Asset Allocation under the Basel Accord Risk Measures

Zaiwen Wen†  Xianhua Peng‡  Xin Liu§  Xiaodi Bai¶  Xiaoling Sun∥

First version January 2013

Abstract

Financial institutions are now required by regulators to meet much more stringent capital requirements than they were before the recent financial crisis; in particular, the capital requirement for a large bank’s trading book under the Basel 2.5 Accord more than double that under the Basel II Accord. The dramatic increase of capital requirements render it necessary for banks to take into account the stringent constraint of capital requirement when they make asset allocation decisions. In this paper, we propose asset allocation models which incorporate the capital requirements under the Basel 2.5 risk measure and the most recently proposed Basel III risk measure. A unified method is developed based on the alternating direction augmented Lagrangian method in the case of finite discrete distributions. Although our targeted model is essentially equivalent to mixed-integer programming problems, each step of our algorithm only consists of the minimization of convex quadratic programming subproblems which either have

---

*We are grateful to Steven Kou for his insightful comments to the paper. Zaiwen Wen was partially supported by the NSFC grant 11101274, research fund for the Doctoral Program of Higher Education of China, Shanghai Pujiang Program 12PJ1404800. Xianhua Peng was partially supported by a grant from School-Based-Initiatives of HKUST (Grant No. SBII1SC03) and Hong Kong RGC Direct Allocation Grant (Project No. DAG12SC05-3). Xin Liu was partially supported by the NSFC grants 10831006 and 11101409. Xiaodi Bai and Xiaoling Sun were supported by the NSFC grant 10971034 and the Joint NSFC/RGC grant 71061160506.

†Department of Mathematics, MOE-LSC and Institute of Natural Sciences, Shanghai Jiaotong University, Shanghai 200240, China, zw2109@sjtu.edu.cn.
‡Department of Mathematics, Hong Kong University of Science and Technology, Hong Kong, maxhpeng@ust.hk.
§Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing 100080, China, liuxin@lsec.cc.ac.cn.
¶Department of Management Science, School of Management, Fudan University, Shanghai 200433, China, xdbai@fudan.edu.cn.
∥Department of Management Science, School of Management, Fudan University, Shanghai 200433, China, xls@fudan.edu.cn.
closed-form solutions or can be relatively easily solved; hence, the method is able
to find good quality suboptimal solutions efficiently. Numerical experiments on both
simulated and real market data show that the method compares favorably with other
existing approximation methods.

Keywords: Asset Allocation, Basel Accords, Capital Requirements, Value-at-Risk,
Expected Shortfall, Alternating Direction Methods

1 Introduction

One of the major aftermaths of the recent financial crisis starting in 2007 is that financial
institutions are now required to meet much more stringent capital requirements than they
were before the crisis. The dramatic increase of capital requirements has been imposed in
the Basel Accords, which have been undergoing significant revisions since the inception
of the financial crisis. The framework of the latest version of the Basel Accord, the Basel
III Accord (Basel Committee on Banking Supervision, 2010), was announced in December
2010 and will be implemented in major countries, including the United States (Board of

In particular, the capital requirements for banks’ trading books, which are calculated
by the Basel Accord risk measure for the trading book, have been increased substantially.
Before the 2007 financial crisis, the Basel II risk measure (Basel Committee on Banking
Supervision, 2006) was used in the calculation. During the crisis, it was found that the
Basel II risk measure had serious drawbacks such as being procyclical and not being con-
servative enough. In response to the financial crisis, the Basel committee revised the Basel
II market risk framework and imposed the “Basel 2.5” risk measure (Basel Committee on
Banking Supervision, 2009) in July 2009. It has been estimated that the Basel 2.5 capi-
tal requirement for a large bank’s trading book on average more than double the capital
requirement under Basel II (Basel Committee on Banking Supervision, 2012, p.11).

The substantial increase in the capital requirements for the trading book makes it more
important for banks to take into account the constraint of capital requirements when they
construct investment portfolio. In this paper, we address this issue by proposing a new
asset allocation model that incorporates the capital requirement imposed by Basal Accords.
More precisely, we propose the “mean-\(\rho\)-Basel” asset allocation model, where \(\rho\) denotes
the risk measure used for measuring the risk of the investment portfolio, such as variance,
value-at-risk (VaR), or conditional value-at-risk (CVaR), etc.; \(\rho\) can be freely chosen by
the portfolio manager; and “Basel” denotes the constraint that the regulatory capital of the
portfolio calculated by the Basel Accord risk measure should not exceed a certain upper limit.

The complexity of the Basel Accord risk measures for calculating the capital requirements poses a challenge to solving the proposed “mean-\(\rho\)-Basel” asset allocation model. Basel Accords use VaR or CVaR with scenario analysis as the risk measure for calculating the capital requirements for a bank’s trading book. A scenario refers to a specific economic regime such as an economic boom and a financial crisis; scenario analysis is the approach to analyze the behavior of the random loss under different scenarios. The Basel II risk measure involves the calculation of VaR under 60 different scenarios. The Basel 2.5 risk measure involves the calculation of VaR under 120 scenarios, including 60 stressed scenarios. Most recently, in May 2012, the Basel Committee released a consultative document (Basel Committee on Banking Supervision, 2012) that presents the initial policy proposal of a new risk measure to replace the Basel 2.5 risk measure for the trading book; the new risk measure involves the calculation of CVaR under stressed scenarios. Currently, this new proposal is still under discussion and has not been finalized. Apparently it is beyond the scope of this paper to discuss whether the newly proposed risk measure is better than the Basel 2.5 risk measure; instead, we will consider both the Basel 2.5 and the newly proposed Basel risk measure in the mean-\(\rho\)-Basel model. See Section 2.2 for details about Basel Accord risk measures.

There have been a lot of studies on the single-period portfolio selection model of “mean-\(\rho\)”, where \(\rho\) is the measure of portfolio risk such as variance, VaR, or CVaR. For recent development of the mean-variance portfolio selection models and algorithms, see, e.g., Chairawongse et al. (2012). In a recent interesting paper, Lim, Shanthikumar, and Vahn (2011) evaluated CVaR as the risk measure in data-driven portfolio optimization and show that portfolios obtained by solving mean-CVaR problems are unreliable due to estimation errors of CVaR and/or the mean asset returns. To address the issue of estimation risk, Karoui, Lim, and Vahn (2011) introduced a new approach called performance-based regularization to data-driven mean-CVaR portfolio optimization problem. Rockafellar and Uryasev (2002) developed a method to reduce the data-driven mean-CVaR portfolio selection problem to a linear programming (LP) problem. The problem of mean-VaR is more difficult than mean-CVaR because of the nonconvexity of the problem. Softwares such as CPLEX can be used to solve such problems of small-to-medium size. Bai et al. (2012) proposed a penalty decomposition method for probabilistically constrained convex programs which include the mean-VaR problem as a special case.
It seems more challenging to solve the mean-$\rho$-Basel model than the mean-$\rho$ model because of the complexity of the Basel Accords that involve multiple VaRs or CVaRs under different scenarios. In this paper, we develop an unified and computationally efficient method to solve the mean-$\rho$-Basel problem. The method is based on the alternating direction method (ADM), or, equivalently, block coordinate descent scheme. The method reduces the original problem to convex quadratic programming subproblems that may have closed-form solutions if certain variables are fixed. Hence, the method is capable of solving large scale problems.

The proposed method also applies to the mean-$\rho$ problem such as mean-VaR and mean-CVaR problem. In particular, the method applies to the "mean-Basel" problem, where the Basel Accord risk measures are used as the risk measures for quantifying the risk of the portfolio. The Basel Accord risk measures involve multiple VaR or CVaR under different scenarios, which essentially corresponds to different models or distributions of the asset returns. Hence, using the Basel Accord risk measures, or, more generally, VaR or CVaR with scenario analysis, as the portfolio risk measure provides a way to address the problem of model uncertainty.

In summary, the main contribution of the paper is two-fold. (i) We formulate new asset allocation models that involve Basel Accord risk measures, such as the "mean-$\rho$-Basel" and the "mean-Basel" models. To the best of our knowledge, there has been no literature on asset allocation using the Basel Accord risk measures. (ii) We propose an efficient alternating direction method (ADM) for finding suboptimal solutions of the asset allocation models. We also establish the convergence properties of the method under mild conditions. Although there is no theoretical guarantee that the method will converge to the global solution because of the nonconvexity of the problem, numerical experiments on both simulated and real market data show that the method can identify suboptimal solutions that can often be much better than the approximate solutions of the mixed-integer programming formulation computed by CPLEX in one hour.

This remainder of the paper is organized as follows. In Section 2, we review the definition and properties of Basel Accord risk measures as well as some relevant risk measures. In Section 3, we discuss the data-based Basel Accord risk measures and formulate the asset allocation models with the Basel Accord risk measures as the measure of regulatory capital or portfolio risk. In Section 4, we propose the augmented Lagrangian alternating direction methods for solving the asset allocation problem and provide convergence analysis. Section 5 provides the numerical results which demonstrate the accuracy and efficiency of the
proposed methods.

We use standard linear algebra notation in this paper. For a vector \( \mathbf{x} = (x_1, x_2, \ldots, x_n)^\top \in \mathbb{R}^n \), \( x_i^v \) denotes the \( v \)-th power of \( x_i \). Let \((i_1, i_2, \ldots, i_n)\) be a permutation of \( (1, 2, \ldots, n) \) such that \( x_{i_1} \leq x_{i_2} \leq \cdots \leq x_{i_n} \), then we define \( x^{(j)} := x_{i_j} \), \( j = 1, \ldots, n \); hence, both \( x^{(j)} \) and \( x_{i_j} \) mean the \( j \)-th smallest component of \( x \). For \( t \in \mathbb{R} \), \( \lceil t \rceil \) denotes the smallest integer larger than or equal to \( t \). For a vector \( \mathbf{x} \in \mathbb{R}^n \), we use \( x^{(k)} \) to denote the \( k \)-th iteration of \( x \) in the optimization algorithm.

2 Review of Relevant Risk Measures

There is a vast literature on the theoretical frameworks and concrete examples of risk measures; it is beyond the scope of this paper to discuss and compare different risk measures. In this section, we will only review the risk measures that will be used in the asset allocation problems in this paper.

