A HYBRID APPROACH FOR INDEX TRACKING WITH PRACTICAL CONSTRAINTS

YINGJIE LI AND XIAOGUANG YANG
Institute of Systems Science
Chinese Academy of Science, Beijing 100190, China

SHUSHANG ZHU
Department of Finance and Investment
Sun Yat-Sen University, Guangzhou 510275, China

DONG-HUI LI
School of Mathematical Sciences
South China Normal University, Guangzhou, 510631, China

Abstract. Index tracking is a popular way for passive fund management, which aims to reproduce the performance of a market index by investing in a subset of the constituents of the index. The formulation of index tracking with some realistic constraints always leads to an optimization problem that is very hard to solve. In this paper, we propose an approximate formulation to the original optimization problem and analyze the approximation error bound. It is shown that the approximation can be reasonably close to the original problem. We consider both cases where the mean absolute error and mean square error are used as the tracking error measurements. The mean absolute error measurement results in a mixed-integer linear programming problem and can be solved by some standard solvers, such as CPLEX. The mean square error measurement leads to a mixed-integer quadratic programming problem. An efficient hybrid heuristic method is given to solve this problem. We do some numerical experiments by the use of five data sets from OR-Library. The results are promising.

1. Introduction. There are two kinds of basic strategies adopted for portfolio management: active and passive. Active strategies rely on the assumption that excellent investors can, through their expertise and judgement, outperform the market. While passive fund managers believes that the financial markets are efficient, and it is impossible to beat the market return consistently.

In recent years, “index funds” has gained well acceptance, especially among some mutual/pension funds sponsors. An index fund is a typical passive managed fund, which aims at constructing a portfolio to simulate the performance of some specified market index, such as the S&P 500, Dow Jones indices and etc, during a predefined period of time. We refer to ([21]) for more available indices. Compared with active
managed portfolio, index funds have lower fixed fees/transaction costs, and higher diversification. It is particularly favorable in a bull market, when the active managed funds have slim chance to beat the market.

There are two different strategies to track the market index, namely, the full replication and partial replication. Full replication is done by investing all constituents of the index proportional to their shares in the tracked index. It is simple both conceptually and computationally and can achieve a perfect match. However, its disadvantages are also obvious ([3]). It incurs high initial transactions costs, and is difficult to rebalance when changes are made to the index, for example, the issue of a new set of shares. On the contrary, in a partial replication, an investment is made in a small proportion of the shares, while attempting to match the performance of the entire index. This incurs lower initial transaction costs, and is easier to rebalance. It also allows the investor to constrain the choice of investment that is made, by insisting on the inclusion or exclusion of some companies, or by setting the proportion of the capital that is to be invested in others [20]. Consequently, in practice, the partial replication is generally preferred.

In this work, we consider the index tracking problem using partial replication strategy. Our model is to minimize the tracking error subject to some realistic constraints including the cardinality constraint, the transaction cost constraint and other constraints such as the upper and lower bound limits to the position of each asset. The tracking error in our model is defined as the measure of the difference between the return of the index and that of the tracking portfolio. There are some other tracking error measurements in the literature such as the the correlations between the tracking portfolio and the index, and the variance of the difference between the returns of the tracking portfolio and those of the index ([15], [20]).

[3] discussed index tracking problems with realistic constraints. We will propose an index tracking model that is similar to but a little different from that of [3]. The major difference between our model and that of [3] lies in the definition of portfolio/index return. We use compound interest rate, while [3] choose continuous interest rate. The proposed tracking error is a non-convex function of the decision variables. However, the use of compound interest could help us to deduce a convex approximation to the model. In particular, under mean absolute error measurement, we can use a linear programming to approximate the model, and under mean square error measurement, we can derive a convex quadratic programming approximation to the model. Our error bound analysis shows that under mild conditions, the approximation can be locally sufficiently close to the original index tracking model.

It has been theoretically proved that the index tracking problem with cardinality constraint is NP-hard ([19]). So, heuristic method, such as simulated annealing ([13]) and genetic algorithms ([10]) are often suggested. In literature, heuristic methods for cardinality constrained index tracking problem has been addressed by numerous authors. [20] treated the problem of selecting the optimal subset of assets included in the tracking portfolio and the problem of determining the optimal asset amount separately. He proposed a genetic algorithm to generate the subsets, and used a quadratic programming to find the proportion of the available capital that should be invested in each asset. However, the work by [20] did not consider realistic constraints such as the position constraints and transaction cost constraints. [9] presented a threshold accepting algorithm for index tracking. They renewed the solutions by selling a random index in the current selected asset set, denoted as $J$, and buying a random index in the neighborhood of $J$. New solutions are accepted if they
are no worse (by a specified threshold) than the best solution found so far. They tested their proposed algorithm with artificial indices, where the exact solution is known, but give no results on real world market problems. [3] present an evolutionary heuristic method for his proposed non-convex index tracking problems. Like the threshold accepting algorithm proposed by [9], their proposed genetic method has the advantage of imposing no restrictions on the shape of the objective function or the constraints. [19] chose to hold the weights of the assets in the portfolio constant and their proposed tracking model is a mixed-integer quadratic programming. They presented a set coded genetic algorithm to identify the optimal subset of assets and use quadratic programming to determine the optimal asset weights. However, the strategy of holding the assets weights constant is hard to manage in practice and could bring lots of transaction costs.

In this work, we present a new hybrid optimization algorithm for our proposed index tracking model. In the algorithm, we treat the work of selecting the optimal asset subset and the work of determining the optimal amount of assets alternatively. We present a simple heuristic based on simulated annealing ([13]) to identify the optimal subset of assets. We solve a linear/quadratic program (corresponding to the mean absolute error measurement and the mean square error measurement, respectively) to find the corresponding optimal tracking portfolio with the selected asset subset. The resulting candidate solution is accepted according to the rule of simulated annealing. Simulated annealing is one of the most popular heuristic ([13]) for optimization since it provides a mean to escape local optima by allowing hill-climbing moves. We refer to [1], [14] for details about simulated annealing.

The latter part of the paper is organized as follows. We will describe the detail of the index tracking problem in the next section. Section 3 describes our hybrid method and Section 4 presents some numerical results with artificial data and realistic data. We conclude the paper with some remarks in Section 5.

