A SYMMETRIC GAUSS-SEIDEL BASED METHOD FOR A CLASS OF MULTI-PERIOD MEAN-VARIANCE PORTFOLIO SELECTION PROBLEMS

NING ZHANG

Department of Applied Mathematics
The Hong Kong Polytechnic University
Hong Kong, China

(Communicated by Hailiang Yang)

ABSTRACT. It is commonly accepted that the estimation error of asset returns’ sample mean is much larger than that of sample covariance. In order to hedge the risk raised by the estimation error of the sample mean, we propose a sparse and robust multi-period mean-variance portfolio selection model and show how this proposed model can be equivalently reformulated as a multi-block non-smooth convex optimization problem. In order to get an optimal strategy, a symmetric Gauss-Seidel based method is implemented. Moreover, we show that the algorithm is globally linearly convergent. The effectiveness of our portfolio selection model and the efficiency of its solution method are demonstrated by empirical experiments on both the synthetic and real datasets.

1. Introduction. Since the mean-variance portfolio selection theory was introduced by Markowitz in his seminal paper [28], this methodology has been extensively studied. It has been shown that the Markowitz’s mean-variance portfolios are unstable and perform poorly in terms of their out-of-sample mean and variance [14, 30, 9]. One of the most well-accepted reason is that both the sample mean and sample covariance are not easy to estimate, especially the sample mean, see [6, 21, 14, 30, 9]. Due to this reason, the global minimum-variance portfolio becomes more favorable. However, the sample mean can still provide some useful information.

In order to address the uncertainty of asset returns’ sample mean in mean-variance portfolio optimization problem, we take a robust optimization approach to multi-period portfolio management. For a large-scale asset pool, to reduce the complexity of portfolio management, the sparse character of portfolios is encouraged. Therefore, in this paper, we propose the following \( \ell_1 \) norm penalized robust...

2010 Mathematics Subject Classification. 65K05, 90C25, 91G10.

Key words and phrases. Robust portfolio selection, multi-block convex optimization, symmetric Gauss-Seidel decomposition, Q-linear convergence, calmness.

This author’s research is supported in part by the Singapore-ETH Centre(SEC), which was established as a collaboration between ETH Zurich and National Research Foundation(NRF) Singapore (FI 370074011) under the auspices of the NRFs Campus for Research Excellence and Technological Enterprise (CREATE) programme.
multi-period portfolio selection model:
\[
\min_{x_0 \in X} \left\{ \frac{1}{2} x_0^T C x_0 + \rho \|x_0\|_1 + \max_{\xi_1 \in U_1} \left\{ \min_{x_1 \in X(x_0)} f_1(x_1, \xi_1) + \max_{\xi_2 \in U_2} \left\{ \min_{x_2 \in X(x_1)} f_2(x_2, \xi_2) + \cdots + \max_{\xi_n \in U_n} \left\{ \min_{x_n \in X(x_{n-1})} f_n(x_n, \xi_n) \right\} \right\} \right\},
\]
where for \(k = 1, \ldots, n\), uncertainty sets \(U_k\) are given polyhedrons,
\[
f_k(x_k, \xi_k) := -\xi_k^T x_k + \frac{1}{2} x_k^T C_k x_k, \quad k = 1, \ldots, n,
\]
\[
X(x_{k-1}) := \left\{ x_k \in \mathbb{R}^p \mid e^T y = 1, \ W_k x_k + T_k x_{k-1} \geq h_k \right\}, \quad h_k \in \mathbb{R}^m, \quad k = 1, \ldots, n,
\]
\[
X := \left\{ x \in \mathbb{R}^p \mid e^T x = 1 \right\}.
\]

Here, for each \(k \in \{1, \ldots, n\}\), \(x_k \in \mathbb{R}^p\) is the vector of asset weights in the \(k\)-th period, \(\xi_k \in \mathbb{R}^p\) and \(C_k \in \mathbb{S}^p\) are the sample mean and the sample covariance matrix of the asset returns, respectively, \(e \in \mathbb{R}^p\) is a vector of ones, and \(T_k, W_k \in \mathbb{R}^{m \times p}\) are given parameters. Penalty parameter \(\rho > 0\) is used by control the sparsity of the initial period portfolio, and the sets \(^{(2b)}\) can be used to control transaction costs and/or to satisfy some trading rules. If \(U_k \equiv \{0\}\), \(k = 1, \ldots, n\), model \(^{(1)}\) reduces to multi-period minimum-variance portfolio optimization problem.

There exists a large body of literature on mean-variance portfolio selection problem, see e.g. [18, 5, 14, 17, 12, 29, 44, 43] and the references therein. Therefore, in this paper, we review two important papers about robust multi-period portfolio management. The first robust counterpart of multi-period portfolio optimization problem was introduced in [2]. The unknown parameters considered in [2] are modeled as ellipsoidal sets and then their robust counterpart model can be reformulated as a tractable convex conic programming. Compared to [2], polyhedral uncertainty sets are adapted in [3]. They showed that their robust counterpart model can be reformulated as a linear optimization problem. Note that these two works are based on the multi-stage stochastic linear programs for portfolio optimization that presented in [13], which is a different asset allocation framework from Markowitz’s portfolio theory. To know more about the robust portfolio optimization problem, we refer readers to [22] and the references therein.

It is well known that the direct extension of alternating direction method for multipliers (ADMM) for multi-block convex optimization problem with linear constraints may not be convergent [7]. In order to get convergent algorithms, some modifications to the directly extended ADMM have been done. One of the most efficient and easy-to-use modifications is the symmetric Gauss-Seidel based ADMM (sGS-ADMM), which was proposed by [8]. The sGS-ADMM was initiated by [26, Algorithm SCB-SPADMM], where “SCB-SPADMM” is short for Schur complement based semiproximal ADMM (sPADMM). This algorithm was generalized and reformulated to the inexact setting in [8, 27]. To know more about the block symmetric Gauss-Seidel type methods and their applications, we refer readers to [27, 37, 26, 8, 25] and the references therein.

The contributions of this paper are mainly threefold. First, we propose a novel sparse and robust mean-variance portfolio selection model. Under some mild assumptions on uncertainty set, the sparse character can be kept over multiple periods.
Besides, linear constraints are also involved in each period to control transaction cost and to satisfy some other necessary rules. Second, we show that the proposed model in this paper can be equivalently reformulated into a multi-block nonsmooth convex optimization problem. Third, a globally Q-linearly convergent algorithm is proposed to find the optimal investment strategies. Taking the multi-period minimum-variance portfolio as a benchmark, 370 stocks in Standard & Poor’s 500 index are used to demonstrate the effectiveness of the proposed model.

The remainder of this paper is organized as follows. Section 2 shows how the proposed robust multi-period mean-variance portfolio selection model can be reformulated as a multi-block nonsmooth convex optimization problem. In Section 3, a symmetric Gauss-Seidel based multi-block semi-proximal alternating direction method of multipliers is presented. Convergence results and algorithm implementation details are also discussed in this section. We report the results obtained by empirical experiments in Section 4. Finally, we conclude the paper in Section 5.

2. Preliminaries. Let $\phi : \mathbb{R}^n \to \mathbb{R}$ be a closed convex function. The proximal mapping (see e.g. [35]) associated with $\phi$ is defined by
\[
\text{Prox}_\phi(v) := \arg \min_x \left\{ \phi(x) + \frac{1}{2} \|x - v\|^2 \right\}, \quad \forall \ v \in \mathbb{R}^n.
\]
It is the unique solution of the following convex minimization problem:
\[
\min_x \left\{ \phi(x) + \frac{1}{2} \|x - v\|^2 \right\}.
\]
The conjugate function (see e.g. [33]) of $\phi$ is denoted by $\phi^*$, i.e.,
\[
\phi^*(y) = \sup_x \{ \langle x, y \rangle - \phi(x) \}.
\]

**Example 1.** Let $\alpha > 0$ and $\phi(x) := \alpha \|x\|_s$, $(s = 1, 2, \infty)$, then its conjugate function is given by
\[
\phi^*(y) := \mathbb{I}_{B_\rho}(y) \text{ with } B_\rho := \{ y \mid \|y\|_t \leq \alpha \}, \quad (t = \infty, 2, 1).
\]
Let $\phi$ be the indicator function (see e.g. [33]) of convex set $C$, i.e., $\phi(x) := \mathbb{I}_C(x)$, then its conjugate function $\phi^*(y) = \sup_{x \in C} y^T x$.