2.1 Value-at-risk and Conditional Value-at-risk

Value-at-risk (VaR) is one of the most widely used risk measures in risk management. VaR is a quantile of the loss distribution at some pre-defined probability level. More precisely, let \( F_X(\cdot) \) be the distribution function of the random loss \( X \), then for a given \( \alpha \in (0, 1) \), VaR of \( X \) at level \( \alpha \) is defined as

\[
\text{VaR}_\alpha(X) := \inf\{x \mid F_X(x) \geq \alpha\} = F_X^{-1}(\alpha).
\]

(1)

Jorion (2007) provides comprehensive discussion on VaR and risk management.

Conditional value-at-risk (CVaR) proposed by Rockafellar and Uryasev (2002) is another prominent and widely used risk measure. For the random loss \( X \), the \( \alpha \)-tail distribution of \( X \) is defined by the distribution function:

\[
F_{\alpha,X}(x) := \begin{cases} 
0, & \text{for } x < \text{VaR}_\alpha(X), \\
\frac{F_X(x)-\alpha}{1-\alpha}, & \text{for } x \geq \text{VaR}_\alpha(X).
\end{cases}
\]

(2)

And then the CVaR at level \( \alpha \) of \( X \) is defined to be

\[
\text{CVaR}_\alpha(X) := \text{mean of the } \alpha \text{-tail distribution of } X \\
= \int_{-\infty}^{\infty} x dF_{\alpha,X}(x).
\]

(3)
Expected shortfall (ES) is a risk measure that is equivalent to CVaR and is independently introduced in Acerbi and Tasche (2002). CVaR or ES has the subadditivity property and belongs to the class of coherent risk measures (Artzner et al., 1999); VaR does not satisfy the subadditivity property in general and belongs to another class of risk measures called insurance risk measures (Wang, Young, and Panjer, 1997).

2.2 Basel Accords Risk Measures

Basel Accords use VaR or CVaR with scenario analysis as the risk measure for calculating the capital requirements for a bank’s trading book. A scenario refers to a specific economic regime such as an economic boom and a financial crisis. Scenario analysis is necessary because studies have shown that behavior of economic variables is substantially different at different regimes of economy (see, e.g., Hamilton, 1989). In particular, many economic variables exhibit dramatic changes in their behavior during financial crisis or when government monetary or fiscal policy undergo sudden changes (Sims and Zha, 2006). There are also evidence that the volatility and correlation among asset returns increase in economic downturn (see, e.g., Dai et al., 2007).

The Basel II Accord (Basel Committee on Banking Supervision, 2006) specifies that the capital charge for the trading book on any particular day \( t \) for banks using the internal models approach should be calculated by the formula

\[
    c_t = \max \left\{ \text{VaR}_{t-1}^\alpha(X), \frac{k}{60} \sum_{s=1}^{60} \text{VaR}_{s}^\alpha(X) \right\}, \tag{4}
\]

where \( X \) is the loss of the bank’s trading book; \( k \) is a constant that is no less than 3; \( \text{VaR}_{t-1}^\alpha(X) \) is the 10-day VaR of \( X \) at \( \alpha = 99\% \) confidence level calculated on day \( t - s \), \( s = 1, \ldots, 60 \). \( \text{VaR}_{s}^\alpha(X) \) is calculated under the scenario corresponding to information available on day \( t - s \). For example, \( \text{VaR}_{s}^\alpha(X) \) of a portfolio of equity options is calculated conditional on the value of the equity prices, equity volatilities, yield curves, etc. on the day \( t - s \). Therefore, the Basel II risk measure is a VaR with scenario analysis that involves 60 scenarios.

After the 2007 financial crisis, the Basel II risk measure (4) has been criticized for two reasons: (i) The risk measure (4) and other risk measures based on contemporaneous observations are procyclical, i.e., risk measurement obtained by those risk measures tend to be low in booms and high in crises, which is exactly opposite to the goal of effective regulation (Adrian and Brunnermeier, 2008). (ii) Banks’ actual losses during the financial
crisis were significantly higher than the capital requirements calculated by the risk measure (4).

In response to the financial crisis, the Basel Committee revised the Basel II market risk framework and replaced the Basel II risk measure by the “Basel 2.5” risk measure (Basel Committee on Banking Supervision, 2009) in July 2009. The Basel 2.5 risk measure for calculating the capital requirement for market risk is defined by

\[
c_t = \max \left\{ \text{VaR}_{\alpha,t-1}(X), \frac{k}{60} \sum_{s=1}^{60} \text{VaR}_{\alpha,t-s}(X) \right\} \\
+ \max \left\{ s\text{VaR}_{\alpha,t-1}(X), \frac{\ell}{60} \sum_{s=1}^{60} s\text{VaR}_{\alpha,t-s}(X) \right\},
\]

where \( \text{VaR}_{\alpha,t-s}(X) \) is the same as that in (4); \( s\text{VaR}_{\alpha,t-s}(X) \) is called the stressed VaR of \( X \) on day \( t-s \) at confidence level \( \alpha = 99\% \), which is calculated under the scenario that the financial market is under significant stress such as the one that happened during the period from 2007 to 2008; \( k \) and \( \ell \) are constants no less than 3. The additional capital requirements based on stressed VaR help to reduce the procyclicality of the original risk measure (4) and to significantly increase the capital requirements.

In May 2012, the Basel Committee released a consultative document (Basel Committee on Banking Supervision, 2012) that presents the initial policy proposal regarding the Basel Committee’s fundamental review of the trading book capital requirements. In particular, the Committee proposes to replace the Basel 2.5 risk measure by a new risk measure; the new risk measure uses CVaR (or, equivalently, ES) instead of VaR for calculating capital requirements. More precisely, under the new risk measure, the capital requirement for a group of trading desks that share similar major risk factors, such as equity, credit, interest rate, and currency, etc., is defined as the CVaR of the loss that may be incurred by the group of trading desks; for instance, an equity trading desk and an equity option trading desk will be grouped together for the purpose of calculating regulatory capital. The CVaR should be calculated under stressed scenarios rather than under the current market condition. Currently, this proposed new risk measure is under discussion and it is not clear yet whether it is going to be the final version of the Basel III risk measure. In addition, although the proposed new risk measure involves the calculation of CVaR under only stressed scenarios, it might be desirable to include the calculation of CVaR under current market conditions, as in Basel 2.5. Hence, with some abuse of the terminology and for the sake of notational convenience, we will call the following risk measure the “Basel III” risk measure, which is
obtained by replacing VaR in the Basel 2.5 risk measure by CVaR and is defined as
\[ c_t = \max \left\{ \text{CVaR}_{\alpha,t-1}(X), \frac{k}{60} \sum_{s=1}^{60} \text{CVaR}_{\alpha,t-s}(X) \right\} \]
\[ + \max \left\{ \text{sCVaR}_{\alpha,t-1}(X), \frac{\ell}{60} \sum_{s=1}^{60} \text{sCVaR}_{\alpha,t-s}(X) \right\}, \]
where \( X \) is loss at the trading desk; \( \text{CVaR}_{\alpha,t-s}(X) \) is the CVaR of \( X \) calculated on day \( t-s \); and \( \text{sCVaR} \) is the stressed CVaR that should be calculated under stressed scenarios.

### 3 Asset Allocation Models Incorporating Basel Accords Risk Measures

Consider a portfolio composed of \( d \) assets and let \( u = (u_1, u_2, \ldots, u_d)^T \in \mathbb{R}^d \) denote the portfolio weights of these assets, which are the percentage of initial wealth invested in the assets. Let \( R = (R_1, R_2, \ldots, R_d)^T \in \mathbb{R}^d \) be the random vector of simple returns of these assets over a specified time horizon, e.g., one day. Then the simple return of the portfolio is \( R^T u \) and the loss of the portfolio (per $1 of investment) is \(-R^T u\). Let \( \mu \in \mathbb{R}^d \) be the (estimated) expected return of the \( d \) assets. Then \( \mu^T u \) is the expected return of the portfolio.

The risk of the portfolio is measured by \( \rho(-R^T u) \), where \( \rho \) is a properly chosen risk measure. There are generally two approaches to the computation of \( \rho(-R^T u) \): (i) one first assumes and estimates a (parametric) probability model for the joint distribution of \( R \) and then compute \( \rho(-R^T u) \); (ii) one estimates the risk \( \rho(-R^T u) \) directly from the historical observations of \( R \) without assuming any hypothetical model for \( R \).

As discussed in Section 2.2, the return vector \( R \) is usually observed under different scenarios such as economic booms and finance crisis. Suppose there are \( m \) scenarios. For each \( s = 1, \ldots, m \), let \( \tilde{R}^{[s]} \in \mathbb{R}^{n_s \times d} \) be the collection of \( n_s \) observations of \( R \) under the \( s \)-th scenario, where each row of the matrix \( \tilde{R}^{[s]} \) represents one observation of \( R^T \). Then we define the matrix \( \tilde{R} \) and the observations of the portfolio loss \( x(u) \) as follows:
\[
\tilde{R} := \begin{pmatrix} \tilde{R}^{[1]} \\ \tilde{R}^{[2]} \\ \vdots \\ \tilde{R}^{[m]} \end{pmatrix} \in \mathbb{R}^{n \times d}, \quad x(u) := -\tilde{R}u = \begin{pmatrix} -\tilde{R}^{[1]}u \\ -\tilde{R}^{[2]}u \\ \vdots \\ -\tilde{R}^{[m]}u \end{pmatrix} = \begin{pmatrix} x^{[1]}(u) \\ x^{[2]}(u) \\ \vdots \\ x^{[m]}(u) \end{pmatrix} \in \mathbb{R}^n, \quad n := \sum_{s=1}^{m} n_s,
\]
where \( x^{[s]}(u) = -\tilde{R}^{[s]}u \in \mathbb{R}^{n_s} \) denotes the observations of the portfolio loss under the \( s \)-th scenario, \( s = 1, 2, \ldots, m \).
In this paper, we are going to directly estimate $\rho(-R^T u)$ from the return observations $\tilde{R}$, as this approach does not require a subjective model for $R$ and hence greatly reduces model misspecification error.

### 3.1 Sample Versions of Measures of Portfolio Risk

For each risk measure $\rho$ discussed in Section 2.2, $\rho(-R^T u)$ can be estimated from the return observations $\tilde{R}$ by their corresponding sample versions shown below.