2. Problem formulation.

2.1. Preliminary. Suppose that there are $N$ assets in the market from which a tracking portfolio should be found. Let the time series be defined for equally spaced intervals. We observe the performance of the $N$ assets as well as the market index at time $0, 1, \ldots, T$ and re-balance the current portfolio at time $T$. We are interested in deciding the best tracking portfolio to hold for some future investment period, denoted as $\{t_0, \ldots, t_0 + L\}$. The mechanism behind our work is that history would repeat itself. We are going to find the portfolio that has best tracked the index over time interval $[0, T]$ and expect it would do a good job during the investment period.

Let $S_i(t)$ be the price of asset $i$ ($i = 1, \ldots, N$) at time $t$, and $I(t)$ be the value of the index at time $t$. The return of the index over the period $[t, t + 1)$ is defined as

$$r_I(t) = \frac{I(t + 1) - I(t)}{I(t)}, \quad t = 0, 1, \ldots, T - 1.$$ 

Let $c_i(t)$ be the number of units of asset $i$ in the tracking portfolio at time $t$ that remains constant in the interval $[t, t + 1)$. Define

$$J_t = \{i | c_i(t) \neq 0\}$$
as the set of indices of assets appearing in the tracking portfolio. The value of the tracking portfolio at time $t$, denoted by $v(t)$, is then calculated via

$$v(t) = \sum_{i=1}^{N} c_i(t)S_i(t) = \sum_{j \in J(t)} c_j(t)S_j(t).$$

The tracking portfolio could also be characterized by defining return $r_i(t)$ and the weights $w_i(t)$ of each of the assets in the portfolio at time $t$,

$$r_i(t) = \frac{S_i(t + 1) - S_i(t)}{S_i(t)}; \quad w_i(t) = \frac{c_i(t)S_i(t)}{v(t)}, \quad t = 0, 1, ..., T - 1.$$  \hspace{1cm} (1)

Then the return of the portfolio at time $t$ is

$$r_P(t) = \frac{v(t + 1) - v(t)}{v(t)} = \sum_{i=1}^{N} w_i(t)r_i(t), \quad t = 0, 1, ..., T - 1.$$ \hspace{1cm} (2)

Several possible strategies can be used to replicate a financial index ([18]). Here are some examples. The buy and hold strategies determine the quantities in each asset $c_i(t)$ and maintain these value constant throughout the tracking period. Constant mix strategies hold a constant fraction of wealth in asset. Buy assets as the market falls and sell them as it rises. There are also some other strategies such as constant proportion portfolio insurance, option based portfolio insurance, etc. In our method, we consider the buy and hold strategies. They are simple and could be followed by all investors easily.

As we have supposed that the quantity $c_i(t)$ remains unchanged through the future investment period. In the latter part of the paper, we will simply use $c_i$ instead of $c_i(t)$.

### 2.2. Practical Constraints.
From a practical point of view, there are a number of constraints that should be considered. In this work, we consider the restrictions on the positions to each asset, the transaction costs and the number of assets in the portfolio.

#### 2.2.1. Transaction cost constraint.
There are several kinds of transaction costs in portfolio management, such as brokerage fees, taxes, fund loads, etc. We will also consider the transaction costs $C_{\text{trans}}$ incurred at re-balancing the portfolio at time $T$. Let $\bar{c} = (c_1, c_2, ..., c_N)$ be the old portfolio before re-balance. We assume that the transaction cost $G_i$ incurred in asset $i$ is proportional to its absolute value change, that is,

$$G_i = \begin{cases} f_s^i(c_i - \bar{c}_i)S_i(T), & c_i < \bar{c}_i, \\ f_b^i(c_i - \bar{c}_i)S_i(T), & c_i \geq \bar{c}_i, \end{cases}$$ \hspace{1cm} (3)

where $f_s^i$, $f_b^i$ denote the fractional cost of selling and buying one unit of asset $i$ at time $T$, respectively. A rational investor would like to limit the transaction in order to prevent unduly buying and selling. Let $C$ be the total wealth that could be invested in the new portfolio, and $\gamma \ (0 \leq \gamma < 1)$ be the upper bound proportion limited to the transaction costs. Then we get the transaction cost constraint

$$C_{\text{trans}} = \sum_{i=1}^{N} G_i \leq \gamma C.$$ \hspace{1cm} (4)

We see from (3) that as a function of $c_i$, $G_i$ is discontinuous, which makes the problem be computationally very difficult. In order to make the problem be easy to
handle, we are going to approximate the constraints using the idea in [6]. Specifically, we use the following set of constraints instead of constraints (3) and (4)

\[
\begin{align*}
G_i & \geq f_i'(c_i - c_i)S_i(T), \quad i = 1, 2, \ldots, N, \\
G_i & \geq f_i'(c_i - c_i)S_i(T), \quad i = 1, 2, \ldots, N, \\
\sum_{i=1}^{N} G_i & \leq \gamma C.
\end{align*}
\]

2.2.2. Position constraint. In general, short selling is prohibited, which means \(c_i \geq 0\). We suppose that all the capital \(C\) is invested in the portfolio, which results in the constraint

\[
\sum_{i=1}^{N} c_i S_i(T) + C_{\text{trans}} = C.
\]

In order to avoid small trades, we require for each \(i\), \(\frac{c_i S_i(T)}{C} \geq \epsilon_i\) with some prescribed constant \(\epsilon_i \in [0, 1]\). Meanwhile, to guarantee a certain level of diversity of the tracking portfolio, we impose an upper bound of the weight for each \(i\). Consequently, we have the following constraints

\[
\epsilon_i \leq \frac{c_i S_i(T)}{C} \leq \delta_i, \quad i \in J_T,
\]

with constant \(\delta_i\) satisfying \(0 \leq \epsilon_i \leq \delta_i \leq 1\).

In practice, we may have some capital concentration constraints in the tracking portfolio. For example, we may require \(0.1 \leq \frac{c_1 S_1(T) + c_2 S_2(T) + c_5 S_5(T)}{C} \leq 0.5\). This is computationally easy to handle. So, in the sequel, we will neglect these constraints, without loss of generality.

2.2.3. Cardinality constraint. In index tracking, the portfolio manager would like to restrict the number of assets. This results in the following cardinality constraint

\[
\# \{J_T\} = K,
\]

where \(\# \{\cdot\}\) denotes the size of the set \(\{\cdot\}\) and \(K\) is an integer that stands for the number of assets of the portfolio. Related work about portfolio selection with cardinality constraint can be found in, for example, [3], [5] and [9].