Next, we consider the following two-stage optimization problem:
\[
\min_{x \in \mathcal{X}} \left\{ f(x) + \max_{\xi \in \mathcal{U}} \mathcal{Q}(x, \xi) \right\}, \quad (3)
\]
where $f(x) := \frac{1}{2} x^T C x + \rho \|x\|_1$, and $\mathcal{Q} : \mathbb{R}^p \times \mathcal{U} \to \mathbb{R} \cup \{ \pm \infty \} :$
\[
\mathcal{Q}(x, \xi) = \left\{ \min_{y \in \mathbb{R}^p} -\xi^T y + \frac{1}{2} y^T D y \mid W y + T x \geq h, \ e^T y = 1 \right\}. \quad (4)
\]
Here, we assume that $\mathcal{U}$ is a polyhedron, both $C$ and $D$ are positive semidefinite matrices.

Denote the feasible set of (4) by $\mathcal{X}(x) := \{ y \in \mathbb{R}^p \mid W y + T x \geq h, \ e^T y = 1 \}$. The function $\mathcal{Q}(x, \xi)$ is defined with the extended reals $\mathbb{R} \cup \{ \pm \infty \}$ for range, because it takes the values $+\infty$ or $-\infty$ if the optimization problem in brackets is infeasible or unbounded below, respectively.

Let $\mathcal{F}$ be the feasible set of (3), i.e.,
\[
\mathcal{F} := \mathcal{X} \cap \{ x \mid \exists y \text{ such that } W y \geq h - T x, \ e^T y = 1 \}. \quad (5)
\]
optimization problem:

multi-period mean-variance model

Theorem 2.2. Suppose that \( F \neq \emptyset \), then the value function \( Q(x, \xi) \) is convex in \( x \) for all \( \xi \in U \) and the value function \( Q(x, \xi) \) is concave in \( \xi \) for each \( x \in X \).

**Proof.** For any given \( \xi^* \in U \), suppose that \( x_1, x_2 \in X \). If there exist \( y_1 \in \mathcal{X}(x_1) \) and \( y_2 \in \mathcal{X}(x_2) \) such that \( Q(x_1, \xi^*) = -(\xi^*)^T y_1 + \frac{1}{2} y_1^T D y_1, \ i = 1, 2, \) then it is obviously that \( \bar{y} := \alpha y_1 + (1 - \alpha) y_2, \ \forall \alpha \in [0, 1] \) lies in the feasible set \( \mathcal{X}(\bar{x}) \) with \( \bar{x} := \alpha x_1 + (1 - \alpha) x_2 \). Consequently, it holds that

\[
Q(\bar{x}, \xi^*) \leq \alpha Q(x_1, \xi^*) + (1 - \alpha) Q(x_2, \xi^*), \ \forall \alpha \in [0, 1]. \tag{6}
\]

If either \( \mathcal{X}(x_1) \) or \( \mathcal{X}(x_2) \) is an empty set, inequality (6) holds automatically. Next, we show that for any given \( x^* \in X \), it holds that

\[
Q(x^*, \bar{\xi}) \geq \alpha Q(x^*, \xi_1) + (1 - \alpha) Q(x^*, \xi_2), \ \forall \alpha \in [0, 1], \tag{7}
\]

where \( \bar{\xi} := \alpha \xi_1 + (1 - \alpha) \xi_2 \). If \( x^* \notin F \), inequality (7) holds automatically. Suppose that \( x^* \in F \), then there exists \( \bar{y} \in \mathcal{X}(x^*) \) such that

\[
Q(x^*, \bar{\xi}) = -\bar{\xi}^T \bar{y} + \frac{1}{2} \bar{y}^T D \bar{y} \geq \alpha Q(x^*, \xi_1) + (1 - \alpha) Q(x^*, \xi_2).
\]

This completes the proof. \( \square \)

Next, we will show that the robust multi-period mean-variance portfolio optimization problem (1) can be equivalently written as a multi-block convex optimization problem.

**Theorem 2.2.** Suppose that the feasible set of (1) is nonempty. Then the robust multi-period mean-variance model (1) can be reformulated as the following concave optimization problem:

\[
\max_{\xi_1 \in U_{k}} \left\{ \min_{x_k \in \mathcal{X}(x_{k-1})} \frac{1}{2} x_0^T C x_0 + \rho \| x_0 \|_1 + \sum_{k=1}^{n} f_k(x_k, \xi_k) \right\}, \tag{8}
\]

where \( \mathcal{X}(x_{k-1}) := \mathcal{X} \). Moreover, model (1) can be further reformulated as follows:

\[
\min_{x_0, x_1, \ldots, x_n} \frac{1}{2} x_0^T C x_0 + \rho \| x_0 \|_1 + \sum_{k=1}^{n} \left( \frac{1}{2} x_k^T C_k x_k + 2 I_k (-x_k) \right), \tag{9}
\]

s.t. \( e^T x_k = 1, \ k = 0, 1, \ldots, n, \)
\[
W_k x_k + T_k x_{k-1} \geq h_k, \ k = 1, \ldots, n.
\]

**Proof.** Without loss of generality, we assume that \( n = 2 \). Then, model (1) can be explicitly rewritten as

\[
\min_{x_0 \in X} \frac{1}{2} x_0^T C x_0 + \rho \| x_0 \|_1 + \max_{\xi_2 \in U_2} \left\{ \min_{x_1 \in \mathcal{X}(x_0)} \frac{1}{2} x_1^T C_1 x_1 - \xi_1^T x_1 + \max_{\xi_2 \in U_2} Q_2(x_1, \xi_2) \right\},
\]

where

\[
Q_2(x_1, \xi_2) = \left\{ \min_{x_2 \in \mathcal{X}(x_1)} \frac{1}{2} x_2^T C_2 x_2 - \xi_2^T x_2 \right\}.
\]

Directly from Lemma 2.1 and minimax theorem in [36], we can obtain that

\[
\min_{x_0 \in X} \frac{1}{2} x_0^T C x_0 + \rho \| x_0 \|_1 + \max_{\xi_1 \in U_1} \left\{ \min_{x_1 \in \mathcal{X}(x_0)} \frac{1}{2} x_1^T C_1 x_1 - \xi_1^T x_1 + \frac{1}{2} x_2^T C_2 x_2 - \xi_2^T x_2 \right\}.
\]
By using Lemma 2.1 and minimax theorem in [36] again, we have that
\[
\max_{\xi_k \in \mathcal{U}_k, k=1,2} \min_{x_k \in X(k-1)} \left\{ \frac{1}{2} x_k^T C_0 x_0 + \frac{1}{2} x_{k-1}^T C_1 x_{k-1} + \xi_k^T x_1 + \frac{1}{2} x_2^T C_2 x_2 - \xi_2^T x_2 \right\},
\]
and consequently, model (1) can be equivalently reformulated as (8).

Using minimax theorem in [36] once more, (8) can also be equivalently written as
\[
\min_{x_0, x_k \in X(k-1)} \left\{ \frac{1}{2} x_0^T C_0 x_0 + \rho \|x_0\|_1 + \sum_{k=1}^n \frac{1}{2} x_k^T C_k x_k + \left\{ \max_{\xi_k} \langle \xi_k^T, - x_k \rangle - \mathbb{I}_{\mathcal{U}_k}(\xi_k) \right\} \right\}
\]

Then one has (9) from the definition of conjugate function. The proof is completed. \(\Box\)

Note that if the uncertainty sets \(\mathcal{U}_k, k=1,\ldots,n\) are \(\ell_\infty\) balls, i.e., for \(k=1,\ldots,n\),
\[
\mathcal{U}_k := \{ \xi \in \mathbb{R}^p \colon \|\xi\|_\infty = \max_{i=1,\ldots,p} |\xi_i| \leq \lambda_k \}, \lambda_k \geq 0.
\]
Then, we have that \(\mathbb{I}_{\mathcal{U}_k}^*(x_k) = \lambda_k \|x_k\|_1\). This implies that the reformulated problem (9) is an \(\ell_1\) regularized model which can yield a sparse solution.

