leq\rho\right\}\\&=\mathbb{P}\left\{\max_{\mathbf{0}\leq\mathbf{y}\leq\mathbf{w}^d}\{(\mathbf{w}+G\mathbf{y})^T\mathbf{r}+\mathbf{b}^T\mathbf{y}\}\leq\rho\right\}.\end{aligned}\tag{5.4}$$

Lemma 5.2. *For the left hand side of the probability constraint in problem (R2), we have*

$$\max_{(\boldsymbol{\mu},V)\in\mathcal{S}}\mathbb{P}\{\widetilde{\mathbf{w}}^T\widetilde{\mathbf{r}}\leq\rho\}\leq\min_{\mathbf{0}\leq\mathbf{y}\leq\mathbf{w}^d}\max_{(\boldsymbol{\mu},V)\in\mathcal{S}}\mathbb{P}\left\{(\mathbf{w}+G^T\mathbf{y})^T\mathbf{r}+\mathbf{b}^T\mathbf{y}\leq\rho\right\}.$$

Proof. For any given $\mathbf{w}$, $\mathbf{w}^d$, G, $\mathbf{b}$ and ρ, let

$$\varphi(\mathbf{y})=\mathbb{P}\left\{(\mathbf{w}+G\mathbf{y})^T\mathbf{r}+\mathbf{b}^T\mathbf{y}\leq\rho\right\}.$$

Then for any $\mathbf{y}$ with $\mathbf{0}\leq\mathbf{y}\leq\mathbf{w}^d$, we have

$$\mathbb{P}\left\{\max_{\mathbf{0}\leq\mathbf{y}\leq\mathbf{w}^d}\{(\mathbf{w}+G\mathbf{y})^T\mathbf{r}+\mathbf{b}^T\mathbf{y}\}\leq\rho\right\}\leq\varphi(\mathbf{y}).\tag{5.5}$$

The continuity of probability implies that there exists at least a $\mathbf{y}^* \in [\mathbf{0}, \mathbf{w}^d]$ such that $\varphi(\mathbf{y}^*)$ attains the minimum value of $\varphi(\mathbf{y})$ over $[\mathbf{0}, \mathbf{w}^d]$. Then we have

$$\mathbb{P}\left\{\max_{\mathbf{0}\le\mathbf{y}\le\mathbf{w}^d}\{(\mathbf{w}+G\mathbf{y})^T\mathbf{r}+\mathbf{b}^T\mathbf{y}\}\le\rho\right\}\le\varphi(\mathbf{y}^*)=\min_{\mathbf{0}\le\mathbf{y}\le\mathbf{w}^d}\varphi(\mathbf{y})$$
$$=\min_{\mathbf{0}\le\mathbf{y}\le\mathbf{w}^d}\mathbb{P}\left\{(\mathbf{w}+G\mathbf{y})^T\mathbf{r}+\mathbf{b}^T\mathbf{y}\le\rho\right\},$$

and hence, from (5.4)

$$\begin{aligned}\max_{(\boldsymbol{\mu},V)\in\mathcal{S}}\mathbb{P}\{\widetilde{\mathbf{w}}^T\widetilde{\mathbf{r}}\le\rho\} &= \max_{(\boldsymbol{\mu},V)\in\mathcal{S}}\mathbb{P}\left\{\max_{\mathbf{0}\le\mathbf{y}\le\mathbf{w}^d}\{(\mathbf{w}+G^T\mathbf{y})^T\mathbf{r}+\mathbf{b}^T\mathbf{y}\}\le\rho\right\}\\ &\le \max_{(\boldsymbol{\mu},V)\in\mathcal{S}}\min_{\mathbf{0}\le\mathbf{y}\le\mathbf{w}^d}\mathbb{P}\left\{(\mathbf{w}+G^T\mathbf{y})^T\mathbf{r}+\mathbf{b}^T\mathbf{y}\le\rho\right\}\\ &\le \min_{\mathbf{0}\le\mathbf{y}\le\mathbf{w}^d}\max_{(\boldsymbol{\mu},V)\in\mathcal{S}}\mathbb{P}\left\{(\mathbf{w}+G^T\mathbf{y})^T\mathbf{r}+\mathbf{b}^T\mathbf{y}\le\rho\right\}\end{aligned}$$

where the last inequality comes from the max–min theorem of non-convex function (e.g. see [25] for detail). This completes the proof. □

Inequalities (5.2) and (5.3) give a tight lower bound to the objective function of problem (R2), while Lemma 5.2 gives an approximation to the chance constraint in problem (R2) when the right hand side of the inequality in Lemma 5.2 is not greater than the given probability level α. Based on these facts, an approximation to the hard problem (R2) is given by

$$\begin{aligned}(\mathrm{R3}):\quad &\max_{\mathbf{w},\mathbf{w}^d,\mathbf{y}} && \min_{(\boldsymbol{\mu},V)\in\mathcal{S}}\{(\mathbf{w}+G^T\mathbf{y})^T\boldsymbol{\mu}+\mathbf{b}^T\mathbf{y}\}\\ &\text{s.t.} && \max_{(\boldsymbol{\mu},V)\in\mathcal{S}}\mathbb{P}\left\{(\mathbf{w}+G^T\mathbf{y})^T\mathbf{r}+\mathbf{b}^T\mathbf{y}\le\rho\right\}\le\alpha\\ &&& (\mathbf{w},\mathbf{w}^d)\in\widetilde{\mathcal{W}},\ \mathbf{0}\le\mathbf{y}\le\mathbf{w}^d,\end{aligned}$$

here we use $(\mathbf{w}, \mathbf{w}^d) \in \widetilde{\mathcal{W}}$ to express $\widetilde{\mathbf{w}} \in \widetilde{\mathcal{W}}$. It is clear that the optimal solution of problem (R3) is feasible for problem (R2) and will be accepted as an approximate solution of problem (R2). If $G = 0$ and $\mathbf{b} = \mathbf{0}$, then problem (R3) is reduced to problem (R1), the robust portfolio selection problem without options. Using a similar method to (2.4), the chance constraint in problem (R3) can be written as

$$\max_{(\boldsymbol{\mu},V)\in\mathcal{S}}\left\{\mathcal{F}_{\xi}^{-1}(1-\alpha)\sqrt{\boldsymbol{\phi}^T(V^TFV+D)\boldsymbol{\phi}}-\boldsymbol{\phi}^T\boldsymbol{\mu}-\mathbf{b}^T\mathbf{y}\right\}\le-\rho, \tag{5.6}$$

where

$$\boldsymbol{\phi}=\mathbf{w}+G^T\mathbf{y}=\begin{bmatrix}I_n & G^T\end{bmatrix}\begin{pmatrix}\mathbf{w}\\ \mathbf{y}\end{pmatrix}, \tag{5.7}$$

I_n is the n dimensional unit matrix. Let $\lambda_0 = \mathcal{F}_{\xi}^{-1}(1-\alpha)$ and

$$P=\begin{bmatrix}I_n\\ G\end{bmatrix}(V^TFV+D)\begin{bmatrix}I_n & G^T\end{bmatrix}.$$

Then, inequality (5.6) can be rewritten as

$$\max_{(\boldsymbol{\mu},V)\in\mathcal{S}}\left\{\lambda_0\sqrt{(\mathbf{w}^T,\mathbf{y}^T)P\begin{pmatrix}\mathbf{w}\\ \mathbf{y}\end{pmatrix}}-(\mathbf{w}^T,\mathbf{y}^T)\begin{pmatrix}\boldsymbol{\mu}\\ G\boldsymbol{\mu}+\mathbf{b}\end{pmatrix}\right\}\le-\rho. \tag{5.8}$$

Notice that the positive definiteness of the matrix $V^TFV + D$ implies that the $(n + q) \times (n + q)$ matrix P is positive semi-definite. Because the objective function of problem (R3) is independent of the variables V, the worst-case expected return within the joint ellipsoidal uncertainty set $\mathcal{S}$ is in fact equivalent to the worst-case expected return within the uncertainty set $\mathcal{S}_{\mu}$, that is, problem (R3) can be rewritten as

$$\begin{aligned}(\mathrm{R4}):\quad &\max_{\mathbf{w},\mathbf{w}^d,\mathbf{y}} && \min_{\boldsymbol{\mu}\in\mathcal{S}_{\mu}}\left\{(\mathbf{w}+G^T\mathbf{y})^T\boldsymbol{\mu}+\mathbf{b}^T\mathbf{y}\right\}\\ &\text{s.t.} && (5.8),\quad(\mathbf{w},\mathbf{w}^d,\mathbf{y})\in\mathcal{W}_1,\end{aligned}$$

where

$$\mathcal{W}_1=\{(\mathbf{w},\mathbf{w}^d,\mathbf{y}):(\mathbf{w},\mathbf{w}^d)\in\widetilde{\mathcal{W}},\mathbf{0}\le\mathbf{y}\le\mathbf{w}^d\}$$

is the new feasible region with respect to $(\mathbf{w}, \mathbf{w}^d, \mathbf{y})$. The uncertainty set $\mathcal{S}_\mu$ is called a marginal uncertainty set relative to the set $\mathcal{S}$. The worst-case objective function of problem (R4) is defined in the marginal uncertainty set $\mathcal{S}_\mu$, while the worst-case chance constraint (5.8) is defined in the joint ellipsoidal uncertainty set $\mathcal{S}$. This is the main characteristic of problem (R4), and the same way can be used to problem (R1′) without options. We call (R4) and (R1′) the robust portfolio selection problems under "marginal + joint" ellipsoidal uncertainty set $(\mathcal{S}_\mu, \mathcal{S})$.

We close this section by giving some remarks on problem (R4). Let $\boldsymbol{\mu}_1$ be the worst-case estimate of $\boldsymbol{\mu}$ in the objective function of (R4) and $\boldsymbol{\mu}_2$ be the worst-case estimate of $\boldsymbol{\mu}$ in constraint (5.8). Generally speaking, $\boldsymbol{\mu}_1 \neq \boldsymbol{\mu}_2$, but this does not impact the solution of the proposed problem, moreover, in a sense it will improve the performance of solution. In fact, $\boldsymbol{\mu}_1$ is the *real* worst-case estimate of $\boldsymbol{\mu}$ in uncertainty set $\mathcal{S}$, this leads to

$$(\mathbf{w} + G^T\mathbf{y})^T\boldsymbol{\mu}_1 + \mathbf{b}^T\mathbf{y} \leq (\mathbf{w} + G^T\mathbf{y})^T\boldsymbol{\mu}_2 + \mathbf{b}^T\mathbf{y}$$

since $\boldsymbol{\mu}_2 \in \mathcal{S}_\mu$. Thus, we in fact maximize the *real* worst-case expectation return for the given uncertainty set $\mathcal{S}$ which is the same as [8]. But it can not obtain the *real* worst-case expectation return to the solo use of the joint uncertainty set as in Lu (2011), [17,18] since it only can obtain the estimate $\boldsymbol{\mu}_2$. However, the confidence level of $(\boldsymbol{\mu}, V) \in \mathcal{S}$ in problem (R4) is consistent with Lu (2011), [17,18] (see page 8 in [17]). Hence, problem (R4) in fact shares the advantages of models in, [8], Lu (2011) [17,18].

