Fairness and Optimality in Trading

by

Van Vinh Nguyen

Submitted to the School of Engineering
in partial fulfillment of the requirements for the degree of

Master of Science in Computation for Design and Optimization

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

September 2010

Author ....
School of Engineering
August 5, 2010

Certified by ....
Dimitris Bertsimas
Boeing Professor of Operations Research
Thesis Supervisor

Accepted by ....
Karen Willcox
Associate Professor of Aeronautics and Astronautics
Codirector, Computation for Design and Optimization Program

Fairness and Optimality in Trading

by

Van Vinh Nguyen

Submitted to the School of Engineering
on August 5, 2010, in partial fulfillment of the
requirements for the degree of
Master of Science in Computation for Design and Optimization

Abstract

This thesis proposes a novel approach to address the issues of efficiency and fairness when multiple portfolios are rebalanced simultaneously. A fund manager who rebalances multiple portfolios needs to not only optimize the total efficiency, i.e., maximize net risk-adjusted return, but also guarantee that trading costs are fairly split among the clients. The existing approaches in the literature, namely the Social Welfare and the Competitive Equilibrium schemes, do not compromise efficiency and fairness effectively. To this end, we suggest an approach that utilizes popular and well-accepted resource allocation ideas from the field of communications and economics, such as Max-Min fairness, Proportional fairness and α-fairness.

We incorporate in our formulation a quadratic model of market impact cost to reflect the cumulative effect of trade pooling. Total trading costs are split fairly among accounts using the so-called *pro rata* scheme. We solve the resulting multi-objective optimization problem by adopting the Max-Min fairness, Proportional fairness and α-fairness schemes. Under these schemes, the resulting optimization problems have non-convex objectives and non-convex constraints, which are NP-hard in general. We solve these problems using a local search method based on linearization techniques. The efficiency of this approach is discussed when we compare it with a deterministic global optimization method on small size optimization problems that have similar structure to the aforementioned problems.

We present computational results for a small data set (2 funds, 73 assets) and a large set (6 funds, 73 assets). These results suggest that the solution obtained from our model provides a better compromise between efficiency and fairness than existing approaches. An important implication of our work is that given a level of fairness that we want to maintain, we can always find Pareto-efficient trade sets.

Thesis Supervisor: Dimitris Bertsimas
Title: Boeing Professor of Operations Research

Acknowledgments

First and foremost, I am very grateful to my advisor, Professor Dimitris Bertsimas, for his invaluable support and guidance. I learned a lot from his broad knowledge and deep insight in various technical topics. He also constantly provided me with useful advices for career advancement and personal development. Under his supervision, I was really inspired to work on challenging research problems.

I am also thankful to Dan A. Iancu and Nikolaos K. Trichakis from MIT Operations Research Center for their thorough guidance and useful discussions during the course of this project. They introduced me to the topic of fair trading, helped me get familiar with computational tools and supported me with the financial data. Without their help, this thesis would never be completed.

My special thanks go to the Singapore-MIT Alliance program for awarding me a Graduate Fellowship. The program has provided me with the invaluable opportunity to experience an enjoyable time at MIT.

Finally, I would like to dedicate this thesis to my family for their unconditional love and continuous encouragement.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Motivation and objectives

This thesis proposes a novel approach to address the issues of efficiency and fairness when multiple portfolios are rebalanced simultaneously. A fund manager who manages multiple accounts of diverse clients needs to periodically rebalance the portfolios to maintain the desired levels of return and risk exposure. In practice, the manager often wants to pool and execute trades from these portfolios simultaneously to increase efficiency. In this process, the nonlinear trading impact costs depend on the size of the cumulative trade. Thus, the individual problems of rebalancing the accounts are in fact interdependent. As a central decision maker, the fund manager needs to decide how to distribute the total trading costs fairly among the clients. The existing approaches in the literature, namely the Social Welfare and Competitive Equilibrium schemes, do not address the issue of fair distribution of trading costs appropriately. The objective of this research is to contribute towards this direction. Specifically, we want to build a mathematical model that captures the interdependence between different portfolios under the presence of market impact cost. We also suggest a way to split the trading costs fairly among the clients. Finally, we apply popular resource allocation ideas from the fields of communications and economics, such as Max-Min fairness, Proportional fairness and α-fairness, to our problem to get solutions that trade off efficiency and fairness.

1.2 Approaches and contributions

We propose here a model to the multi-account rebalancing problem which incorporates quadratic market impact cost to reflect the cumulative effect of trade pooling. Total trading cost is split fairly among accounts using the so-called *pro rata* scheme. We then solve the resulting multi-objective optimization problem under the Max-Min fairness, Proportional fairness and α-fairness schemes. Under these schemes, we need to solve either a single or a series of non-convex problems with non-convex constraints, which are NP-hard in general. We solve these problems using a local search method based on linearization techniques. The efficiency of this approach is discussed when we compare it with a deterministic global optimization method on small size optimization problems that have similar structure to the aforementioned problems.

We present computational results for a small data set (2 funds, 73 assets) and a large data set (6 funds, 73 assets). These results suggest that the solution obtained from our model provides a better compromise between efficiency and fairness than existing approaches. An important implication of our work is that given a level of fairness that we want to maintain, we can always find Pareto-efficient trade sets.

1.3 Thesis outline

The outline of this thesis is as follows. In Chapter 2, we formulate the multi-account rebalancing problem as a multiple objective optimization problem and review the existing approaches to solve this problem. In the last section of this chapter, we review some important fairness concepts from the field of communications that we want to use in our problem. In Chapter 3, we explain in details how we applied these fairness concepts to the multiple portfolio trading problem. We also propose in this chapter a local search algorithm that can be applied to solve the resulting nonlinear and non-convex optimization programs. This chapter concludes with a description of a deterministic global optimization algorithm based on branch-and-bound strategy that we utilize as a benchmark to our proposed approach. In Chapter 4, we present the numerical experiments based on two case studies that we generate from real-world financial data. We compare the performance of different approaches and show that our proposed method is superior to existing approaches in the literature in term of compromise between efficiency and fairness. We also argue that our local search algorithm could possibly provides near optimal solutions by comparing the outcomes to those from the branch and bound algorithm on small size problems. Chapter 5 summarizes this thesis and provides some directions for future development of this research.

Chapter 2

Model Formulation and Fairness Concepts

2.1 Introduction

The foundations of modern portfolio theory were laid by Markowitz with his *mean-variance analysis* framework [8]. Under this framework, the asset allocation problem is solved via means of an optimization problem, in which the asset weights are chosen so as to maximize the portfolio expected return given a specified level of portfolio return variance. Alternatively, the asset weights can be chosen to minimize the portfolio return variance given a desired level of expected return. Under the *risk aversion formulation* (see for e.g. [2]), the mean-variance portfolio optimization problem takes the form

$$\max\{\boldsymbol{\mu}'\mathbf{w} - \lambda \mathbf{w}'\boldsymbol{\Sigma}\mathbf{w} \mid \mathbf{w}'\mathbf{e} = 1\}, \tag{2.1}$$

where $\mathbf{w} \in \mathbf{R}^n$ is the vector of asset weights, $\boldsymbol{\mu} \in \mathbf{R}^n$ and $\boldsymbol{\Sigma} \in \mathbf{R}^{n \times n}$ are the expected values and covariance matrix of asset returns, respectively, $\mathbf{e} = [1, \ldots, 1] \in \mathbf{R}^n$ and λ is the risk aversion parameter reflecting the investor's risk preference.

In practice, portfolio managers often introduce additional constraints to the original formulation to reflect specific institution's features or investment policies. Examples include long-only constraints, turnover constraints or sector exposure constraints

[2]. Practical portfolio optimization problems may also involve complicated mixed integer constraints such as minimum holding requirement [14] or restrictions on the number of assets we invest on [1]. The inclusion of transaction costs in the asset allocation problem was first studied by Pogue [17] and further expanded and modified by other authors (see for e.g. [7, 9]). The principal idea in these studies is the adjustment of the objective function of (2.1) to take into account various transaction costs such as brokerage fees/commissions or market impact cost. Specifically, the portfolio optimization problem is rewritten as

$$\max\{\boldsymbol{\mu}'\mathbf{w} - \lambda\mathbf{w}'\boldsymbol{\Sigma}\mathbf{w} - \lambda_C C(\mathbf{w}) \mid \mathbf{w}'\mathbf{e} = 1\}, \tag{2.2}$$

where λ_C is the transaction cost aversion parameter and C is the transaction cost function. The transaction cost function C can be a complicated nonlinear function and it is commonly approximated by a linear or quadratic function for tractability. In Section 2.2, we discuss in more detail the quadratic model approximation of the transaction cost function that we use in this thesis.

The multi-account portfolio rebalancing problem with transaction costs was introduced by O'Cinneide et al. [11]. In their study, O'Cinneide et al. noticed that when a fund manager optimizes a client's portfolio individually, the calculated trades are suboptimal in the presence of concurrent trades of other accounts. The authors also suggested two approaches to tackle the multi-account portfolio rebalancing problem by applying ideas of common good distribution from microeconomics, namely the Social Welfare scheme and the Competitive Equilibrium scheme. The main idea of the Social Welfare scheme is to maximize the total goodness of participating clients. The problem with the Social Welfare idea is that the resulting allocation might favor several clients over others. As we will see in Section 2.4, the resulting allocation vector from Social Welfare scheme is Pareto optimal but need not satisfy justified fairness criteria. Furthermore, the justification for Competitive Equilibrium scheme is that it corresponds to how the clients would behave if they traded independently in an efficient market, given that each client had complete information about the trades of others at execution time. However, this is a heuristic scheme and the resulting allocation vector might not be Pareto optimal. The lack of an approach that guarantees both fairness and Pareto optimality in the current literature motivates us to tackle this problem with the application of fairness ideas from the fields of communications and economics, namely the Max-Min fairness, Proportional fairness and α-fairness scheme.

The structure of this chapter is as follows. In Section 2.2, we introduce the notation and the multi-account portfolio rebalancing problem that we consider in this thesis. Section 2.3 reviews current approaches in the literature, including the Independent scheme, the Social Welfare scheme and the Competitive Equilibrium scheme. Section 2.4 introduces the concepts of fair rational preference relation and fair dominance. It also describes the α fairness scheme as a fair aggregation problem using a class of parametric utility functions. Proportional fairness and Max-Min fairness appear as two special cases of the α fairness scheme.

2.2 Model formulation

Notation

In this thesis, we will consider a scenario in which a fund manager has to manage n portfolios (or funds) that are invested in the same market with m risky assets (or stocks). The funds and assets are indexed by $i = 1, ..., n$ and $j = 1, ..., m$ respectively. Each fund i is assumed to have an initial investment (in dollar value), denoted by $\mathbf{w}_i \in \mathbf{R}^m$. We want to rebalance by making changes to the positions of the funds. Let $\mathbf{x}_i \in \mathbf{R}^m$ be the dollar change in the position of the ith fund and let x_{ij} be the change in dollar value held in the jth asset. We model the utility of the ith fund as the risk-adjusted mean return of its position:

$$U_i(\mathbf{x}_i) = \boldsymbol{\mu}^T(\mathbf{w}_i + \mathbf{x}_i) - \lambda_i(\mathbf{w}_i + \mathbf{x}_i)^T\boldsymbol{\Sigma}(\mathbf{w}_i + \mathbf{x}_i), \tag{2.3}$$

where $\boldsymbol{\mu} \in \mathbf{R}^m$ is the vector of the mean returns of the assets, $\boldsymbol{\Sigma} \in \mathbf{R}^{m \times m}$ is the covariance matrix, and λ_i is the risk aversion factor of the ith fund.