2.3. Objective function and approximation. The objective of our index tracking is to minimizing the tracking error. There are many kinds of tracking errors, among which, the following tracking error given by [3] is very popular

\[
TR(c) = \frac{1}{T} \left( \sum_{t=0}^{T-1} \left| \frac{r_P(t) - r_I(t)}{v(t)} \right|^\alpha \right)^\frac{1}{\alpha}, \quad \alpha > 0,
\]

where \(\alpha > 0\) is a constant. In this work, we consider the cases \(\alpha = 1\) and \(\alpha = 2\). The case \(\alpha = 1\) corresponds to the so-called mean absolute error measurement and \(\alpha = 2\) corresponds to the mean square error measurement, respectively. We refer to [4] for alternative tracking errors.

The tracking error defined by (7) is generally not convex due to the nonconvexity of \(r_P(t)\). In fact, substituting (1) and (2) to (7), we obtain

\[
TR(c) = \frac{1}{T} \left( \sum_{t=0}^{T-1} \left( \sum_{i=1}^{N} \frac{c_i S_i(t)}{v(t)} \left( r_i(t) - r_I(t) \right) \right)^\alpha \right)^\frac{1}{\alpha}, \quad \alpha = 1 \text{ or } 2.
\]
Since $v(t)$ relies on $c_i$, the function $TR(c)$ above is not convex. In what follows, we are going to find a convex approximation to $TR(c)$, so that we can solve the problem more easily.

It follows from (2) that for any $t \in \{0, 1, \cdots, T - 1\}$,

$$v(t + 1) = v(t)(1 + r_P(t)).$$

This implies

$$v(t) = \frac{v(T)}{\prod_{j=t}^{T-1}(1 + r_P(j))}.$$  

Notice that $r_P(t)$ is expected to be very close to $r_I(t)$, $t = 0, 1, \ldots, T - 1$, near the optimal tracking portfolio. This motivates us to replace $r_P(t)$ by $r_I(t)$, which yields the approximate relation

$$v(t) \approx v(T) \prod_{j=t}^{T-1}(1 + r_I(j)).$$  

Observing by (4) that $C_{trans}$ is bounded from above by $\gamma C$, we may approximate $v(T)$ by a constant (see also [6]) and get

$$v(T) = C - C_{trans} \approx (1 - \gamma)C \triangleq \bar{v}(T).$$

Define

$$\bar{v}(t) \triangleq \frac{\bar{v}(T)}{\prod_{j=t}^{T-1}(1 + r_I(j))} = \frac{(1 - \gamma)C}{\prod_{j=t}^{T-1}(1 + r_I(j))},$$  

$$\bar{r}_P(t) \triangleq \sum_{i=1}^{N} \left( \frac{c_i S_i(T)}{\bar{v}(t)} \right) r_i(t).$$

Since $\bar{r}_P(t)$ is a linear approximation to $r_P(t)$, we get the following convex approximation to the objective function given by (7):

$$\overline{TR}(c) = \frac{1}{T} \left( \sum_{t=0}^{T-1} |\bar{r}_P(t) - r_I(t)|^\alpha \right)^{\frac{1}{\alpha}}, \ \alpha = 1 \text{ or } 2,$$  

which is much easier to handle than the original one.

2.4. Optimization problem. To formulate the problem in a more tractable way, we introduce an auxiliary vector $z = (z_1, z_2, \ldots, z_N)$ whose elements are binary variables. More specifically, for each $i = 1, 2, \ldots, N$, $z_i = 1$ if and only if asset $i$ is included in the tracking portfolio, otherwise, $z_i$ takes the value 0. Then minimizing the tracking error defined by (10) with $\alpha = 2$ results in the following mixed-integer
quadratic program

\[(P_{r2}) \quad \min_{c,z} \quad \frac{1}{T} \sum_{t=0}^{T-1} (\bar{r}_P(t) - r_I(t))^2 \]

\[
\text{s.t.} \quad \sum_{i=1}^{N} c_i S_i(T) + \sum_{i=1}^{N} G_i = C, \quad \sum_{i=1}^{N} G_i \leq \gamma C, \\
G_i \geq f^i_i(c_i - \bar{c}_i), \ G_i \geq f^i_i(\bar{c}_i - c_i), \ i = 1, 2, \ldots, N, \\
\epsilon_i z_i C \leq c_i S_i(T) \leq \delta_i z_i C, \ 0 \leq \epsilon_i \leq \delta_i \leq 1, \ i = 1, 2, \ldots, N, \\
\sum_{i=1}^{N} z_i = K, \ z_i \in \{0, 1\}, \ i = 1, 2, \ldots, N.
\]

Replacing the objective function of \((P_{r2})\) with

\[
\frac{1}{T} \sum_{t=0}^{T-1} |\bar{r}_P(t) - r_I(t)|
\]

yields the tracking problem with mean absolute error measure (the case for \(\alpha = 1\)). By introducing auxiliary vector \(f=(f_0, f_1, \ldots, f_{T-1})\), the problem can be equivalently reformulated as the following mixed-integer linear program

\[(P_{r1}) \quad \min_{c,z} \quad \frac{1}{T} \sum_{t=0}^{T-1} f_t \]

\[
\text{s.t.} \quad \sum_{i=1}^{N} c_i S_i(T) + \sum_{i=1}^{N} G_i = C, \quad \sum_{i=1}^{N} G_i \leq \gamma C, \\
G_i \geq f^i_i(c_i - \bar{c}_i), \ G_i \geq f^i_i(\bar{c}_i - c_i), \ i = 1, 2, \ldots, N, \\
f_t = \bar{r}_P(t) - r_I(t), \ f_t \geq r_I(t) - \bar{r}_P(t), \ t = 0, 1, \ldots, T - 1, \\
\epsilon_i z_i C \leq c_i S_i(T) \leq \delta_i z_i C, \ 0 \leq \epsilon_i \leq \delta_i \leq 1, \ i = 1, 2, \ldots, N, \\
\sum_{i=1}^{N} z_i = K, \ z_i \in \{0, 1\}, \ i = 1, 2, \ldots, N.
\]