6. A SOCP relaxation with LMI constraints

In this section, we will formulate problems (R1′) and (R4) into convex programs with linear matrix inequality constraints. Because problem (R1′) is a special case of problem (R4) when $G = 0$ and $\mathbf{b} = \mathbf{0}$, in what follows, we will focus on problem (R4) and all the results in this section can be applied to problem (R1′).

To this end, we need an explicit form of the set $\mathcal{S}_\mu$. It can be observed from (4.1) that the joint ellipsoidal uncertainty set $\mathcal{S}$ is symmetric with respect to $\overline{\mathbf{x}}_i = (\overline{\mu}_i, \overline{V}_{1i}, \ldots, \overline{V}_{mi})^T$, $i = 1, \ldots, n$. Thus the projection $\mathcal{S}_\mu$ of $\mathcal{S}$ on the subspace in which $\boldsymbol{\mu}$ exists can be obtained by setting $\mathbf{x}_i = (\mu_i, \overline{V}_{1i}, \ldots, \overline{V}_{mi})^T$, $i = 1, \ldots, n$, that is,

$$\mathcal{S}_\mu = \left\{\boldsymbol{\mu} : \mu_i = \overline{\mu}_i + \zeta_i, \sum_{i=1}^{n}\left(\frac{\zeta_i}{s_i}\right)^2 \leq \kappa\right\}, \tag{6.1}$$

where $\kappa = (m+1)c(\omega)$ and s_i is given by (4.3). Note that $\mathcal{S}_\mu$ is still an ellipsoidal uncertainty set on the subspace of the vector $\boldsymbol{\mu}$.

Now we consider the worst-case objective function of problem (R4) for any given $(\mathbf{w}, \mathbf{w}^d, \mathbf{y}) \in \mathcal{W}_1$, i.e.

$$\begin{aligned} \min_{\boldsymbol{\mu}} \quad & (\mathbf{w} + G^T\mathbf{y})^T\boldsymbol{\mu} + \mathbf{b}^T\mathbf{y} \\ s.t. \quad & \boldsymbol{\mu} \in \mathcal{S}_\mu. \end{aligned} \tag{6.2}$$

Lemma 6.1. *Problem (6.2) has an unique optimal solution $\boldsymbol{\mu}^*$ with*

$$\mu_i^* = \overline{\mu}_i - \sqrt{\frac{\kappa}{\widetilde{\mathbf{s}}^T\widetilde{\mathbf{s}}}}\,\widetilde{s}_i s_i, \quad i = 1, 2, \ldots, n, \tag{6.3}$$

and its optimal value is

$$\boldsymbol{\phi}^T\overline{\boldsymbol{\mu}} - \sqrt{\kappa}\|S\boldsymbol{\phi}\| + \mathbf{b}^T\mathbf{y},$$

where

$$S = \begin{pmatrix} s_1 & & \\ & \ddots & \\ & & s_n \end{pmatrix}$$

is a positive definite diagonal matrix, $\widetilde{\mathbf{s}} = (\widetilde{s}_1, \ldots, \widetilde{s}_n)^T$ with $\widetilde{s}_i = \phi_i s_i$, $i = 1, \ldots, n$ and $\boldsymbol{\phi} = (\phi_1, \ldots, \phi_n)^T$ given by (5.7).

Proof. Denote

$$\nu_i = \frac{\mu_i}{s_i}, \qquad \overline{\nu_i} = \frac{\overline{\mu}_i}{s_i}, \quad i = 1, \ldots, n. \tag{6.4}$$

Then the objective function of (6.2) can be written as

$$\sum_{i=1}^{n}\phi_i s_i \nu_i + \mathbf{b}^T\mathbf{y} = \sum_{i=1}^{n}\widetilde{s}_i\nu_i + \mathbf{b}^T\mathbf{y} = \widetilde{\mathbf{s}}^T\boldsymbol{\nu} + \mathbf{b}^T\mathbf{y}.$$

Define

$$S_\nu = \{\boldsymbol{\nu} = (\nu_1, \ldots, \nu_n)^T : \|\boldsymbol{\nu} - \overline{\boldsymbol{\nu}}\|^2 \le \kappa\}.$$

Then problem (6.2) is equivalent to

$$\begin{array}{ll} \min\limits_{\boldsymbol{\nu}} & \widetilde{\mathbf{s}}^T \boldsymbol{\nu} + \mathbf{b}^T \mathbf{y} \\ s.t. & \boldsymbol{\nu} \in \mathcal{S}_\nu. \end{array} \tag{6.5}$$

Straightforward calculation gives the KKT point of (6.5)

$$\boldsymbol{\nu}^* = \overline{\boldsymbol{\nu}} - \sqrt{\frac{\kappa}{\widetilde{\mathbf{s}}^T \widetilde{\mathbf{s}}}}\, \widetilde{\mathbf{s}}.$$

Combining (6.4) gives the optimal solution $\boldsymbol{\mu}^*$ with μ_i^* given by (6.3). Substituting $\boldsymbol{\mu}^*$ into (6.2) generates the second conclusion of this lemma. The proof is completed. □

In what follows, we formulate inequality constraint (5.8) in problem (R4) into some equivalent linear matrix inequalities (LMIs). To this end, the following lemmas are required.

Lemma 6.2 (*Schur Complement, See [26]*). *Let $A \in \mathbb{S}_n^{++}$, $C \in \mathbb{S}_n$, and $B \in \mathbb{M}_{m,n}$, the $m \times n$ matrix space. Then*

$$\begin{bmatrix} A & B \\ B^T & C \end{bmatrix} \succeq 0 \quad \text{if and only if} \quad C - B^T A^{-1} B \succeq 0.$$

Lemma 6.3 (*$\mathcal{S}$-Procedure, See [26]*). *Let $F_i(\mathbf{x}) = \mathbf{x}^T A_i \mathbf{x} + 2\mathbf{b}_i^T \mathbf{x} + \widetilde{c}_i$, $i = 0, 1, 2, \ldots, \widetilde{p}$ be quadratic functions of $\mathbf{x} \in \mathbb{R}^n$. Then $F_0(\mathbf{x}) \le 0$ for all $\mathbf{x} \in \{\mathbf{x} | F_i(\mathbf{x}) \le 0, i = 1, 2, \ldots, \widetilde{p}\}$, if there exists $\theta_i \ge 0$, $i = 1, 2, \ldots, \widetilde{p}$ such that*

$$\sum_{i=1}^{\widetilde{p}} \theta_i \begin{pmatrix} \widetilde{c}_i & \mathbf{b}_i \\ \mathbf{b}_i^T & A_i \end{pmatrix} - \begin{pmatrix} \widetilde{c}_0 & \mathbf{b}_0 \\ \mathbf{b}_0^T & A_0 \end{pmatrix} \succeq 0.$$

Moreover, if $\widetilde{p} = 1$, then the converse holds if there exists $\mathbf{x}_0 \in \mathbb{R}^n$ such that $F_1(\mathbf{x}_0) \le 0$.

The following result is a property of the standard Kronecker product (denoted by $\otimes$) of two matrices.

Lemma 6.4 (*See [27]*). *If $H_1 \succeq 0$ and $H_2 \succeq 0$, then $H_1 \otimes H_2 \succeq 0$.*

For any given $(\mathbf{w}, \mathbf{w}^d, \mathbf{y}) \in \mathcal{W}_1$, and ρ, we define a function of $\boldsymbol{\mu}$ and V

$$\begin{aligned} \mathcal{Q}(\boldsymbol{\mu}, V) &= \lambda_0 \sqrt{(\mathbf{w}^T, \mathbf{y}^T) P \begin{pmatrix} \mathbf{w} \\ \mathbf{y} \end{pmatrix}} - (\mathbf{w}^T, \mathbf{y}^T) \begin{pmatrix} \boldsymbol{\mu} \\ G\boldsymbol{\mu} + \mathbf{b} \end{pmatrix} + \rho \\ &= \lambda_0 \sqrt{(\mathbf{w} + G\mathbf{y})^T (V^T F V + D)(\mathbf{w} + G\mathbf{y})} - (\mathbf{w} + G\mathbf{y})^T \boldsymbol{\mu} - \mathbf{b}^T \mathbf{y} + \rho. \end{aligned}$$

Then inequality constraint (5.8) means that $\mathcal{Q}(\boldsymbol{\mu}, V) \le 0$ holds for any $(\boldsymbol{\mu}, V) \in \mathcal{S}$. From (4.1), the uncertainty set $\mathcal{S}$ can be expressed as

$$\mathcal{S} = \left\{ (\boldsymbol{\mu}, V) : \sum_{i=1}^n \mathbf{x}_i^T \left(\frac{A^T A}{(m+1)s_i^2} \right) \mathbf{x}_i - 2 \sum_{i=1}^n \left(\frac{A^T A \overline{\mathbf{x}}_i}{(m+1)s_i^2} \right)^T \mathbf{x}_i + \sum_{i=1}^n \overline{\mathbf{x}}_i^T \left(\frac{A^T A}{(m+1)s_i^2} \right) \overline{\mathbf{x}}_i - c(\omega) \le 0 \right\}. \tag{6.6}$$

Let

$$\mathbf{u} = \begin{pmatrix} -\dfrac{A^T A \overline{\mathbf{x}}_1}{(m+1)s_1^2} \\ \vdots \\ -\dfrac{A^T A \overline{\mathbf{x}}_n}{(m+1)s_n^2} \end{pmatrix} \in \mathbb{R}^{(m+1)n}, \qquad \eta = \sum_{i=1}^n \overline{\mathbf{x}}_i^T \left(\frac{A^T A}{(m+1)s_i^2} \right) \overline{\mathbf{x}}_i - c(\omega), \tag{6.7}$$

$$R = \begin{pmatrix} \dfrac{A^T A}{(m+1)s_1^2} & & \\ & \ddots & \\ & & \dfrac{A^T A}{(m+1)s_n^2} \end{pmatrix} \in \mathbb{R}^{[(m+1)n] \times [(m+1)n]}, \tag{6.8}$$

and

$$\mathbf{x} = \begin{pmatrix} \mathbf{x}_1 \\ \vdots \\ \mathbf{x}_n \end{pmatrix} \in \mathbb{R}^{(m+1)n}, \tag{6.9}$$

where $\mathbf{x}_i = (\mu_i, V_{1i}, V_{2i}, \ldots, V_{mi})^T \in \mathbb{R}^{m+1}$, $i = 1, 2, \ldots, n$. Then, the uncertainty set $\mathcal{S}$ can be rewritten using the following quadratic function constraint

$$\mathcal{S} = \left\{ \mathbf{x} : \mathbf{x}^T R \mathbf{x} + 2\mathbf{u}^T \mathbf{x} + \eta \le 0 \right\}. \tag{6.10}$$

The following lemma gives LMI expressions for the inequality constraint (5.8).