Allocation of trading cost

Given the desired rebalancing changes of all the funds in the jth asset, $x_{1j}, x_{2j}, \ldots, x_{nj}$, the *net trade volume* on asset j is defined as

$$z_j = x_{1j} + x_{2j} + \cdots + x_{nj}. \tag{2.4}$$

Trading costs include commissions, fees and market impact cost. In this thesis, we will focus mainly on the market impact cost because it dominates commissions and fees for large transactions.

A common assumption is that the charge for trading an amount z_j on asset j is independent from the charge for trading an amount z_k on asset k $(j \neq k)$. Thus, for each asset $j = 1, \ldots, m$, we can define the trading cost function $t_j : \mathbf{R} \to \mathbf{R}$ such that given a net trade volume z_j, $t_j(z_j)$ is the *total market impact cost* due to trading activities of all the funds on asset j. This trading cost is then split among the funds according to the volumes they trade on that particular asset. Let $\tau_j : \mathbf{R} \times \mathbf{R} \to \mathbf{R}$ be the *trading cost splitting function* corresponding to asset j. A fund i that trades an amount x_{ij} on asset j will be charged $\tau_j(x_{ij}, z_j)$, which depends only on x_{ij} and the net trade z_j. The functions $\tau_j(x, z)$ should satisfy:

$$\sum_{i=1}^{n} \tau_j(x_{ij}, z_j) = t_j(z_j), \quad \forall j = 1, \ldots, m. \tag{2.5}$$

We use the *pro rata* allocation scheme [18] to split the trading costs fairly among the individual funds. As shown in [18], the pro-rata split costs are justified in the sense that they correspond to expected costs that the clients would pay if they traded in the market simultaneously without any information of other clients' trades. In this scheme, each fund receives an amount proportional to its own rebalancing change in a particular asset:

$$\tau_j(x_{ij}, z_j) = \frac{x_{ij}}{z_j} t_j(z_j) = \frac{x_{ij}}{\sum_{i=1}^{n} x_{ij}} t_j\Big(\sum_{i=1}^{n} x_{ij}\Big). \tag{2.6}$$

Since the marginal market impact cost increases with respect to the trade volume, the market impact cost $t_j(z_j)$ is often a nonlinear function of the trade volume z_j. Hence, to simplify the computation, piecewise linear or quadratic approximation are common choices in practice. Even though a piecewise linear function is considered a good option in the single account portfolio optimization problem, it introduces fractional terms under the pro rata scheme. Thus, we will consider the quadratic model of market impact cost:

$$t_j(z_j) = \alpha_j z_j^2, \tag{2.7}$$

where α_j is *price impact parameter* of asset jth. If we substitute (2.7) into (2.6), we obtain

$$\tau_j(x_{ij}, z_j) = \alpha_j x_{ij}(x_{1j} + \cdots + x_{nj}). \tag{2.8}$$

We will see later in this section how the utility functions of clients are modified to take into account the trading costs of the form in (2.8). However, we first discuss our specifications for the sets of feasible rebalancing trades in the next section.

Constraints

For each fund i, let $\mathcal{C}_i$ be the set of feasible rebalancing vectors $\mathbf{x}_i$. We use the following constraints to specify $\mathcal{C}_i$:

1. *No short-selling.* Many funds and institutional investors are prohibited from selling stocks short. Hence, we require that

$$\mathbf{w}_i + \mathbf{x}_i \geq 0.$$

2. *Self-financing.* We assume that there are no cash flows in or out of each portfolio (except for trading costs)

$$\mathbf{e}^T\mathbf{x}_i = 0.$$

3. *Turnover:* Portfolios with high turnover often incur large transaction costs. Hence, we want to limit the total trades of fund i to within some fraction γ_i of the initial investment amount:

$$\| \mathbf{x}_i \|_1 \leq \gamma_i \mathbf{e}^T \mathbf{w}_i.$$

4. *Sector exposure.* The exposure of a fund after rebalancing in any sector S should remain within a percentage δ_{is} of the initial exposure

$$(1-\delta_{is})\sum_{j\in\mathcal{S}} w_{ij} \leq \sum_{j\in\mathcal{S}}(w_{ij}+x_{ij}) \leq (1+\delta_{is})\sum_{j\in\mathcal{S}} w_{ij},$$

where $\mathcal{S}$ is the set of assets belonging to sector S.

Multiobjective optimization problem formulation

In this section, we will formulate the multi-account portfolio rebalancing problem as a multi-objective optimization problem. We first adjust the utility function of each fund to incorporate the trading costs, forming the *effective utility function*:

$$\bar{U}_i(\mathbf{x}) = U_i(\mathbf{w}_i + \mathbf{x}_i) - \sum_{j=1}^{m} \tau_j(x_{ij}, z_j). \tag{2.9}$$

Let $\mathbf{x} = (\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n) \in \mathbf{R}^{mn}$, $\mathbf{A} = \operatorname{diag}(\alpha_1, \alpha_2, \ldots, \alpha_m)$. Using (2.8), we can rewrite (2.9) as

$$\bar{U}_i(\mathbf{x}) = U_i(\mathbf{w}_i + \mathbf{x}_i) - \mathbf{x}_i^T \mathbf{A}\mathbf{z}. \tag{2.10}$$

The multi-account portfolio optimization problem facing fund managers can then be written as the following vector optimization problem:

$$\begin{aligned} \text{maximize} \quad & (\bar{U}_1(\mathbf{x}), \bar{U}_2(\mathbf{x}), \ldots, \bar{U}_n(\mathbf{x})) \\ \text{s.t.} \quad & \mathbf{x}_i \in \mathcal{C}_i, \quad \forall i = 1, \ldots, n. \end{aligned} \tag{2.11}$$

Because of the coupling market impact costs, the objectives are interdependent and the problem (2.11) is non-separable. A feasible trade vector $\mathbf{x}^*$ of problem (2.11) is called *Pareto-optimal* (or *Pareto-efficient*) if there does not exist any feasible trade vector $\mathbf{x} \neq \mathbf{x}^*$ such that $\bar{U}_i(\mathbf{x}) \geq \bar{U}_i(\mathbf{x}^*)$, $\forall i = 1, \ldots, n$ (we will revisit the concept of Pareto-optimal solution in Section 2.4 when we discuss about rational preference relations). There could be many feasible trade vectors that are Pareto-optimal, and the set of those Pareto-optimal vectors is called the *Pareto-efficient frontier*. When solving problem (2.11), the fund manager wants to find a Pareto-optimal solution that reflects his or her beliefs about the tradeoff between different clients' objective. In the next section, we will review some existing approaches of tackling problem (2.11) and discuss the advantages and drawbacks of these approaches.

2.3 Existing approaches

2.3.1 Independent scheme

In practice, many fund managers do recognize the presence of market impact cost but they often ignore concurrent trades from other portfolios. In other words, they try to optimize each fund's utility independently, taking into account only the impact cost resulting from that particular fund's trades. That is, they decide on the rebalancing vector $\mathbf{x}_i$ by solving the problem

$$\begin{aligned} \text{maximize} \quad & \boldsymbol{\mu}^T(\mathbf{w}_i + \mathbf{x}_i) - \lambda_i(\mathbf{w}_i + \mathbf{x}_i)^T \boldsymbol{\Sigma}(\mathbf{w}_i + \mathbf{x}_i) - \mathbf{x}_i^T \mathbf{A} \mathbf{x}_i \\ \text{s.t.} \quad & \mathbf{w}_i + \mathbf{x}_i \geq 0 \\ & \mathbf{1}^T \mathbf{x}_i = 0 \\ & \| \mathbf{x}_i \|_1 \leq \gamma_i \mathbf{1}^T \mathbf{w}_i \\ & (1 - \delta_{is}) \sum_{j \in S} w_{ij} \leq \sum_{j \in S} (w_{ij} + x_{ij}) \leq (1 + \delta_{is}) \sum_{j \in S} w_{ij}, \quad \forall S. \end{aligned} \tag{2.12}$$

Notice that the last term in the objective of (2.12) is $\mathbf{x}_i^T \mathbf{A} \mathbf{x}_i$, instead of $\mathbf{x}_i^T \mathbf{A} \mathbf{z}$ as in the effective utility function (equation (2.10)). This is because the fund manager only uses the trade information of fund i. Let $\mathbf{x}_i^b$ be the optimal solution of (2.12). The optimal net trade vector is then

$$\mathbf{z}^b = \mathbf{x}_1^b + \mathbf{x}_2^b + \cdots + \mathbf{x}_n^b.$$

The resulting effective utility for the ith fund is then given as:

$$\bar{U}_i^b(\mathbf{x}^b) = U_i(\mathbf{w}_i + \mathbf{x}_i^b) - \sum_{j=1}^{m} \tau_j(x_{ij}^b, z_j^b).$$

The performance of each fund under this scheme is not efficient due to the minimal usage of trading information of other funds. In particular, the resulting vector of effective utilities $(\bar{U}_1^b, \bar{U}_2^b, \ldots, \bar{U}_n^b)$ is not Pareto-optimal. From a computational view point, the optimization problems involved are convex and can be solved efficiently using quadratic optimization solvers.

2.3.2 Social Welfare scheme

In this scheme, we determine $\mathbf{x}_i$'s simultaneously by optimizing the sum of effective utilities of n funds:

$$\begin{aligned}
\text{maximize} \quad & \sum_{i=1}^{n}(\boldsymbol{\mu}^T(\mathbf{w}_i+\mathbf{x}_i) - \lambda_i(\mathbf{w}_i+\mathbf{x}_i)^T\boldsymbol{\Sigma}(\mathbf{w}_i+\mathbf{x}_i)) - \mathbf{z}^T A\mathbf{z} \\
\text{s.t.} \quad & \mathbf{w}_i + \mathbf{x}_i \geq 0, \quad \forall i = 1,\ldots,n \\
& \mathbf{1}^T\mathbf{x}_i = 0, \quad \forall i = 1,\ldots,n \\
& \| \mathbf{x}_i \|_1 \leq \gamma_i \mathbf{1}^T\mathbf{w}_i, \quad \forall i = 1,\ldots,n \\
& (1-\delta_{is})\sum_{j\in S} w_{ij} \leq \sum_{j\in S}(w_{ij}+x_{ij}) \leq (1+\delta_{is})\sum_{j\in S} w_{ij}, \quad \forall i, S \\
& \mathbf{z} = \sum_{i=1}^{n}\mathbf{x}_i.
\end{aligned} \tag{2.13}$$

Due to the aggregation, the net trading costs, $\mathbf{z}^T A\mathbf{z}$, are given by convex quadratic functions. Thus, (2.13) is a convex problem and can be solved efficiently. The sum of effective utilities reflects the social welfare idea which is popular in the microeconomics literature (see for e.g. [15]). The resulting vector of effective utilities

$$(\bar{U}_1^{soc}, \bar{U}_2^{soc}, \ldots, \bar{U}_n^{soc})$$

is Pareto-optimal, i.e., we can not increase a client's utility by reallocation trades without worsening the utility of another [11]. However, under this scheme, some clients can gain at the expense of others. One might argue that a modified version of Social Welfare scheme can be applied with different weights assigned to the effective utilities. However, such choice of weights is not clear and can be hardly justified.