3. A symmetric Gauss-Seidel based ADMM. In this section, we discuss the implementation details of the sGS-ADMM for finding the optimal solution of the multi-period mean-variance portfolio problems.

By introducing auxiliary variables, problem (9) can be equivalently rewritten as
\[
\min_{x,v} \vartheta(w_p) := \frac{1}{2} x_0^T C_0 x_0 + \varphi(x_0) + \sum_{k=1}^n \left( \frac{1}{2} x_k^T C_k x_k + \mathbb{I}_{\mathcal{U}_k}^*(-x_k) + \mathbb{I}_{\geq 0}(v_k) \right)
\]
s.t. \(e^T x_k = 1, k=0,1,\ldots,n\),
\(W_k x_k + T_k x_{k-1} - v_k = h_k, k=1,\ldots,n\),
(10)
where \(\varphi(x_0) := \rho \|x_0\|_1\), \(x := (x_0, x_1,\ldots,x_n), v := (v_1,\ldots,v_n)\), and \(w_p := (x,v)\).

Its Lagrangian dual problem takes the following form:
\[
\max_{y,\lambda,\mu,\chi,\eta,\zeta} \varrho(w_d) := -\sum_{k=0}^n \left[ \frac{1}{2} y_k^T C_k y_k + \lambda_k \right] - \sum_{k=1}^n \left[ -h_k^T \mu_k + \mathbb{I}_{\geq 0}(\chi_k) \right] - \varphi^*(\eta)
\]
s.t. \(C_0 y_0 + e \lambda_0 - T_0^T \mu_1 + \eta = 0,\)
\(C_k y_k + e \lambda_k - W_k^T \mu_k - T_{k+1}^T \mu_{k+1} + \xi_k = 0, k=1,\ldots,n-1,\)
\(C_n y_n + e \lambda_n - W_n^T \mu_n + \xi_n = 0,\)
\(\mu_k - \chi_k = 0, k=1,\ldots,n,\)
\(\xi_k \in \mathcal{U}_k, k=1,\ldots,n,\)
(11)
where $y := (y_0, y_1, \ldots, y_n)$, $\lambda := (\lambda_0, \lambda_1, \ldots, \lambda_n)$, $\mu = (\mu_1, \ldots, \mu_n)$, $\chi := (\chi_1, \ldots, \chi_n)$, $\xi := (\xi_1, \ldots, \xi_n)$, and $w_\sigma := (y, \lambda, \mu, \chi, \eta, \xi)$. For a given parameter $\sigma > 0$, the augmented Lagrangian function associated with (11) is given by

$$L_\sigma(y, \lambda, \mu, \chi, \eta, \xi; x, v)$$

$$= \frac{1}{2} \sum_{k=0}^{n} \left[ y_k^T C_k y_k + \lambda_k \right] + \sum_{k=1}^{n} \left[ -h_k^T \mu_k + \mu_k \right] + \varphi^*(\eta)$$

$$- \langle C_0 y_0 + e\lambda_0 - T_1^T \mu_1, y_0 \rangle - \langle C_n y_n + e\lambda_n - W_n^T \mu_n + \xi_n, x_n \rangle$$

$$- \sum_{k=1}^{n-1} \langle C_k y_k + e\lambda_k - W_k^T \mu_k - T_{k+1}^T \mu_{k+1} + \xi_k, x_k \rangle - \sum_{k=1}^{n} \langle \mu_k - \chi_k, v_k \rangle$$

$$+ \frac{\sigma}{2} \left( \|C_0 y_0 + e\lambda_0 - T_1^T \mu_1 + \eta\|^2 + \|C_n y_n + e\lambda_n - W_n^T \mu_n + \xi_n\|^2 \right)$$

$$+ \frac{\sigma}{2} \left( \sum_{k=1}^{n} \|\mu_k - \chi_k\|^2 + \sum_{k=1}^{n-1} \|C_k x_k + e\lambda_k - W_k^T \mu_k - T_{k+1}^T \mu_{k+1} + \xi_k\|^2 \right).$$

Now, we are ready to state the symmetric Gauss-Seidel based multi-block semi-proximal ADMM (sGS-sPADMM) for the large scale multi-period portfolio selection problem.

**Algorithm 1: An sGS-sPADMM**

Let $\sigma > 0$ and $\tau \in (0, (1 + \sqrt{5})/2)$ be given parameters. Let $S_k$, $k = 0, 1, \ldots, n$ be given self-adjoint, semidefinite, linear operators. Input initial point $(y^0, \lambda^0, \mu^0, \chi^0, \eta^0, \xi^0, x^0, v^0)$. Set $l := 0$.

**Step 1a:** For $k = 1, \ldots, n$, compute

$$\mu_k^{l+\frac{1}{2}} := \min_{\mu_k} L_\sigma(y^l, \lambda^l, \mu_k^{l+\frac{1}{2}}, \mu_k, \mu_{k+1}^{l+\frac{1}{2}}, \chi^l, \eta^l, \xi^l; x^l, v^l).$$

**Step 1b:** Compute

$$(\chi^{l+1}, \xi^{l+1}, \eta^{l+1}) := \min_{\chi, \xi, \eta} L_\sigma(y^l, \lambda^l, \mu^{l+\frac{1}{2}}, \chi^l, \eta^l, \xi^l; x^l, v^l).$$

**Step 1c:** For $k = n, n-1, \ldots, 1$, compute

$$\mu_k^{l+1} := \min_{\mu_k} L_\sigma(y^l, \lambda^l, \mu_k^{l+\frac{1}{2}}, \mu_{k-1}^{l+\frac{1}{2}}, \mu_k, \mu_{k+1}^{l+\frac{1}{2}}, \chi^{l+1}, \xi^{l+1}, \eta^{l+1}; x^l, v^l).$$

**Step 2:** For $k = 0, 1, \ldots, n$, compute

$$(y^{l+1}, \chi^{l+1}) := \min_{y, \lambda} L_\sigma(y, \lambda, \mu^{l+\frac{1}{2}}, \chi^{l+1}, \xi^{l+1}, \eta^{l+1}; x^l, v^l) + \frac{1}{2} \left\| \left( \frac{y - y^l}{\lambda - \lambda^l} \right) \right\|^2 S,$$

where

$$S := \text{Diag}(S_0, S_1, \ldots, S_n).$$

**Step 3:** Compute

$$x_0^{l+1} = x_0^l - \tau \sigma(C_0 y_0^{l+1} + e^0 - T_1^T \mu_1^{l+1} + \eta^{l+1}),$$

$$x_k^{l+1} = x_k^l - \tau \sigma(C_k y_k^{l+1} + e\lambda_k^{l+1} - W_k^T \mu_k^{l+1} - T_{k+1}^T \mu_{k+1}^{l+1} + \xi_k^{l+1}), k = 1, \ldots, n - 1,$

$$x_n^{l+1} = x_n^l - \tau \sigma(C_n y_n^{l+1} + e\lambda_n^{l+1} - W_n^T \mu_n^{l+1} + \xi_n^{l+1}),$$

$$v_k^{l+1} = v_k^l - \tau \sigma(\mu_k^{l+1} - \chi_k^{l+1}), k = 1, \ldots, n.$$

**Step 4:** If a termination criterion is not met, set $l := l + 1$ and go to Step 1.
Remark 1. (a): Note that both in the Step 1a and Step 1c of Algorithm 1 have to solve the following subproblems: for $k = 1, \ldots, n$,

$$
\mu_k := \min_{\mu_k} \mathcal{L}_\sigma(y^l, \lambda^l, \hat{\mu}_{\leq k-1}, \mu_k, \hat{\mu}_{\geq k-1}, \chi^l, \eta^l, \xi^l; u^l, v^l).
$$

Specifically, for $k = 0, 1, \ldots, n$,

$$(I + T_k^T T_k + W_k^T W_k) \mu_k = R_k,$$

where

$$
R_k = -\sigma^{-1} (-h_k + T_k x_{k-1} + W_k x_k - v_k) - T_k (C_{k-1} y_{k-1} - e \lambda_{k-1} + W_{k-1} \hat{\mu}_{k-1} - \xi_{k-1}) - W_k (C_k y_k - e \lambda_k + T_{k+1} \hat{\mu}_{k+1} - \xi_k) + \chi_k, \ k = 1, \ldots, n,
$$

and $W_k^T \hat{\mu}_0 = 0, T_{n+1}^T \hat{\mu}_{n+1} = 0$.