Lemma 6.5. *Let $\mathcal{S}$ be a ω-confidence uncertainty set given by (6.6) (hence by (4.1)) for any given $\omega \in (0, 1)$. Then there exists a positive semi-definite matrix $X \in \mathbb{R}^{n\times n}$ and a $\theta \ge 0$ such that inequality constraint (5.8) in problem* (R4) *is equivalent to the following linear matrix inequalities (LMIs)*

$$\begin{pmatrix} \theta\eta - 2(\rho - \mathbf{b}^T\mathbf{y}) & \theta\mathbf{u}^T + \mathbf{v}^T \\ \theta\mathbf{u} + \mathbf{v} & \theta R - \lambda_0 X \otimes \begin{pmatrix} 0 & 0 \\ 0 & F \end{pmatrix} \end{pmatrix} \succeq 0, \tag{6.11}$$

$$\begin{pmatrix} \tau_1 & \boldsymbol{\phi}^T \\ \boldsymbol{\phi} & X \end{pmatrix} \succeq 0, \qquad \theta \ge 0, \tag{6.12}$$

where

$$\mathbf{v} = -(\phi_1, \mathbf{0}_m^T, \ldots, \phi_n, \mathbf{0}_m^T)^T \in \mathbb{R}^{(m+1)n} \tag{6.13}$$

$\mathbf{0}_m$ is the m-dimension all-zero vector, and $\tau_1 = \sqrt{\boldsymbol{\phi}^T D \boldsymbol{\phi}}$.

Proof. Applying (5.7), we can write $\mathcal{Q}(\boldsymbol{\mu}, V)$ as

$$\begin{aligned} \mathcal{Q}(\boldsymbol{\mu}, V) &= \lambda_0\sqrt{\boldsymbol{\phi}^T(V^TFV + D)\boldsymbol{\phi}} - \boldsymbol{\phi}^T\boldsymbol{\mu} - \mathbf{b}^T\mathbf{y} + \rho \\ &= \lambda_0\sqrt{\boldsymbol{\phi}^T(V^TFV)\boldsymbol{\phi} + \boldsymbol{\phi}^T D\boldsymbol{\phi}} - \boldsymbol{\phi}^T\boldsymbol{\mu} - \mathbf{b}^T\mathbf{y} + \rho. \end{aligned}$$

Let $\boldsymbol{\phi} \ne 0$[1] and $\tau_1 = \sqrt{\boldsymbol{\phi}^T D\boldsymbol{\phi}}$. Then $\tau_1 > 0$ since D is positive definite. Thus, the second-order derivatives of $\mathcal{Q}(\boldsymbol{\mu}, V)$ exist at $(\boldsymbol{\mu}, V) = (\mathbf{0}_n, \mathbf{0}_{m\times n})$, and the first-order and second-order derivatives of $\mathcal{Q}(\boldsymbol{\mu}, V)$ at $(\boldsymbol{\mu}, V) = (\mathbf{0}_n, \mathbf{0}_{m\times n})$ are given by

$$\left.\frac{\partial \mathcal{Q}}{\partial \mathbf{x}_i}\right|_{(\boldsymbol{\mu},V)=(\mathbf{0}_n,\mathbf{0}_{m\times n})} = \begin{pmatrix} -\phi_i \\ \mathbf{0}_m \end{pmatrix}, \quad i = 1, \ldots, n,$$

and

$$\left.\frac{\partial^2 \mathcal{Q}}{\partial \mathbf{x}_i \partial \mathbf{x}_j}\right|_{(\boldsymbol{\mu},V)=(\mathbf{0}_n,\mathbf{0}_{m\times n})} = \begin{pmatrix} 0 & \mathbf{0}_m^T \\ \mathbf{0}_m & \lambda_0 \dfrac{\phi_i\phi_j}{\tau_1} F \end{pmatrix}, \quad i, j = 1, \ldots, n,$$

where $\mathbf{x}_i = (\mu_i, V_{1i}, V_{2i}, \ldots, V_{mi})^T \in \mathbb{R}^{m+1}$, $i = 1, 2, \ldots, n$. When $\sum_{i=1}^n \|\mathbf{x}_i\|^2$ is small enough, the function $\mathcal{Q}(\boldsymbol{\mu}, V)$ can be approximated by its second-order Taylor series expansion at $(\boldsymbol{\mu}, V) = (\mathbf{0}_n, 0_{m\times n})$

$$\begin{aligned} \mathcal{Q}(\boldsymbol{\mu}, V) &= \mathcal{Q}(\mathbf{0}_n, \mathbf{0}_{m\times n}) + \sum_{i=1}^n \left(\left.\frac{\partial \mathcal{Q}}{\partial \mathbf{x}_i}\right|_{(\boldsymbol{\mu},V)=(\mathbf{0}_n,\mathbf{0}_{m\times n})} \right) \mathbf{x}_i + \frac{1}{2}\sum_{i,j=1}^n \mathbf{x}_i^T \left(\left.\frac{\partial^2 \mathcal{Q}}{\partial \mathbf{x}_i \partial \mathbf{x}_j}\right|_{(\boldsymbol{\mu},V)=(\mathbf{0}_n,\mathbf{0}_{m\times n})} \right) \mathbf{x}_j + o\left(\sum_{i=1}^n \|\mathbf{x}_i\|^2 \right) \\ &\cong \frac{1}{2}\sum_{i,j=1}^n \mathbf{x}_i^T \begin{pmatrix} 0 & \mathbf{0}_m^T \\ \mathbf{0}_m & \lambda_0 \dfrac{\phi_i\phi_j}{\tau_1} F \end{pmatrix} \mathbf{x}_j + \sum_{i=1}^n \begin{pmatrix} -\phi_i \\ \mathbf{0}_m \end{pmatrix} \mathbf{x}_i - \mathbf{b}^T\mathbf{y} + \rho. \end{aligned} \tag{6.14}$$

Using notations of (6.9) and (6.13), (6.14) can be expressed as a standard quadratic function in matrix-vector form

$$\widetilde{\mathcal{Q}}(\mathbf{x}) = 2\mathcal{Q}(\boldsymbol{\mu}, V) = \mathbf{x}^T M \mathbf{x} + 2\mathbf{v}^T \mathbf{x} + 2(\rho - \mathbf{b}^T\mathbf{y}) \tag{6.15}$$

[1] If there exists matrix G, such that $\boldsymbol{\phi} = \mathbf{w} + G^T\mathbf{y} = 0$ for certain $(\mathbf{w}, \mathbf{y}) \in \mathcal{W}_1$, then function $\mathcal{Q}(\boldsymbol{\mu}, V) = -\mathbf{b}^T\mathbf{y} + \rho$ that is independent of $(\boldsymbol{\mu}, V)$. Thus, the derivatives of $\mathcal{Q}(\boldsymbol{\mu}, V)$ with respect to $(\boldsymbol{\mu}, V)$ at $(\boldsymbol{\mu}, V) = (\mathbf{0}_n, \mathbf{0}_{m\times n})$ is zeros. We do not consider this trivial case.

where

$$M = (M_{ij}) \in \mathbb{R}^{[(m+1)n]\times[(m+1)n]}, \tag{6.16}$$

with

$$M_{ij} = \begin{pmatrix} 0 & \mathbf{0}_m^T \\ \mathbf{0}_m & \lambda_0 \dfrac{\phi_i\phi_j}{\tau_1} F \end{pmatrix} \in \mathbb{R}^{(m+1)\times(m+1)}, \quad i, j = 1, \ldots, n.$$

In view of the definition of joint uncertainty set $\mathcal{S}$ and $\omega > 0$, it follows that $c(\omega) > 0$ (see also Proposition 3.1 and (13) in Lu (2011) [17] for detail). Clearly, $\bar{\mathbf{x}} \in \mathcal{S}$ since $\bar{\mathbf{x}}^T R\bar{\mathbf{x}} + 2\mathbf{u}^T\bar{\mathbf{x}} + \eta = -c(\omega) < 0$ holds, where

$$\bar{\mathbf{x}} = \begin{pmatrix} \bar{\mathbf{x}}_1 \\ \vdots \\ \bar{\mathbf{x}}_n \end{pmatrix} \in \mathbb{R}^{(m+1)n}.$$

Thus, the uncertainty set $\mathcal{S}$ is nonempty. Then, by (6.10) and (6.15) and Lemma 6.3, $\widetilde{\mathcal{Q}}(\mathbf{x}) \leq 0$ for all $\mathbf{x} \in \mathcal{S}$ if and only if there exists a $\theta \geq 0$ such that

$$\theta \begin{pmatrix} \eta & \mathbf{u}^T \\ \mathbf{u} & R \end{pmatrix} - \begin{pmatrix} 2(\rho - \mathbf{b}^T\mathbf{y}) & \mathbf{v}^T \\ \mathbf{v} & M \end{pmatrix} \succeq 0, \tag{6.17}$$

where $\mathbf{u}$, η and R are given by (6.7) and (6.8),

$$M = \lambda_0 \left(\frac{\boldsymbol{\phi}\boldsymbol{\phi}^T}{\tau_1} \right) \otimes \begin{pmatrix} 0 & \mathbf{0}_m^T \\ \mathbf{0}_m & F \end{pmatrix}$$

where $\otimes$ denotes the Kronecker product of matrices. Applying the fact that $F \succ 0$ and Lemma 6.4, (6.17) holds if and only if there exists X such that

$$\begin{pmatrix} \theta\eta - 2(\rho - \mathbf{b}^T\mathbf{y}) & \theta\mathbf{u}^T - \mathbf{v}^T \\ \theta\mathbf{u} - \mathbf{v} & \theta R - \lambda_0 X \otimes \widetilde{F} \end{pmatrix} \succeq 0, \quad X \succeq \frac{\boldsymbol{\phi}\boldsymbol{\phi}^T}{\tau_1}, \theta \geq 0, \tag{6.18}$$

where $\widetilde{F} = \begin{pmatrix} 0 & \mathbf{0}_m^T \\ \mathbf{0}_m & F \end{pmatrix}$. Hence, for any $\mathbf{x} \in \mathcal{S}$, $\widetilde{\mathcal{Q}}(\mathbf{x}) \leq 0$ (or $\mathcal{Q}(\boldsymbol{\mu}, V) \leq 0$) holds if and only if there exist $\theta \geq 0$ and a matrix X such that conditions (6.18) hold. In view of $\tau_1 > 0$ and Lemma 6.2, we have that $X \succeq \frac{\boldsymbol{\phi}\boldsymbol{\phi}^T}{\tau_1}$ holds if and only if the matrix X satisfies

$$\begin{pmatrix} \tau_1 & \boldsymbol{\phi}^T \\ \boldsymbol{\phi} & X \end{pmatrix} \succeq 0.$$

This obtain the equivalent conditions given by (6.11) and (6.12). The proof is finished. □

Based on Lemmas 6.1 and 6.5, problem (R4) can be expressed as a second-order cone programming (SOCP), i.e. we have the following result.