2.3.3 Competitive Equilibrium scheme

This iterative scheme (proposed in [11]) borrows the idea of market competitive equilibrium from game theory. It stimulates a game in which players are competing for a scarce resource and are trying to maximize their own utilities given the information of other players. In the multiple account rebalancing problem, the scarce resource is the market liquidity. Under this scheme, we sequentially determine $\mathbf{x}_i$ by optimizing the ith fund's effective utility, considering the trades of other funds as constants. Specifically, let $\mathcal{N} = \{1, 2, \ldots, n\}$ and $i \in \mathcal{N}$. Given a trade vector $\bar{\mathbf{x}} = (\bar{\mathbf{x}}_1, \bar{\mathbf{x}}_2, \ldots, \bar{\mathbf{x}}_n) \in \mathbf{R}^{mn}$, we define a subproblem $\mathbf{P}\big(i, \sum_{i' \neq i} \bar{\mathbf{x}}_{i'}\big)$ as

$$
\begin{aligned}
\text{maximize} \quad & \boldsymbol{\mu}^T(\mathbf{w}_i + \mathbf{x}_i) - \lambda_i(\mathbf{w}_i + \mathbf{x}_i)^T \boldsymbol{\Sigma}(\mathbf{w}_i + \mathbf{x}_i) - \mathbf{x}_i^T A \mathbf{z} && (2.14) \\
\text{s.t.} \quad & \mathbf{w}_i + \mathbf{x}_i \geq 0 \\
& \mathbf{1}^T \mathbf{x}_i = 0 \\
& \| \mathbf{x}_i \|_1 \leq \gamma_i \mathbf{1}^T \mathbf{w}_i \\
& (1 - \delta_{is}) \sum_{j \in S} w_{ij} \leq \sum_{j \in S} (w_{ij} + x_{ij}) \leq (1 + \delta_{is}) \sum_{j \in S} w_{ij}, \quad \forall S && (2.15) \\
& \mathbf{z} = \mathbf{x}_i + \sum_{i' \neq i} \bar{\mathbf{x}}_{i'}.
\end{aligned}
$$

The iterative algorithm for finding the optimal trade vector under the Competitive Equilibrium scheme is as follows:

Algorithm 1: Competitive Equilibrium Algorithm

Step 1 Initialize $k = 0$, $\mathbf{x}^{(0)} = \mathbf{0} \in \mathbf{R}^{mn}$ and $\epsilon > 0$

Step 2 Let $\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)}$.

Step 3 For $i = 1, 2, \ldots, n$,

Solve $\mathbf{P}\left(i, \sum_{i' \neq i} \mathbf{x}_{i'}^{(k+1)}\right)$ and denote the optimal solution as $\mathbf{x}^*$.

Update $\mathbf{x}_i^{(k+1)} = \mathbf{x}^*$.

End for

Step 4 If $\| \mathbf{x}^{(k+1)} - \mathbf{x}^{(k)} \| < \epsilon$ then stop.

Else, increment $k := k + 1$ and go to Step 2.

If Algorithm 1 converges, the resulting optimal solution is the Nash-equilibrium solution of the market and thus is argued to be a fair solution. However, as discussed in [11], this is an ad-hoc scheme. There is no mathematical guarantee for convergence and convergence rate of the algorithm. In particular, the resulting effective utilities need not be Pareto optimal. As we will see later in the numerical examples, the outcome utility vector from this scheme is not Pareto optimal, even though it is fairer than the social scheme and more efficient than the independent scheme.

2.4 Fairness concepts

2.4.1 Happiness levels

Before our discussions of fairness concepts, we want to rewrite the multicriteria optimization problem (2.11) in a more appropriate form, which has the objectives representing the happiness levels of each clients. Clearly, the effective utilities are not good measures of clients' satisfaction of a trading cost allocation. For example, given a feasible trade vector $\mathbf{x}$ such that $\bar{U}_i(\mathbf{x}) > \bar{U}_k(\mathbf{x})$ for some $i \neq k$, it is inappropriate to conclude that client i is happier than client k. Indeed, client i could have incurred an amount of trading costs that exceeds what he or she normally pays when trading independently, while client k only receives a small portion of this, in which case client k should be happier than client i. The fact that $\bar{U}_i(\mathbf{x}) > \bar{U}_k(\mathbf{x})$ might be simply because the initial holdings of client i is much larger than the initial holdings of client k. A client's happiness level is intuitively the proportion of maximum allowable effective utility that the client achieves. The formal definition of the happiness level is as follows. First, we assume that the clients are not happy unless they get at least as much as they would get under the Independent scheme. We then refine the set $\mathcal{F}$ of feasible allocation vectors $\mathbf{x}$ as

$$\mathcal{F} = \left\{\mathbf{x} \in \mathbf{R}^{mn} | \mathbf{x}_i \in \mathcal{C}_i, \bar{U}_i(\mathbf{x}) \geq \bar{U}_i^b, \forall i = 1, \ldots, n\right\}, \tag{2.16}$$

where $\bar{U}_i^b = \bar{U}_i^b(\mathbf{x}^b)$ is the resulting effective utility of fund i under the Independent scheme (Section 2.3).

When an allocation is feasible, each client will compare his utility to the maximum utility he or she could get in the feasible set. The client is happier when his or her utility is closer to this maximum level. Let $h_i(\mathbf{x})$ denote the *happiness level* of player

i when allocation $\mathbf{x}$ is chosen. We define $h_i(\mathbf{x})$ as

$$h_i(\mathbf{x}) = \frac{\bar{U}_i(\mathbf{x}) - \bar{U}_i^b}{\bar{U}_i^{max} - \bar{U}_i^b}, \tag{2.17}$$

where the maximum achievable utilities $\bar{U}_i^{max} = \max_{\mathbf{x}\in\mathcal{F}} \bar{U}_i(\mathbf{x})$ can be found by solving

$$\begin{aligned} \text{maximize} \quad & \boldsymbol{\mu}^T(\mathbf{w}_i + \mathbf{x}_i) - \lambda_i(\mathbf{w}_i + \mathbf{x}_i)^T \boldsymbol{\Sigma}(\mathbf{w}_i + \mathbf{x}_i) - \mathbf{x}_i^T \mathbf{A}\mathbf{z} \\ \text{s.t.} \quad & \boldsymbol{\mu}^T(\mathbf{w}_k + \mathbf{x}_k) - \lambda_k(\mathbf{w}_k + \mathbf{x}_k)^T \boldsymbol{\Sigma}(\mathbf{w}_k + \mathbf{x}_k) - \mathbf{x}_k^T \mathbf{A}\mathbf{z} \geq \bar{U}_k^b, \ \forall k = 1, \ldots, n \\ & \mathbf{w}_k + \mathbf{x}_k \geq 0, \quad \forall k = 1, \ldots, n \\ & \mathbf{1}^T \mathbf{x}_k = 0, \quad \forall k = 1, \ldots, n \\ & \| \mathbf{x}_k \|_1 \leq \gamma_k \mathbf{1}^T \mathbf{w}_k, \quad \forall k = 1, \ldots, n \\ & (1 - \delta_{ks}) \sum_{j\in S} w_{kj} \leq \sum_{j\in S} (w_{kj} + x_{kj}) \leq (1 + \delta_{ks}) \sum_{j\in S} w_{kj}, \quad \forall S, \forall k \\ & \mathbf{z} = \sum_{i=1}^{n} \mathbf{x}_i. \end{aligned} \tag{2.18}$$

2.4.2 Equitable efficient solution

In this section, we will review the concept of *equitably efficient solution* [12]. First, we modify the multiple objective problem (2.11) with the objectives being replaced by the happiness levels and the feasible set being replaced by $\mathcal{F}$ as defined in previous section:

$$\begin{aligned} \text{maximize} \quad & (h_1(\mathbf{x}), h_2(\mathbf{x}), \ldots, h_n(\mathbf{x})) \\ \text{s.t.} \quad & \mathbf{x} \in \mathcal{F}. \end{aligned} \tag{2.19}$$

Intuitively, the outcome solution of (2.19) is fair if all the happiness levels are the same. Traditionally, fairness is quantified by *inequality measures.* For example, given an outcome vector $\mathbf{h} = (h_1, h_2, \ldots, h_n)$, the *mean absolute difference* is defined as:

$$\Gamma(\mathbf{h}) = \frac{1}{2n^2} \sum_{i=1}^{n} \sum_{k=1}^{n} |h_i - h_k|, \tag{2.20}$$

or the *maximum absolute difference* is defined as:

$$d(\mathbf{h}) = \max_{i,k} |h_i - h_k| . \tag{2.21}$$

More examples on inequality measures can be found in [12] and the references therein. An outcome $\mathbf{h}$ is often considered fair if inequality measures such as $\Gamma(\mathbf{h})$ or $d(\mathbf{h})$ are small. Unfortunately, simple minimization of the inequality measures to ensure fairness often results in inefficient outcomes for individuals. Thus, we need an allocation strategy that takes into account efficiency but still ensures fairness criteria of the allocation vector.

We will consider the formal definition of an equitably efficient solution. First of all, a *preference relation* is a model that specifies which one among two arbitrary outcome vectors is preferred. Given two outcome vectors $\mathbf{h}_1$ and $\mathbf{h}_2$, we denote $\mathbf{h}_1 \succeq \mathbf{h}_2$ if $\mathbf{h}_1$ is preferred to $\mathbf{h}_2$. In addition, the strict preference $\succ$ and indifference $\cong$ are defined as:

$$\mathbf{h}_1 \succ \mathbf{h}_2 \Leftrightarrow (\mathbf{h}_1 \succeq \mathbf{h}_2 \quad \text{and} \quad \mathbf{h}_2 \not\succeq \mathbf{h}_1)$$
$$\mathbf{h}_1 \cong \mathbf{h}_2 \Leftrightarrow (\mathbf{h}_1 \succeq \mathbf{h}_2 \quad \text{and} \quad \mathbf{h}_2 \succeq \mathbf{h}_1).$$

Definition 1. *A* **rational preference relation** *is a preference relation that satisfies the following properties:*

1. *Reflexive:*

$$\mathbf{h} \succeq \mathbf{h}.$$

2. *Transitive:*

$$(\mathbf{h}' \succeq \mathbf{h}'' \quad \textit{and} \quad \mathbf{h}'' \succeq \mathbf{h}''') \Rightarrow \mathbf{h}' \succeq \mathbf{h}'''.$$

3. *Strictly monotonic*

$$\mathbf{h} + \epsilon \mathbf{e}_i \succ \mathbf{h}, \quad \textit{for } \epsilon > 0, i = 1, \ldots, n,$$

where $\mathbf{e}_i \in \mathbf{R}^n$ *is the vector with its* i^{th} *element equals to one and other elements equal to zero.*

An outcome vector $\mathbf{h}_1$ is said to *rationally dominate* $\mathbf{h}_2$ iff $\mathbf{h}_1 \succ \mathbf{h}_2$ for all rational preference relations $\succ$. A feasible solution $\mathbf{x}$ to (2.19) is called *Pareto-efficient* iff the corresponding output vector $\mathbf{h}(\mathbf{x})$ is rationally nondominated. This is a more general definition of Pareto-efficiency than the one in Section 2.2. We can see that if the preference relation $\succ$ is simply an elemenwise comparison, i.e, $\mathbf{h} \succeq \mathbf{h}'$ iff $h_i \geq h'_i$, $\forall i = 1, \ldots, n$, then the two definitions become equivalent. In this thesis, we will mainly use the elemenwise comparison preference relation when talking about Pareto-efficiency.