**VaR:** Suppose that there is only one scenario, i.e., $m = 1$. For a given $\alpha \in (0, 1)$, let $p = \lceil \alpha n \rceil$. Then the sample version of the VaR at level $\alpha$ of the portfolio is

$$\rho_{\text{VaR}}(x(u)) := x(u)_p = (-\tilde{R}u)_p. \quad (8)$$

**CVaR:** Suppose that $m = 1$. For a given $\alpha \in (0, 1)$, let $p = \lceil \alpha n \rceil$. Then the sample version of CVaR at level $\alpha$ of the portfolio is

$$\rho_{\text{CVaR}}(x(u)) := \frac{p - \alpha n}{(1 - \alpha)n} x(u)_p + \frac{1}{(1 - \alpha)n} \sum_{i=p+1}^n x(u)_i. \quad (9)$$

**Basel 2.5:** For a given $\alpha \in (0, 1)$, let $p_s = \lceil \alpha n_s \rceil$, $s = 1, \ldots, m$. Then $x^s(u)_{p_s}$ is the sample version of VaR at level $\alpha$ of the portfolio estimated from the data set $\tilde{R}^s$ corresponding to the $s$-th scenario. Let $m_1 = m_2 = 60$ and $m = 120$. Suppose the first $m_1$ scenarios correspond to the current market condition and the last $m_2$ scenarios correspond to stressed scenarios. Then the sample version of Basel 2.5 Accord risk measure is given by

$$\rho_{\text{Basel2.5}}(x(u)) := \max \left\{ x^1(u)_{p_1}, \frac{k}{m_1} \sum_{s=1}^{m_1} x^s(u)_{p_s} \right\}$$

$$+ \max \left\{ x^{m_1+1}(u)_{p_{m_1+1}}, \frac{\ell}{m_2} \sum_{s=m_1+1}^m x^s(u)_{p_s} \right\}. \quad (10)$$

Because the two constants $k$ and $\ell$ in the above formula are no less than 3, one would expect that in many situations $x^1(u)_{p_1} \leq \frac{k}{m_1} \sum_{s=1}^{m_1} x^s(u)_{p_s}$ and $x^{m_1+1}(u)_{p_{m_1+1}} \leq \frac{\ell}{m_2} \sum_{s=m_1+1}^m x^s(u)_{p_s}$. Hence, in many cases $\rho_{\text{Basel2.5}}(x(u))$ would be equal to the following “simplified Basel 2.5” risk measure, which is defined as

$$\rho_{\text{Basel2.5s}}(x(u)) := \frac{k}{m_1} \sum_{s=1}^{m_1} x^s(u)_{p_s} + \frac{\ell}{m_2} \sum_{s=m_1+1}^m x^s(u)_{p_s}. \quad (11)$$
**Basel III:** Let $\alpha$ and $p_s$ be defined the same as before. Then

$$\rho_{\text{CVaR}}(x^{[s]}(u)) := \frac{p_s - \alpha n_s}{(1 - \alpha)n_s} x^{[s]}(u)_{(p_s)} + \frac{1}{(1 - \alpha)n_s} \sum_{i=p_s+1}^{n_s} x^{[s]}(u)_{(i)}$$

is the sample version of CVaR at level $\alpha$ of the portfolio estimated from the data set $\tilde{R}^{[s]}$. Suppose the first $m_1 = 60$ scenarios correspond to the current market condition and the last $m_2 = 60$ scenarios correspond to stressed scenarios. Then the sample version of Basel-III risk measure is

$$\rho_{\text{BaselIII}}(x(u)) := \max \left\{ \rho_{\text{CVaR}}(x^{[1]}(u)), \frac{k}{m_1} \sum_{s=1}^{m_1} \rho_{\text{CVaR}}(x^{[s]}(u)) \right\}$$

$$+ \max \left\{ \rho_{\text{CVaR}}(x^{[m_1+1]}(u)), \frac{\ell}{m_2} \sum_{s=m_1+1}^{m} \rho_{\text{CVaR}}(x^{[s]}(u)) \right\},$$

where $m_1 = m_2 = 60$ and $m = 120$. Similarly, the simplified Basel III risk measure can be defined:

$$\rho_{\text{BaselIII}}(x(u)) := \frac{k}{m_1} \sum_{s=1}^{m_1} \rho_{\text{CVaR}}(x^{[s]}(u)) + \frac{\ell}{m_2} \sum_{s=m_1+1}^{m} \rho_{\text{CVaR}}(x^{[s]}(u)).$$

### 3.2 The “Mean-\(\rho\)-Basel” and “Mean-Basel” Asset Allocation Models

Suppose a portfolio manager in a financial institution attempts to construct a portfolio composed of the $d$ assets and to choose the portfolio weights $u \in \mathbb{R}^d$ to optimize the portfolio performance. The manager can freely choose the risk measure $\rho$ to measure the risk of the portfolio, such as variance, VaR, or CVaR; in addition, he or she has the freedom to choose a model for the asset returns $R$ or a data set $\tilde{Y} \in \mathbb{R}^{n' \times d}$, which has a similar structure as $\tilde{R}$ defined in (7) and contains observations of the asset returns, to estimate the portfolio risk. Hence, the portfolio risk will be given by $\rho(y(u))$, where $y(u) := -\tilde{Y} u$. Furthermore, the manager can specify that the expected portfolio return should be no less than a target return $r_0$, namely, the portfolio weights $u$ should satisfy

$$u \in \mathcal{U}_{r_0} = \{ u \in \mathbb{R}^d \mid \mu^\top u \geq r_0, 1^\top u = 1, u \geq 0 \}.$$  

Here, it is assumed that the portfolio is long only; this assumption can be relaxed or removed without incurring additional technical difficulty in solving the portfolio selection problem specified below.
At the same time, the manager has to satisfy the constraint that the regulatory capital for his or her portfolio should not exceed certain upper limit $C_0$, which is allocated to him or her by the financial institution’s top management. The capital requirement for the portfolio is calculated by the Basel Accord risk measure $\rho_{\text{Basel}}$, which is specified by the regulators; in addition, the data set $\tilde{R}$ used for calculating the capital requirements should also satisfy certain criteria and cannot be freely chosen by the portfolio manager. For example, the Basel 2.5 risk measure requires that $\tilde{R}$ should include 60 normal scenarios and 60 stressed scenarios. Hence, the data set $\tilde{R}$ may be different from the data set $\tilde{Y}$, and the capital requirement for the portfolio is $\rho_{\text{Basel}}(x(u))$, where $x(u) = -\tilde{R}u$.

To summarize, the portfolio manager is facing the following “mean-$\rho$-Basel” asset allocation problem:

$$\min_{u \in \mathcal{U}_0} \rho(y(u))$$

subject to

$$\rho_{\text{Basel}}(x(u)) \leq C_0,$$

where $y(u) = -\tilde{Y}u$, $x(u) = -\tilde{R}u$, $\rho_{\text{Basel}}$ is the Basel Accord risk measure for calculating regulatory capital, i.e., $\rho_{\text{Basel2.5}}$ or $\rho_{\text{Basel3}}$; $C_0$ is the upper bound of the available capital; $\rho$ is the risk measure for calculating the risk level of the portfolio.

In Section 4, we will develop a unified method for solving the mean-$\rho$-Basel problem. The method can also be applied to solve the following “mean-$\rho$” problem:

$$\min_{u \in \mathcal{U}_0} \rho(x(u))$$

where $\rho$ can be any risk measure chosen by the portfolio manager. The “mean-$\rho$” problem does not take into account the Basel Accord capital constraint and is hence simpler than the problem (15).

Alternatively, the portfolio manager can construct the portfolio by maximizing the expected return of the portfolio subject to the constraint that the portfolio risk measured by $\rho$ does not exceed a pre-specified risk budget $b_0$. The corresponding asset allocation problem is

$$\min_{u \in \mathcal{U}} -\mu^\top u,$$

subject to

$$\rho(x(u)) \leq b_0,$$

where $\mathcal{U} = \{u \in \mathbb{R}^d \mid 1^\top u = 1, u \geq 0\}$.

The asset allocation problems (15), (16) and (17) with $\rho \in \{\rho_{\text{VaR}}, \rho_{\text{Basel2.5}}\}$ or with $\rho_{\text{Basel}} = \rho_{\text{Basel2.5}}$ are mixed-integer programming (MIP) problems. In fact, the problem
problem \((16)\) with \(\rho = \rho_{\text{VaR}}\) can be formulated as
\[
\min_{u \in \mathcal{U}_0, \gamma \in \mathbb{R}, y \in \mathbb{R}^n} \gamma
\]
\[
\text{s.t.} \quad -\tilde{R}u \leq \gamma + \eta y, \quad 1^\top y = n - p, y \in \{0, 1\}^n,
\]
where \(\eta\) is a large constant, for example, \(\eta = \max_{u \in \mathcal{U}_0} \max_{j=1,\ldots,n} (-\tilde{R}u)_j\). Similarly, problem \((16)\) with \(\rho = \rho_{\text{Basel2.5}}\) can be rewritten as:
\[
\min \gamma_1 + \gamma_2
\]
\[
\text{s.t.} \quad -\tilde{R}^{[s]}u \leq \beta_s \mathbf{1} + \eta_z^{[s]}, \quad s = 1, \ldots, m,
\]
\[
1^\top z^{[s]} \leq n_s - p_s, \quad z^{[s]} \in \{0, 1\}^{n_s}, \quad s = 1, \ldots, m,
\]
\[
\beta_1 \leq \gamma_1, \quad \frac{k}{m_1} \sum_{s=1}^{m_1} \beta_s \leq \gamma_1,
\]
\[
\beta_{m_1+1} \leq \gamma_2, \quad \frac{\ell}{m_2} \sum_{s=m_1+1}^{m} \beta_s \leq \gamma_2,
\]
\[
u \in \mathcal{U}_0.
\]

On the other hand, by using a similar proof as Theorem 16 in Rockafellar and Uryasev (2002), problem \((16)\) with \(\rho = \rho_{\text{Basel3}}\) can be formulated as a linear programming problem:
\[
\min \gamma_1 + \gamma_2
\]
\[
\text{s.t.} \quad t_1 + \frac{1}{(1 - \alpha)n_1} \sum_{i=1}^{n_1} y_i^{[1]} \leq \gamma_1,
\]
\[
\frac{k}{m_1} \sum_{s=1}^{m_1} \left( t_s + \frac{1}{(1 - \alpha)n_s} \sum_{i=1}^{n_s} y_i^{[s]} \right) \leq \gamma_1,
\]
\[
t_{m_1+1} + \frac{1}{(1 - \alpha)n_{m_1+1}} \sum_{i=1}^{n_{m_1+1}} y_i^{[m_1+1]} \leq \gamma_2,
\]
\[
\frac{\ell}{m_2} \sum_{s=m_1+1}^{m} \left( t_s + \frac{1}{(1 - \alpha)n_s} \sum_{i=1}^{n_s} y_i^{[s]} \right) \leq \gamma_2,
\]
\[
y_i^{[s]} \geq 0, \quad y_i^{[s]} \geq -\tilde{R}_i^{[s]}u - t_s, \quad s = 1, \ldots, m,
\]
\[
u \in \mathcal{U}_0.
\]

### 4 The Alternating Direction Augmented Lagrangian Method

We develop the alternating direction augmented Lagrangian method (ADM) for solving both the mean-\(\rho\)-Basel and the mean-\(\rho\) problem. Although the ADM approach has been
widely used in convex optimization (He et al., 2002; Wen et al., 2010; Yang and Zhang, 2011), it appears that its use for solving optimization problems involving VaR or Basel Accords risk measures is new. In addition, the proposed method is different from the penalty decomposition methods in Bai et al. (2012) as their division of blocks of variables leads to subproblems that are more expensive to solve.