Theorem 6.6. *Let $\mathcal{S}$ given by (4.1) be a ω-confidence uncertainty set for any given $\omega \in (0, 1)$. Then robust portfolio selection problem* (R4) *can be relaxed as the following optimization problem with LMI constraints and second-order cone constraints*

$$\begin{aligned} \max_{\tau,\tau_1,\tau_2,\theta,\mathbf{w},\mathbf{w}^d,\mathbf{y},\boldsymbol{\phi},X} \quad & \tau \\ \text{s.t.} \quad & \sqrt{\kappa}\|S\boldsymbol{\phi}\| \leq \tau_2 \\ & \tau + \tau_2 - \boldsymbol{\phi}^T\overline{\boldsymbol{\mu}} - \mathbf{b}^T\mathbf{y} \leq 0 \\ & \begin{pmatrix} \theta\eta - 2(\rho - \mathbf{b}^T\mathbf{y}) & \theta\mathbf{u}^T + \mathbf{v}^T \\ \theta\mathbf{u} + \mathbf{v} & \theta R - \lambda_0 X \otimes \widetilde{F} \end{pmatrix} \succeq 0, \\ & \begin{pmatrix} \tau_1 & \boldsymbol{\phi}^T \\ \boldsymbol{\phi} & X \end{pmatrix} \succeq 0, \\ & \|D^{1/2}\boldsymbol{\phi}\| \leq \tau_1 \\ & \boldsymbol{\phi} - \mathbf{w} - G^T\mathbf{y} = 0 \\ & (\mathbf{w}, \mathbf{w}^d, \mathbf{y}) \in \mathcal{W}_1, \qquad \theta \geq 0, \end{aligned} \tag{6.19}$$

where scales $\tau, \tau_1, \tau_2, \theta$, vectors $(\mathbf{w}, \mathbf{w}^d, \mathbf{y})$, $\boldsymbol{\phi}$, and matrix X are unknown variables, and $\eta, \mathbf{u}, \mathbf{v}$, and matrices $R, \widetilde{F}$ are coefficients.

Proof. By introducing an auxiliary variable τ, problem (R4) can be rewritten as

$$\begin{array}{ll} \max\limits_{\mathbf{w},\mathbf{w}^d,\mathbf{y},\tau} & \tau \\ \text{s.t.} & \min\limits_{\boldsymbol{\mu}\in\mathcal{S}_\mu} \left\{(\mathbf{w}+G^T\mathbf{y})^T\boldsymbol{\mu}+\mathbf{b}^T\mathbf{y}\right\} \geq \tau \\ & (5.8), \quad (\mathbf{w},\mathbf{w}^d,\mathbf{y}) \in \mathcal{W}_1. \end{array} \tag{6.20}$$

Using the result of Lemma 6.1, the first inequality of problem can be replaced by

$$\boldsymbol{\phi}^T\overline{\boldsymbol{\mu}} - \sqrt{\kappa}\|S\boldsymbol{\phi}\| + \mathbf{b}^T\mathbf{y} \geq \tau. \tag{6.21}$$

Introducing one more auxiliary variable τ_2, inequality (6.21) can be replaced by the following two constraints

$$\sqrt{\kappa}\|S\boldsymbol{\phi}\| \leq \tau_2, \qquad \tau + \tau_2 - \boldsymbol{\phi}^T\overline{\boldsymbol{\mu}} - \mathbf{b}^T\mathbf{y} \leq 0$$

where the first one is a second order cone constraint while the second one is a linear inequality constraint. This gives the first two inequality constraints in problem (6.19). Let

$$\|D^{1/2}\boldsymbol{\phi}\|_2 = \tau_1.$$

Then the condition $\mathcal{Q}(\boldsymbol{\mu}, V) \leq 0$ can be written as

$$\mathcal{Q}(\boldsymbol{\mu}, V) = \lambda_0\sqrt{\boldsymbol{\phi}^T(V^TFV)\boldsymbol{\phi} + \tau_1^2} - \boldsymbol{\phi}^T\boldsymbol{\mu} - \mathbf{b}^T\mathbf{y} + \rho \leq 0.$$

The equality constraint $\|D^{1/2}\boldsymbol{\phi}\|_2 = \tau_1$ can be further relaxed as $\|D^{1/2}\boldsymbol{\phi}\|_2 \leq \tau_1$ so that it becomes a second order cone constraint too. Then, the third, fourth and fifth inequality constraints in problem (6.19) come from the result of Lemma 6.5. The proof is completed. □

Problem (6.19) is a tight approximation to problem (R4). When the results of Lemma 6.1, 6.5 and Theorem 6.6 are applied to problem (R1′), that is the case of the robust portfolio selection without options, a similar SOCP form of problem (R1′) can be obtained, that is, when $G = 0$, $\mathbf{b} = \mathbf{0}$, problem (6.19) reduces to the following SOCP

$$\begin{array}{ll} \max\limits_{\tau,\tau_1,\tau_2,\theta,\mathbf{w},\mathbf{w}^d,\mathbf{y},\boldsymbol{\phi},X} & \tau \\ \text{s.t.} & \sqrt{\kappa}\|S\mathbf{w}\| \leq \tau_2 \\ & \tau + \tau_2 - \mathbf{w}^T\overline{\boldsymbol{\mu}} \leq 0 \\ & \begin{pmatrix} \theta\eta - 2\rho & \theta\mathbf{u}^T + \mathbf{v}^T \\ \theta\mathbf{u} + \mathbf{v} & \theta R - \lambda_0 X \otimes \widetilde{F} \end{pmatrix} \succeq 0, \\ & \begin{pmatrix} \tau_1 & \mathbf{w}^T \\ \mathbf{w} & X \end{pmatrix} \succeq 0, \\ & \|D^{1/2}\mathbf{w}\| \leq \tau_1 \\ & \mathbf{w} \in \mathcal{W}, \qquad \theta \geq 0. \end{array} \tag{6.22}$$

This is a tight relaxation of problem (R1′). We call τ in both problems (6.19) and (6.22) the *worst-case return*. We will compare the performance of robust portfolio selection problems (6.19) and (6.22) on simulated and real market data in the next section. The duality of problem (6.19) can be obtained easily by introducing some dual variables, see Appendix for detail.

7. Test results

This section reports test results to compare performance of problem (R1) under the "marginal + joint" ellipsoidal uncertainty set $(\mathcal{S}_\mu, \mathcal{S})$ with that of the Goldfarb and Iyengar [8] model (denoted by RGI, see problem (47) in [8]) under the separable uncertainty set $\mathcal{S}_v \times \mathcal{S}_m$, and performance changes from problems (R1) and problem (R2) to problems (6.19) and (6.22). All computational tests are performed using SeDuMi V1.3 [28] on a 2.0 GHz Core 2 Duo machine.

Three classes of test problems are used in numerical experiments. The first class of test problems are generated based on simulated stocks data. The second class of test problems are obtained based on real stocks data from the Chinese stock market and the third class of test problems are generated based on simulated options data. The first two classes of test problems will be used to compare the performance of models (R1) and RGI, while the third class of problems will be used to test the performance of model (R2) by implementing models (6.19) and (6.22).

7.1. Test results on simulated stocks data

In this subsection, the performance of proposed model (R1) and Goldfarb and Iyengar [8] model (denoted by RGI) on simulated stocks data will be reported. Data on stock returns are randomly generated using the same way as described in [8] (see also [17]). $n = 40$ (the number of stocks) and $m = 5$ (the number of factors) are selected. The expected return vector $\boldsymbol{\mu} \in \mathbb{R}^n$ of n assets are independently generated according to a uniform distribution on interval [0.5%, 2.5%]. Both the symmetric positive semi-definite factor covariance matrix F and the factor loading matrix V are randomly generated using the MATLAB function randn(n, m) and the elements of both the matrices are located in the interval [0.1%, 5%]. The covariance matrix D of the residual returns $\boldsymbol{\epsilon}$ is assumed to be certain and sets to $D = 0.1\text{diag}(V^T FV)$, that is, the linear model explains 90% of the asset variance.

Based on the generated V, F and D, sequences of asset return vector $\mathbf{r}$ and factor return vector $\mathbf{f}$ are generated according to the normal distributions $\mathcal{N}(\boldsymbol{\mu}, V^T FV + D)$ and $\mathcal{N}(0, F)$ for an investment period of length $p = 90$, respectively. Then $\bar{\mathbf{x}}_i$ and $s_i^2, i = 1, \ldots, n$ are calculated from the resulting sequences $\mathbf{r}$ and $\mathbf{f}$ using (4.2) and (4.3). The joint uncertainty set $\mathcal{S}$ is determined by (4.1) for a given confidence level $\omega > 0$ (the computation of $c(\omega) > 0$ in (4.1) can refer to [17]) and the corresponding marginal uncertainty set $\mathcal{S}_\mu$ is determined by (6.1) for the same value of ω. Note that using a similar method to [8], the matrices D and F are not updated in generating sequences data of $\mathbf{r}$ and $\mathbf{f}$. A risk averse investor will choose a value of ρ larger than -0.05. Three values of ρ, i.e. $\rho = -0.05, -0.03$ or -0.01[2] will be used to test the performance of the models.