(b): In Step 1b, the optimization problem (14) can be specifically written as

$$
\chi_{k+1} = \Pi_{U_0} (\mu_{k+1}^{\frac{1}{2}} - w^l/\sigma), \ k = 1, \ldots, n, \quad (17a)
$$

$$
\eta_{k+1} = \arg\min_{\eta} \left\{ \varphi^* (\eta) + \frac{\sigma}{2} \| \eta + C_k y_k^l + e \lambda_k^l - T_k^T \mu_1^{\frac{1}{2}} + x_1^l \|_2^2 \right\}, \quad (17b)
$$

$$
\xi_k^{l+1} = \Pi_{U_k} (C_k y_k^l - e \lambda_k^l + W_k^T \mu_k^{l+1} + T_{k+1} \hat{\mu}_{k+1} + x_k^l), \ k = 1, \ldots, n - 1, \quad (17c)
$$

$$
\xi_n^{l+1} = \Pi_{U_k} (C_n y_n^l - e \lambda_n^l + W_n^T \mu_n^{l+1} + x_n^l), \quad (17d)
$$

Note that for each $k \in \{1, \ldots, n\}$, if $U_k$ is an $\ell_\infty$ ball $B_\infty$, then the projection $\Pi_{U_k}$ has explicit formula, i.e.,

$$
\Pi_{B_\infty} (x) = \max\{-\alpha, \min(x, \alpha)\},
$$

If decision makers take an $\ell_1$ ball or a general polyhedron as uncertainty set $U_k$, the projections (17c) and (17d) can be efficiently calculated by using the algorithms discussed in [11, 19].

(c): In Step 2, the optimal solution $(y_k^{l+1}, \lambda_k^{l+1})$ of problem (15) is solutions of the following linear equations:

$$
V_0 \begin{bmatrix} y_0 \\ \lambda_0 \end{bmatrix} = - \begin{bmatrix} 0 \\ 1 \end{bmatrix} + \begin{bmatrix} C_0 \\ e \end{bmatrix} \begin{bmatrix} x_0 - \sigma (-T_1^T \mu_1^{l+1} + \eta^{l+1}) + S_0 \begin{bmatrix} y_0 \\ \lambda_0 \end{bmatrix} \end{bmatrix},
$$

$$
V_k \begin{bmatrix} y_k \\ \lambda_k \end{bmatrix} = - \begin{bmatrix} 0 \\ 1 \end{bmatrix} + \begin{bmatrix} C_k \\ e \end{bmatrix} \begin{bmatrix} x_k - \sigma (-T_{k+1}^T \mu_{k+1}^{l+1} - W_k^T \mu_k^{l+1} + \xi_k^{l+1}) + S_k \begin{bmatrix} y_k \\ \lambda_k \end{bmatrix} \end{bmatrix},
$$

$$
V_n \begin{bmatrix} y_n \\ \lambda_n \end{bmatrix} = - \begin{bmatrix} 0 \\ 1 \end{bmatrix} + \begin{bmatrix} C_n \\ e \end{bmatrix} \begin{bmatrix} x_n - \sigma (-W_n^T \mu_n^{l+1} + \xi_n^{l+1}) + S_n \begin{bmatrix} y_n \\ \lambda_n \end{bmatrix} \end{bmatrix},
$$

where

$$
V_k := \begin{bmatrix} C_k & 0 \\ 0 & 0 \end{bmatrix} + \sigma \begin{bmatrix} C_k \\ e \end{bmatrix} ^T \begin{bmatrix} C_k & e \end{bmatrix} + S_k, \ k = 0, 1, \ldots, n. \quad (18)
$$

Assumption 1. For $k = 0, 1, \ldots, n$, the self-adjoint, semidefinite, linear operators $S_k$ are chosen such that matrices $V_k$, which are defined by (18), are positive definite.
Theorem 3.1. Suppose that Assumption 1 holds. Let \( \bar{w} := (w_p, w_d) \). Define the proximal type Karush-Kuhn-Tucker (KKT) solution mapping associated with (10) and (11) as

\[
\mathcal{R}(w) = \begin{pmatrix}
e^T x_k - 1, \ k = 0, 1, \ldots, n \\
W_k x_k + T_k x_{k-1} - v_k - h_k, \ k = 1, \ldots, n, \\
C_0 y_0 + e\lambda_0 - T_1^T \mu_1 + \eta, \\
C_k y_k + e\lambda_k - W_k^T \mu_k - T_{k+1}^T \mu_{k+1} + \xi_k, \ k = 1, \ldots, n - 1 \\
C_n y_n + e\lambda_n - W_n^T \mu_n + \xi_n, \\
\mu_k - \chi_k, \ k = 1, \ldots, n \\
\chi_k - \Pi_{\geq 0}(v_k - \chi_k), \ k = 1, \ldots, n \\
\eta - \text{Prox}_{\rho, \|\cdot\|_1}(x_0 - \eta) \\
\xi_k - \Pi_{\Omega_k}(x_k - \xi_k), \ k = 1, \ldots, n
\end{pmatrix}.
\]

It is obvious from the expressions of \( \Pi_{\geq 0}(\cdot) \), \( \text{Prox}_{\rho, \|\cdot\|_1}(\cdot) \), and \( \Pi_{\Omega_k}(\cdot) \), \( k = 1, \ldots, n \) that the multifunction \( \mathcal{R} \) is piecewise polyhedral and so is its inverse mapping \( \mathcal{R}^{-1} \). Moreover, the following set is exactly the solution set to the KKT system associated with (10) and (11):

\[
\Omega_{\text{KKT}} := \{ w | \mathcal{R}(w) = 0 \}.
\]

For the rest of this paper, we always assume the following assumption, which can be guaranteed by the Slater condition (see e.g. [4]), holds:

**Assumption 2.** Suppose that the KKT solution set is nonempty, i.e. \( \Omega_{\text{KKT}} \neq \emptyset \).

It follows from [31, Propoision 1], we know that \( \mathcal{R}^{-1} \) is calm (see, e.g. [15, 41, 42]) at \((0, \bar{w}) \in \text{gph} \mathcal{R}^{-1}\). Furthermore, from e.g. [15, Theorem 3H.3], we know that \( \mathcal{R}^{-1} \) is calm at \((0, \bar{w}) \in \text{gph} \mathcal{R}^{-1}\) if and only if \( \mathcal{R} \) is metrically subregular at \((\bar{w}, 0) \in \text{gph} \mathcal{R} \), i.e., there exist positive scalars \( \kappa, r \) such that

\[
\text{dist}(w, \Omega_{\text{KKT}}) \leq \kappa \|\mathcal{R}(w)\|, \quad \forall w \in \{ w | \|w - \bar{w}\| \leq r \}.
\]

**Theorem 3.1.** Suppose that Assumption 1 holds. Let \( \{(w_p^l, w_d^l)\} \) be the sequence generated by Algorithm 1. Then, the following results hold:

(a): The sequence \( \{w_d^l\} \) converges to an optimal solution of problem (11), and the sequence \( \{w_p^l\} \) converges to an optimal solution of problem (10).

(b): The sequence \( \{(w_p^l, w_d^l)\} \) converges globally \( Q \)-linearly.