Notice that rational preference relations only deal with efficiency. We now incorporate some fairness criteria into our preference model.

Definition 2. *A* **fair rational preference relation** *is a rational preference relation that satisfies two conditions:*

1. *If outcomes are interchanged between clients, the new outcome must be indifferent in terms of the preference relation. In other words, outcomes must be equally desirable under permutations:*

$$(h_{\pi(1)}, h_{\pi(2)}, \ldots, h_{\pi(n)}) \cong (h_1, h_2, \ldots, h_n),$$

   where $(\pi(1), \pi(2), \ldots, \pi(n))$ *is a permutation of* $(1, 2, \ldots, n)$.

2. *If by transferring outcomes between two clients, we achieve a smaller difference in outcomes between those clients then it will be more desirable compared to the previous outcome. This property is known as the* **principle of transfers***:*

$$h_{i'} > h_{i''} \Rightarrow \mathbf{h} - \epsilon \mathbf{e}_i + \epsilon \mathbf{e}_{i''} \succ \mathbf{h} \quad \textit{for } 0 < \epsilon < h_{i'} - h_{i''}.$$

Based on a fair preference relation, we call an outcome vector $\mathbf{h}_1$ *fairly (or equitably) dominates* $\mathbf{h}_2$ (denoted as $\mathbf{h}_1 \succ_e \mathbf{h}_2$) if $\mathbf{h}_1 \succ \mathbf{h}_2$ for all fair rational preference relation $\succ$.

Definition 3. *A feasible solution of* (2.19) *is called a fairly (equitably) efficient solution if the corresponding outcome vector* $\mathbf{h}(\mathbf{x})$ *is equitably nondominated.*

2.4.3 Max-Min fairness, Proportional fairness and α-fairness

One way to obtain equitably efficient solutions of (2.19) is to solve the corresponding *fair aggregation* problem:

$$\begin{aligned} \text{maximize} \quad & \sum_{i=1}^{n} u(h_i(\mathbf{x})) \\ \text{s.t.} \quad & \mathbf{x}_i \in \mathcal{C}_i, \quad \forall i = 1, \ldots, n, \end{aligned} \tag{2.22}$$

where $u : R \to R$ is a strictly concave, increasing utility function. It is shown in [5] that any Pareto-efficient solution of (2.22) is an equitably efficient solution of (2.19).

In this thesis, we will use a class of parametric utility function $u(h, \alpha)(\alpha > 0)$ defined as

$$u(h, \alpha) = \begin{cases} \dfrac{h^{1-\alpha}}{1-\alpha} & \text{if } \alpha \neq 1 \\ \log(h) & \text{if } \alpha = 1. \end{cases} \tag{2.23}$$

The fair aggregation problem (2.22) with utility function u as defined in (2.23) is called the **α-fairness scheme**. The parameter α quantifies the degree of fairness we want to ensure. As α increases, the solution gets fairer (the inequality measures of outcome vector are smaller). For the case of $\alpha = 1$, the problem (2.22) corresponds to the so-called **Proportional fairness (PF) scheme** proposed in [4]. The solution of PF maximizes the product of additional utilities compared to the status-quo (this is commonly known as *Nash criterion* [10]). In particular, the solution to the PF is fair and Pareto-optimal.

For the case of $\alpha \to \infty$, problem (2.22) becomes the so-called **Max-Min fairness (MMF) scheme**. Intuitively speaking, MMF maximizes the smallest happiness level, and then the second smallest happiness level, and the third smallest one, and so on. This is a lexicographical optimization problem of the sorted happiness level vector $\mathbf{h}$. The MMF problem is mathematically defined as follows. Let X be the set of feasible $\mathbf{x}$ in (2.19). A vector $\mathbf{h}$ is called lexicographically greater than $\mathbf{h}'$ if there exists an index $k, 0 \leq k < n$, such that $h_i = h'_i \; \forall i \leq k$ and $h_{k+1} > h'_{k+1}$. We will write $\mathbf{h} \succeq_{lex} \mathbf{h}'$ to indicate that $\mathbf{h}$ is lexicographically greater than $\mathbf{h}'$. Let $\langle \mathbf{h}(\mathbf{x}) \rangle$ be the vector obtained by sorting $\mathbf{h}(\mathbf{x})$ in non-decreasing order. The MMF problem is

$$\text{Find} \quad \mathbf{x}^* \in X \quad \text{s.t.} \quad \langle \mathbf{h}(\mathbf{x}^*) \rangle \succeq_{lex} \langle \mathbf{h}(\mathbf{x}) \rangle, \quad \forall \mathbf{x} \in X. \tag{2.24}$$

The solution $\mathbf{x}^*$ of the MMF problem is considered absolutely fair in the sense that if we deviate from $\mathbf{x}^*$, smaller happiness levels decreases while the higher happiness levels increases (which is a more unfair outcome). The lexicographical optimization problem (2.24) is not a mathematical programming problem and finding $\mathbf{x}^*$ directly is not easy. In Chapter 3, we will show how MMF is adopted in our problem and present an iterative algorithm to find the optimal solution $\mathbf{x}^*$.

Finally, we want to emphasize that in the α-fairness scheme above, the parameter α is strictly positive. If $\alpha = 0$, we can easily see that $u(h) = h$ and problem (2.22) is simply the maximization of the sum of happiness levels. This is a similar version of the Social Welfare scheme (because $h_i(\mathbf{x})$ is a linear function of $\bar{U}_i(\mathbf{x}), \; \forall i = 1, 2, \ldots, n$), except that $h_i(\mathbf{x})$'s are better measures of clients' satisfaction of an allocation vector. The optimal solution to (2.22) when $\alpha = 0$ is a Pareto-optimal solution of (2.19) but it needs not be an equitably efficient solution because the strict concavity condition of $u(h)$ is violated. Similarly, the solution under the Social Welfare scheme needs not be an equitably efficient solution.

In the next chapter, we will adopt the fair aggregation problem (2.22) to our problem. We show that the resulting optimization problems have non-convex objectives and non-convex constraints, which are NP-hard in general. We then provide a solution method to the non-convex problems based on a local search algorithm and discuss its performance on these problems.

Chapter 3

Fairness in Trading

In this chapter, we discuss the adoption of Max-Min fairness, Proportional fairness and α-fairness to the multiple portfolios rebalancing problem with quadratic market impact cost. The structure of this chapter is as follows. In Section 3.1, we discuss the formulation of the multi-account portfolio rebalancing problem under the Max-Min fairness scheme and introduce an iterative algorithm to solve this problem. Section 3.2 follows with a discussion on the local search method that we use for solving the resulting nonconvex quadratically constrained quadratic programs. We then extend this solution method to solve the resulting concave maximization problems with nonconvex constraints under Proportional fairness and α-fairness schemes in Section 3.3. Finally, in Section 3.4, we provide a deterministic global optimization algorithm based on branch-and-bound method that we will use to compare with our local search method.

3.1 Max-Min fairness

Given the definition of $\mathbf{h}(\mathbf{x})$ as in (2.17) and the refined feasible set $\mathcal{F}$ of allocation vectors $\mathbf{x}$, we rewrite the MMF problem (2.24) as

$$\begin{aligned} \text{lexmax} \quad & \langle \mathbf{h}(\mathbf{x}) \rangle \\ \text{s.t.} \quad & \mathbf{x} \in \mathcal{F}, \end{aligned} \tag{3.1}$$

where $\langle \mathbf{h}(\mathbf{x}) \rangle$ is the vector $\mathbf{h}(\mathbf{x})$ sorted in non-decreasing order.

In order to solve (3.1), we adopt an iterative algorithm introduced by Pioro et al. [16]. Let $\mathcal{N} = \{1, 2, \ldots, n\}$ and $B \subseteq \mathcal{N}$. Also, let $\mathbf{t}^B = \{t_i^B : i \in B\}$. The B-vector $\mathbf{t}^B$ represents the *blocking happiness levels* and B represents the set of clients with happiness levels at least $\mathbf{t}^B$. Let $B' = \mathcal{N} \setminus B$ be the complement of B. Given a set B and the blocking vector $\mathbf{t}^B$, we define a subproblem $\mathbf{P}(B, \mathbf{t}^B)$ that will be used in the iterative algorithm as:

$$\begin{aligned} \text{maximize} \quad & y \\ \text{s.t.} \quad & h_i(\mathbf{x}) \geq y, \quad \forall i \in B' \\ & h_i(\mathbf{x}) \geq t_i^B, \quad \forall i \in B \\ & \mathbf{x} \in \mathcal{F}. \end{aligned} \tag{3.2}$$

Clearly, if we have already known the set of clients B whose happiness levels attain maximum values at $\mathbf{t}^B$ then problem (3.2) maximizes the $(|B|+1)^{\text{th}}$ smallest happiness level. Also, for a given value y^0 of y in (3.2) and an iteration index k, we define the test problem $\mathbf{T}(B, \mathbf{t}^B, y^0, k)$ as:

$$\begin{aligned} \text{maximize} \quad & h_k(\mathbf{x}) \\ \text{s.t.} \quad & h_i(\mathbf{x}) \geq y^0, \quad \forall i \in B' \setminus \{k\} \\ & h_i(\mathbf{x}) \geq t_j^B, \quad \forall i \in B \\ & \mathbf{x} \in \mathcal{F}. \end{aligned} \tag{3.3}$$

The iterative algorithm for solving (3.1) is shown in Algorithm 2. In Step 2 of this iterative algorithm, we find the maximum value for the smallest happiness level and add this value to $\mathbf{t}^B$. We then perform a test to determine which clients correspond to this smallest happiness level and add these clients to the blocking set B. In the next iteration, we solve the problem $\mathbf{P}(B, \mathbf{t}^B)$ to get the maximum value of the second smallest happiness level. Similarly, we then solve $\mathbf{T}(B, \mathbf{t}^B, y^0, k)$ and update the blocking set B as well as the blocking levels $\mathbf{t}^B$.

Algorithm 2: MMF Algorithm

Step 1 Initialize $B = \emptyset$ and $\mathbf{t}^B = \emptyset$

Step 2 If $B = \mathcal{N}$ then stop, $\mathbf{t}^B$ is the optimal outcome vector.

Else, solve $\mathbf{P}(B, \mathbf{t}^B)$ and denote the optimal solution as $(\mathbf{x}^0, y^0)$

Step 3 For each $k \in B'$ such that $h_k(\mathbf{x}^0) = y^0$, solve the test problem $\mathbf{T}(B, \mathbf{t}^B, y^0, k)$. Let $\mathbf{x}^1$ be the optimal solution to $\mathbf{T}(B, \mathbf{t}^B, y^0, k)$.