We first compare the mean Sharp ratios of the proposed model (R1) with model RGI [8] and Lu's [17,18] (denoted by RLU) model based on the simulated data, see Fig. 1. In Fig. 1 and the following figures, RLX1 denotes the results of problem (R1) with the optimal solution obtained by solving model (6.22). The mean Sharp ratio is given by[3]

$$MSR = \frac{(\overline{\boldsymbol{\mu}} - r_f \mathbf{e}_n)^T \mathbf{w}^*}{\sqrt{(\mathbf{w}^*)^T (\overline{V}^T F \overline{V} + D)\mathbf{w}^*}}$$

where $\overline{\boldsymbol{\mu}}, \overline{V}$ is estimated by (4.2), $r_f = 2.5\%$ is taken as the risk-free annual return rate and $\mathbf{w}^*$ is the optimal solutions obtained respectively by models (R1), RGI and RLU. We find from Fig. 1 that RLX1 has the best mean Sharp ratio for different confidence level and RGI has clearly the better mean Sharp ratio than RLU for most of values ω. Indeed, it is not surprising, from (2.4) and the inequality constraint of problem (R1′), that the mean Sharp ratio of problem (R1) is not less than a constant, that is

$$\frac{(\overline{\boldsymbol{\mu}} - r_f \mathbf{e}_n)^T \mathbf{w}^*}{\sqrt{(\mathbf{w}^*)^T (\overline{V}^T F \overline{V} + D)\mathbf{w}^*}} \geq \min_{(\boldsymbol{\mu}, V) \in \mathcal{S}} \frac{(\boldsymbol{\mu} - r_f \mathbf{e}_n)^T \mathbf{w}^*}{\sqrt{(\mathbf{w}^*)^T (V^T F V + D)\mathbf{w}^*}}$$
$$\geq \mathcal{F}_\xi^{-1}(1 - \alpha) + \frac{\rho - r_f \mathbf{e}_n^T \mathbf{w}^*}{\sqrt{(\mathbf{w}^*)^T (\widehat{V}^T F \widehat{V} + D)\mathbf{w}^*}}$$

which is also shared by model RGI but not shared by model RLU, where $\widehat{V}$ is the worst-case estimate of V but can not be obtained explicitly. Based on the fact, in the rest of this section, we only consider the performance comparisons of model (R1), (R4) and RGI since these models have a similar constraint, i.e. a chance constraint that is different from the mean-variance considered by Lu (2011) [17]. Another interesting result is found from Fig. 1. MSR increases as the confidence level ω when investors choose high risk expected returns ($\rho = -0.05$) and decreases as the confidence level ω for relatively conservative expected returns ($\rho = -0.01$). It is not hard to understand this phenomenon, because both ρ and ω can reflect the investors' attitude to risk. Too high risk or a conservative portfolio may not be the best Sharp ratio portfolio. The result in Fig. 1 suggests that in order to get a portfolio with good MSR, one may choose either a relatively small ρ and large ω or a relatively large ρ and small ω.

In the following Figs. 2–4, we give the comparisons of the worst-case return and diversification number of the resulting portfolio[4] of model (R1) with those of model RGI in [8]. The worst-case return in the three figures is τ that is the resulting maximum value of the objective function in problem (6.22).

The following observations can be obtained from the three figures. (1) For a fixed value of ρ, the worst-case returns of two models decrease and diversification numbers of resulting portfolios increase as the uncertainty set becomes larger, i.e. when the value of ω increases. It is clear that a larger value of ω indicates that the investor is more conservative and will select a portfolio to diversify market risks. (2) For a fixed value of ω, the worst-case return of two models also decrease and diversification numbers of resulting portfolios increase as the preset value of parameter ρ increases from -0.05 to -0.01. This is because, as mentioned above, the parameter ρ is used to reflect risky tolerance of investors. A value of ρ close to zero

---

[2] ρ here is corresponding to α of problem (40) in [8]. The returns rate $\mathbf{r}$ of risky assets in our framework express the gross returns rate per annum, instead of the pure returns used in [8].

[3] The worst-case Sharp ratio

$$WCSR = \min_{(\boldsymbol{\mu}, V) \in \mathcal{S}_m \times \mathcal{S}_v} \frac{(\boldsymbol{\mu} - r_f \mathbf{e}_n)^T \mathbf{w}^*}{\sqrt{(\mathbf{w}^*)^T (V^T F V + D)\mathbf{w}^*}}$$

is also used to measure the performance of optimal portfolio in [8]. But unlike the separable uncertainty set $\mathcal{S}_m \times \mathcal{S}_v$ in [8], we can not obtain directly the worst-case Sharp ratio in joint uncertainty set $\mathcal{S}$. Thus, for simplicity, we only compare the mean Sharp ratio which in fact can also reflect the performance of the optimal solution.

[4] See Definition 1 in [18] on the detail definition of diversification number of a portfolio.

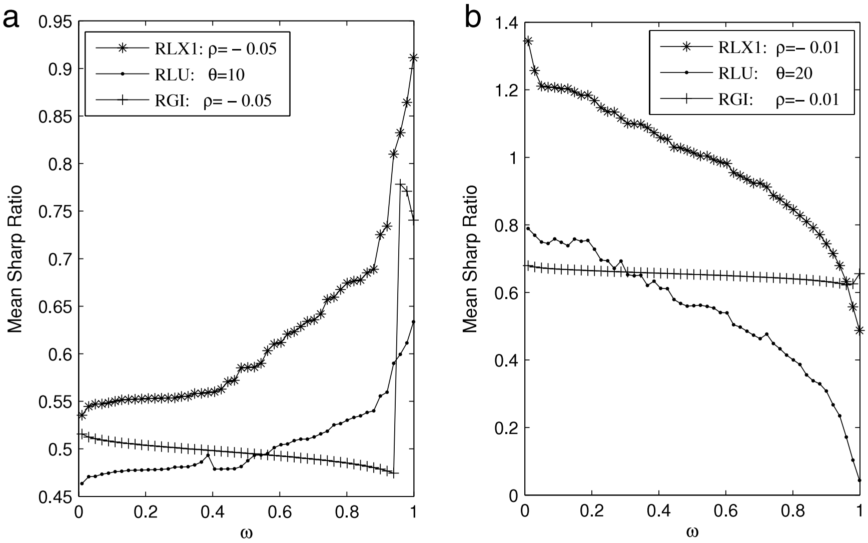

Fig. 1. Comparisons of RLX1, RGI with RLU on the mean Sharp ratio. θ is the risk averse parameter of [17,18].

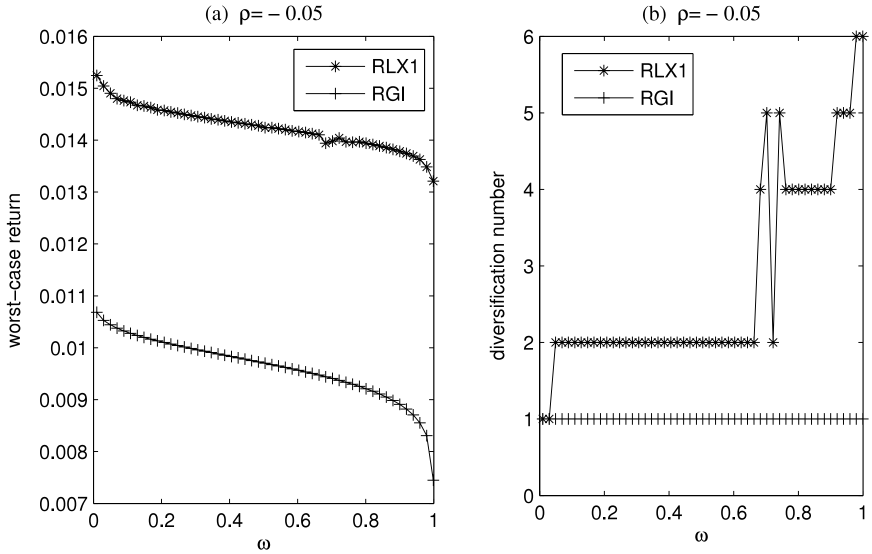

Fig. 2. Comparisons of RLX1 and RGI on the worst-case return and diversification number of resulting portfolio.

indicates that the investor is risk averse and hence, a portfolio with good diversification will be preferred to him. (3) The model (R1) proposed by this paper generates portfolios with worst-case return and diversification number greater than that generated by the model RGI of [8]. This is not surprising, since for a given value of ω, the confidence level of the "marginal + joint" uncertainty set $(\mathcal{S}_\mu, \mathcal{S})$ is exactly the same as ω, while the confidence level of the separable uncertainty set $\mathcal{S}_v \times \mathcal{S}_m$ is much higher than ω. It can be observed from these three figures that for small values of ω, model (R1) is more robust than model RGI, and that for large values of ω, model RGI displays overconfidence. Hence, the proposed "marginal + joint" uncertainty set possesses the advantages of both the separable uncertainty set $\mathcal{S}_v \times \mathcal{S}_m$ and the joint uncertainty set $\mathcal{S}$.

7.2. Test results with real market stocks and simulated options data

In this subsection, data from the real stock market is used to test the performance of the proposed models. Tests are performed by the following rolling-horizon procedure (RHP).

RHP:

1. Select n stocks and m market factors. Collect data of these stocks and factors, and then calculate on the return vectors $\mathbf{r}$ of stocks and the return vectors $\mathbf{f}$ of factors. Partition these data into T equal periods, and each period contains p trading days. Set $t = 1$.

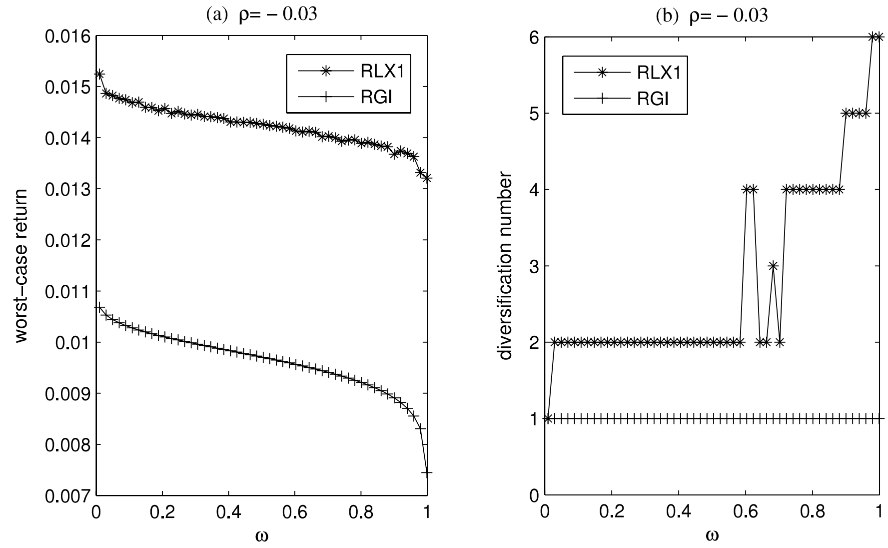

Fig. 3. Comparisons of RLX1 and RGI on the worst-case return and diversification number.