**Proof.** Without loss of generality, we assume that \( n = 2 \). Let

\[
\theta := [y_0; \lambda_0; y_1; \lambda_1; y_2; \lambda_2; r_1; r_2], \quad \zeta := [\eta; \xi_1; \xi_2; \chi_1; \chi_2],
\]

\[
f(\theta) = \frac{1}{2} \theta^T \Sigma \theta + c^T \theta, \quad \text{and} \quad \psi(\zeta) := \phi^*(\eta) + \mathbb{I}_{U_1}(\xi_1) + \mathbb{I}_{U_2}(\xi_2) + \mathbb{I}_{\geq 0}(\chi_1) + \mathbb{I}_{\geq 0}(\chi_2),
\]

where

\[
\Sigma = \text{Diag}(C_0, 0, C_1, 0, C_2, 0, 0, 0) \in \mathbb{S}^{3(p+1)+2m},
\]

\[
c = [0; 1; 0; 1; 0; 1; -h_1; -h_2] \in \mathbb{R}^{3(p+1)+2m}.
\]

Therefore, problem (11) can be rewritten into the following form:

\[
\min_{\theta, \zeta} f(\theta) + \psi(\zeta) \quad \text{s.t.} \quad A\theta + B\zeta = 0.
\]

(21)
where

\[
A := \begin{pmatrix}
C_0 & e & 0 & 0 & 0 & -T_1^T & 0 \\
0 & 0 & C_1 & e & 0 & -W_1^T & -T_2^T \\
0 & 0 & 0 & 0 & C_2 & e & 0 & -W_2^T \\
0 & 0 & 0 & 0 & 0 & 0 & I & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & I
\end{pmatrix},
\]

and \(B = \text{Diag}(I, I, -I, -I) \in \mathbb{S}^{3(p+1)+2m}\). In order to get the convergence result (a), from [8, Theorem 5.1], it is sufficient to show that

\[
\Sigma + \sigma A^T A + S \succ 0,
\]

where \(S := \text{Diag}(S_0, S_1, S_2, 0) \in \mathbb{S}^{3(p+1)+2m}\). By some elementary calculation, we can get

\[
\Sigma + \sigma A^T A + S = \begin{pmatrix}
\Theta_{11} & -\Theta_{12} \\
-\Theta_{12}^T & \Theta_{22}
\end{pmatrix} \in \mathbb{S}^{3(p+1)+2m},
\]

where \(\Theta_{11} := \text{Diag}(V_0, V_1, V_2) \in \mathbb{S}^{3(p+1)}, \Theta_{22} := \sigma (\Gamma \Gamma^T + I) \in \mathbb{S}^{2m}\), and

\[
\Theta_{12}^T := \sigma \Gamma \begin{pmatrix}
C_0 & e & 0 & 0 & 0 \\
0 & 0 & C_1 & e & 0 \\
0 & 0 & 0 & 0 & C_2 & e
\end{pmatrix} \in \mathbb{R}^{2m,3(p+1)},
\]

with

\[
\Gamma = \begin{pmatrix}
T_1 & W_1 & 0 \\
0 & T_2 & W_2
\end{pmatrix}.
\]

It is obvious that \(\Theta_{22}\) is positive definite. Under Assumption 1, matrix \(\Theta_{11}\) is also a positive definite matrix. By using Schur complement lemma, to get (22), it is sufficient to show that

\[
\Theta_{22} - \Theta_{12}^T \Theta_{11}^{-1} \Theta_{12} \succ 0.
\]

Elementary calculation shows that

\[
\Theta_{22} - \Theta_{12}^T \Theta_{11}^{-1} \Theta_{12} = \sigma I + \sigma \Gamma \begin{pmatrix}
\Xi_1 & \Xi_1 \\
\Xi_1 & \Xi_1
\end{pmatrix} \Gamma^T,
\]

where, for \(k = 0, 1, 2\),

\[
\Xi_k := I - \sigma (C_k e) (V_k)^{-1} \begin{pmatrix}
C_k \\
e_k^T
\end{pmatrix} \succeq 0.
\]

This implies that (23) holds. Therefore, result (a) holds. Result (b) holds directly from (a), (20) and [20, Theorem 2]. The proof is completed.

**Remark 2.** Note that the reformulated multi-block convex optimization problem (10) can also be solved by the two-block semi-proximal ADMM (sPADMM) [16]. However, the “most appropriate” semi-proximal terms in sPADMM are not easy to construct.
3.2. Parameters and stopping criteria. Let \( \mathcal{F}_P \) and \( \mathcal{F}_D \) be the feasible regions of (10) and (11), respectively. The essential objective functions (see [32]) of (10) and (11) are given by

\[
\bar{\vartheta}(w_p) := \begin{cases} \vartheta(w_p), & w_p \in \mathcal{F}_P, \\ +\infty, & \text{otherwise}, \end{cases} \quad \bar{\varphi}(w_d) := \begin{cases} \varphi(w_d), & w_p \in \mathcal{F}_D, \\ -\infty, & \text{otherwise}. \end{cases}
\]

Furthermore, denote set-valued functions \( \mathcal{T}_p := \partial \bar{\vartheta} \) and \( \mathcal{T}_d := \partial \bar{\varphi} \) (the notation “\( \partial \)” stands for the generalized subdifferential in the sense of Clarke [10]). Under Assumption 2, it holds that both \( \mathcal{T}_p \) and \( \mathcal{T}_d \) are piecewise polyhedral. Let \( (\bar{w}_d, \bar{w}_p) \in \Omega_{KKT} \). From [31, Proposition 1] and [15, Theorem 3H.3], we know that there exist positive scalars \( \kappa'_1 \) and \( \kappa'_2 \) such that

\[
\text{dist}(w_p, \Omega_P) \leq \kappa'_1 \text{dist}(0, \mathcal{T}_p(w_p)), \quad \forall w_p \in \{ w_p : \| w_p - \bar{w}_p \| \leq r_1 \} \tag{24} \\
\text{dist}(w_d, \Omega_D) \leq \kappa'_2 \text{dist}(0, \mathcal{T}_d(w_d)), \quad \forall w_d \in \{ w_d : \| w_d - \bar{w}_d \| \leq r_2 \} \tag{25}
\]

where \( \Omega_P \) and \( \Omega_D \) are the optimal solution sets of (10) and (11), respectively.

**Proposition 1.** Let \( \{(w^l_p, w^l_d)\} \) be the sequence generated by Algorithm 1. Then, there exist integer \( \bar{l} > 0 \), positive scalars \( \kappa_1 \) and \( \kappa_2 \) such that for all \( l \geq \bar{l} \),

\[
\vartheta(w^l_p) - \varphi(w^l_d) \geq \kappa_1 \text{dist}^2(w^l_p, \Omega_P) + \kappa_2 \text{dist}^2(w^l_d, \Omega_D). \tag{26}
\]

**Proof.** Suppose that the sequence \( \{(w^l_p, w^l_d)\} \) converges to some \( (\bar{w}_d, \bar{w}_p) \in \Omega_{KKT} \). From (24), (25) and [1, Theorem 3.3], one has that, for all \( l \geq \bar{l} \),

\[
\vartheta(w^l_p) \geq \vartheta(\bar{w}_p) + \kappa_1 \text{dist}^2(w^l_p, \Omega_P), \\
-\varphi(w^l_d) \geq -\varphi(\bar{w}_d) + \kappa_2 \text{dist}^2(w^l_d, \Omega_D).
\]

Add the above two inequalities together, and we can get (26) from strong duality theorem. The proof is completed. \( \square \)