If $h_k(\mathbf{x}^1) = y^0$ then put $B = B \cup \{k\}$ and $t_k^B = y^0$

Step 4 Go to Step 2

When h_i's are concave functions and $\mathcal{F}$ is a convex set, this process terminates when $B = \mathcal{N}$, in which case the obtained solution $\mathbf{x}^0$ is the optimal solution to the MMF problem and the optimal happiness level vector is given as $\mathbf{t}^B$. In addition, under the most favorable scenario when all the happiness levels are the same, the MMF algorithm requires solving $n + 1$ subproblems (problem $\mathbf{P}(\emptyset, \emptyset)$ and n test problems $\mathbf{T}(B, \mathbf{t}^B, y^0, k)$). However, in the worst case, the MMF algorithm requires solving $O(n^2/2)$ subproblems. Because h_i's are non-concave functions and $\mathcal{F}$ is a non-convex set in our case, the adoption of this algorithm is purely a heuristic approach that has no guarantee for convergence. However, from our experimental experiences, the algorithm performs well when we have good convex approximations to the resulting subproblems, as we will discuss shortly. Interested readers are referred to [13] for alternative sequential algorithms that can be applied to general non-convex MMF problems.

3.2 Local search method

The problem of finding the maximum achievable effective utility $\bar{U}_i^{max}$ (problem (2.18)) and the subproblems $\mathbf{P}(B, \mathbf{t}^B)$ and $\mathbf{T}(B, \mathbf{t}^B, y^0, k)$ used in the MMF algorithm (problems (3.2) and (3.3)) are classified as *non-convex Quadratically Constrained Quadratic Programs (QCQPs).* The non-convexity comes from indefinite trading cost terms $\mathbf{x}_i^T \mathbf{A}\mathbf{z}$'s, which make $\bar{U}_i(\mathbf{x})$'s and $\mathbf{h}_i(\mathbf{x})$'s neither convex nor concave functions of $\mathbf{x}$. The general non-convex QCQP can be written in the form:

$$\begin{aligned} \text{maximize} \quad & \mathbf{x}^T Q_0 \mathbf{x} + b_0^T \mathbf{x} + c_0 \\ \text{s.t.} \quad & \mathbf{x}^T Q_i \mathbf{x} + b_i^T \mathbf{x} + c_i \geq 0, \quad \forall i = 1, \ldots, m \\ & \mathbf{x} \in \mathcal{G}, \end{aligned} \tag{3.4}$$

where $\mathbf{x} \in \mathbf{R}^n$, $Q_i \in \mathbf{S}^{n\times n}$ (the set of symmetric matrices of dimension $n \times n$) are indefinite matrices, $b_i \in \mathbf{R}^n$, $c_i \in \mathbf{R}$ and $\mathcal{G}$ is a polyhedral which can be specified by general linear constraints. For convenience in discussion of the methodology, we will use $\mathbf{x}, m, n$ in this section as anonymous variables and parameters instead of the definitions given in Section 2.2. It is commonly known that non-convex QCQP is NP-hard, which means the computation time typically grows exponentially with problem dimension. Therefore, it is often extremely hard to solve a non-convex QCQP globally, especially when the problem involves several hundreds of variables.

There are several ways to deal with the non-convex QCQP problem. Relaxation methods such as semidefinite relaxations or Lagrangian relaxations can provide upper bounds for problem (3.4). For example, the semidefinite relaxation of (3.4) is given as:

$$\begin{aligned} \text{maximize} \quad & \text{Trace}(\mathbf{X} Q_0) + b_0^T \mathbf{x} + c_0 \\ \text{s.t.} \quad & \text{Trace}(\mathbf{X} Q_i) + b_i^T \mathbf{x} + c_i \geq 0, \quad \forall i = 1, \ldots, m \\ & \begin{bmatrix} \mathbf{X} & \mathbf{x} \\ \mathbf{x}^T & 1 \end{bmatrix} \succeq 0, \\ & \mathbf{x} \in \mathcal{G} \\ & \mathbf{X} \in \mathbf{S}^{n\times n}. \end{aligned} \tag{3.5}$$

However, such convex relaxation approaches do not provide any good feasible point as an outcome. We adopt here a local search algorithm that starts at some feasible point $\mathbf{x}_0$ and works sequentially to improve the objective until it converges to a local maximum. At each iteration k, we solve the convex approximation of problem (3.4) around the current solution $\mathbf{x}_k$ and assign $\mathbf{x}_{k+1}$ as the maximizer of this approximation problem. The convex approximation of (3.4) can be obtained by using the so-called *linearization techniques*. For example, consider the objective of (3.4), we can write the eigenvalue decomposition of matrix Q_0 as $Q_0 = HDH^{-1}$, with $D \in \mathbf{R}^n$ is the diagonal matrix of eigenvalues. We then decompose D as $D = D_1 - D_2$ where D_1 and D_2 are diagonal matrices with nonnegative diagonal elements. Matrix Q_0 can then be written as the difference between two positive semidefinite matrices:

$$\begin{aligned} Q_0 &= H(D_1 - D_2)H^{-1} = HD_1H^{-1} - HD_2H^{-1} \\ Q_0 &= Q_0^+ - Q_0^-. \end{aligned} \tag{3.6}$$

We want to linearize the convex part $\mathbf{x}^T Q_0^+ \mathbf{x}$ in the objective function by using the first order Taylor's series approximation around some current feasible solution $\mathbf{x}_k$:

$$\begin{aligned} \mathbf{x}^T Q_0 \mathbf{x} + b_0^T \mathbf{x} + c_0 &= \mathbf{x}^T Q_0^+ \mathbf{x} - \mathbf{x}^T Q_0^- \mathbf{x} + b_0^T \mathbf{x} + c_0 \\ &\geq \mathbf{x}_k^T Q_0^+ \mathbf{x}_k^T + 2\mathbf{x}_k^T Q_0^+ (\mathbf{x} - \mathbf{x}_k) - \mathbf{x}^T Q_0^- \mathbf{x} + b_0^T \mathbf{x} + c_0. \end{aligned} \tag{3.7}$$

Similarly, the indefinite quadratic constraints in (3.4) can be approximated by concave quadratic lower bounds as

$$\mathbf{x}^T Q_i \mathbf{x} + b_i^T \mathbf{x} + c_i \geq \mathbf{x}_k^T Q_i^+ \mathbf{x}_k^T + 2\mathbf{x}_k^T Q_i^+ (\mathbf{x} - \mathbf{x}_k) - \mathbf{x}^T Q_i^- \mathbf{x} + b_i^T \mathbf{x} + c_i. \tag{3.8}$$

The convex approximation $\mathbf{P}(k)$ of problem (3.4) on a small region $B(\mathbf{x}_k, \epsilon)$ around the feasible solution $\mathbf{x}_k$ is then given as

$$\begin{aligned} \text{maximize} \quad & \mathbf{x}_k^T Q_0^+ \mathbf{x}_k^T + 2\mathbf{x}_k^T Q_0^+ (\mathbf{x} - \mathbf{x}_k) - \mathbf{x}^T Q_0^- \mathbf{x} + b_0^T \mathbf{x} + c_0 \\ \text{s.t.} \quad & \mathbf{x}_k^T Q_i^+ \mathbf{x}_k^T + 2\mathbf{x}_k^T Q_i^+ (\mathbf{x} - \mathbf{x}_k) - \mathbf{x}^T Q_i^- \mathbf{x} + b_i^T \mathbf{x} + c_i \geq 0, \quad \forall i = 1, \ldots, m \\ & \mathbf{x} \in \mathcal{G} \cap B(\mathbf{x}_k, \epsilon). \end{aligned} \tag{3.9}$$

The local search method for solving the non-convex QCQP problem using linearization technique is defined formally below:

Algorithm 3: Local Search Algorithm

Step 1 Initialize $k = 0$, $\xi > 0$, $\mathbf{x}_0$ is a feasible solution to (3.4)

Step 2 If $\| \mathbf{x}_{k+1} - \mathbf{x}_k \| < \xi$ then stop

Else solve $\mathbf{P}(k)$ and let $\mathbf{x}_{k+1}$ be the optimal solution

Step 3 Increment k and go to Step 2

Theorem 1. *The solutions $\mathbf{x}_0, \mathbf{x}_1, \ldots, \mathbf{x}_k, \ldots$ produced by the above local search algorithm are feasible to the problem* (3.4) *and satisfy*

$$f(\mathbf{x}_0) \leq f(\mathbf{x}_1) \leq \cdots \leq f(\mathbf{x}_k) \leq \ldots ,$$

where $f(\mathbf{x})$ is the objective function of (3.4).

Proof. We will prove by induction. The conclusion is obvious for $k = 0$. Assume that the conclusion holds for $k \geq 0$. Let $\mathbf{x}_{k+1}$ be the optimizer of problem $\mathbf{P}(k)$. Because $\mathbf{x}_{k+1}$ is feasible to $\mathbf{P}(k)$, we have

$$\mathbf{x}_k^T Q_i^+ \mathbf{x}_k^T + 2\mathbf{x}_k^T Q_i^+ (\mathbf{x}_{k+1} - \mathbf{x}_k) - \mathbf{x}_{k+1}^T Q_i^- \mathbf{x}_{k+1} + b_i^T \mathbf{x}_{k+1} + c_i \geq 0, \quad \forall i = 1, \ldots, m \quad (3.10)$$

Using (3.8) with $\mathbf{x} = \mathbf{x}_{k+1}$ and (3.10), we get

$$\mathbf{x}_{k+1}^T Q_i \mathbf{x}_{k+1} + b_i^T \mathbf{x}_{k+1} + c_i \geq 0, \quad \forall i = 1, \ldots, m \quad (3.11)$$

Furthermore, we have $\mathbf{x}_{k+1} \in \mathcal{G}$. Hence, $\mathbf{x}_{k+1}$ is a feasible solution to the problem (3.4). In addition, let $f_k(\mathbf{x})$ be the objective function of the convex approximation problem $\mathbf{P}(k)$ (3.9). From the optimality of $\mathbf{x}_{k+1}$, we have

$$f(\mathbf{x}_{k+1}) \geq f_k(\mathbf{x}_{k+1}) \geq f_k(\mathbf{x}_k) = f(\mathbf{x}_k).$$

Thus, the conclusion also holds for $k + 1$. By induction, the conclusion holds for all

$k \geq 0$. □

It is clear from Theorem 1 that given a proper choice of ϵ and ξ, the optimal solution $\mathbf{x}^*$ obtained from the local search algorithm is at least a local maximum of problem (3.4). An advantage of this approach over the Newton-like methods is that it does not stop at a saddle point. Furthermore, as we can see in the next section, the algorithm can be easily extended to solve the resulting problems under Proportional fairness and α-fairness schemes.