3. Validity of the approximation. As pointed out in the last section, the models \((P_{r1})\) and \((P_{r2})\) are not the exact index-tracking models but their approximations. In this section, we analyze the relationships between the approximation models and their exact versions. We first give some simple but useful relations among \(r_I(t)\), \(\bar{v}(t)\), \(\bar{v}(t + 1)\) and \(\bar{r}_P(t)\). Indeed, we get from (8) that \(\bar{v}(t) = \bar{v}(t + 1)/(1 + r_I(t))\). This implies

\[
r_I(t) = \frac{\bar{v}(t + 1) - \bar{v}(t)}{\bar{v}(t)}, \ t = 0, 1, \ldots, T - 1. \quad (11)
\]

We also have from (2) and (9)

\[
\bar{r}_P(t) = \frac{r_P(t) v(t)}{\bar{v}(t)} = \frac{v(t+1) - v(t)}{\bar{v}(t)}, \ t = 0, 1, \ldots, T - 1. \quad (12)
\]

The following two propositions show that under some reasonable conditions, the approximations can be very close to the exact models near the solution.
Proposition 1. Suppose that the portfolio \((c_1, c_2, \ldots, c_N)\) satisfies
\[
\frac{1}{T} \sum_{t=0}^{T-1} |\bar{r}_P(t) - r_I(t)|^\alpha = \epsilon_1^\alpha, \quad \alpha = 1 \text{ and } 2,
\]
where \(\epsilon_1 > 0\) is sufficiently small. Suppose further that the corresponding return sequence \(\{r_P(t)\}\) satisfies \(\max_t |r_P(t)| \leq A\) with some constant \(A > 0\) and that
\[
\frac{v(T)}{\bar{v}(T)} = 1 + \epsilon_2 \quad \text{and} \quad \frac{\max_t \bar{v}(t)}{\min_t \bar{v}(t)} \leq C(T),
\]
where \(\epsilon_2\) is a constant and \(C(T)\) is a constant related to the time window \(T\). Then we have
\[
\frac{1}{T} \sum_{t=0}^{T-1} |r_P(t) - r_I(t)| \leq \epsilon_1 + C(T)A(\epsilon_1 + \epsilon_2)
\]
and
\[
\frac{1}{T} \sum_{t=0}^{T-1} (r_P(t) - r_I(t))^2 \leq 2 \left( C(T)^2 A^2 (\epsilon_1 + \epsilon_2)^2 + \epsilon_1^2 \right).
\]

Proof. Let \(\beta = v(T)/\bar{v}(T)\). By the assumption, we have
\[
1 \leq \frac{C - C_{trans}}{C - \gamma C} = \beta = \frac{v(T)}{\bar{v}(T)} = 1 + \epsilon_2,
\]
which implies \(\epsilon_2 > 0\).

We get from (11) and (12)
\[
v(t + 1) - v(t) = \bar{r}_P(t)\bar{v}(t),
\]
\[
\bar{v}(t + 1) - \bar{v}(t) = r_I(t)\bar{v}(t).
\]
Then we obtain for any \(t \in \{0, 1, \ldots, T - 1\},\)
\[
\bar{v}(t) - v(t) = (v(T) - v(t)) - (\bar{v}(T) - \bar{v}(t)) - (v(T) - \bar{v}(T))
\]
\[
= \sum_{j=t}^{T-1} \bar{r}_P(j)\bar{v}(j) - \sum_{j=t}^{T-1} r_I(j)\bar{v}(j) - (v(T) - \bar{v}(T))
\]
\[
= \sum_{j=t}^{T-1} (\bar{r}_P(j) - r_I(j)) \bar{v}(j) - (\beta - 1)\bar{v}(T).
\]
We also have by the definition of \(r_P(t)\) and \(\bar{r}_P(t)\)
\[
|\bar{r}_P(t) - r_P(t)| = |r_P(t)| \left| \frac{v(t) - \bar{v}(t)}{\bar{v}(t)} \right|
\]
\[
= |r_P(t)| \left| \frac{\sum_{j=t}^{T-1} (\bar{r}_P(j) - r_I(j)) \bar{v}(j) - (\beta - 1)\bar{v}(T)}{\bar{v}(t)} \right|
\]
\[
\leq |r_P(t)|C(T) \left| \sum_{j=0}^{T-1} (\bar{r}_P(j) - r_I(j)) + (\beta - 1) \right|
\]
\[
\leq C(T)A (T\epsilon_1 + \epsilon_2),
\]
where the first inequality follows from (14).
Then we have
\[
\frac{1}{T} \sum_{t=1}^{T-1} |\tilde{r}_p(t) - r_p(t)| \leq \frac{1}{T} \sum_{t=1}^{T} C(T)A(T\epsilon_1 + \epsilon_2) = C(T)A(T\epsilon_1 + \epsilon_2). \tag{16}
\]

It follows from (13) and (16) that
\[
\frac{1}{T} \sum_{t=1}^{T-1} |r_p(t) - r_I(t)| \leq \frac{1}{T} \sum_{t=1}^{T-1} |\tilde{r}_p(t) - r_p(t)| + \frac{1}{T} \sum_{t=1}^{T-1} |\tilde{r}_p(t) - r_I(t)|
\leq \epsilon_1 + C(T)A(T\epsilon_1 + \epsilon_2).
\]

For the case \(\alpha = 2\), we can derive from (13) and (15)
\[
|\tilde{r}_p(t) - r_p(t)|^2 \leq |r_p(t)|^2 C(T)^2 \left( \sum_{j=0}^{T-1} |\tilde{r}_p(j) - r_I(j)| + (\beta - 1) \right)^2
\leq |r_p(t)|^2 C(T)^2 \left( \left( \sum_{j=0}^{T-1} |\tilde{r}_p(j) - r_I(j)|^2 \right)^{1/2} + (\beta - 1) \right)^2
= A^2 C(T)^2 (T\epsilon_1 + \epsilon_2)^2,
\tag{17}
\]
where last inequality follows from Cauchy inequality. So, we get from (17)
\[
\frac{1}{T} \sum_{t=1}^{T} |\tilde{r}_p(t) - r_p(t)|^2 \leq A^2 C(T)^2 (T\epsilon_1 + \epsilon_2)^2.
\]

This together with (13) imply
\[
\frac{1}{T} \sum_{t=1}^{T} |r_p(t) - r_I(t)|^2 \leq 2 \left( \frac{1}{T} \sum_{t=1}^{T} |\tilde{r}_p(t) - r_I(t)|^2 + \frac{1}{T} \sum_{t=1}^{T} |\tilde{r}_p(t) - r_p(t)|^2 \right)
\leq 2 \left( A^2 C(T)^2 (T\epsilon_1 + \epsilon_2)^2 + \epsilon_2^2 \right).
\]

The proof is completed. \(\square\)

As a direct corollary of Proposition 1, we have the following results immediately.