For simplicity of presentation, we first present the method for solving the mean-$\rho$ problem (16) in Section 4.1 and provide convergence analysis for the method in Section 4.2. We then demonstrate the method for solving the mean-$\rho$-Basel model (15) in Section 4.3.

4.1 The ADM Algorithm for Solving Problem (16)

The asset allocation problem (16) with $\rho = \rho_{\text{VaR}}$ or $\rho = \rho_{\text{Basel2.5}}$ is usually difficult because of its connection to the MIP. By introducing certain intermediate variables, we can exploit the structures of the risk measures and design an efficient approximate algorithm.

Our algorithmic framework is adapted from the alternating direction augmented Lagrangian method (Wen, Goldfarb, and Yin, 2010). Treating the implicit relationship $x(u)$ as a variable $x$ and imposing $x = -\tilde{R}u$ explicitly, we can write (16) as

$$\min_{u \in U_0, x \in \mathbb{R}^n} \rho(x)$$

s.t. $x + \tilde{R}u = 0$.  \hspace{1cm} (21)

The augmented Lagrangian function for (21) is defined as

$$\mathcal{L}(x, u, \lambda) := \rho(x) + \lambda^\top (x + \tilde{R}u) + \frac{\sigma}{2} \|x + \tilde{R}u\|^2,$$  \hspace{1cm} (22)

where $\sigma > 0$ is the penalty parameter and $\lambda \in \mathbb{R}^n$ is the Lagrangian multiplier associated with the equality constraint $x + \tilde{R}u = 0$. Using (22), we can devise an ADM that minimizes (22) with respect to $x$ and $u$ in an alternating fashion. To be specific, given some initial guess $u^{(0)} \in \mathbb{R}^d$ and $\lambda^{(0)} := 0$, the simplest ADM solves the following two subproblems sequentially in each iteration:

$$x^{(j+1)} = \arg\min_{x \in \mathbb{R}^n} \mathcal{L}(x, u^{(j)}, \lambda^{(j)}),$$ \hspace{1cm} (23)

$$u^{(j+1)} = \arg\min_{u \in U_0} \mathcal{L}(x^{(j+1)}, u, \lambda^{(j)}),$$ \hspace{1cm} (24)

and then updates the Lagrange multiplier $\lambda$ by

$$\lambda^{(j+1)} = \lambda^{(j)} + \beta \sigma (x^{(j+1)} + \tilde{R}u^{(j+1)}),$$ \hspace{1cm} (25)
where $\beta > 0$ is an appropriately chosen step length.

The subproblem (24) is a quadratic programming (QP) problem:

$$u^{(j+1)} = \arg \min_{u \in U} \frac{1}{2} u^\top \tilde{R}^\top \tilde{R} u + b^\top u,$$

(26)

where $b = \tilde{R}^\top \left( \frac{1}{\sigma} \lambda^{(j)} + x^{(j+1)} \right)$. Since the dimension $d$ of $u$ is often much smaller than a few hundreds in practice, problem (26) can be solved quite efficiently by standard QP solvers, such as CPLEX.

The subproblem (23) can be simplified as

$$x^{(j+1)} = \arg \min_x \psi(x) = \rho(x) + \frac{\sigma}{2} \|x - v^{(j)}\|^2,$$

(27)

where $v^{(j)} = - \left( \tilde{R} u^{(j)} + \frac{1}{\sigma} \lambda^{(j)} \right)$. Although $\rho(x)$ is a nonsmooth function, it can be shown that the above problem is essentially equivalent to a simple QP whose Hessian is an identity.

For simplicity, we now temporarily omit the superscripts in $x^{(j)}$ and $u^{(j)}$, and denote them by $x$ and $u$, respectively. Similar notation is used for $v$. In addition, we let $v_0 = -\infty$.

For the VaR case, the subproblem (27) is

$$\min_x \psi(x) = x(p) + \frac{\sigma}{2} \|x - v\|^2,$$

whose closed-form solution is given in the following lemma.

**Lemma 4.1.** Consider (16) with $\rho(x(u)) = \rho_{\text{VaR}}(x(u))$. Suppose, without loss of generality, that $v_1 \leq v_2 \leq \ldots \leq v_n$. The optimal solution $x$ of (27) are given by

$$x(i) = \begin{cases} 
\gamma_i^* & \text{if } i^* \leq i \leq p; \\
 v_i & \text{otherwise},
\end{cases}$$

(28)

where

$$i^* := \max \{i \mid i \leq p, \ v_{i-1} < \gamma_i\} \text{ and } \gamma_i = \frac{\sigma \sum_{j=i}^p v_j - 1}{\sigma(p - i + 1)}.$$

(29)

**Proof.** See Appendix A. \qed

For the Basel 2.5 case, the subproblem (27) can be re-written as

$$\min_x \max \left\{ x^{[1]}(p_1), \frac{k}{m_1} \sum_{s=1}^{m_1} x^{[s]}(p_1), \frac{\ell}{m_2} \sum_{s=m_1+1}^{m} x^{[s]}(p_1 + 1) \right\} + \max \left\{ x^{[m_1+1]}(p_1 + 1), \frac{\ell}{m_2} \sum_{s=m_1+1}^{m} x^{[s]}(p_1 + 1) \right\} + \frac{\sigma}{2} \sum_{s=1}^{m} \|x^{[s]} - v^{[s]}\|^2,$$

(30)
where \( v_s = - \left( \tilde{R}_s u + \frac{1}{2} \lambda_s \right) \). Note that (30) is separable in terms of two group of variables \( \{x^{[1]}, \cdots, x^{[m_1]}\} \) and \( \{x^{[m_1+1]}, \cdots, x^{[m]}\} \):

\[
\begin{align*}
& \min_{x^{[1]}, \cdots, x^{[m_1]}} \quad \max \left\{ \frac{k}{m_1} \sum_{s=1}^{m_1} x^{[s]}_{(p_s)} \right\} + \frac{\sigma}{2} \sum_{s=1}^{m_1} \| x^{[s]} - v^{[s]} \|^2, \\
& \min_{x^{[m_1+1]}, \cdots, x^{[m]}} \quad \max \left\{ \frac{\ell}{m_2} \sum_{s=m_1+1}^{m} x^{[s]}_{(p_s)} \right\} + \frac{\sigma}{2} \sum_{s=m_1+1}^{m} \| x^{[s]} - v^{[s]} \|^2.
\end{align*}
\]

(31)

Since (31) and (32) are the same except the constant terms, it suffices to consider one of them.

**Lemma 4.2.** Suppose, without loss of generality, that \( v_1^{[s]} \leq v_2^{[s]} \leq \cdots \leq v_{n_s}^{[s]} \), \( s = 1, \ldots, m_1 \). The optimal solution \( x^{[s]} \) of (31) are given by

\[
x^{[s]}_{(i)} = \begin{cases} 
 y^{[s]}_{(i)} & \text{if } 1 \leq i \leq p_s; \\
 v^{[s]}_{(i)} & \text{otherwise,}
\end{cases}
\]

(33)

where \( y \) in (33) are determined by the optimal solutions of the following QP:

\[
\begin{align*}
& \min_{y, \gamma} \psi(y, \gamma) = \gamma + \frac{\sigma}{2} \sum_{s=1}^{m_1} \| y^{[s]} - w^{[s]} \|^2 \\
& \text{s.t.} \quad y^{[s]} \leq y_{p_s}^{[s]} 1, \quad s = 1, \cdots, m_1; \\
& \quad y^{[1]} \leq \gamma; \\
& \quad \frac{k}{m_1} \sum_{s=1}^{m_1} y_{p_s}^{[s]} \leq \gamma.
\end{align*}
\]

(34)

Here \( w^{[s]} \) are vectors consisting of the first \( p_s \) smallest components of \( v^{[s]} \), \( s = 1, \ldots, m_1 \).

**Proof.** The equivalence between (31) and (34) can be proved by using similar arguments as that for Lemma 4.1. \[\Box\]

**Remark 4.1.** The subproblem (27) corresponding to CVaR can also be derived.

**Remark 4.2.** Suppose that the risk is defined as \( \rho(x(u)) = \rho_{\text{Basel2.5s}}(x(u)) \). Then the subproblem (27) is

\[
\begin{align*}
& \min_{x} \quad \frac{k}{m_1} \sum_{s=1}^{m_1} x^{[s]}_{(p_s)} + \frac{\ell}{m_2} \sum_{s=m_1+1}^{m} x^{[s]}_{(p_s)} + \frac{\sigma}{2} \| x^{[s]} - v^{[s]} \|^2, \\
& \text{which is separable in terms of the variables } x^{[s]}, \text{ and each of the optimal } x^{[s]} \text{ is determined by Lemma 4.1.}
\end{align*}
\]

15
4.2 Convergence Analysis

Since VaR and Basel 2.5 risk measure are nonconvex functions, the convergence of the ADM scheme (23)-(25) to a global optimal solution is not guaranteed, at least not in theory. However, empirical evidence suggests that ADM seems to converge from any starting point. In the following, we show that ADM can find a first-order stationary point of (16) or (21) under some mild conditions.

We first review the definitions of locally Lipschitz functions and the Clarke’s generalized gradient.