3.3 Proportional fairness and α-fairness

In this section, we will write the explicit form of multi-account rebalancing problem under Proportional fairness and α-fairness scheme. First, we introduce additional decision variables h_i. The problem formulation under Proportional fairness scheme is then given as

$$\begin{aligned} \text{maximize} \quad & \sum_{i=1}^{n} \log(h_i) \\ \text{s.t.} \quad & h_i \leq \frac{\bar{U}_i(\mathbf{x}) - \bar{U}_i^b}{\bar{U}_i^{max} - \bar{U}_i^b}, \quad \forall i = 1, \ldots, n \\ & \mathbf{x} \in \mathcal{F}. \end{aligned} \tag{3.12}$$

Similarly, the problem formulation under α-fairness scheme is

$$\begin{aligned} \text{maximize} \quad & \sum_{i=1}^{n} \frac{h_i^{1-\alpha}}{1-\alpha} \\ \text{s.t.} \quad & h_i \leq \frac{\bar{U}_i(\mathbf{x}) - \bar{U}_i^b}{\bar{U}_i^{max} - \bar{U}_i^b}, \quad \forall i = 1, \ldots, n \\ & \mathbf{x} \in \mathcal{F}. \end{aligned} \tag{3.13}$$

We notice that the objectives of (3.12) and (3.13) are concave functions. The first constraints in these two problems are non-convex because $\bar{U}_i(\mathbf{x})$'s are non-concave functions of $\mathbf{x}$. The set $\mathcal{F}$ is a non-convex set for the same reason. However, we notice that the feasible sets of (3.12) and (3.13) have the same form as in the problem (3.4). Thus, the local search algorithm proposed in Section 3.2 can be similarly applied for problems (3.12) and (3.13). The only difference is that at each iteration, we only need to apply linearization technique to the non-convex constraints.

3.4 Global optimization

To further facilitate our argument that the local search algorithm proposed in Section 3.2 is an efficient approach for solving the non-convex problems in our application, we compare it with a global optimization algorithm. Global optimization is a branch of nonlinear programming that deals with absolutely optimal solution of complicated non-convex programs. It has wide application in various fields such as engineering design, process control, biotechnology, etc. The most popular techniques in global optimization are deterministic approaches (such as branch and bound or interval analysis) and stochastic heuristics (such as genetic algorithm, simulated annealing, etc). In this section, we will review briefly a branch and bound algorithm for solving non-convex problem (3.4) that we use to compare with our local search algorithm.

The main idea of branch and bound (BNB) algorithm is to recursively divide a problem into subproblems until a solution with a desired level of optimality is obtained. A general BNB algorithm includes four processes: *Branching*, *Selection*, *Bounding* and *Elimination*. Branching involves dividing the box containing the feasible region of the current problem into smaller boxes. The new subproblems on these smaller boxes are created and added to the current list Λ of subproblems. Selection refers to the process of selecting the appropriate subproblem in Λ to process next. Bounding involves solving the relaxation of current subproblem to find an upper bound (assuming we are solving the maximization problem). Finally, Elimination is the process of deleting subproblems that are either infeasible or suboptimal. The details of the BNB algorithm are provided in Algorithm 4 [3].

Algorithm 4: Branch and Bound Algorithm

Step 1 Determine a set B_1 enclosing feasible region $\mathcal{X}$ of (3.4)

Step 2 Determine an upper bound f_1^U on B_1 and a feasible point $\mathbf{x}_1 \in B_1 \cap \mathcal{X}$

Step 3 If $\nexists \mathbf{x}_1$ then STOP.

Else let $f^L := f(\mathbf{x}_1)$, store B_1 in Λ, $r := 1$

Step 4 If $\Lambda = \emptyset$ then STOP.

Step 5 Remove a node $B \in \Lambda$ and split it into smaller nodes $B_{r+1}, \ldots, B_{r+h}$

Step 6 Determine upper bounds $f_{r+1}^U, \ldots, f_{r+h}^U$

Step 7 For $p := r+1$ to $r+h$ do

if($B_p \cap \mathcal{X} = \emptyset$)

$f_p^U := -\infty$

if($f_p^U > f_L$)

determine a feasible point $\mathbf{x}_p$ and $f_p := f(\mathbf{x}_p)$

if($f_p > f_L$)

$f_L := f_p$

remove all $B_k \in \Lambda$ with $f_k^U < f^L$

if ($f_p < f_L - \delta$)

save $\mathbf{x}_p$ as approximation of the optimum

elseif(size(B_p) $\geq \epsilon$) store B_p in Λ

End for

Step 8 Increment $r := r + h$ and go to Step 4

Even though BNB method is appealing in terms of the ability to provide global optimality, it is often intractable for large problems with hundreds of variables and constraints as in our case. However, we can use the BNB algorithm as a good benchmark to test the performance of our proposed algorithm for instances of small size. The analysis of the local search algorithm's performance on small size problems with similar structures could give hints to the behaviors of the method on large scale problems in practice.

Chapter 4

Computational Results

In this chapter, we present computational experiments with real world financial data. We will show that by applying the fairness schemes discussed in Chapter 3, we can obtain solutions that dominate those obtained from the Independent and Competitive Equilibrium schemes. In addition, we show that our approach gives a better tradeoff between fairness and efficiency as compared to the Social Welfare scheme. Finally, we compare our local search approach and the BNB algorithm in terms of optimality and running time for small size problems with similar structures. We show that in these particular cases, our method actually can provide an optimal solution in a much shorter computation time. All numerical experiments are implemented using YALMIP [6], a MATLAB-based advanced modeling language for convex and non-convex optimization problems. All convex optimization problems are modeled in YALMIP and solved by calling the semidefinite programing solver SEDUMI. The BNB algorithm is carried out by a specialized global optimization module in YALMIP, which calls linear solver GLPK for linear relaxation problems, and the general purpose solver SNOPT for finding good feasible solutions.

The structure of this chapter is as follows. In Section 4.1, we describe the financial data that are used in our computational experiments. Section 4.2 follows with the results obtained from Scenario I that has 2 funds and 73 stocks. We present in Section 4.3 the results for a more complicated scenario, which has 6 funds and 73 stocks. Finally, in Section 4.4 we show the comparison of our proposed local search algorithm and the BNB method as described in Chapter 3.

4.1 Data description

Our computational experiments are based on financial data used in [18]. Specifically, we have historical information for a universe of 73 stocks from January 2005 to February 2007. The mean returns $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$ are estimated from these historical data. All turnover parameters γ_i and δ_{is} are set to 5%. Risk aversion parameters are generated randomly in the interval $\lambda_i \in [5 \cdot 10^{-6}, 30 \cdot 10^{-6}]$. Market price impact parameters α_j are confidential information which are usually not available to public so we set α_j to an arbitrary (but meaningful) value of 0.125.

Using the above settings, we generate two scenarios to test our algorithms. In Scenario I, there are 2 funds invested in 73 stocks with total initial investment of \$500 million and \$1 billion. We assume for this case that all stocks are in one sector. In Scenario II, there are 6 funds with initial investment of \$10 million for fund 6 and \$1 billion for other funds. There are 6 non-overlapping sectors $S_1, \ldots, S_6$. Funds 1, 2 invested in S_1, S_2, S_3; funds 3, 4, 5 invested in S_4, S_5, S_6 and fund 6 invested in all sectors. In this setting, we want to see the behavior of different fairness schemes when there are two common types of resources (i.e., two sector groups) to be allocated to 6 clients. The presence of fund 6 introduces a coupling effect between the two groups of sectors.

4.2 Scenario I

Table 4.1 shows the effective utilities of 2 funds under different fairness schemes. The general α-fairness scheme is solved for $\alpha = 0.1, 0.5, 1(\text{Proportional fairness}), 2, 4, 6$. The simple setting of this first scenario enables us to see the behavior of the outcomes under different fairness schemes. We can see that in terms of efficiency, the Independent scheme is the least favorable scheme. Meanwhile, the Social Welfare scheme gives the best total throughput. For the general α-fairness scheme (including Proportional and MMF fairness), we can see that as α increases, efficiency decreases. It is interesting to note that all the effective utility vectors obtained from α-fairness, Proportional fairness and MMF dominate that of the Competitive Equilibrium scheme. The observation suggests that market equilibrium does not guarantee an efficient solution in this case. This is mainly because of the present of a central decision maker (the fund manager) who has complete trading information of all the clients.

Table 4.1: Effective utilities (in million \$) under Scenario I.

Scheme	Fund 1	Fund 2	Total
Independent	31.7836	9.5213	41.3049
Max. allowable	31.9316	9.5998	N/A
Social	31.9236	9.5596	41.4832
Comp. Equi.	31.8609	9.5606	41.4215
$\alpha = 0.1$	31.9221	9.5606	41.4826
$\alpha = 0.5$	31.9125	9.5648	41.4772
$\alpha = 1$ (PF)	31.9042	9.5681	41.4723
$\alpha = 2$	31.8971	9.5707	41.4678
$\alpha = 4$	31.8918	9.5725	41.4643
$\alpha = 6$	31.8896	9.5733	41.4629
MMF	31.8846	9.5749	41.4595

In Table 4.2, we show the happiness levels of the two funds under different fairness schemes. Due to our assumption that the Independent scheme is the worst case, happiness levels of clients under this scheme are zeros. Under the Social Welfare scheme, even though the outcome vector is Pareto-optimal, we can see that the first client is much more favored than the second client, which is an unfair situation. We can also see in Table 4.2 that the clients are almost equally happy under the Competitive Equilibrium scheme. However, this observation might not be true in a more complicated scenario, as we will see in Section 4.3. For the general α-fairness scheme, as α increases, the total efficiency is traded off with equity among clients. In particular, under the MMF scheme, the two clients are equally happy. The tradeoff between efficiency and fairness is shown in Figure 4-1. We can see that MMF is an extreme point on this tradeoff curve that corresponds to the most equitable outcome. On the other hand, the Social Welfare scheme is at the other extreme of the tradeoff curve, which corresponds to the most efficient outcome, but with no fairness properties at all. From Figure 4-1, we can see the price that one needs to pay to guarantee a desired level of equity among the clients. For example, one needs to sacrifice approximately 4% of efficiency to achieve an absolutely fair solution under the MMF scheme. Such a price of fairness can be reduced if the decision maker's preference toward fairness criteria is less extreme. Figure 4-1 also shows the clear dominance of the α-fairness scheme to the Competitive Equilibrium scheme. At the same level of fairness obtained by the Competitive Equilibrium scheme, the solution obtained from our proposed α-fairness scheme can significantly improve the efficiency (approximately 17% more efficient).

Table 4.2: Happiness levels (in %) under Scenario I.