**Corollary 1.** Suppose that the portfolio \((\epsilon_1, \epsilon_2, ..., \epsilon_N)\) satisfies
\[
\frac{1}{T} \sum_{t=0}^{T-1} |\tilde{r}_p(t) - r_I(t)|^\alpha \to 0, \quad \alpha = 1 \text{ and } 2.
\]

Suppose further that the corresponding return sequence \(\{r_p(t)\}\) satisfies \(\max_i |r_p(t)| \leq A\), with some constant \(A > 0\) and that
\[
\frac{\nu(T)}{\tilde{\nu}(T)} \to 1 \quad \text{and} \quad \frac{\max_i \tilde{\nu}(t)}{\min_i \tilde{\nu}(t)} \leq C(T),
\]
where \(C(T)\) is a constant related to the time window \(T\). Then we have \(\frac{1}{T} \sum_{t=0}^{T-1} |r_p(t) - r_I(t)| \to 0\) and \(\frac{1}{T} \sum_{t=0}^{T-1} (r_p(t) - r_I(t))^2 \to 0\).
Proposition 1 and Corollary 1 indicate that if the approximated portfolio tracks the index sufficiently, then the real portfolio will also track the index well.

**Proposition 2.** Suppose that the portfolio \((c_1, c_2, ..., c_N)\) satisfies

\[
\frac{1}{T} \sum_{t=0}^{T-1} |r_P(t) - r_I(t)|^\alpha \to 0, \quad \alpha = 1 \text{ and } 2.
\]

(18)

Suppose further that the return sequence \(\{r_P(t)\}\) generated by the corresponding portfolio satisfies

\[
\frac{v(T)}{\bar{v}(T)} \to 1, \quad \min_t r_P(t) \geq A_1 \quad \text{and} \quad \max_t r_P(t) \leq A_2,
\]

where \(-1 < A_1 < A_2, A_2 > 0\). Then we have

\[
\frac{1}{T} \sum_{t=0}^{T-1} |\bar{r}_P(t) - r_I(t)| \to 0 \quad \text{and} \quad \frac{1}{T} \sum_{t=0}^{T-1} (\bar{r}_P(t) - r_I(t))^2 \to 0.
\]

**Proof.** We only need to show

\[
\frac{1}{T} \sum_{t=0}^{T-1} |\bar{r}_P(t) - r_P(t)|^\alpha \to 0, \quad \alpha = 1 \text{ and } 2.
\]

Let \(\beta = v(T)/\bar{v}(T)\). By assumption, we have that \(1 \leq \beta \to 1\). From the definition of \(r_P(t), r_I(t), v(t), \bar{v}(t)\), we have for \(\alpha = 1 \text{ or } 2\)

\[
\frac{1}{T} \sum_{j=0}^{T-1} |\bar{r}_P(j) - r_P(j)|^\alpha = \frac{1}{T} \sum_{j=0}^{T-1} |r_P(j)|^\alpha \left| 1 - \frac{v(j)}{\bar{v}(j)} \right|^\alpha.
\]

Denote

\[
r_I(t) - r_P(t) = \epsilon'_t, \quad t = 0, 1, ..., T - 1,
\]

and

\[
A'_1 \triangleq 1 + A_1 > 0.
\]

It follows from (18) that

\[
\epsilon \triangleq \frac{1}{T} \sum_{t=0}^{T} |\epsilon'_t|^\alpha \to 0, \quad \alpha = 1, 2.
\]

For any \(t \in \{0, 1, \cdots, T - 1\}\), it holds that

\[
|\bar{r}_P(t) - r_P(t)| = |r_P(t)| \left| 1 - \frac{v(t)}{\bar{v}(t)} \right| = |r_P(t)| \left| 1 - \beta \prod_{j=t}^{T-1} \left( \frac{1 + r_I(j)}{1 + r_P(j)} \right) \right| \\
= |r_P(t)| \left| 1 - \beta \prod_{j=t}^{T-1} \left( 1 + \frac{\epsilon'_j}{1 + r_P(j)} \right) \right|.
\]
If $\beta \prod_{j=t}^{T-1} \left(1 + \frac{\epsilon'_j}{1 + r_P(j)}\right) > 1$, then we have

$$
|\bar{r}_P(t) - r_P(t)| = |r_P(t)| \left(\beta \prod_{j=t}^{T-1} \left(1 + \frac{\epsilon'_j}{1 + r_P(j)}\right) - 1\right)
$$

$$
\leq |r_P(t)| \left(\beta \prod_{j=t}^{T-1} \left(1 + \frac{|\epsilon'_j|}{A'_t}\right) - 1\right)
$$

$$
\leq |r_P(t)| \left(\beta \left(1 + \sum_{j=t}^{T-1} \frac{|\epsilon'_j|}{A'_1}\right)^{T-t} - 1\right)
$$

$$
= |r_P(t)| \left(\beta \left(1 + \sum_{j=t}^{T-1} \frac{|\epsilon'_j|}{A'_1} + o\left(\frac{\sum_{j=t}^{T-1} |\epsilon'_j|}{A'_1}\right)\right) - 1\right), (19)
$$

where the last inequality follows from the well-known geometric inequality.