Definition 4.1. A function \( f(x) : \mathbb{R}^n \rightarrow \mathbb{R} \) is Lipschitz near a point \( x_0 \in \text{int}(\text{dom } f) \) if there exist \( K \geq 0 \) such that \( |f(x) - f(x')| \leq K\|x - x'\| \) for all \( x, x' \in B_\delta(x_0) \), where \( \delta > 0 \) sufficiently small so as to have \( B_\delta(x_0) := \{ x \in \mathbb{R}^n : |x - x_0| < \delta \} \subset \text{dom } f \). A locally Lipschitz function is a function that is Lipschitz near every point in \( \mathbb{R}^n \).

Definition 4.2. Let \( f(x) : \mathbb{R}^n \rightarrow \mathbb{R} \). The Clarke’s generalized gradient of \( f(x) \) at \( x \in \text{dom } f \) is defined as

\[
\bar{\partial} f(x) \triangleq \{ \xi \in \mathbb{R}^n \mid \xi^\top d \leq f^\circ(x; d), \forall d \in \mathbb{R}^n \},
\]

where, for any \( d \in \mathbb{R}^n \), \( f^\circ(x; d) \) is the Clarke’s generalized directional derivatives

\[
f^\circ(x; d) \triangleq \lim_{y \to x, t \to 0^+} \sup \frac{f(y + td) - f(y)}{t}.
\]

Using Proposition 2.1.2 in Clarke (1990), we have

Proposition 4.1. Let \( f(x) : \mathbb{R}^n \rightarrow \mathbb{R} \) be Lipschitz near \( x \in \text{dom } f \), then \( \partial f(x) \) is nonempty and convex. Moreover, if \( f \) is locally Lipschitz, then \( \partial f(x) \) exists over \( \mathbb{R}^n \).

Next, we show that the sample versions of risk measures, such as VaR and Basel 2.5 risk measure, are locally Lipschitz. Hence, it follows from Proposition 4.1 that their Clarke generalized gradients exist over \( \mathbb{R}^n \).

Lemma 4.3. The functions \( \rho_{\text{VaR}}(x) \) and \( \rho_{\text{Basel2.5}}(x) \) defined in (8) and (10) are locally Lipschitz function.

Proof. See Appendix B. \( \square \)
It can also be verified that the Clarke’s generalized gradient of $\rho_{\text{VaR}}(x)$ is

$$
\tilde{\partial}\rho_{\text{VaR}}(x) = \text{conv}\left\{ e_i \mid i \in E_{x(\rho)} \right\},
$$

(37)

where $e_i$ is the $i$-th unit vector in $\mathbb{R}^n$.

By using the fact that $\rho_{\text{Basel2.5}}(x)$ is composed of a few $\rho_{\text{VaR}}(x)$ and maximization, Theorem 2.3.3 on a finite sum of functions and Theorem 2.3.10 on the chain rule in Clarke (1990), we can obtain the Clarke’s generalized gradient of $\rho_{\text{Basel2.5}}(x)$:

$$
\tilde{\partial}\rho_{\text{Basel2.5}}(x) \subseteq \text{conv}\left( \bigcup_{i \in I_1(x)} \tilde{\partial}f_i(x) \right) + \text{conv}\left( \bigcup_{i \in I_2(x)} \tilde{\partial}f_i(x) \right),
$$

(38)

where

1. $f_1 = x^{[1]}(u)_{(p_1)}$, and $\tilde{\partial}f_1(x) = \text{conv}\left\{ e_i \mid i \in E_{x^{[1]}_{(p_1)}} \right\}$;

2. $f_2 = \frac{k}{m_1} \sum_{s=1}^{m_1} x^{[s]}(u)_{(p_s)}$, and $\tilde{\partial}f_2(x) = \frac{k}{m_1} \sum_{s=1}^{m_1} \text{conv}\left\{ e_i \mid i \in E_{x^{[s]}_{(p_s)}} \right\}$;

3. $I_1(x) := \{ i \mid \max\{ f_1(x), f_2(x) \} = f_i(x), \ i = 1, 2 \}$;

4. $f_3 = x^{[1]}(u)_{(p_1)}$, and $\tilde{\partial}f_3(x) = \text{conv}\left\{ e_i \mid i \in E_{x^{[1]}_{(p_1)}} \right\}$;

5. $f_4 = \frac{l}{m_2} \sum_{s=m_1+1}^{m} x^{[s]}(u)_{(p_s)}$, and $\tilde{\partial}f_4(x) = \frac{l}{m_2} \sum_{s=m_1+1}^{m} \text{conv}\left\{ e_i \mid i \in E_{x^{[s]}_{(p_s)}} \right\}$;

6. $I_2(x) := \{ i \mid \max\{ f_3(x), f_4(x) \} = f_i(x), \ i = 3, 4 \}$.

By applying the corollary of Proposition 2.4.3 and Theorem 2.3.10 (chain rule) in Clarke, 1990), we obtain the following first-order optimality (KKT) conditions of (16).

**Proposition 4.2.** Suppose that $\rho(x)$ is locally Lipschitz, $x(u) = -\bar{R}u$. If $u$ is a local minimizer of (16), it holds

$$
0 \in -\bar{R}^T \tilde{\partial}\rho(x(u)) + N_{U_{r_0}}(u),
$$

(39)

where $N_{U_{r_0}}(u) = \text{cl} \left\{ \cup_{t \geq 0} \partial \text{dist}_{U_{r_0}}(u) \right\}$ is the normal cone to $U_{r_0}$ at $u$.

Finally, we are able to establish the following convergence results.

**Theorem 4.1.** Suppose that $\bar{R}^T \bar{R} > 0$. Let $\{(x^{(j)}, u^{(j)}, \lambda^{(j)})\}$ be a sequence generated by scheme (23)-(25). Assume that $\sum_{j \to \infty} \| \lambda^{(j+1)} - \lambda^{(j)} \|^2 < \infty$ and $\{\lambda^{(j)}\}$ is bounded. Then any accumulation point of $\{u^{(j)}\}$ is a first-order stationary point.
Proof. See Appendix C. □

**Remark 4.3.** Suppose that \( \rho = \rho_{\text{VaR}} \). The first-order optimality conditions of (16) can also be described in terms of the so called order-value optimization (OVO) problem studied in Andreani et al. (2005). For all \( u \in U_{r_0} \), we define two sets

\[
L_\rho(u) := \{ i \in \{1, \ldots, n\} \mid x_i(u) < \rho(x(u)) \}, \\
E_\rho(u) := \{ i \in \{1, \ldots, n\} \mid x_i(u) = \rho(x(u)) \}.
\]

Given \( u \in U_{r_0} \), a feasible sequence \( u^{(j)} \subset U_{r_0} \) is called a descent sequence for a function \( \phi \) if \( \lim_{j \to \infty} u^{(j)} = u \) and there exists \( j_0 \) such that \( \phi(u^{(j)}) < \phi(u) \), \( \forall j > j_0 \). A vector \( d \in \mathbb{R}^d \) is called an unitary tangent direction to the set \( U_{r_0} \) at \( u \in U_{r_0} \) if there exists a feasible sequence \( u^{(j)} \) such that \( \lim_{j \to \infty} u^{(j)} = u \) and \( d = \lim_{j \to \infty} \frac{u^{(j)} - u}{\|u^{(j)} - u\|_2} \). Let \( u \in U_{r_0} \). By applying Theorems 2.1 and 2.2 in Andreani et al. (2005) to (16), we obtain the corresponding first-order optimality conditions.

1. \( u \) is a local minimizer of (16) if and only if for all feasible sequence \( u^{(j)} \) that converges to \( u \),

\[
\# \{ i \in E_\rho(u) \mid u^{(j)} \text{ is a descent sequence for } x_i(u) \} < p - \#L_\rho(u).
\]

2. Assume that \( u \) is a local minimizer of (16). Then, for all unitary tangent directions \( d \),

\[
\# \left\{ i \in E_\rho(u) \mid \left\langle d, \tilde{R}_i \right\rangle < 0 \right\} < p - \#L_\rho(u).
\]

Hence, a convergence result similar to Theorem 4.1 can also be established.

### 4.3 The ADM for Solving Models (15) and (17)

The ADM framework for solving the two models are as follows.

**ADM for Model (17):** Let \( x = -\tilde{R}u \). The corresponding augmented Lagrangian function for (17) is defined as

\[
L_c(x, u, \lambda) := -\mu^T u + \lambda^T (x + \tilde{R}u) + \frac{\sigma}{2} \|x + \tilde{R}u\|^2,
\]

where \( \mu \) and \( \lambda \) are the Lagrange multipliers.

18
where \( \sigma > 0 \) is the penalty parameter and \( \lambda \in \mathbb{R}^n \) is the Lagrangian multiplier. Then the ADM method becomes

\[
x^{(j+1)} = \arg \min_{x \in \mathbb{R}^n} \|x - v^{(j)}\|^2, \quad \text{s.t.} \quad \rho(x) \leq b_0, \tag{43}
\]

\[
u^{(j+1)} = \arg \min_{u \in U} \frac{1}{2} u^\top \tilde{R}^\top \tilde{R} u + b_c^\top u, \tag{44}
\]

\[
\lambda^{(j+1)} = \lambda^{(j)} + \beta \sigma (x^{(j+1)} + \tilde{R} u^{(j+1)}), \tag{45}
\]

where \( v^{(j)} = - (\tilde{R} u^{(j)} + \frac{1}{\sigma} \lambda^{(j)}) \), \( b_c = \tilde{R}^\top (\frac{1}{\sigma} \lambda^{(j)} + x^{(j+1)}) - \tilde{n} \), and \( \beta > 0 \). Note that the subproblem (44) is exactly the same as (26). It can also be proved that solving (43) is equivalent to solving a QP. When \( \rho(x(u)) = \rho_{\text{VaR}}(x(u)) \), the closed form solution of (43) is given by (28), where \( \gamma_{v_*} \) is set to \( b_0 \).