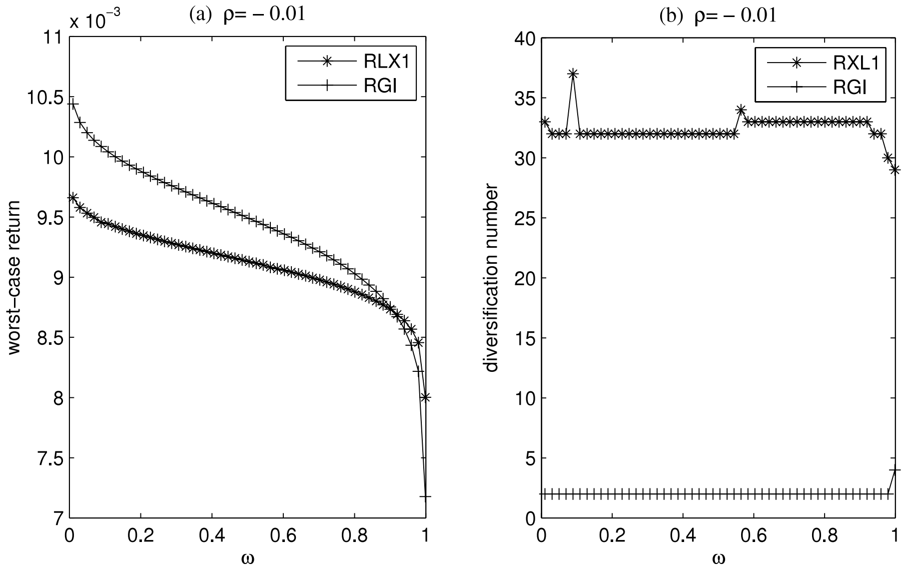

Fig. 4. Comparisons of RLX1 and RGI on the worst-case return and diversification number.

2. Estimate the following parameters using the data of p trading days in the t-th period.
   (a) Evaluate the least square estimates $\bar{\mathbf{x}}$ (i.e. $\overline{\boldsymbol{\mu}}$ and $\overline{V}$) using (4.2).
   (b) Evaluate s_i^2, the estimation of variance σ_i^2 of residual ϵ_i, using (4.3), set $d_i = s_i^2 (i = 1, \ldots, n)$, and then calculate the maximum likelihood estimate F on covariance matrix of market factors by the following formula

$$F = \frac{1}{p-1}\left[BB^T - \frac{1}{p}(B\mathbf{e}_p)(B\mathbf{e}_p)^T\right].$$

3. Give a confidence threshold ω,[5] to set up both the joint ellipsoidal uncertainty set $\mathcal{S}$ and the corresponding marginal uncertainty set $\mathcal{S}_\mu$ using (4.1) and (6.1), respectively.
4. Solve SOCP problem (6.22) to obtain a robust portfolio $\mathbf{w}^t$ for problem (R1).
5. Simulate prices of call and put options with mature in half a year using Black–Scholes formula [29] and calculate matrix G and vector $\mathbf{b}$ in (3.3).

5 Set $\widetilde{\omega} = \omega^{1/n}$, we can build the separable uncertainty $\mathcal{S}_m \times \mathcal{S}_v$ using (56) and (57) in [8], and the confidence level of $\mathcal{S}_m \times \mathcal{S}_v$ in fact is greater than $\widetilde{\omega}^n = \omega$, see [8,17] for detail.

Table 1
The stock codes of chosen 21 stocks from SSE50.

600,000	600,005	600,015	600,016	600,019	600,028	600,030
600,036	600,050	600,089	600,104	600,362	600,383	600,489
600,519	600,547	600,550	600,739	600,795	600,837	600,900

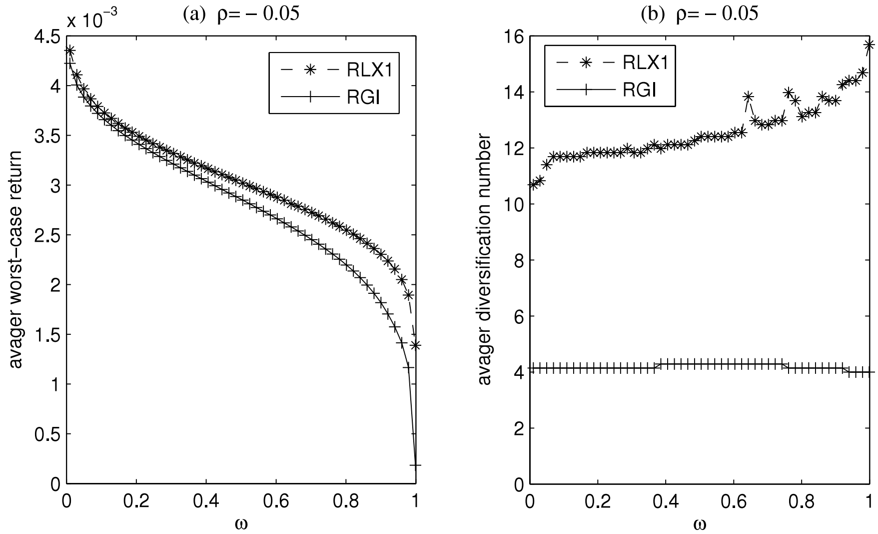

Fig. 5. Comparisons of model RLX1 and RGI on the average worst-case return and diversification number.

6. Solve SOCP problem (6.19) to obtain a robust portfolio $\widetilde{\mathbf{w}}^t$ for problem (R2).
7. Use $\mathbf{w}^t$ and $\widetilde{\mathbf{w}}^t$ as investment strategies in the next period and calculate their results. Set $t = t + 1$ and go to step 2 if $t < T$.

We select $n = 21$ stocks from Shanghai Stock 50 index (SSE50) in tests, and Table 1 lists the codes of these stocks. $m = 5$ market factors are considered consisting of Shanghai Stock index (SSE), Shanghai Stock 50 index (SSE50), Shanghai Stock index 180 index (SSE180), Shanghai Stock A Share index (SSAS) and Shanghai Stock Industry Index (SSI). The collected price time-series of these stocks and factors cover the period from May 8, 2006 to June 25, 2010 (a total of 1008 trading days). The period is then partitioned into 8 subperiods of length $p = 126$ (half a year) so that there are 7 investment subperiods in the test and each subperiod contains $p = 126$ trading days.

RHP is first applied to the 21 stocks data, and then comparisons of model (R1) with model RGI are performed on the average worst-case return, average diversification number and wealth growth rate at each subperiod. Let $\mathbf{w}^t$ be the robust portfolio generated by solving model (6.22) based on the t-th subperiod data, τ^t and $N^t (t = 1, \ldots, 7)$ the resulting worst-case return and diversification number. Then, the average worst-case return is calculated by $(\sum_{t=1}^{7} \tau^t)/7$ and the average diversification number is calculated by $(\sum_{t=1}^{7} N^t)/7$. Assume that $\mathbf{w}^t$ is used as the investing strategy over the $(t + 1)$-th subperiod. Then the wealth growth rate of $\mathbf{w}^t$ over the $(t + 1)$-th subperiod is calculated by

$$\left(\prod_{1 \le k \le p} \left(\mathbf{e}_n + \widehat{\mathbf{r}}_k^{t+1} \right) \right)^T \mathbf{w}^t - 1, \tag{7.1}$$

where $\widehat{\mathbf{r}}_k^{t+1} \in \mathbb{R}^n$ is the k-th column of matrix $\mathbf{r}^{t+1} = [\widehat{\mathbf{r}}_1^{t+1}, \ldots, \widehat{\mathbf{r}}_k^{t+1}, \ldots, \widehat{\mathbf{r}}_p^{t+1}]$, the estimations of return rate $\mathbf{r}$ at the $(t + 1)$-th subperiod.

Fig. 5 gives results of models RLX1 and RGI on the average worst-case returns and diversification numbers for fixed value of $\rho = -0.05$. It can be observed from the figure that the average worst-case return of model RLX1 is greater than that of the model RGI, and the robust portfolio obtained from model RLX1 is fairly diversified. Results in Fig. 6 show that the wealth growth rate generated from model RLX1 is greater than that generated from model RGI for all subperiods. Fig. 7 gives the wealth growth rates of model RLX1 for different values of ω. It can be observed from the results in Fig. 7 that the variances of the wealth growth rates in all investment subperiods decrease as the value of ω increases, and that robust portfolios with relative steady and conservative returns can be obtained when the uncertainty set is enlarged.

The next test is performed for model (R2) and its SOCP approximation (6.19) (denoted by RLX2) based on real market stocks and simulated options. It is assumed that there are 15 put and 15 call options which mature in half of year on each stock, and hence there exist 21 stocks and 630 options available for portfolio selection. It is assumed that the price process

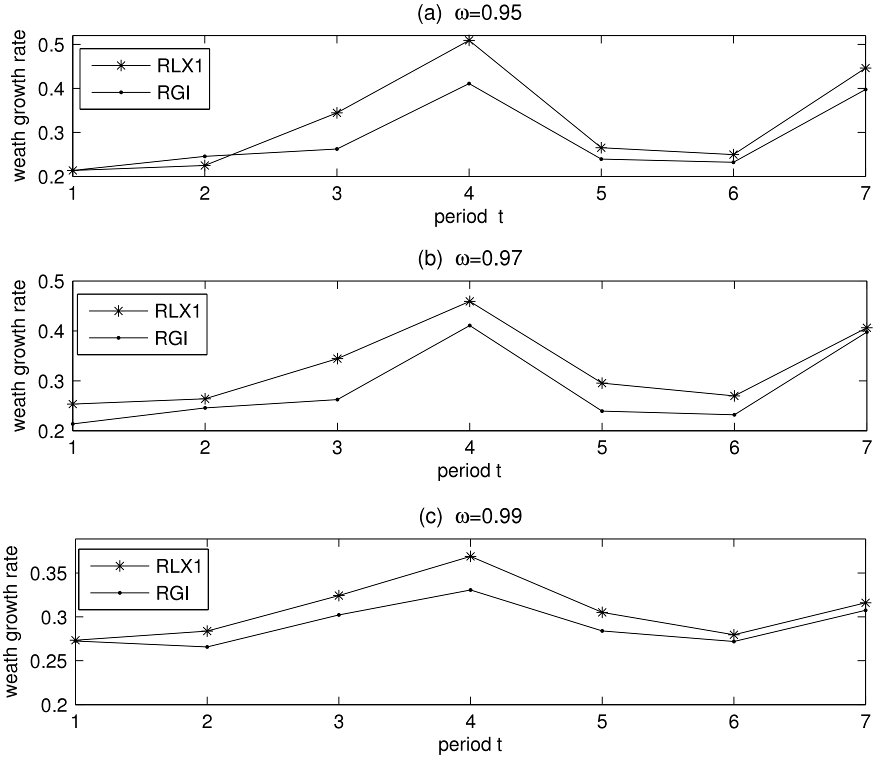

Fig. 6. Comparisons of model RLX1 and RGI on wealth growth rate for $\omega = 0.95$, 0.97 and 0.99.