For the case $\alpha = 1$, we have from (19)

$$
|\bar{r}_P(t) - r_P(t)| \leq |r_P(t)| \left(\beta \left(1 + \sum_{j=0}^{T-1} \frac{|\epsilon'_j|}{A'_1}\right) + o\left(\frac{\sum_{j=0}^{T-1} |\epsilon'_j|}{A'_1}\right)\right) - 1
$$

The last inequality implies

$$
\frac{1}{T} \sum_{t=0}^{T-1} |r_P(t) - \bar{r}_P(t)| \to 0.
$$

For the case $\alpha = 2$, we also have from (19)

$$
|\bar{r}_P(t) - r_P(t)|^2 \leq |r_P(t)|^2 \left(\beta \left(1 + \sum_{j=0}^{T-1} \frac{|\epsilon'_j|}{A'_1}\right) + o\left(\frac{\sum_{j=0}^{T-1} |\epsilon'_j|}{A'_1}\right)\right)^2
$$

$$
\leq |r_P(t)|^2 \left((\beta - 1) + \beta \left(\frac{T \sqrt{\epsilon}}{A'_1} + o\left(\frac{T \sqrt{\epsilon}}{A'_1}\right)\right)\right)^2,
$$

where the last inequality follows from Cauchy inequality. This also implies

$$
\frac{1}{T} \sum_{t=0}^{T-1} |r_P(t) - \bar{r}_P(t)|^2 \to 0.
$$
If $\beta \prod_{j=t}^{T-1} \left(1 + \frac{\epsilon_j}{1 + r_P(j)}\right) < 1$, then we obtain

$$|\bar{r}_P(t) - r_P(t)| = |r_P(t)| \left(1 - \beta \prod_{j=t}^{T-1} \left(1 + \frac{\epsilon_j}{1 + r_P(j)}\right)\right)$$

$$\leq |r_P(t)| \left(1 - \beta \prod_{j=t}^{T-1} \left(1 - \frac{\epsilon_j}{A_1}\right)\right)$$

$$\leq |r_P(t)| \left(1 - \left(1 - \sum_{j=t}^{T-1} \frac{\epsilon_j}{A_1}\right)\right)$$

$$= |r_P(t)| \left(1 - \beta + \frac{\sum_{j=t}^{T-1} \epsilon_j}{A_1}\right), \quad (20)$$

where the last inequality follows from first-order Taylor expansion.

For $\alpha = 1$, we get from (20)

$$|\bar{r}_P(t) - r_P(t)| \leq |r_P(t)| \left(\beta \sum_{j=0}^{T-1} \left|\frac{\epsilon_j}{A_1}\right| + (1 - \beta)\right)$$

$$= |r_P(t)| \left(\beta \frac{T}{A_1} (1 - \beta)\right).$$

This implies

$$\frac{1}{T} \sum_{t=0}^{T-1} |r_P(t) - \bar{r}_P(t)| \to 0.$$

For $\alpha = 2$, we also have from (20)

$$|\bar{r}_P(t) - r_P(t)|^2 \leq |r_P(t)|^2 \left(\beta \sum_{j=0}^{T-1} \left|\frac{\epsilon_j}{A_1}\right| + (1 - \beta)\right)^2$$

$$\leq |r_P(t)|^2 \left(\beta \frac{T\sqrt{\epsilon}}{A_1} (1 - \beta)\right)^2.$$

This also implies

$$\frac{1}{T} \sum_{t=0}^{T-1} |r_P(t) - \bar{r}_P(t)|^2 \to 0.$$

The proof is completed.

Proposition 2 shows that if there exists a portfolio that sufficiently close the index, then the associated approximate portfolio will also track the index well.

4. A hybrid method. In this section, we propose a hybrid method to solve the proposed index tracking formulations ($P_{r1}$) and ($P_{r2}$). Since ($P_{r1}$) and ($P_{r2}$) are very similar, we only construct algorithm for solving ($P_{r2}$). Algorithm for solving ($P_{r1}$) can be constructed in a similar way. The mixed quadratic programming ($P_{r2}$) is generally NP hard ([19]). Numerical methods for the exact solution of the problem might be unaffordable. For example, if we solve the problem with the universe assets $N = 225$ and cardinality constraint $K = 10$ by exhaustive search, it will cost a PC millions of years to get a solution ([16]). Other exact techniques like Branch-and-Bound methods ([11]) might be used to handle this combinational
problem. However, the computational complexity still grows exponentially with the size of the universe assets. Consequently, heuristics, such as the simulated annealing ([13]) and genetic algorithms ([10]) are practically welcome.

Heuristics contribute to a reasonable extent in solving global optimization problems, mainly combinatorial problems ([17]). Simulated annealing (SA) is one of the most popular heuristics [13], that have been employed in a wide variety of problems. The key points of an SA to escape from local optima relies on its acceptance rule. Under the SA acceptance rule, improved solutions are always accepted, while a fraction of non-improved (inferior) solutions are accepted in the hope of escaping local optima in the search of a global optima. The probability of accepting non-improved solutions depends on a temperature parameter, which is typically non-increasing with each iteration of the algorithm. If the temperature parameter is decreased sufficiently slow, then the sequence generated by an SA will be asymptotic convergent to a global optima ([7]). We refer to ([12]), [1], [14] for details about simulated annealing. However, SA, like other meta-heuristics, suffers from the slow convergence that brings about the high computational cost.

In what follows, we are going to develop a hybrid method to find some approximate solution to the problem. We propose a heuristic procedure to identify the appropriate asset set that should be invested, and use a quadratic programming solver to find the optimal tracking portfolio that invests only in the selected asset set. The idea of hybridizing quadratic programming with heuristic algorithm has been successfully applied in portfolio selection ([5], [19]). We follow this idea and propose a hybrid approach for solution of problem (Pr2).

4.1. A hybrid optimization method for index tracking. We see from (Pr2) that it is the cardinality constraint that makes the problem NP hard. Once we have decided which assets to be invested, (Pr2) reduces to a standard convex quadratic programming which could be effectively solved ([8]). Based on this observation, we consider to treat (Pr2) using two steps, first to choose the asset set to be invested and then to decide the investment weight of each asset. We propose a simulated annealing based heuristic to explore the investment asset set and resort to quadratic programming to find the investment weight.

Figure 1 gives the flowchart of our proposed hybrid heuristic. We first use some initialization procedure to produce an initial asset set. Then quadratic programming are used to determine the best tracking portfolio that are invested only in the selected asset set. The produced tracking portfolio is accepted as the next candidate solution through some simulated annealing rule, called SAm for simplify. If the produced portfolio fails to be accepted, adjust the asset set and resolve the corresponding investment weights. The process repeats until we get some acceptable candidate. We repeat the above procedure until it satisfies some predetermined termination conditions.

In what follows, we describe our proposed hybrid method in detail.