Scheme	Fund 1	Fund 2
Social	94.61	48.82
Comp. Equi.	52.22	50.05
$\alpha = 0.1$	93.59	50.01
$\alpha = 0.5$	87.09	55.36
$\alpha = 1$ (PF)	81.49	59.61
$\alpha = 2$	76.69	62.95
$\alpha = 4$	73.09	65.29
$\alpha = 6$	71.63	66.21
MMF	68.26	68.26

4.3 Scenario II

Tables 4.3 and 4.4 show the effective utilities and happiness levels under different fairness schemes in Scenario II. Due to the similarity of results from the α-fairness scheme, we only present in Tables 4.3 and 4.4 the effective utilities and happiness levels under Proportional fairness and MMF schemes. We can see from Tables 4.3 and 4.4 the similar behaviors as observed in Scenario I. As we can see from Table 4.4, Social Welfare is clearly an unfair scheme because fund 2 is extremely favored (7.36% better than the maximum achievable level) while fund 1 is not happy at all (9.43% worse than the baseline level). Another interesting observation is that the Competitive Equilibrium scheme does not scale well when there are more funds. In

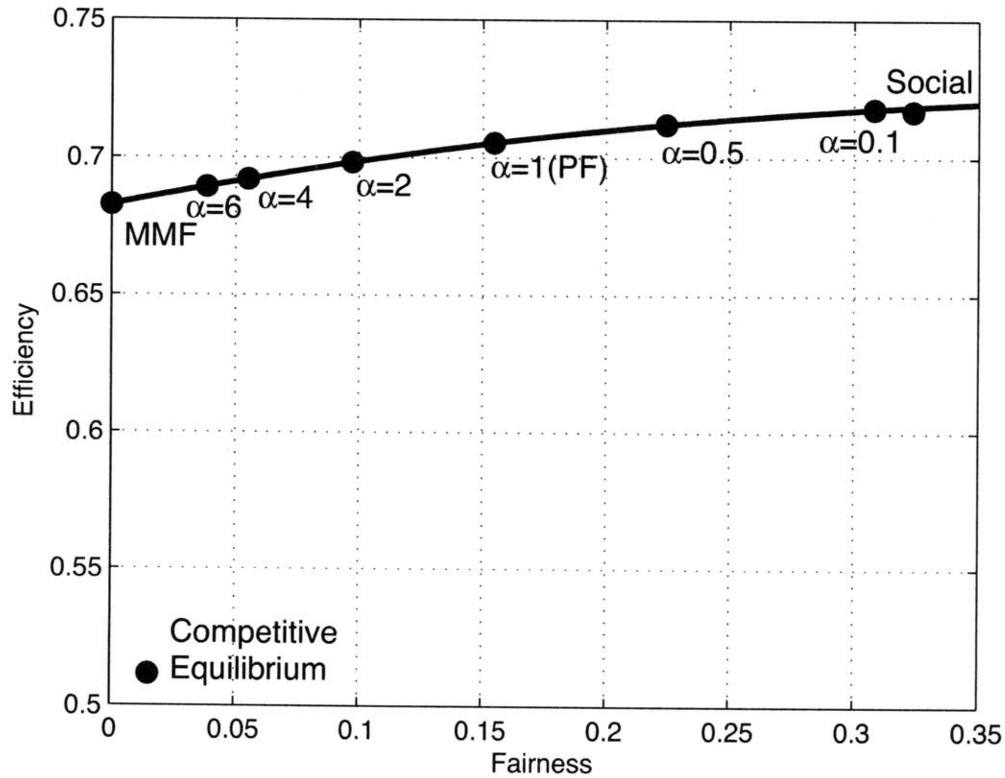

Figure 4-1: Tradeoff between efficiency and fairness under Scenario I. x-axis is the standard deviation of the happiness levels and y-axis is the mean of the happiness levels. The solid curve is the fitting for α-fairness data points.

contrast to the fairly flat distribution of happiness levels obtained from the Competitive Equilibrium scheme under Scenario I, the happiness level distribution under Scenario II for this fairness scheme is substantially diverse. From Table 4.4, we can also see the happiness levels of clients under Proportional fairness and MMF are very close to each other, especially for the case of MMF. We recall that, under the settings of this scenario, there are 2 sector groups (namely $\{S_1, S_2, S_3\}$ and $\{S_4, S_5, S_6\}$), and all the funds (except for fund 6) are investing in either of these sector groups. Intuitively, under the absolute fair situation, the outcome happiness levels of all the funds investing in the same sector group should be the same. This is exactly what we can observe from the results show in Table 4.4. Funds 1 and 2 are invested in the first sector group $\{S_1, S_2, S_3\}$ and have equal happiness levels under MMF scheme. Similarly, funds 3, 4 and 5 have equal happiness levels under MMF scheme because they are invested in the same sector group $\{S_4, S_5, S_6\}$. In Figure 4-2, we plot the bar chart of the happiness levels under different schemes. This helps illustrate further the observations that we discussed above.

Table 4.3: Effective utilities (in million \$) under Scenario II.

Scheme	Fund 1	Fund 2	Fund 3	Fund 4	Fund 5	Fund 6	Total
Independent	12.2478	12.2264	14.5260	12.4109	12.2004	0.1145	63.7260
Max. allowable	12.2637	12.2427	14.5854	12.4828	12.2651	0.1162	N/A
Social	12.2463	12.2439	14.5515	12.4564	12.2163	0.1147	63.8291
Comp. Equi.	12.2485	12.2295	14.5508	12.4309	12.2255	0.1154	63.8006
Proportional	12.2553	12.2340	14.5510	12.4449	12.2260	0.1156	63.8268
MMF	12.2547	12.2335	14.5513	12.4415	12.2279	0.1152	63.8241

Table 4.4: Happiness levels (in %) under Scenario II.

Scheme	Fund 1	Fund 2	Fund 3	Fund 4	Fund 5	Fund 6
Social	-9.43	107.36	42.93	63.28	24.57	11.76
Comp. Equi	4.40	19.02	41.75	27.82	38.79	52.94
Proportional	47.17	46.63	42.09	47.29	39.57	64.71
MMF	43.40	43.56	42.59	42.56	42.50	41.18

Figure 4-3 shows the tradeoff between efficiency and fairness in different fairness

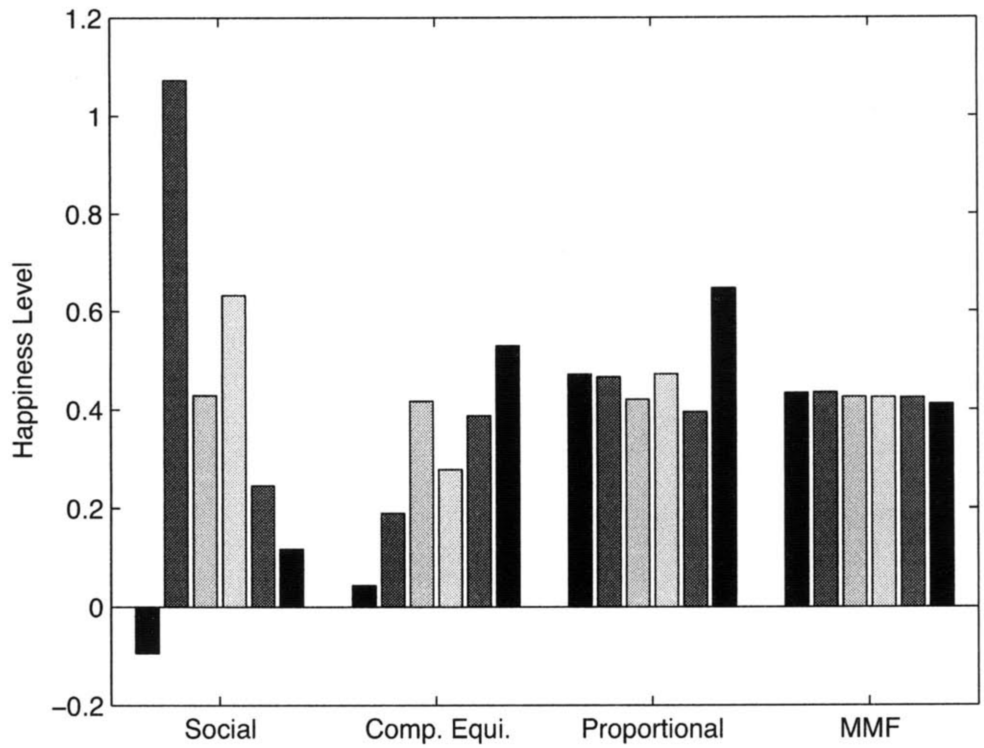

Figure 4-2: Happiness levels of 6 funds obtained from different fairness schemes under Scenario II .

schemes under Scenario II. We can see from this figure similar behaviors as observed in Scenario I. However, under the more complicated settings of Scenario II, the dominance of the α-fairness scheme over the existing approaches in the literature such as the Social Welfare scheme and the Competitive Equilibrium scheme is amplified. Indeed, in this case, all outcomes obtained from the α-fairness scheme dominate those from the Social Welfare scheme and the Competitive Equilibrium scheme in terms of both fairness and optimality. For example, under the α-fairness scheme with $\alpha = 0.1$, the resulting outcome is approximately 17% more efficient than the outcome from the Competitive Equilibrium scheme, and the distribution of happiness levels is slightly flatter under this α-fairness scheme. On the other hand, from Figure 4-3, we see that the Social Welfare scheme is totally dominated by α-fairness scheme in terms of fairness and optimality. Indeed, in contrast to the high total efficiency (i.e., total happiness level) obtained under Scenario I, the Social Welfare scheme is less efficient than any of the outcomes from the α-fairness scheme, including the most equitable MMF scheme. Under the Social Welfare scheme, the total effective utility achieves the maximum value (Table 4.3), but the mean happiness level is much lower than those from the α-fairness scheme. In addition, the happiness level distribution diverges the most under this scheme. This could be the result of optimizing the unscaled happiness levels under the Social Welfare scheme (i.e., combining the happiness levels with the unfair weights). In terms of the price of fairness, if we consider the most efficient solution is that of the α-fairness scheme with $\alpha = 0.1$ (see Figure 4-3), one needs to pay approximately 5.9% of efficiency to get a solution which is completely fair (i.e., the MMF scheme). We can see that, in both Scenarios I and II, such price for fairness is relatively small compare to the benefit that all clients are absolutely equally happy.

4.4 Comparison with BNB algorithm

In this section, we discuss the performance of our proposed algorithm in comparison with the BNB algorithm described in Section 3.4 on small instances of non-convex

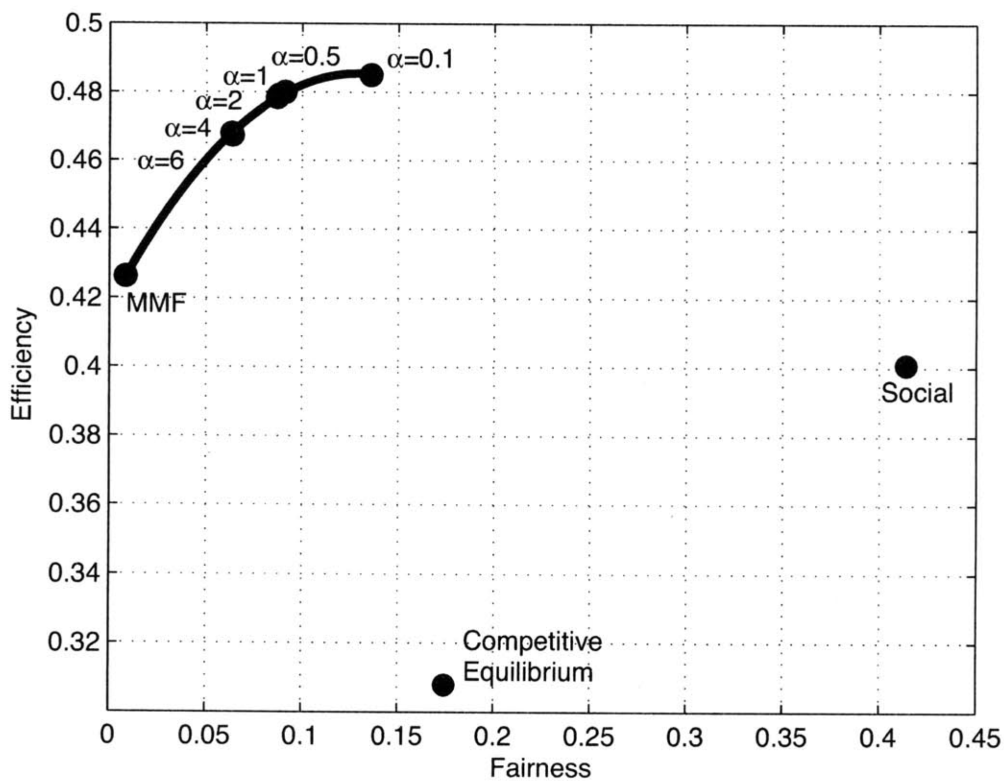

Figure 4-3: Tradeoff between efficiency and fairness under Scenario II. x-axis is the standard deviation of the happiness levels and y-axis is the mean of the happiness levels. The solid curve is the fitting of α-fairness data points.