In order to measure the accuracy of approximated optimal solution, we first define the following notations,

\[
\text{Res}_{D_1} = \max \left\{ \frac{\| C_0 y_0 + e \lambda_0 - T^T \mu - \xi \|}{1 + \| C_0 \|}, \frac{\| C_n y_0 + e \lambda_n - W^T \mu - \xi \|}{1 + \| C_n \|} \right\}, \\
\text{Res}_{D_2} = \max_{k = 1, \ldots, n-1} \left\{ \frac{\| C_k y_k + e \lambda_k - W^T \mu_k - T^T \mu_{k+1} + \xi \|}{1 + \| C_k \| + \| W_k \| + \| T_k \|} \right\}, \\
\text{Res}_{D_3} = \max_{k = 1, \ldots, n-1} \left\{ \frac{\| \mu_k - \chi_k \|}{1 + \| \mu_k \|} \right\}, \\
\text{Res}_{P_1} = \max_{k = 0, \ldots, n-1} \left\{ \| e^T x_k - 1 \| \right\}, \quad \text{Res}_{P_2} = \max_{k = 1, \ldots, n} \left\{ \frac{\| W_k x_k + T_k x_{k-1} - v_k - h_k \|}{1 + \| h_k \|} \right\}, \\
\text{Dobj} = -\frac{1}{2} \sum_{k = 0}^n \| y_k C_k y_k + \lambda_k \| + \sum_{k = 1}^n h_k^T \mu_k, \\
\text{Pobj} = \frac{1}{2} x_0^T C_0 x_0 + \rho \| x_0 \| + \sum_{k = 1}^n \left( \frac{1}{2} x_k^T C_k x_k + h_k^* (-x_k) \right). 
\]

Furthermore, define

\[
\text{Res}_D = \max \{ \text{Res}_{D_1}, \text{Res}_{D_2}, \text{Res}_{D_3} \}, \\
\text{Res}_P = \max \{ \text{Res}_{P_1}, \text{Res}_{P_2} \}, \\
\text{gap} = \frac{\| \text{Pobj} - \text{Dobj} \|}{1 + \| \text{Pobj} \| + \| \text{Dobj} \|}.
\]
From Proposition 1, it's reasonable to terminate Algorithm 1 when
\[ \max\{\text{Res}_P, \text{Res}_D, \text{gap}\} < \varepsilon, \]
or the maximum number of iterations is achieved.

In order to improve the convergence speed of the algorithm, we adapt the following rule, which is greatly inspired by \[24, 40\], to update penalty parameter \( \sigma \): given \( \varsigma > 0 \) and \( \varpi \geq 1 \),

\[ \sigma_{k+1} = \begin{cases} 
\min(\varpi \sigma_k, 10^6), & \text{if } \text{Res}_P/\text{Res}_D < \varsigma, \\
\max(\sigma_k/\varpi, 10^{-6}), & \text{if } \text{Res}_P/\text{Res}_D > 1/\varsigma, \\
\sigma_k, & \text{otherwise}.
\end{cases} \] (27)

In all our numerical experiments, we choose \( \tau = 1.618 \), \( \varepsilon = 10^{-5} \), the maximum number of iterations is set to 5,000, and \( \varsigma = 0.1 \).

4. Numerical Experiments. This section aims to evaluate the numerical performance of sGS-sPADMM for solving the robust multi-period mean-variance portfolio selection model (1). We test the efficiency of Algorithm 1 for solving (11) by comparing its performance with the directly extended ADMM (see Appendix A.1) and a two-block ADMM with semi-proximal terms (sPADMM, see Appendix A.2).

Though the directly extended ADMM may not be convergent \[7\], it can also be viewed as a benchmark for solving multi-block convex optimization problems by first-order algorithms. In our numerical experiment, these three algorithms use the same parameters and stopping criteria that described in subsection 3.2. All computational results are obtained by running Matlab R2015b on Mac OS X 10.10.5 (2.9 GHz Intel Core i5 16GB 1867 MHz DDR3).

4.1. Performance on synthetic data. Given assets number \( p \), we generated the matrix \( C_0, C_1, C_2 \) randomly as

\[ \text{data} = -1 + 2 \times \text{rand}(60, p), \quad \tilde{C} = \text{cov(data)}, \quad \text{data}_1 = -0.2 + 0.4 \times \text{rand}(30, p), \]
\[ \text{data}_2 = -0.2 + 0.4 \times \text{rand}(30, p), \quad \text{data}_3 = -0.2 + 0.4 \times \text{rand}(30, p), \]
\[ C_0 = \tilde{C} + \text{cov(data}_1), \quad C_1 = \tilde{C} + \text{cov(data}_1), \quad C_2 = \tilde{C} + \text{cov(data}_1). \] (28)

For \( k = 1, 2 \), the uncertainty set \( U_k := \{ \xi | \|\xi\|_\infty \leq \alpha \} \) and \( \mathcal{X}(x_{k-1}) := \{ x_k | |x_k - x_{k-1}| \leq 0.1 \} \). Set the parameters \( \alpha = 0.5 \) and \( \varpi = 2 \) in (27). By using these synthetic data, we compare our sGS-sPADMM with directly extended ADMM. In order to even out the measurement, we repeat both algorithms 10 times for each choice of \( p \) and report average results over these 10 instances in Table 1 and Table 2.

Table 1 and Table 2 report the detailed numerical results of sGS-sPADMM and directly extended ADMM (ADMM-d) for solving two-period (\( n=1 \)) and three period (\( n=2 \)) portfolio problem, respectively. From these two tables, we can see that our sGS-sPADMM is not only a algorithm with a guaranteed convergence, but also practically efficient for synthetic data generated as (28).

4.2. Performance on “Standard & Poor’s 500” stocks. In this part, all the numerical experiments are conducted on data taken from “Standard & Poor’s 500” (SPX) stock market index. This dataset consists of 370 stocks. From the proof of Theorem 3.1, we know that problem (11) can be viewed as a two-block convex optimization problem, and the sPADMM (Appendix A.2) can also be applied.
Table 1. Comparison between the performance of sGS-sPADMM and directly extended ADMM on synthetic data. \( n = 1, \alpha = 0.5, \rho = 0.005 \lambda_{\text{max}}(C_0) \). All the results are averaged over 10 instances.

| \( P \) | sGS-sPADMM | | | ADMM-d | | |
|---|---|---|---|---|---|---|
| \( P \) | Iter | Time (s) | Res | | Iter | Time (s) | Res |
| 200 | 348.8 | 1.6 | 9.95e-06 | | 345.9 | 1.6 | 9.91e-06 |
| 300 | 187.3 | 1.6 | 9.76e-06 | | 325.9 | 2.6 | 9.63e-06 |
| 500 | 175.6 | 4.8 | 9.60e-06 | | 275.3 | 7.2 | 9.45e-06 |
| 1000 | 250.2 | 41.5 | 9.25e-06 | | 296.6 | 48.3 | 9.43e-06 |
| 1500 | 346.7 | 143.2 | 9.66e-06 | | 469.4 | 193.8 | 9.42e-06 |
| 2000 | 406.0 | 300.6 | 4.97e-06 | | 585.2 | 428.0 | 9.47e-06 |

Therefore, in this part we aim to compare the numerical performance of the sGS-sPADMM with directly extended ADMM and sPADMM. Besides, we also evaluate the effectiveness of the robust two-period mean-variance portfolio selection model.
by taking the widely used global minimum-variance portfolio selection model (see e.g. [21]) as a benchmark.

We first focus on testing the performance the these three algorithms for solving two-period portfolio selection problems by using the daily returns in 2006. In order to determine parameters $C_0$ and $C_1$, we use the past 60 daily returns. For example, in the period of “Jan.”, $C_0$ and $C_1$ are the covariance matrices generated by the daily returns from 6 Oct. 2005 to 29 Dec. 2005, and from 1 Nov. 2005 to 27 Jan. 2006, respectively. In Table 3, we report the detailed numerical results for sGS-sPADMM, directly extended ADMM and sPADMM in solving the multi-period portfolio selection problem with different $\rho$. From this table, we can see that our sGS-sPADMM outperforms the other two algorithms for all six different datasets. The efficient frontiers obtained from robust multi-period mean-variance and global minimum-variance portfolio selection models are presented in Figure 1. Note that there are numbers of measures to characterize the risk level, such as the standard deviation [28], mean-absolute deviation [23], conditional value-at-risk [34], probabilistic risk measure [38]. Since the model proposed in this paper is based on the Markowitz mean-variance model, we use the standard deviation of portfolio’s return to represent the level of risk. We can see from Figure 1 that strategies generated by our proposed robust two-period mean-variance portfolios considerably reduce risk with respect to the one generated by two-period minimum-variance model (the uncertainty set $U = \{0\}$).