4.1.1. Initialization. At the beginning, we need to get an initial asset set. We determine the initial asset based on the following observation. First, we consider to fully replicate the market. In other words, we solve the quadratic program (Pr2) by dropping the cardinality constraint and get a portfolio \( c = (c_1, c_2, \ldots, c_N) \). Intuitively, we think that the assets that possess large proportions in that portfolio should be selected as candidates to get into the initial asset set. The following process gives the detail to find the initial asset set.
Consider the following relaxation formulation of the index tracking model \((P_{2})\),

\[
\begin{align*}
\min_{c} & \quad \frac{1}{T} \sum_{t=0}^{T-1} (\bar{r}_P(t) - r_I(t))^2 \\
\text{s.t.} & \quad \sum_{i=1}^{N} c_i S_i(T) + \sum_{i=1}^{N} G_i = \sum_{i=1}^{N} \bar{c}_i S_i(T), \quad \sum_{i=1}^{N} G_i \leq \gamma C, \\
& \quad G_i \geq f^b_i(c_i - \bar{c}_i), \quad G_i \geq f^s_i(\bar{c}_i - c_i), \quad i = 1, 2, \ldots, N, \\
& \quad c_i S_i(T) \leq \delta_i C, \quad 0 \leq \delta_i \leq 1, \quad i = 1, 2, \ldots, N.
\end{align*}
\]  

(21)

Figure 1. Flowchart for the proposed hybrid heuristic
It was obtained by ignoring the cardinality constraints in \((P_r2)\). Let \(c = (c_1, c_2, \ldots, c_N)\) be the solution vector obtained by solving the convex quadratic problem \((21)\) via CPLEX. Relabel \(c\) so that it satisfies,

\[
S_1(T)c_1 \geq S_2(T)c_2 \geq \ldots \geq S_N(T)c_N.
\] (22)

We use the following procedure to construct an initial asset subset.

**Procedure 0 (Construction of the initial asset subset \(J^0\))**

1. **Initialization.** Let the solution \(c = (c_1, c_2, \ldots, c_N)\) to \((21)\) satisfy \((22)\) and 
   \[ J_{\text{candi}} = \{1, 2, \ldots, K\} \]
   1. **Feasibility test.** If \((P_r2)\) is feasible with
      \[
      \begin{cases}
      z_j = 1, & j \in J_{\text{candi}}, \\
      z_j = 0, & j \notin J_{\text{candi}},
      \end{cases}
      \] (23)
   then we set \(J^0 = J_{\text{candi}}\) and stop.
   2. **Updating.** Consider to update \(J_{\text{candi}}\) by replacing asset \(K\) by \(K+1, K+2, \ldots\), until \((P_r2)\) could yield some feasible solution with \(J_{\text{candi}}\). If it fails to get a feasible solution, try to replace \(K\) by asset \(K-1, K-2, \ldots\). If it still fails, consider to replace asset \((K-1, K)\) by the pair of assets \((K+1, K+2), (K+1, K+3), (K+2, K+3), (K+2, K+4), \ldots\), and then consider to replace \((K-2, K), (K-3, K), \ldots, (1, K), (K-2, K-1), (K-3, K-1), \ldots\). If it still fails to obtain a feasible solution, consider to update \(J_{\text{candi}}\) by replacing 3, 4, \ldots, \(K-1\) assets, successively.

   As the Update procedure is essentially exhaustive search, it could always yield some well-defined initial asset set as long as \((P_r2)\) is feasible. In many cases, if \((P_r2)\) has feasible solutions, \(J_{\text{candi}} = \{1, 2, \ldots, K\}\) could yield a feasible solution.

4.1.2. A heuristic strategy to adjust the asset subset. Let the asset set at iteration \(l\) be denoted by \(J^l = \{i_1, i_2, \ldots, i_K\}\). Our heuristic aims to find a \(J^{l+1}\) by extracting some asset, say \(i_{\text{out}} \in J^l\), and finding a new asset, say \(i_{\text{in}} \notin J^l\), from \(J^l\). That is, \(J^{l+1} = J^l \cup \{i_{\text{in}}\} \setminus \{i_{\text{out}}\}\), where \(i_{\text{in}} \notin J^l\) and \(i_{\text{out}} \in J^l\).

   We first consider the choice of the out asset \(i_{\text{out}}\). We regard the asset with least investment amount in the current portfolio to be least important. Based on this point, the out-asset is determined according to the amount of the scalars \(S_i(T)c(i)\), \(i \in J^l\). Asset with smaller \(S_i(T)c(i)\) has more priority to be considered.

   For a given out asset \(i_{\text{out}}\), we use a simulated annealing procedure, denoted as \(SA_q\), to generate a possible in-asset. More precisely, for \(\bar{J}^l = \{j | j \notin J^l\}\), we assign a fitness value for each \(j \in \bar{J}^l\), and accept asset \(j\) as a candidate in asset according to \(SA_q\) one by one, through the order of its fitness value. Asset with better fitness value have more priority to be considered.

   Suppose asset \(j\) is accepted as \(i_{\text{in}}\) by \(SA_q\). Then we could get a candidate asset subset with asset \(i_{\text{out}}\) out and asset \(i_{\text{in}}\) in. Otherwise, if all \(\bar{J}^l\) fails to be accepted by \(SA_q\), we find a next candidate out asset and an in asset. This procedure continues until we could find some accepted pair of out asset and in asset, denoted by \((i_{\text{out}}, i_{\text{in}})\).

   In the case the above process could not find a pair \((i_{\text{out}}, i_{\text{in}})\), we let \(i_1 = \arg\max\{S_i(T)c(i), i \in J^l\}\), and randomly select an asset \(j \in \bar{J}^l\) as the \(i_{\text{in}}\), and set \(J^{l+1} = J^l \cup \{i_{\text{in}}\} \setminus \{i_1\}\). Otherwise, set \(J_{\text{candi}} = J^l \cup \{i_{\text{in}}\} \setminus \{i_{\text{out}}\}\), and use another simulated annealing procedure, called \(SA_m\), to accept \(J_{\text{candi}}\) as \(J^{l+1}\).
This procedure repeats until some $J_{\text{candi}}$ could be accepted by $SA_m$. If no generated $J_{\text{candi}}$ could be accepted, randomly select a asset $j \in \mathcal{A}$ as the $i_{\text{in}}$ and set $J^{l+1} = J^l \cup \{i_{\text{in}}\} \setminus \{i_1\}$.