**ADM for Model (15):** Let \( x = - \tilde{R} u \) and \( y = - \tilde{Y} u \). Then problem (15) is equivalent to

\[
\begin{aligned}
\min_{u \in U_0, x, y} & \rho(y), \\
\text{s.t.} & \quad \rho_{\text{Basel}}(x) \leq C_0, \\
& \quad x + \tilde{R} u = 0, \\
& \quad y + \tilde{Y} u = 0.
\end{aligned}
\]

The augmented Lagrangian function for the above problem is defined as

\[
\mathcal{L}_e(x, u, \lambda) := \rho(y) + \lambda^\top (x + \tilde{R} u) + \frac{\sigma_1}{2} \|x + \tilde{R} u\|^2 + \pi^\top (y + \tilde{Y} u) + \frac{\sigma_2}{2} \|y + \tilde{Y} u\|^2, \tag{46}
\]

where \( \sigma_1, \sigma_2 > 0 \). Then the corresponding ADM method becomes

\[
x^{(j+1)} = \arg \min_{x} \|x - v^{(j)}\|^2, \quad \text{s.t.} \quad \rho_{\text{Basel}}(x) \leq C_0, \tag{47}
\]

\[
y^{(j+1)} = \arg \min_{y} \rho(y) + \frac{\sigma_2}{2} \|y - w^{(j)}\|^2, \tag{48}
\]

\[
u^{(j+1)} = \arg \min_{u \in U_0} \frac{1}{2} u^\top (\sigma_1 \tilde{R}^\top \tilde{R} + \sigma_2 \tilde{Y}^\top \tilde{Y}) u + b_c^\top u, \tag{49}
\]

\[
\lambda^{(j+1)} = \lambda^{(j)} + \beta_1 \sigma_1 (x^{(j+1)} + \tilde{R} u^{(j+1)}), \tag{50}
\]

\[
\pi^{(j+1)} = \pi^{(j)} + \beta_2 \sigma_2 (y^{(j+1)} + \tilde{Y} u^{(j+1)}), \tag{51}
\]

where \( v^{(j)} = - (\tilde{R} u^{(j)} + \frac{1}{\sigma} \lambda^{(j)}) \), \( w^{(j)} = - (\tilde{Y} u^{(j)} + \frac{1}{\sigma} \pi^{(j)}) \), \( b_c = \tilde{R}^\top (\lambda^{(j)} + \sigma_1 x^{(j+1)}) + \tilde{Y}^\top (\pi^{(j)} + \sigma_2 y^{(j+1)}) \), and \( \beta_1, \beta_2 > 0 \). The subproblem (47) can be solved in a similar fashion as Lemma 4.2. Solving the subproblem (48) depends on the specific risk measure \( \rho \), see, for example, Lemmas 4.1 and 4.2. The subproblem (49) can be solved by standard QP solvers.
5 Numerical results

In this section, we demonstrate the effectiveness of the ADM method for the mean-Basel and the mean-$\rho$-Basel problem by comparing the ADM with the MIP using both simulated and real market data.

5.1 Data Description

In our experiments, the real market data and simulated data sets are generated as follows.

- **S&P 500 Data Set.** The S&P 500 data set comprises the daily returns of 359 stocks that have ever been included in the S&P 500 index and do not have missing data during the following specified time periods. Let $t_0 = 03/01/2012$. For $s = 1, \ldots, 60$, $\tilde{R}_{[s]}$ denotes the trailing five year daily returns of the stocks on the day $t_0 - s + 1$ (i.e., the daily returns of the stocks during the period from the day $t_0 - s - 2058$ to the day $t_0 - s + 1$). Let $l = 06/01/2007$ and $u = 06/01/2009$. For $s = 61, \ldots, 120$, $\tilde{R}_{[s]}$ is defined as the daily returns of the stocks during the stressed period from the day $l + 120 - s$ to the day $u - s + 61$.

- **Simulated Data.** We simulate the prices of 350 stocks based on a multi-dimensional version of the double-exponential jump diffusion model (Kou, 2002)

$$dS_i(t) = \mu_i dt + \sigma_i dW_i(t) + d \left( \sum_{k=1}^{N_i(t)} (e^{V_{ik}} - 1) \right), i = 1, \ldots, n, \quad (52)$$

where $W_1(t), \ldots, W_n(t)$ are $n$ correlated Brownian motions with $dW_i(t)dW_j(t) = \rho_{ij} dt$; $N_i(t)$ is a Poisson process with intensity $\lambda_i$ and $N_i(t)$ is independent of $N_j(t)$ for $i \neq j$; $\{V_{i1}, V_{i2}, \ldots\}$ are i.i.d. log jump sizes with a double-exponential distribution with density:

$$f_i(x) = \begin{cases} p_i \eta_{iu} e^{-\eta_{iu} x}, & x \geq 0, \\ (1 - p_i) \eta_{id} e^{\eta_{id} x}, & x < 0, \end{cases} \text{ where } 0 < p_i < 1, \eta_{iu} > 0, \eta_{id} > 0;$$

$V_{ik}$ and $V_{jl}$ are independent for $i \neq j$; the Brownian motions, Poisson processes and the jump sizes are mutually independent. Under model (52), the log return of the $i$-th stock from time $t$ to $t + \Delta t$ is given by

$$(\mu_i - \frac{1}{2} \sigma_i^2) \Delta t + \sigma_i (W_i(t + \Delta t) - W_i(t)) + \sum_{k=N_i(t)+1}^{N_i(t+\Delta t)} V_{ik}.$$
In the simulation, we choose \( \Delta t := 1/252 \) (one day) and the parameters \( \{\mu_i, \sigma_i, \lambda_i, p_i, \eta_{iu}, \eta_{id}, \rho_{ij}\} \) are set to be the estimates based on the historical data of some large-cap stocks. The stock returns generated in the above model have the same tail heaviness as those generated by the negative exponential tail model considered in Lim, Shanthikumar, and Vahn (2011).

5.2 Parameter Settings of the ADM and MIP

Our method is implemented in MATLAB. All the experiments were performed on a Dell Precision Workstation T5500 with Intel Xeon CPU E5620 at 2.40GHz and 12GB of memory running Ubuntu 12.04 and MATLAB 2011b. All the quadratic programming subproblems in the ADM method are solved by the QP solvers in CPLEX 12.4 with Matlab interface; and the mixed integer programming (MIP) reformulations of the asset allocation models are solved by the MIP solvers in CPLEX 12.4. In our test, the parameters \( \sigma \) and \( \beta \) in (23)-(25) are set to be \( 10^{-3} \) and 0.1, respectively. The initial Lagrangian multiplier is \( \lambda = 0 \). The method is terminated if \( \frac{\|u^{(j+1)}-u^{(j)}\|_2}{\max(1,\|u^{(j)}\|_2)} \leq 10^{-4} \) or after a maximal number of 1000 iterations. The default setting of the MIP solver in CPLEX 12.4 is used. The maximum CPU time limit for all solvers is set to 3600 seconds.

5.3 Comparing ADM with MIP by Varying the Number of Stocks in Model (16)

In this subsection, we evaluate the performance of the ADM and MIP methods on both real market data and simulated data sets by varying the number of stocks in the portfolio. The first experiment is to solve

\[
\min_{u \in \mathcal{U}_{r_0}} \rho_{\text{VaR}}(x(u)).
\]

Specifically, we randomly select \( d \in \{100, 150, 200, 250, 300, 350\} \) columns from \( \tilde{R} \) and still denote the resulted matrix by \( \tilde{R} \). The mean \( \mu \) in \( \mathcal{U}_{r_0} \) is set as the sample mean of \( \tilde{R} \). The prescribed return level \( r_0 \) is set to be the 80% quantile of the cross-sectional expected returns of the \( d \) assets. The confidence level \( \alpha \) of VaR is set to 0.99. The first four columns of Table 1 report the number of binary variables, continuous variables, and linear constraints, denoted by “0-1”, “cont.”, and “cons.”, respectively, in the MIP reformulation of (53). The results on \( \rho_{\text{VaR}}(x(u)) \) computed at the final iteration and the CPU time used in seconds are...
Table 1: The computed $\rho_{\text{VaR}}(x(u))$ and the gap between the ADM and MIP methods.

<table>
<thead>
<tr>
<th>Problem information</th>
<th>Simulated data</th>
<th>Real market data</th>
</tr>
</thead>
<tbody>
<tr>
<td>$d$</td>
<td>ADM</td>
<td>MIP</td>
</tr>
<tr>
<td>100</td>
<td>4379</td>
<td>0.0267</td>
</tr>
<tr>
<td>150</td>
<td>4379</td>
<td>0.0253</td>
</tr>
<tr>
<td>200</td>
<td>4379</td>
<td>0.0243</td>
</tr>
<tr>
<td>250</td>
<td>4379</td>
<td>0.0231</td>
</tr>
<tr>
<td>300</td>
<td>4379</td>
<td>0.0228</td>
</tr>
<tr>
<td>350</td>
<td>4379</td>
<td>0.0233</td>
</tr>
</tbody>
</table>

The detailed values of $\rho_{\text{VaR}}(x(u))$ and the relative error defined by

$$\text{rel.err} = 100 \frac{\rho_{\text{VaR}}(x(u_{\text{ADM}})) - \rho_{\text{VaR}}(x(u_{\text{MIP}}))}{\rho_{\text{VaR}}(x(u_{\text{MIP}}))}$$

are reported in Table 1, where $u_{\text{ADM}}$ and $u_{\text{MIP}}$ are solutions computed by the ADM and MIP methods, respectively. From the figure and table, we can observe that the ADM method computed good approximation solutions whose relative errors to MIP were less than 5% at most scenarios. However, MIP reached the maximum CPU time 3600 seconds in all cases while ADM only used CPU time on the order of 100 seconds.

Figure 1: Comparison on minimizing $\rho_{\text{VaR}}(x(u))$ for different number of stocks using real market and simulated data.

The second experiment is to solve

$$\min_{u \in \mathcal{U}_0} \rho_{\text{Basel2.5}}(x(u)).$$

(54)
Table 2: The computed $\rho_{\text{Basel2.5}}(x(u))$ and the gap between the ADM and MIP methods.