QCQP problems. The main purpose of the comparison is to give an idea of how our method performs in the settings of our application. Specifically, we generate 6 scenarios similar to Scenario I but with smaller number of funds. The reason why we test on smaller numbers of funds is because the running time of BNB algorithm grows exponentially with the problem dimension and current computational tools prohibit us from experimenting on large problems. All 6 scenarios have 2 funds and the number of stocks are 5, 10, 15, 20, 25 and 30, respectively (the stocks are extracted from the original universe of 73 stocks). Under each scenario, we solve the problem of finding $\bar{U}_1^{max}$. We think that it is sufficient to compare the methods on this single problem because all other subproblems encountered in our application bear similar structures. For each method, we record the optimal objective obtained and the computational time. We also record the optimality gaps resulting from the BNB algorithm. The BNB algorithm is implemented in YALMIP [6], which calls the linear program solver GLPK for solving relaxation problems and the general purpose solver SNOPT to find good feasible solutions.

Table 4.5 shows the comparison between our proposed local search approach with the BNB algorithm. We can see clearly that the optimal objectives obtained from these two methods are the same. In addition, the local search algorithm is superior to the BNB algorithm in term of computation time for $n \geq 20$. Thus, we believe that the proposed local search algorithm is a tractable approach that could possibly lead to near global optimal solutions of the non-convex optimization problems arise in our application.

Table 4.5: Comparison of local search and BNB algorithms ($n = 2$).

	Local search		BNB		
m	Optimal obj.	Time (s)	Optimal obj.	Optimality gap (%)	Time (s)
5	13.5363	18.26	13.5363	0.0378	0.73
10	13.8372	3.36	13.8372	0.0003	5.75
15	10.1860	38.17	10.1861	0.9894	26.33
20	11.8041	32.39	11.8041	0.0000	171.84
25	12.2679	12.19	12.2679	0.7623	356.17
30	16.0859	15.20	16.0859	0.8231	492.07

Chapter 5

Conclusion

5.1 Summary

In this thesis, we approached a practical problem faced by many fund managers when they execute trades for multiple portfolios of diverse clients in a short period of time. Due to the presence of trading costs, these clients have conflicting objectives. As a central decision maker, the fund manager needs to make a decision of how to distribute the market liquidity among the clients by charging them based on the amount that they trade in the market. This process poses a challenge to the fund manager who needs not only to optimize the total efficiency of the system but also to ensure that the participating clients are equally happy. Existing approaches such as the Social Welfare and the Competitive Equilibrium schemes not compromise the optimality and fairness criteria effectively.

We proposed in this thesis a novel approach that fills up the gap in the current literature. We utilized the pro rata allocation scheme to split the total trading cost among the clients. The pro rata scheme is justified because under this scheme, the trading cost incurred by each client coincides with the cost when they trade independently in the open market without information of other's trade. Our formulation of this problem incorporated a quadratic model of market impact cost as the primary source of trading costs when trading is performed at large scale. The effective utility of each clients is then defined as the expected return adjusted for the risk and expected market impact cost. To address the fairness issue among clients, we first normalized the clients' effective utilities to the same range of [0, 1] and called the scaled version of effective utilities the happiness levels of clients. Based on these happiness levels, we formulated our problem as a multiple objective optimization problem in which fairness of the outcome vector is an important criterion in addition to the total system optimality. We then discussed a possible way to obtain solutions to this multiple objective problem that are both Pareto-efficient and equitable. The proposed approach involved solving the fair aggregation problem which utilizes a class of parametric utility function of the happiness levels. Under this so-call α-fairness scheme, we could obtain outcomes that are Pareto-optimal and satisfy the important properties of equitable outcomes such as the principle of transfer or being equally desirable under permutation. The Max-Min fairness and Proportional fairness schemes which are popular in communications are the two special cases of this α-fairness scheme.

We also justified our approach by conducting computational experiments based on real-world financial data. The computational experiments involved solving the non-convex and nonlinear problems, which are typically impossible to solve globally in practice. We proposed a local search algorithm based on linearization techniques that could provide good feasible solutions in efficient running time. The results obtained indicated that the outcomes of the α-fairness scheme are superior to the existing approaches in the literature in compromising the optimality and fairness criteria. In particular, outcomes from the α-fairness scheme is about 17% more efficient than outcomes of the Competitive Equilibrium scheme, given the same level of fairness (quantified by the standard deviation of the happiness levels) is maintained. Our computational experiments also suggested that even though the Social Welfare scheme could sometimes produce solution with good total system efficiency, the unfairness among clients is often far from being acceptable. We also quantified the price that one needs to pay for to obtain an absolutely fair solution, as given under the MMF scheme. The computational results suggested that the price of fairness often ranges from 4% to 6% of system efficiency (measured as the average of the happiness levels). Our belief is that such a price is relatively small compare to the benefit of having all clients equally happy and thus the MMF scheme could be a good choice for fund managers in practice.

5.2 Future directions

This thesis provides a good example of how ideas of resource allocation from the field of communications can be adopted to finance and trading. We will discuss here several directions for future development in this topic.

First of all, we utilized in our formulation a quadratic model of market impact cost (equation 2.7). In practice, to accurately capture the nonlinear nature of trading cost with respect to the total trade, it is sometimes desirable to have a more accurate trading cost model, such as the piecewise linear function or the function of the form $t_j(z_j) = \alpha_j z_j^{p_j}$, where $p_j \in (1, 2]$ $\forall j = 1, \ldots, m$. However, such models of trading cost will pose computational challenges because when we incorporate them into the pro rata scheme, the resulting optimization problems have either fractional or non-integer power terms. In such cases, it is difficult to apply the proposed local search algorithm because the decomposition of non-concave utility function into difference of convex functions is not obvious.

Another possible development in our research is the modification of the mathematical model to reflect the real-world scenarios. An example could be the inclusion of the minimum holding requirement in the constraint set, which makes the resulting optimization problems mixed-integer problems. Even though mixed-integer nonlinear programs are difficult to deal with in practice, they could be possibly solved using heuristics that utilize the special structures of the problem at hand. It is also interesting to see how the proposed fair trading approach could fit into the framework of robust portfolio optimization. Robust optimization is appealing because it enables the decision maker to take into account the risk caused by the change in price of assets when optimizing his portfolio. A critical issue that one can address is that how could we compute the price for fairness and robustness in combination in such cases.

The multi-account portfolio rebalancing we consider here is an instance of multiple objective optimization programs. The principal difference between our problem and the typical multiple objective problems is that we only focus on a subset of Pareto-efficient outcomes that satisfy stated fairness criteria. In the literature of vector optimization, a popular method to generate the Pareto-efficient outcome is to combine different objectives with corresponding weights into a single objective. Clearly, given that the set of equitably efficient outcomes is a subset of the Pareto-efficient outcomes, there exist "fair" combinations of weights such that by optimizing the scalarized objective, we can get an equitably efficient solution. Hence, a critical issue that one might want to address is how to characterize those fair combinations of weights. In addition, even though we suggest in this thesis an approach to obtain the equitably efficient solution by solving the fair aggregation problems, it might not be the best choice. For example, as we could see in Figure 4-3, the outcomes from α-fairness scheme do not distribute evenly on the tradeoff curve when the problem is at large scale. The issue of generating evenly spaced fair and efficient frontier might thus pose interesting challenges for future work in this topic.

Bibliography

[1] T.-J. Chang, N. Meade, J. E. Beasley, and Y. M. Sharaiha. Heuristics for cardinality. *Computers and Operations Research*, 27(13):1271–1303, November 2000.

[2] Frank J. Fabozzi, Petter N. Kolm, Dessislava Pachamanova, and Sergio M. Focardi. *Robust Portfolio Optimization and Management.* Wiley, 2007.

[3] Eligius M.T. Hendrix and Boglarka G.-Toth. *Introduction to Nonlinear and Global Optimization.* Springer, 2010.

[4] Frank Kelly. Charging and rate control for elastic traffic. *European Transactions on Telecommunications*, 8:33–37, 1997.

[5] Michael M. Kostreva, Wlodzimierz Ogryczak, and Adam Wierzbicki. Equitable aggregations and multiple criteria analysis. *European Journal of Operation Research*, 158:362–377, 2004.

[6] J. Lfberg. Yalmip : A toolbox for modeling and optimization in MATLAB. In *Proceedings of the CACSD Conference*, Taipei, Taiwan, 2004. URL `http://users.isy.liu.se/johanl/yalmip`.

[7] Miguel Sousa Lobo, Maryam Fazel, and Stephen Boyd. Portfolio optimization with linear and fixed transaction costs and bounds on risk. Information systems, technical report, Stanford University, 2000.

[8] Harry M. Markowitz. Portfolio selection. *Journal of Finance*, 7:77–91, 1952.

[9] John E. Mitchell and Stephen Braun. Rebalancing an investment portfolio in the presence of transaction costs. Technical report, Department of Mathematical Sciences, Rensselaer Polytechnic Institute, 2002.

[10] J.F. Nash. The bargaining problem. *Econometrica*, 18:155–162, 1950.

[11] Colm O'Cinneide, Bernd Scherer, and Xiaodong Xu. Pooling trades in a quantitative investment process. *The Journal of Portfolio Manegement*, 2006.

[12] Wlodzimierz Ogryczak. Multicriteria models for fair resource allocation. *Control and Cybernetics*, 36(2), 2007.

[13] Wlodzimierz Ogryczak and Tomasz Sliwinski. Sequential algorithms for max-min fair bandwidth allocation. In *Proceedings of the European Computing Conference*, volume 27 of *Lecture Notes Electrical Engineering*, pages 511–522. Springer US, 2009.

[14] Andre F. Perold. Large-scale portfolio optimization. *Management Science*, 30 (10):1143–1160, October 1984.

[15] Robert S. Pindyck and Daniel L. Rubinfeld. *Microeconomics.* New York: Prentice Hall, sixth edition, 2005.

[16] M. Pioro, M. Dzida, E. Kubilinskas, and P. Nilsson. Applications of the max-min fairness principle in telecommunication network design. *Next Generation Internet Networks*, pages 219–225, April 2005.

[17] Gerry A. Pogue. An extension of the Markowitz portfolio selection model to include variable transactions's costs, short sales, leverage policies and taxes. *Journal of Finance*, 25(5):1005–1027, 1970.

[18] Nikolaos Trichakis and Dan Iancu. Operation research in the real world - Final Report. Technical report, MIT, 2008.