**Table 3.** Comparison among the performance of sGS-sPADMM, directly extended ADMM, and sPADMM for solving two-period portfolio selection problem with different $\rho$. The parameter $\varpi = 5$.

| Period | $\rho$ | mzn | sGS-sPADMM | Iter | Time (s) | Res | ADMM-d | Iter | Time (s) | Res | sPADMM | Iter | Time (s) | Res |
|--------|-------|-----|-------------|------|----------|-----|---------|------|----------|-----|---------|------|----------|-----|
| Jan.   | 1.0e-06 | 272 | 284 | 4.0 | 9.91e-06 | 283 | 4.1 | 9.98e-06 | 539 | 31.6 | 1.00e-05 |
|       | 5.0e-06 | 171 | 248 | 3.3 | 9.88e-06 | 373 | 4.5 | 1.00e-05 | 715 | 42.1 | 9.99e-06 |
|       | 1.0e-05 | 112 | 246 | 3.3 | 9.46e-06 | 500 | 6.2 | 9.98e-06 | 850 | 52.0 | 9.93e-06 |
| Feb.   | 1.0e-06 | 288 | 249 | 3.4 | 9.69e-06 | 301 | 3.7 | 9.97e-06 | 464 | 26.6 | 9.99e-06 |
|       | 5.0e-06 | 173 | 290 | 3.3 | 9.48e-06 | 321 | 4.0 | 9.98e-06 | 568 | 32.2 | 9.95e-06 |
|       | 1.0e-05 | 125 | 258 | 3.6 | 9.03e-06 | 373 | 4.7 | 9.96e-06 | 701 | 40.2 | 9.98e-06 |
| Mar.   | 1.0e-06 | 269 | 223 | 3.1 | 9.91e-06 | 253 | 3.2 | 9.97e-06 | 484 | 27.8 | 9.88e-06 |
|       | 5.0e-06 | 153 | 229 | 3.1 | 9.88e-06 | 310 | 3.8 | 9.93e-06 | 596 | 34.0 | 9.98e-06 |
|       | 1.0e-05 | 122 | 238 | 3.3 | 9.43e-06 | 358 | 4.5 | 9.88e-06 | 594 | 34.1 | 9.96e-06 |
| Apr.   | 1.0e-06 | 272 | 272 | 3.7 | 9.20e-06 | 316 | 4.0 | 9.98e-06 | 573 | 33.1 | 9.99e-06 |
|       | 5.0e-06 | 144 | 285 | 3.8 | 9.84e-06 | 477 | 5.8 | 9.94e-06 | 902 | 51.7 | 9.88e-06 |
|       | 1.0e-05 | 108 | 264 | 3.7 | 9.77e-06 | 548 | 6.9 | 9.87e-06 | 915 | 52.4 | 9.91e-06 |
| May    | 1.0e-06 | 244 | 290 | 3.9 | 9.80e-06 | 510 | 6.3 | 9.84e-06 | 991 | 53.6 | 9.79e-06 |
|       | 5.0e-06 | 127 | 357 | 4.5 | 9.70e-06 | 635 | 7.9 | 9.99e-06 | 1110 | 63.8 | 9.90e-06 |
| Jun.   | 1.0e-06 | 274 | 281 | 3.9 | 9.51e-06 | 448 | 5.6 | 9.73e-06 | 805 | 45.7 | 9.99e-06 |
|       | 5.0e-06 | 151 | 239 | 3.2 | 9.95e-06 | 478 | 4.9 | 9.95e-06 | 911 | 51.6 | 9.78e-06 |
|       | 1.0e-05 | 105 | 386 | 5.2 | 9.64e-06 | 509 | 6.4 | 9.97e-06 | 983 | 55.5 | 9.96e-06 |

Furthermore, we also evaluate the numerical performance by using different numbers of periods and data in different years. The results are presented in Tables 4 and.

---

1 From Yahoo finance, we can observe that the S&P 500 index was stable from Jan. to Apr. 2006, down from May to July 2006, and up from Aug. to Dec. 2006. Therefore, we use the daily returns in 2006 in this section. We have tried data from the second half of 2006, the numerical performance are similar to the first half of 2006 and the effectiveness of our proposed robust two-period mean-variance model is also superior to the minimum-variance portfolio selection model, thus, we report the results by using the first half of 2006.
Figure 1. Two-period efficient frontiers obtained from robust multi-period mean-variance (RMV) and global minimum-variance (MV) portfolio selection models. $\rho = 1e^{-5}$.

5. In these two tables, we report the number of iterations, runtime and residual of three different first-order algorithms for finding optimal portfolios. We can see from these two tables that both the iteration number and runtime of our sGS-sPADMM outperform the other two algorithms for all cases. By observing the performance of three algorithms in Tables 4 and 5, we evaluate the performance of our sGS-sPADMM and directly extended ADMM on larger numbers of periods. This result is displayed in Figure 2.

To summarize, our sGS-sPADMM is an efficient first-order algorithm for solving robust multi-period mean-variance portfolio selection problem (1) on SPX dataset.

Table 4. Comparison among the performance of sGS-sPADMM, directly extended ADMM, and sPADMM for solving three/four-period portfolio selection problem with different $\rho$. Based on 2005 data, the parameter $\varpi = 5$. “nnz” stands for the number of active positions of portfolio.

| n  | $\rho$ | nnz | sGS-sPADMM | ADMM-d | sPADMM |
|----|-------|-----|------------|--------|--------|
|    |       |     | Iter | Time (s) | Res   | Iter | Time (s) | Res   |
| 2  | 1.0e-06 | 270 | 475  | 11.5 | 9.49e-06 | 1189 | 27.6 | 9.93e-06 |
|    | 5.0e-06 | 128 | 663  | 14.8 | 9.86e-06 | 1025 | 21.9 | 9.57e-06 |
|    | 1.0e-05 | 109 | 585  | 14.2 | 9.99e-06 | 1356 | 30.4 | 9.29e-06 |
| 3  | 1.0e-06 | 254 | 623  | 20.6 | 9.13e-06 | 1327 | 40.4 | 9.74e-06 |
|    | 5.0e-06 | 125 | 939  | 31.2 | 9.68e-06 | 1522 | 45.5 | 9.57e-06 |
|    | 1.0e-05 | 103 | 675  | 23.6 | 9.49e-06 | 1484 | 42.5 | 1.00e-05 |
|    | 1.0e-06 | 1319 | 694.8 | 9.97e-06 |
Table 5. Comparison among the performance of sGS-sPADMM, directly extended ADMM, and sPADMM for solving three/four-period portfolio selection problem with different $\rho$. Based on 2006 data, parameter $\varpi = 5$. “nnz” stands for the number of active positions of portfolio.

| $n$ | $\rho$ | Iter | Time (s) | Res     | Iter | Time (s) | Res     | Iter | Time (s) | Res     |
|-----|--------|------|----------|---------|------|----------|---------|------|----------|---------|
| 2   | 1.0e-06| 237  | 11.3     | 9.57e-06| 1146 | 25.5     | 9.89e-06| 1924 | 499.4 | 9.83e-06|
|     | 5.0e-06| 138  | 14.8     | 9.82e-06| 1297 | 26.0     | 9.99e-06| 1684 | 338.2 | 1.0e-05  |
|     | 1.0e-05| 120  | 13.2     | 9.86e-06| 1213 | 24.5     | 9.99e-06| 1569 | 331.5 | 9.92e-06 |
| 3   | 1.0e-06| 225  | 22.7     | 9.87e-06| 810  | 23.2     | 9.94e-06| 1895 | 979.8 | 1.0e-05  |
|     | 5.0e-06| 133  | 25.5     | 9.74e-06| 1057 | 29.0     | 9.97e-06| 1828 | 925.8 | 9.89e-06 |
|     | 1.0e-05| 120  | 23.6     | 9.75e-06| 1377 | 39.5     | 9.99e-06| 2292 | 1161.0 | 1.0e-05  |

Figure 2. Comparison between the performance of sGS-sPADMM and directly extended ADMM on different numbers of periods.