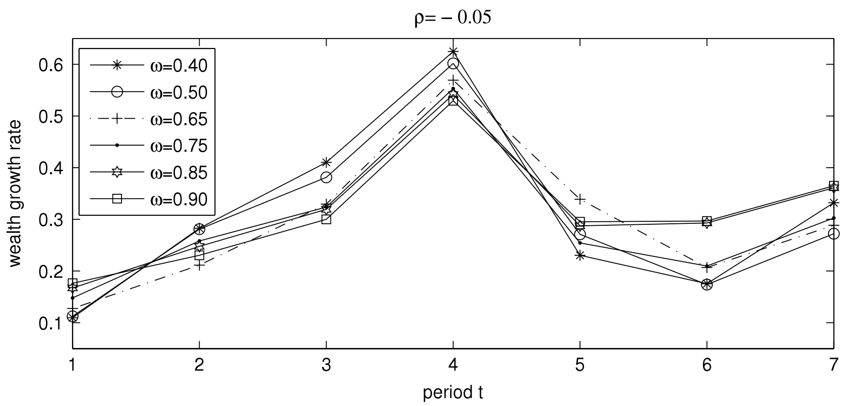

Fig. 7. Wealth growth rate of model RLX1 for different ω.

$S_t^i, t \geq 0$ of stock i follows a geometric Brownian motion so that the well-known Black–Scholes foumula [29] can be used to simulate prices of all put and call options and formulas (3.1) and (3.2) can be used to evaluate the elements of the vector **b** and the matrix G in (3.3)

Let $S_t^i, t = 1, 2, \ldots, 7$ be the price of stock i at the beginning of the t-th investment subperiod (half a year), and $h = (110\% - 95\%)S_t^i/14$. Then the strike prices of all 15 put and 15 call options on stock i for the t-th investment subperiod takes the values of $95\% S_t^i + kh, k = 0, 1, \ldots, 14$. Let $\widehat{r}_t^i$ denote the annual return of stock i in $(t-1)$-th subperiod, and

$$\widehat{\mu}_t^i = \mathbb{E}[\widehat{r}_t^i], \qquad (\widehat{\sigma}_t^i)^2 = \mathrm{var}(\widehat{r}_t^i)$$

be its expectation and variance. It follows from the factor model (2.1) that the expectation and variance can be computed by the following equalities

$$\widehat{\mu}_t^i = \mathbb{E}[\widehat{r}_t^i] = \overline{\mu}_t^i, \qquad (\widehat{\sigma}_t^i)^2 = \mathrm{var}(\widehat{r}_t^i) = \overline{V}_{ti}^T F_t \overline{V}_{ti} + d_{ti}, \quad i = 1, \ldots, n, \tag{7.2}$$

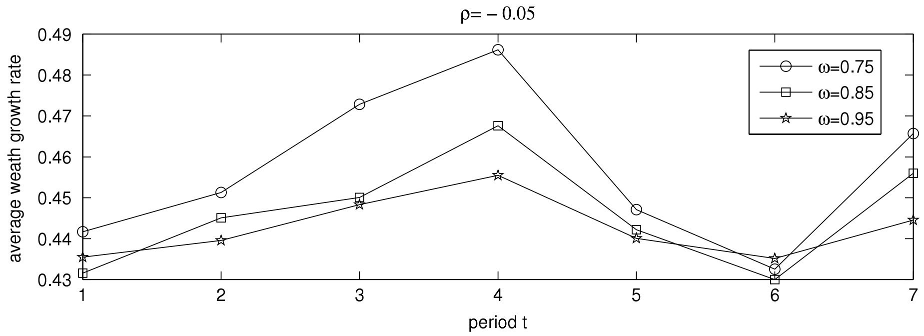

Fig. 8. Wealth growth rate of RLX2 for $\omega = 0.75, 0.85$ and 0.95.

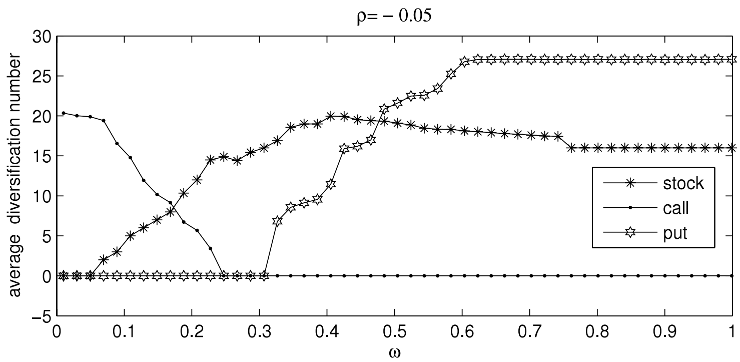

Fig. 9. Average diversification number of RLX2 on call, put options and stocks as a function of ω.

where $\overline{\mu}_t^i$, $\overline{V}_t$, F_t and d_{ti} are estimated based on step 2 in RHP and $(t-1)$-th subperiod data, and $\overline{V}_{ti}$ is the i-th column of the factor loading matrix $\overline{V}_t$, then

$$\widetilde{\mu}_t^i = \frac{1}{2}\widehat{\mu}_t^i, \qquad \widetilde{\sigma}_t^i = \widehat{\sigma}_t^i\sqrt{\frac{1}{2}} \tag{7.3}$$

will be used in Black–Scholes formula to calculate prices c_{tj}^i of call options and prices p_{tj}^i of put options ($j = 1, 2, \ldots, 15$) used in t-th investment subperiod. While

$$r_t^i = \widetilde{\mu}_t^i \tag{7.4}$$

will be used in (3.1) and (3.2) to calculate the elements of the vector $\mathbf{b}$ and the matrix G that are used to build the robust portfolio selection model for the t-th subperiod.

Fig. 8 gives the results of wealth growth rate obtained from RLX2 with real stocks and the simulated options. Let $\widetilde{\mathbf{w}}^t = (\mathbf{w}^t, \widetilde{\mathbf{w}}^{d,t})$ be the robust portfolio of stocks and options obtained from RLX2 at the t-th subperiod. Then the wealth growth rate is calculated by

$$\left(\prod_{1\le k\le p}\left(\mathbf{e}_n + \widehat{\mathbf{r}}_k^{t+1}\right)\right)^T \mathbf{w}^t + \left(\widetilde{\mathbf{r}}_p^d\right)^T \widetilde{\mathbf{w}}^{d,t} - 1, \tag{7.5}$$

where $\widetilde{\mathbf{r}}_p^d = \max\{0, \mathbf{b}^{t+1} + G^{t+1}\widehat{\mathbf{r}}_p^{t+1}\} \in \mathbb{R}^q$. It can be observed by comparing results in Figs. 6 and 8 that the portfolio generated by RLX2 is more robust than those generated by RLX1 and RGI, since the difference between the largest and the lowest wealth growth rates in Fig. 8 is not greater than 0.05 while the same difference for RLX1 and RGI is not less 0.1 (see the case of $\omega = 0.99$ in Fig. 6). Moreover, results of Fig. 8 also show that robust portfolios generated by model RLX2 is not too conservative. This is because the lowest wealth growth rate of RLX2 is not less 0.43 which is greater than the average wealth growth rate of RLX1 exhibited in Fig. 7.

Fig. 9 exhibits the diversification numbers of the robust portfolios generated from model RLX2 as function of the parameter value of ω. When investors assign a fair small value to ω, i.e. the uncertainty set is very small, model RLX2 will generate a portfolio with either pure call options ($\omega \le 0.063$) or call options plus stocks ($0.063 \le \omega \le 0.242$). This is not surprising, because a small uncertainty set implies that the investor expects a high return. It is known that the leverage effect of call options guarantees a desired return when they mature in-the-money. It can also be observed from Fig. 9 that the average numbers of call options in resulting portfolios decrease to zero as the value of ω increases close to 0.242. When the value of ω increases close to 0.311, components of resulting portfolios are changed from entire stocks to stocks plus put options. Then the average volume of stocks and put options contained in resulting portfolios increase as the value of ω continue to increase, and when the value of ω approaches 0.772, the average volume of stocks and put options contained in resulting portfolios stays constant. This implies that a uncertainty set with value of $\omega \geq 0.772$, instead of an uncertainty set with $\omega = 0.99$ in model RLX1 and RGI, can be chosen for risk averse investors and the wealth growth rate of resulting portfolio is still very robust.

8. Conclusions

A robust portfolio selection model under a "marginal + joint" uncertainty set is proposed in this paper. The model processes the advantages of both the separable uncertainty set and joint ellipsoidal uncertainty set. Furthermore, one more robust portfolio selection model with option protection is proposed by combining options into the robust portfolio selection model above. Convex programming approximations with LMIs constraints to both models are formulated. The proposed robust portfolio selection model with options can hedge risks and generates robust portfolios with well wealth growth rate, when either an extreme event such as a market crash occurs or returns of stocks are outside of the uncertainty set. This property of the model has been confirmed by tests on real data from the Chinese stock market during a period involving October 17, 2007 when the SSE index had a large drop (over 200 points) and simulated options.

Acknowledgments

The authors are grateful to two anonymous referees and the Editors for very valuable comments, remarks and suggestions which helped considerably improve this paper.