<table>
<thead>
<tr>
<th>Problem information</th>
<th>Simulated data</th>
<th>Real market data</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ADM</td>
<td>MIP</td>
</tr>
<tr>
<td>$d$</td>
<td>0-1</td>
<td>cont.</td>
</tr>
<tr>
<td>100</td>
<td>58020</td>
<td>222</td>
</tr>
<tr>
<td>150</td>
<td>58020</td>
<td>272</td>
</tr>
<tr>
<td>200</td>
<td>58020</td>
<td>322</td>
</tr>
<tr>
<td>250</td>
<td>58020</td>
<td>372</td>
</tr>
<tr>
<td>300</td>
<td>58020</td>
<td>422</td>
</tr>
<tr>
<td>350</td>
<td>58020</td>
<td>472</td>
</tr>
</tbody>
</table>

Some rows of the data matrix $\tilde{R}$ are deleted so that the number $n_s$ of historical observations is 504 and 463 for $s \in \{1, \ldots, 60\}$ and $s \in \{61, \ldots, 120\}$, respectively. As a result, the number of binary variables and linear constraints are 58020 and 58146, respectively (see the first four columns of Table 2). The computed risk measures $\rho_{\text{Basel2.5}}(x(u))$ and the CPU time used in seconds are presented in Figure 2, and the relative errors are reported in Table 2. From the figure and table, we can observe that the ADM method can compute better solutions than MIP within the same amount of time.

![Figure 2: Comparison on minimizing $\rho_{\text{Basel2.5}}(x(u))$ for different number of stocks using real market and simulated data.](image)

In summary, we can safely conclude that the ADM method is quite competitive to the MIP solver CPLEX 12.4 in terms of both solution quality and computational time.
5.4 Comparing ADM with MIP by Varying the Number of Samples in Model (16)

In this subsection, we evaluate the performance of the ADM method by varying the number of samples; specifically, the number of samples \( n \in \{500, 1000, \ldots, 3000\} \) for S&P 500 data and \( n \in \{2000, 3000, \ldots, 7000\} \) for simulated data. The data sets \( \tilde{R} \) are generated in a similar fashion as in Section 5.3. To simplify the experiments and without loss of generality, we only compare risk measures \( \rho_{\text{VaR}} \) and \( \rho_{\text{Basel2.5s}} \) with two scenarios \( R^{[1]} \) and \( R^{[61]} \) and all other \( R^{[s]} \) are empty. The number of stocks is fixed at \( d = 350 \).

Then we consider two cases as follows: i) solving (16) with \( \rho_{\text{VaR}}(x(u)) \) by using the ADM and MIP method, respectively; ii) solving (16) with \( \rho_{\text{Basel2.5s}}(x(u)) \) by using the ADM, MIP, and the penalty decomposition (PD) method proposed in Bai et al. (2012), respectively.

The problem information, such as the number of binary variables and linear constraints, is described in Table 3. The computed risk measures \( \rho(x(u)) \) and the CPU time in seconds are presented in Figures 3 and 4. We can see from these figures that the ADM method is also quite competitive to the MIP solver CPLEX 12.4 in terms of both solution quality and computational time. Both ADM and PD can provide solutions of almost the same quality. However, PD consumed much more CPU time as the size of the problem increased, whereas the CPU time of ADM remained almost the same no matter what the problem size was.

| Table 3: Information of the problems studied in Section 5.4. |
|----------------------------------|-----------------|-----------------|
| Simulated data                   | Real market data |
| sample                           | 0-1  cont.  cons. | sample          | 0-1  cont.  cons. |
| 2000                             | 2522 351 2525    | 500             | 1022 351 1025    |
| 3000                             | 3522 351 3525    | 1000            | 1522 351 1525    |
| 4000                             | 4522 351 4525    | 1500            | 2022 351 2025    |
| 5000                             | 5522 351 5525    | 2000            | 2522 351 2525    |
| 6000                             | 6522 351 6525    | 2500            | 3022 351 3025    |
| 7000                             | 7522 351 7525    | 3000            | 3522 351 3525    |
5.5 Comparing ADM with MIP on the Mean-$\rho$-Basel Model (15)

In this subsection, we evaluate the performance of the ADM methods on the mean-VaR-Basel model:

$$
\min_{u \in U_0} \rho_{\text{VaR}}(y(u)) \\
\text{s.t.} \quad \rho_{\text{Basel}}(x(u)) \leq C_0,
$$

(55)

where $y(u) = -\tilde{Y}u; x(u) = -\tilde{R}u; \rho_{\text{Basel}}$ is the Basel Accord risk measure $\rho_{\text{Baseline}}$ or $\rho_{\text{Baseline3}}$. The experiments on the real market data were carried out using the same data matrix $\tilde{R}$ as those in Section 5.3; and the data matrix $\tilde{Y}$ was obtained by deleting the duplicated rows in $\tilde{R}$. Because the minimal value of $\rho_{\text{Baseline}}(x(u))$ ranged from 0.1 to 0.14 (see Figure 2) and the MIP approach usually cannot achieve as small objective function value as that of ADM in one hour, we set $C_0$ to 0.2 to guarantee that all methods can find a feasible solution to (55).

We consider the following two cases: i) solving model (55) with $\rho_{\text{Baseline}}(x(u))$ in the constraints by using the ADM and MIP method, respectively; ii) solving model (55) with $\rho_{\text{Baseline3}}(x(u))$ in the constraints by using the ADM and MIP method, respectively.

The problem information is summarized in Table 4. The computed risk measures $\rho_{\text{VaR}}(y(u))$ and the CPU time in seconds are presented in Figure 5. These values as well as $\rho_{\text{Baseline}}(x(u))$ and $\rho_{\text{Baseline3}}(x(u))$ are reported in Table 5, where ADM and MIP correspond to the results of real market data from Table 1. We can see from these figures and
Table 4: Problem information of model (55) with $\rho_{\text{Basel2.5}}$ or $\rho_{\text{Basel3}}$ constraints. The numbers of binary variables and linear constraints are the same for all $d$, while the numbers of continuous variables are $\{223, 273, 323, 373, 423, 473\}$ for Basel 2.5 constraint and $\{58243, 58293, 58343, 58393, 58443, 58493\}$ for Basel III constraint, for the corresponding values of $d$.

<table>
<thead>
<tr>
<th>$d$</th>
<th>$\rho_{\text{Basel2.5}}(x(u)) \leq C_0$</th>
<th>$\rho_{\text{Basel3}}(x(u)) \leq C_0$</th>
</tr>
</thead>
<tbody>
<tr>
<td>100</td>
<td>62399 223 62527</td>
<td>4379 58243 62407</td>
</tr>
</tbody>
</table>

Table 5 shows that: (i) $\rho_{\text{VaR}}$ computed by ADM is smaller than that of MIP in 75% of the cases, and (ii) in the rest of 25% cases, although $\rho_{\text{VaR}}$ computed by ADM is larger than that of MIP, ADM reached smaller value of $\rho_{\text{Basel2.5}}(x(u))$ or $\rho_{\text{Basel3}}(x(u))$ in the constraints than MIP. The comparison with ADM and MIP on the objective function value $\rho_{\text{VaR}}$ shows that $\rho_{\text{Basel2.5}}(x(u))$ and $\rho_{\text{Basel3}}(x(u))$ constraints are not redundant.

6 Conclusions

A major change in the financial regulation after the recent financial crisis is that financial institutions are now required to meet much more stringent regulatory capital requirements than they were before the crisis. It has been estimated that the capital requirement for a
Table 5: Numerical results on the mean-VaR-Basel model (55)

<table>
<thead>
<tr>
<th>stocks</th>
<th>ADM</th>
<th>MIP</th>
<th>ADM_{\text{Basel}2.5 \leq C_0}</th>
<th>MIP_{\text{Basel}2.5 \leq C_0}</th>
<th>ADM_{\text{Basel}3 \leq C_0}</th>
<th>MIP_{\text{Basel}3 \leq C_0}</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ρ_{VaR}</td>
<td>ρ_{VaR}</td>
<td>ρ_{VaR}</td>
<td>ρ_{Basel2.5}</td>
<td>ρ_{VaR}</td>
<td>ρ_{Basel2.5}</td>
</tr>
<tr>
<td>100</td>
<td>0.024</td>
<td>0.023</td>
<td>0.025</td>
<td>0.133</td>
<td>0.024</td>
<td>0.141</td>
</tr>
<tr>
<td>150</td>
<td>0.024</td>
<td>0.022</td>
<td>0.025</td>
<td>0.143</td>
<td>0.027</td>
<td>0.152</td>
</tr>
<tr>
<td>200</td>
<td>0.021</td>
<td>0.021</td>
<td>0.022</td>
<td>0.129</td>
<td>0.028</td>
<td>0.150</td>
</tr>
<tr>
<td>250</td>
<td>0.020</td>
<td>0.020</td>
<td>0.022</td>
<td>0.134</td>
<td>0.025</td>
<td>0.138</td>
</tr>
<tr>
<td>300</td>
<td>0.020</td>
<td>0.020</td>
<td>0.022</td>
<td>0.132</td>
<td>0.027</td>
<td>0.160</td>
</tr>
<tr>
<td>350</td>
<td>0.020</td>
<td>0.020</td>
<td>0.022</td>
<td>0.131</td>
<td>0.025</td>
<td>0.142</td>
</tr>
</tbody>
</table>
Figure 5: Comparison results of model (55) for different number of stocks using real market data.

large bank’s trading book under the Basel 2.5 Accord on average *more than double* that under the Basel II Accord. The significantly higher capital requirement makes it more important for banks to take into account the capital constraint when they construct their investment portfolio. In this paper, we propose a new asset allocation model called the “mean-\(\rho\)-Basel” model that incorporates the Basel Accord capital requirements as one of the constraints. In this model, the capital requirement is measured by the Basel 2.5 or the Basel III risk measure that is imposed by regulators; the risk level of the portfolio is measured by \(\rho\), such as variance, VaR, and CVaR, which can be freely chosen by the portfolio manager.

The complexity of the Basel 2.5 and Basel III risk measures, which involve risk measurement under multiple scenarios including stressed scenarios, poses significant computational challenge to the proposed asset allocation problem due to its nature of non-smoothness and non-convexity. We propose to solve the problem by using the alternating direction augmented Lagrangian method. Essentially, the computational cost is reduced to solving simple convex quadratic programming problems. Hence, the method enables us to solve large-scale asset allocation problems which are usually difficult to solve by many other solvers. Extensive numerical results suggest that our method is promising for finding approximate optimal solutions with good quality.
A  Proof of Lemma 4.1.