5. Conclusion. In this paper, we considered a class of robust multi-period mean-variance portfolio selection problems. Under the assumption that the estimation error of sample mean belongs to a polyhedron, this class of problems can be equivalently reformulated as multi-block convex optimization problems. A symmetric Gauss-Seidel based semi-proximal alternating direction method of multipliers (sGS-sPADMM) was introduced to solve our proposed model. We proved that this algorithm can achieve a Q-linear rate of convergence. Numerical results showed that the robust multi-period mean-variance portfolio selection model is effective and can be efficiently solved by using the sGS-sPADMM.

Acknowledgments. The authors would like to thank the anonymous reviewers for their valuable comments and suggestions.

Appendix A. Appendix.

A.1. Directly extended ADMM. The directly extended ADMM tested in this paper takes the following iteration scheme:
where $\Sigma := S$ where

From the proof of Theorem 3.1, we know that the sequence \( \exp \) is well defined, \( \theta \)

To know more efficient techniques for solving the above linear equation, we refer readers to [24, Section 3.3].

A.2. Semi-proximal ADMM. Note that dual problem (11) can be equivalently rewritten as (21). And consequently, the augmented Lagrangian function (12) can also be rewritten as

\[
\mathcal{L}_\sigma(\theta, \zeta, \Upsilon) := f(\theta) + \psi(\zeta) - \langle A\theta + B\zeta, \Upsilon \rangle + \frac{\sigma}{2} ||A\theta + B\zeta||^2,
\]

where $\Upsilon := [x_0; x_1; \ldots; x_n; v_1; \ldots; v_n]$.

The iteration scheme of semi-proximal ADMM ([16]) is described as follows

\[
\begin{align*}
\theta^{t+1} &= \min_\theta \mathcal{L}_\sigma(\theta, \zeta^t, \Upsilon^t) + \frac{1}{2} \|\theta - \theta^t\|_{S^t}^2, \\
\zeta^{t+1} &= \min_\zeta \mathcal{L}_\sigma(\theta^{t+1}, \zeta, \Upsilon^t), \\
\Upsilon^{t+1} &= \Upsilon^t - \tau(A\theta^{t+1} + B\zeta^{t+1}),
\end{align*}
\]

where $S^t := \text{Diag}(S, 0)$ and $S$ is defined by (16). Suppose that Assumption 1 holds. From the proof of Theorem 3.1, we know that the sequence \( \{(\theta^t, \zeta^t, \Upsilon^t)\} \) generated by sPADMM is well defined, \( \{(\theta^t, \zeta^t)\} \) converges to an optimal solution of (11) and \( \{\Upsilon^t\} \) converges to an optimal solution of (10).

From the point of view of the experience, the step of updating $\theta$ would be most expensive in the algorithm. Specifically, $\theta^{t+1}$ is the solution of the following linear equation:

\[
(\Sigma + \sigma A^T A + S^t) \theta = A^T \Upsilon + S^t \theta^t - \sigma A^T B \zeta^t - c.
\]

(30)

To know more efficient techniques for solving the above linear equation, we refer readers to [24, Section 3.3].

REFERENCES

[1] F. J. Aragón Artacho and M. H. Geoffroy, Characterization of metric regularity of subdifferentials, Journal of Convex Analysis, 15 (2008), 365–380.
[2] A. Ben-Tal, T. Margalit and A. Nemirovski, Robust modeling of multi-stage portfolio problems, In High Performance Optimization, 33 (2000), 303–328.
[3] D. Bertsimas and M. Sim, Tractable approximations to robust conic optimization problems, Mathematical Programming, 107 (2006), 5–36.
[4] J. Borwein and A. S. Lewis, Convex Analysis and Nonlinear Optimization: Theory and Examples, Springer, New York, 2006.
[5] G. C. Calafiore, Multi-period portfolio optimization with linear control policies, Automatica, 44 (2008), 2463–2473.
[6] L. K. Chan, J. Karceski and J. Lakonishok, On portfolio optimization: Forecasting covariances and choosing the risk model, The Review of Financial Studies, 12 (1999), 937–974.
[7] C. Chen, B. He, Y. Ye and X. Yuan, The direct extension of ADMM for multi-block convex minimization problems is not necessarily convergent, Mathematical Programming, 155 (2016), 57–79.
[8] L. Chen, D. F. Sun and K.-C. Toh, An efficient inexact symmetric Gauss–Seidel based majorized ADMM for high-dimensional convex composite conic programming, Mathematical Programming, 161 (2017), 237–270.
[9] V. K. Chopra and W. T. Ziemba, The effect of errors in means, variances, and covariances on optimal portfolio choice, *The Journal of Portfolio Management*, 19 (1993), 6–11.

[10] F. H. Clarke, *Optimization and Nonsmooth Analysis*, volume 5. SIAM, 1990.

[11] L. Condat, Fast projection onto the simplex and the $\ell_1$ ball, *Mathematical Programming*, 158 (2016), 575–585.

[12] X. Cui, J. Gao, X. Li and D. Li, Optimal multi-period mean–variance policy under no-shorting constraint, *European Journal of Operational Research*, 234 (2014), 459–468.

[13] G. B. Dantzig and G. Infanger, Multi-stage stochastic linear programs for portfolio optimization, *Annals of Operations Research*, 45 (1993), 59–76.

[14] V. DeMiguel and F. J. Nogales, Portfolio selection with robust estimation, *Operations Research*, 57 (2009), 560–577.

[15] A. L. Dontchev and R. T. Rockafellar, *Implicit Functions and Solution Mappings*, Springer Monographs in Mathematics. Springer, Dordrecht, 2009.

[16] M. Fazel, T. K. Pong, D. F. Sun and P. Tseng, Hankel matrix rank minimization with applications to system identification and realization, *SIAM Journal on Matrix Analysis and Applications*, 34 (2013), 946–977.

[17] J. Gao and D. Li, Optimal cardinality constrained portfolio selection, *Operations Research*, 61 (2013), 745–761.

[18] X. Mei, V. DeMiguel and F. J. Nogales, Multiperiod portfolio optimization with multiple risky assets and general transaction costs, *Journal of Banking & Finance*, 69 (2016), 108–120.

[19] R. O. Michaud, The markowitz optimization enigma: Is ‘optimized’ optimal?, *Financial Analysts Journal*, 45 (1989), 31–42.

[20] S. M. Robinson, Some continuity properties of polyhedral multifunctions, *Mathematical Programming at Oberwolfach*, 14 (1981), 206–214.
[37] D. F. Sun, K.-C. Toh and L. Q. Yang, A convergent 3-block semiproximal alternating direction method of multipliers for conic programming with 4-type constraints, *SIAM Journal on Optimization*, 25 (2015), 882–915.

[38] Y. Sun, G. Aw, K. L. Teo and G. Zhou, Portfolio optimization using a new probabilistic risk measure, *Journal of Industrial and Management Optimization*, 11 (2015), 1275–1283.

[39] R. J.-B. Wets, Stochastic programs with fixed recourse: The equivalent deterministic program, *SIAM Review*, 16 (1974), 309–339.

[40] J. Yang, D. F. Sun and K.-C. Toh, A proximal point algorithm for log-determinant optimization with group lasso regularization, *SIAM Journal on Optimization*, 23 (2013), 857–893.

[41] J. J. Ye and X. Y. Ye, Necessary optimality conditions for optimization problems with variational inequality constraints, *Mathematics of Operations Research*, 22 (1997), 977–997.

[42] J. J. Ye and J. Zhang, Enhanced Karush-Kuhn-Tucker condition and weaker constraint qualifications, *Mathematical Programming*, 139 (2013), 353–381.

[43] J. Zhai and M. Bai, Mean-risk model for uncertain portfolio selection with background risk, *Journal of Computational and Applied Mathematics*, 330 (2018), 59–69.

[44] Y. Zhang, X. Li and S. Guo, Portfolio selection problems with markowitz’s mean–variance framework: A review of literature, *Fuzzy Optimization and Decision Making*, 17 (2018), 125–158.

Received May 2018; revised August 2018.

*E-mail address: ningzhang_2008@yeah.net*