Appendix

Now we give the duality of problem (6.19) by introducing some dual variables. The dual variables and their corresponding constraints in problem (6.19) are listed in Table A.1, where $Y_{11}^1, Y_{11}^2 \in \mathbb{R}$ and $Y_{22}^1 \in \mathbb{R}^{(m+1)n\times(m+1)n}$, $Y_{22}^2 \in \mathbb{R}^{n\times n}$, $\mathbf{x}^i = (x_1^i, (\mathbf{x}_2^i)^T)^T$ with $x_1^i (\in \mathbb{R}) \geq 0$, $\mathbf{x}_2^i \in \mathbb{R}^n$ satisfying $\|\mathbf{x}_2^i\| \leq x_1^i$ $(i = 1, 2)$, that is,

$$\mathbf{x}^i \in \mathrm{socp}(n+1) = \left\{ \mathbf{x} = (x_1^i, x_2^i, \dots, x_{n+1}^i)^T \in \mathbb{R}^{n+1} : \sqrt{\sum_{k=2}^{n+1} (x_k^i)^2} \leq x_1^i \right\}, \quad i = 1, 2.$$

Then the dual problem of (6.19) is given by

$$\begin{array}{ll} \min\limits_{\mathbf{x}^1,\mathbf{x}^2,\mathbf{x}^3,\mathbf{x}^4,Y^1,Y^2} & 2\rho Y_{11}^1 + x_1^4 \\ s.t. & 2W^T Y_{12}^1 - 2Y_{12}^2 - D^{1/2}\mathbf{x}_2^1 - S\mathbf{x}_2^2 - \overline{\mu} x_1^3 + \mathbf{x}_4^2 \geq 0 \\ & -2\mathbf{b}Y_{11}^1 - \mathbf{b}x_1^3 - G\mathbf{x}_2^4 \geq 0 \\ & \lambda_0 \widetilde{F} \odot Y_{22}^1 - Y_{22}^2 = 0 \\ & -\begin{pmatrix} \eta & \mathbf{u}^T \\ \mathbf{u} & R \end{pmatrix} \cdot Y^1 \geq 0 \\ & \mathbf{e}_n x_1^4 - \mathbf{x}_2^4 \geq 0, \qquad \mathbf{e}_q x_1^3 - \mathbf{x}_2^3 \geq 0 \\ & -\frac{1}{\sqrt{\kappa}} x_1^2 + x_1^3 \geq 0, \qquad -Y_{11}^2 + x_1^1 \geq 0, \qquad x_1^3 \geq 1 \\ & Y^1 \succeq 0, \qquad Y^2 \succeq 0, \qquad \mathbf{x}^1, \mathbf{x}^2 \in \mathrm{socp}(n+1), \qquad \mathbf{x}^3 \geq 0, \end{array} \tag{A.1}$$

where

$$W = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ \mathbf{0}_m & \mathbf{0}_m & \mathbf{0}_m & \cdots & \mathbf{0}_m \\ 0 & 1 & 0 & 0 & 0 \\ \mathbf{0}_m & \mathbf{0}_m & \mathbf{0}_m & \cdots & \mathbf{0}_m \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & 0 & 0 & 1 \\ \mathbf{0}_m & \mathbf{0}_m & \mathbf{0}_m & \cdots & \mathbf{0}_m \end{pmatrix}$$

is an $n(m+1) \times n$ matrix and

Table A.1
The constraints of (6.19) and its dual variables.

The constraints of (6.19)	$\longrightarrow$	Dual variables
$\begin{pmatrix} \theta\eta - 2(\rho - \mathbf{b}^T\mathbf{y}) & \theta\mathbf{u}^T + \mathbf{v}^T \\ \theta\mathbf{u} + \mathbf{v} & \theta R - \lambda_0 X \otimes \widetilde{F} \end{pmatrix} \succeq 0,$	$\longrightarrow$	$Y^1 = \begin{pmatrix} Y^1_{11} & Y^1_{12} \\ Y^1_{21} & Y^1_{22} \end{pmatrix} \succeq 0$
$\begin{pmatrix} \tau_1 & \boldsymbol{\phi}^T \\ \boldsymbol{\phi} & X \end{pmatrix} \succeq 0,$	$\longrightarrow$	$Y^2 = \begin{pmatrix} Y^2_{11} & Y^2_{12} \\ Y^2_{21} & Y^2_{22} \end{pmatrix} \succeq 0$
$\\|D^{1/2}\boldsymbol{\phi}\\| \le \tau_1$	$\longrightarrow$	$\mathbf{x}^1 \in \mathrm{socp}(n+1)$
$\sqrt{\kappa}\\|S\boldsymbol{\phi}\\| \le \tau_2$	$\longrightarrow$	$\mathbf{x}^2 \in \mathrm{socp}(n+1)$
$\tau + \tau_2 - \boldsymbol{\phi}^T\overline{\boldsymbol{\mu}} - \mathbf{b}^T\mathbf{y} \le 0$	$\longrightarrow$	$x^3_1 \ge 0$
$\mathbf{y} - \mathbf{w}^d \le 0$	$\longrightarrow$	$\mathbf{x}^3_2(\in \mathbb{R}^q) \ge 0$
$\mathbf{e}_n^T\mathbf{w} + \mathbf{e}_q^T\mathbf{w}^d = 1$	$\longrightarrow$	$x^4_1 \in \mathbb{R}$
$\boldsymbol{\phi} - \mathbf{w} - G^T\mathbf{y} = 0$	$\longrightarrow$	$\mathbf{x}^4_2 \in \mathbb{R}^n$
$\tau \ge 0, \tau_1 \ge 0, \tau_2 \ge 0,$ $\mathbf{w} \ge 0, \mathbf{w}^d \ge 0, \mathbf{y} \ge 0,$ $\theta \ge 0, \boldsymbol{\phi} \ge 0$	$\longrightarrow$	$\mathbf{x}^5(\in \mathbb{R}^{2(n+q)+4}) \ge 0$

$$\widetilde{F} \odot Z := \begin{pmatrix} \widetilde{F}\cdot Z_{11} & \cdots & \widetilde{F}\cdot Z_{1n} \\ \widetilde{F}\cdot Z_{21} & \cdots & \widetilde{F}\cdot Z_{21} \\ & \vdots & \\ \widetilde{F}\cdot Z_{n1} & \cdots & \widetilde{F}\cdot Z_{nn} \end{pmatrix} \in \mathbb{R}^{n\times n}$$

$Z = (Z_{ij})$ with $Z_{ij} \in \mathbb{R}^{(m+1)\times(m+1)}$ for $i, j = 1, \ldots, n$, and $\widetilde{F} = \begin{pmatrix} 0 & \mathbf{0}_m^T \\ \mathbf{0}_m & F \end{pmatrix}$.

The following theorem indicates that problem (6.19) and its duality (A.1) satisfy the strong dual theorem and is solvable using some existing primal–dual interior algorithms.

Theorem 1. *Let $\omega \in (0, 1), \theta > 0$. If $F \neq 0$ is positive semi-definite and the matrix $A = (\mathbf{e}_p, B^T)$ is full column rank, then problem (6.19) and its dual (A.1) are both strictly feasible and satisfy the strong dual theorem, i.e. they are solvable and the duality gap is zero.* $\square$

The proof of this theorem is similar to that of Theorem 4.5 in [17], and hence is omitted here.

References

[1] H.M. Markowitz, Portfolio selection, Journal of Finance 7 (1952) 77–91.
[2] H.M. Markowitz, The optimization of a quadratic function subject to linear constraints, Naval Research Logistics Quarterly 3 (1956) 111–133.
[3] R.O. Michaud, Efficient Asset Management: A Practical Guide to Stock Portfolio Management and Asset Allocation, in: Financial Management Association, Survey and Synthesis Series, HBS Press, Boston, MA, 1998.
[4] A.L. Soyster, Convex programming with set-inclusive constraints and applications to inexact linear programming, Operations Research 21 (1973) 1154–1157.
[5] A. Ben-Tal, A. Nemirovski, Robust convex optimization, Mathematics of Operations Research 23 (1998) 769–805.
[6] A. Ben-Tal, A. Nemirovski, Robust solution of uncertain linear programs, Operations Research Letters 25 (1999) 1–13.
[7] L. El Ghaoui, H. Lebret, Robust solutions to least-squares problems with uncertain data, SIAM Journal on Matrix Analysis and Applications 18 (1997) 1035–1064.
[8] D. Goldfarb, G. Iyengar, Robust portfolio selection problems, Mathematics of Operations Research 28 (2003) 1–38.
[9] Yu.E. Nesterov, A. Nemirovsky, Interior-Point Polynomial Methods in Convex Programming, in: Studies in Applied Mathematics, vol. 13, SIAM, Philadelphia, PA, 1994.
[10] L. El Ghaoui, M. Oks, F. Oustry, Worst-case value-at-risk and robust portfolio optimization: a conic programming approach, Operations Research 51 (2003) 543–556.
[11] S. Zhu, M. Fukushima, Worst-case conditional value-at-risk with application to robust portfolio management, Operations Research 57 (2009) 1155–1168.
[12] S. Zhu, D. Li, S. Wang, Robust portfolio selection under downside risk measures, Quantitative Finance 7 (2009) 869–885.
[13] D.-S. Huang, S.-S. Zhu, Frank J. Fabozzi, Masao Fukushima, Portfolio selection with uncertain exit time: a robust CVaR approach, Journal of Economic Dynamics and Control 32 (2008) 594–623.
[14] F.J. Fabozzi, D.-H. Huang, G.-F. Zhou, Robust portfolios: contributions from operations research and finance, Annals of Operations Research 176 (2010) 191–220.
[15] B.Y. Halldórsson, R.H. Tütüncü, An interior-point method for a class of saddle problems, Journal of Optimization Theory and Applications 116 (2004) 559–590.
[16] R. Tütüncü, M. Koenig, Robust asset allocation, Annals of Operations Research 132 (2004) 157–187.
[17] Z. Lu, Robust portfolio selection based on a joint ellipsoidal uncertainty set, Optimization Methods and Software 26 (1) (2011) 89–104.
[18] Z. Lu, A computational study on robust portfolio selection based on a joint ellipsoidal uncertainty set, Mathematical Programming, Series A 126 (2011) 193–201. http://dx.doi.org/10.1007/s10107-009-0271-z.
[19] S. Ceria, R. Stubbs, Incorporating estimation errors into portfolio selection: robust portfolio construction, Journal of Asset Management 7 (2) (2006) 109–127.
[20] F. Lutgens, S. Sturm, A. Kolen, Robust one-period option hedging, Operations Research 54 (6) (2006) 1051–1062.

[21] S. Zymler, B. Rustem, D. Kuhn, Robust portfolio optimization with derivative insurance guarantees, European Journal of Operational Research 210 (2) (2011) 410–424.
[22] W.F. Sharpe, Mutual fund performance, Journal of Business 39 (S1) (1966) 119–138.
[23] S. Zymler, D. Kuhn, B. Rustem, Worst-case value-at-risk of non-linear portfolios, Aug. 2009, manuscript. Available: http://www.optimization-online.org/DB_FILE/2009/08/2379.pdf.
[24] G.H. Hardy, J.E. Littlewood, G. Pólya, Some Theorems Concerning Monotonic Functions, second ed., in: Section 3.14 in Inequalities, Cambridge University Press, Cambridge, England, United Kingdom, 1988, pp. 83–84.
[25] M.S. Bazaraa, H.D. Sherali, C.M. Shetty, Nonlinear Programming: Theory and Algorithms, second ed., John Wiley & Sons, New York, 1993.
[26] S. Boyd, L. El Ghaoui, E. Feron, V. Balakrishnan, Linear Matrix Inequalities in System and Control Theory, in: Studies in Applied Mathematics, vol. 15, Society for Industrial and Applied Mathematics, Philadelphia, PA, 1994.
[27] R.A. Horn, C.R. Johnson, Matrix Analysis, Cambridge University Press, NewYork, 1985.
[28] J.F. Sturm, Using SeDuMi 1.02, a MATLAB toolbox for optimization over symmetric cones, Optimization Methods and Software 11–12 (1999) 625–653.
[29] F. Black, M.S. Scholes, The pricing of options and corporate liabilities, Journal of Political Economy 81 (3) (1973) 637–654.