Proof. Let $x$ be an optimal solution of (27). Let

$$q = \arg \max_{1 \leq q \leq p} x_q \quad \text{and} \quad h = \arg \min_{h \geq p+1 \text{ or } h=q} x_h.$$  \hfill (56)

Suppose that $x_q$ is not the $p$-th smallest component of $x$. Then we obtain

$$x_q > x_h \quad \text{and} \quad v_h \geq v_p \geq v_q.$$  \hfill (57)

Assume that $x_q > v_q$. Define $y$ as $y_q = v_q$ and $y_j = x_j (\forall j \neq q)$. Otherwise, $x_q \leq v_q$. It follows from (57) that $v_h \geq v_q \geq x_q > x_h$ and we define $y$ as $y_q = x_h$, $y_h = v_h$ and $y_j = x_j (\forall j \neq q, h)$. It is clear that $\psi(x) > \psi(y)$, which is a contradiction. Hence, we obtain that $x_i = v_i$, $i = p+1, \ldots, n$, and $(x_1, \ldots, x_p)$ are the optimal solutions of

$$\min x_p + \frac{\sigma}{2} \sum_{i=1}^{p} (x_i - v_i)^2$$

s.t. $x_i \leq x_p$, $i = 1, \ldots, k-1$.  \hfill (58)

Since $\{v_i\}$ are sorted in ascending order, we also have $x_1 \leq x_2 \leq \ldots \leq x_p$. The KKT conditions of (58) are

$$x_i \leq x_p, \ i = 1, \ldots, p-1,$$

$$\sigma(x_i - v_i) + \pi_i = 0, \ i = 1, \ldots, p-1,$$  \hfill (60)

$$\pi_i(x_p - x_i) = 0, \ i = 1, \ldots, p-1,$$  \hfill (61)

$$\sigma(x_p - v_p) + 1 - \sum_{j=1}^{p-1} \pi_j = 0.$$  \hfill (62)

The equations in (60) imply that either $x_i = v_i$ (or equivalently $\pi_i = 0$) or $x_i = x_p$, $i = 1, \ldots, p-1$. Suppose that $x_{i-1} < x_i$ and $x_j = x_p$ for $j = i, \ldots, p$, then we have from (62) that $x_p = \gamma_i$, where $\gamma_i$ is defined in (29). Therefore, it suffices to find the index $i^*$ such that $v_{i^* - 1} < \gamma_i$. This completes the proof. \hfill $\square$

B  Proof of Lemma 4.3.

Proof. We first consider the portfolio risk $\rho_{\text{VaR}}(x) := x_{(p)}$. Given $x \in \mathbb{R}^n$, let $L_{x_{(p)}}$, $E_{x_{(p)}}$ and $G_{x_{(p)}}$ be the set of all indices whose corresponding entries of $x$ are less than, equal to
or greater than \( x_{(p)} \), respectively; that is, \( L_{x_{(p)}} \triangleq \{ i \mid x_i < x_{(p)} \} \), \( E_{x_{(p)}} \triangleq \{ i \mid x_i = x_{(p)} \} \) and \( G_{x_{(p)}} \triangleq \{ i \mid x_i > x_{(p)} \} \). Let \( \delta_x = \min_{i \in L_{x_{(p)}} \cup G_{x_{(p)}}} |x_i - x_{(p)}| \). Using the definitions of \( L_{x_{(p)}} \) and \( G_{x_{(p)}} \), we have \( \delta_x > 0 \). Then for any \( y, z \in B(x, \frac{\delta_x}{3}) \), it can be proved that \( \delta_z \geq \frac{2 \delta_x}{3} \) and \( y \in B(z, \frac{\delta_x}{2}) \). Hence, we have

\[
y_i < y_j, \quad \forall i \in L_{z(p)}, \ j \in E_{z(p)}; \quad y_i > y_j, \quad \forall i \in G_{z(p)}, \ j \in E_{z(p)}.
\]

It follows from (63) that \( E_{y(p)} \subset E_{z(p)} \) and consequently

\[
y_{(p)} - z_{(p)} = y_i - z_j = y_i - z_i + z_i - z_j = y_i - z_i, \quad \forall i \in E_{y(p)}, \ \forall j \in E_{z(p)}.
\]

Furthermore, \( |y_{(p)} - z_{(p)}| = |y_i - z_i| \leq \| y - z \| \), which means that \( \rho_{\text{VaR}}(x) \) is Lipschitz near \( x \).

Note that \( \rho_{\text{Basel2.5}}(x) \) is a composite function of a few finite linear combinations of \( \rho_{\text{VaR}}(x) \) (with different \( p \)) and maximization. Since the maximization of two locally Liptschitz functions and the summation of a finite number of locally Liptschitz functions are still locally Liptschitz, we can prove that \( \rho_{\text{Basel2.5}}(x) \) is locally Liptschitz. \( \square \)

## C Proof of Theorem 4.1.

**Proof.** First, we claim

\[
x^{(j+1)} - x^{(j)} \to 0 \quad \text{and} \quad u^{(j+1)} - u^{(j)} \to 0.
\]

Since \( U_\sigma \) is a closed and bounded set, the sequence \( \{ u^{(j)} \} \) is bounded. The boundedness of \( \{ u^{(j)} \} \) and \( \{ \lambda^{(j)} \} \) shows that \( v^{(j)} = -(\bar{R}u^{(j)} + \frac{1}{\sigma} \lambda^{(j)}) \) is bounded, which together with Lemma 4.1 implies the boundedness of \( \{ x^{(j)} \} \). Hence, the sequence \( \{ \mathcal{L}(x^{(j)}, u^{(j)}, \lambda^{(j)}) \} \) is bounded from the definition of the augmented Lagrangian function (22).

Note that the augmented Lagrangian function \( \mathcal{L} \) is strongly convex with respect to the variable \( u \). It holds for any \( u \) and \( \Delta u \) that

\[
\mathcal{L}(x, u + \Delta u, \lambda) - \mathcal{L}(x, u, \lambda) \geq \partial_u \mathcal{L}(x, u, \lambda)^\top \Delta u + c\| \Delta u \|_2^2.
\]

In addition, \( u^{(j+1)} \) being an minimizer of (24) implies that

\[
\partial \mathcal{L}(x^{(j+1)}, u^{(j+1)}, \lambda^{(j)})^\top \Delta u \geq 0.
\]

Combining (66) and (67), we obtain

\[
\mathcal{L}(x^{(j+1)}, u^{(j)}, \lambda^{(j)}) - \mathcal{L}(x^{(j+1)}, u^{(j+1)}, \lambda^{(j)}) \geq c\| u^{(j+1)} - u^{(j)} \|_2^2.
\]
Using \( x^{(j+1)} \) is a minimizer of (23), (68) and (25), we have
\[
\mathcal{L}(x^{(j)}, u^{(j)}, \lambda^{(j)}) - \mathcal{L}(x^{(j+1)}, u^{(j+1)}, \lambda^{(j+1)})
\]
\[
= \mathcal{L}(x^{(j)}, u^{(j)}, \lambda^{(j)}) - \mathcal{L}(x^{(j+1)}, u^{(j)}, \lambda^{(j)}) + \mathcal{L}(x^{(j+1)}, u^{(j)}, \lambda^{(j)}) - \mathcal{L}(x^{(j+1)}, u^{(j+1)}, \lambda^{(j)})
\]
\[
\geq c\|u^{(j+1)} - u^{(j)}\|^2 - \frac{1}{\beta\sigma}\|\lambda^{(j)} - \lambda^{(j+1)}\|^2.
\]
(69)

Summing up both sides of (69) and recalling that \( \{\mathcal{L}(x^{(j)}, u^{(j)}, \lambda^{(j)})\} \) is bounded yields
\[
\sum_{j=1}^{\infty} c\|u^{(j+1)} - u^{(j)}\|^2 - \frac{1}{\beta\sigma}\|\lambda^{(j)} - \lambda^{(j+1)}\|^2 < \infty.
\]
(70)

Since the second term on the left side of the above inequality is bounded, it follows that
\[
u^{(j+1)} - u^{(j)} \to 0.
\]
(71)

By the assumption \( \lim_{j \to \infty} \lambda^{(j+1)} - \lambda^{(j)} = 0 \) and \( \beta, \sigma > 0 \), we obtain from (25) that
\[
\lim_{j \to \infty} x^{(j+1)} + \tilde{R}u^{(j+1)} = 0.
\]
(72)

Hence, using (71) and (72) yields \( x^{(j+1)} - x^{(j)} \to 0 \).

For any limit point \((\bar{x}, \bar{u}, \bar{\lambda})\) of the sequence \( \{(x^{(j)}, u^{(j)}, \lambda^{(j)})\} \), there exists a subsequence \( \{x^{j_i}, u^{j_i}, \lambda^{j_i}\} \) converging to \((\bar{x}, \bar{u}, \bar{\lambda})\). Clearly, we obtain from (71) and (72) that
\[
0 = \lim_{j_i \to \infty} x^{(j_i)} + \tilde{R}u^{(j_i)} = \lim_{j_i \to \infty} x^{(j_i)} + \tilde{R}u^{(j_i-1)} = \bar{x} + \tilde{R}\bar{u}.
\]
(73)

The first-order optimality condition of (23) in the \( j_i \)-th iteration is
\[
0 \in \partial \rho(x^{(j_i)}) + \lambda^{(j_i-1)} + \sigma(x^{(j_i)}) + \tilde{R}u^{(j_i-1)}.
\]
(74)

Similarly, we obtain the first-order optimality condition of (24) in the \( j_i \)-th iteration
\[
\tilde{R}^\top \lambda^{(j_i-1)} + \sigma\tilde{R}^\top(x^{(j_i)}) + \tilde{R}u^{(j_i)} + \eta^{(j_i)} = 0,
\]
(75)

where \( \eta^{(j_i)} \in \mathcal{N}_{\mathcal{U}_0}(u^{(j_i)}) \) is the normal cone to \( \mathcal{U}_0 \) at \( u^{(j_i)} \).

It follows from the compactness and convexity of \( \mathcal{U}_0 \) that the normal cone \( \mathcal{N}_{\mathcal{U}_0}(u) \) coincides with the cone of normals (proposition 2.4.4, Clarke (1990)). Consequently, we obtain \( \lim_{j_i \to \infty} \mathcal{N}_{\mathcal{U}_0}(u^{(j_i)}) \subseteq \mathcal{N}_{\mathcal{U}_0}(\lim_{j_i \to \infty} u^{(j_i)}) \), which gives
\[
\lim_{j_i \to \infty} \eta^{(j_i)} \in \mathcal{N}_{\mathcal{U}_0}(\bar{u}).
\]
(76)

Taking limit to both sides of (74) and (75), and applying (73) and (76), we obtain (39). This completes the proof. \( \square \)
References