It should be noted that, we have used two simulated annealing procedure, $SA_a$ and $SA_m$, in the asset adjustment process. Procedure $SA_a$ is used to explore some candidate set for $J^{l+1}$, denoted as $J_{\text{candi}}$; while procedure $SA_m$ is used to determine if $J_{\text{candi}}$ could be accepted as $J^{l+1}$.

We summarize the above process as the algorithm below.

Algorithm (A hybrid method for $(P_2)$)

0 Initialization. Use Procedure 0 to generate an initial asset set $J^0$. Set $l := 0$. Set the current asset subset be $J^l = \{i_1^l, i_2^l, ..., i_K^l\}$ and compute the optimal tracking portfolio with assets only invested in $J^l$ to get $c^l \equiv c(J^l) = (c(i_1^l), c(i_2^l), ..., c(i_K^l))$. Let $\mathcal{J} = \{j | j \notin J^l\}$, $c_{\text{best}} := c^l$. Set the cooling schedule parameters: the initial temperature $T_{\text{max}}$, $T_{\text{min}}$, cooling rate $\rho$ in $SA_a$ procedure; and the corresponding parameters $T_{\text{max}}$, $T_{\text{min}}$, $\rho'$ in $SA_m$ procedure. Set the precision parameter $\epsilon > 0$ and the iteration tolerance parameter $\text{iter}_{\text{max}}, \text{iter}_{\text{best}}$.

1 Termination Conditions. Stop the algorithm if one of the following three cases happens: (1) the number of iterations $l$ reaches $\text{iter}_{\text{max}}$; (2) the process fails to get a better tracking portfolio after $\text{iter}_{\text{best}}$ many iterations; (3) the best tracking portfolio ever found satisfies: $TR(c_{\text{best}}) \leq \epsilon$.

2 Asset set update. Let $p := 1$ and resort $J^l$ so that it satisfies

$$S_{i_1}(T)c(i_1) \leq S_{i_2}(T)c(i_2) \leq ... \leq S_{i_K}(T)c(i_K).$$

Set $t := T_{\text{max}}$, $t' := T'_{\text{max}}$.

2.1 Let $i_{\text{out}} := i_p$. For each $j \in \mathcal{J}$, let

$$J_{\text{candi}}(j) = J^l \cup \{j\} \setminus \{i_{\text{out}}\}.$$

Construct a portfolio $P(j)$ that corresponding to $J_{\text{candi}}(j)

$$P(j) = \left\{ \begin{array}{ll} c(i), & i \in J_{\text{candi}}(j) \setminus \{j\}, \\ c(i_{\text{out}})S_{i_{\text{out}}}(T)/S_i(T), & i = j. \end{array} \right.$$ 

Define the fitness value of asset $j$ by $TR(P(j))$ and relabel $\mathcal{J}$ so that it satisfies $TR(P(k_1)) \leq TR(P(k_2)) \leq ... \leq TR(P(k_{N-k}))$.

Set $q := 0$.

2.2 $SA_a$ Procedure. If $q \leq N - k$, set $q := q + 1$, $t := \rho t$ and go to step 2.2.1. Else if $p < K$, set $p := p + 1$ and go to step 2.1; if $p \geq K$, go to step 2.4.

2.2.1 If $t \geq T_{\text{min}}$, then accept $k_q$ according to the simulated annealing rule with the probability $\exp^\left(-TR(c') - TR(P(k_q))/t\right)$. If $k_q$ is accepted, set $i_{\text{in}} := k_q$, $J_{\text{candi}} := J_{\text{candi}}(i_{\text{in}})$, $t = T_{\text{max}}$, and go to step 2.2.3. Else, if $k_q$ is not accepted, go to step 2.2.

2.2.2 If $t < T_{\text{min}}$, choose a random $j \in \mathcal{J}$. Let

$$J_{\text{candi}} = J^l \cup \{j\} \setminus \{i_1\},$$

and set $t = T_{\text{max}}$. 


with cardinality constraint \( K \) of our heuristic algorithm, we first use an artificial index. For a tracking problem folio for the problems with realistic size. To provide a means to test the validity quadratic programming, it is generally hardly to obtain the optimal tracking port-

5.1. Test with artificial index. Since the formulation (\( P_{r2} \)) is a mixed-integer quadric programming, it is generally hardly to obtain the optimal tracking portfolio for the problems with realistic size. To provide a means to test the validity of our heuristic algorithm, we first use an artificial index. For a tracking problem with cardinality constraint \( K \), we define the artificial index \( r_I(t) \) as

\[
r_I(t) = \frac{1}{K} \sum_{i=N-K+1}^{N} \frac{S_i(t)}{\sum_{N-K+1}^{N} S_i(t)} r_i(t), t = 0, 1, \ldots, T - 1.
\]
Table 1. Results for artificial index ($\alpha = 1$)

| Index  | $K$  | TR       | PTR  | Time(s) |
|--------|------|----------|------|---------|
| Hang Seng | 5    | $5.0596E-17$ | 0.00% | 2.31 |
|         | 8    | $3.0202E-16$ | 0.00% | 3.05 |
|         | 10   | $2.7229E-17$ | 0.00% | 2.22 |
| DAX    | 5    | $3.2564E-16$ | 0.00% | 1.59 |
|         | 8    | $1.0164E-17$ | 0.00% | 2.06 |
|         | 10   | $9.6868E-16$ | 0.00% | 3.22 |
| FTSE   | 5    | $1.1148E-16$ | 0.00% | 2.20 |
|         | 8    | $1.5338E-15$ | 0.00% | 3.59 |
|         | 10   | $4.2819E-17$ | 0.00% | 2.70 |
| S&P    | 5    | $2.9754E-17$ | 0.00% | 2.59 |
|         | 8    | $1.5216E-16$ | 0.00% | 3.81 |
|         | 10   | $3.3338E-17$ | 0.00% | 2.63 |
| Nikkei | 5    | $3.8798E-18$ | 0.00% | 2.83 |
|         | 8    | $1.0669E-16$ | 0.00% | 3.95 |
|         | 10   | $1.1864E-16$ | 0.00% | 6.84 |

In other words, the artificial index is generated by investing equal units in the last $K$ assets. In the case where $\epsilon_i = 0$, $\delta_i = 1$, $f^b_i = 0$, $f^s_i = 0$, $\forall i$ and $\gamma = 0$, we could get the exact optimal portfolio, denoted by $c^*$, with tracking error $TR(c^*) = \overline{TR}(c^*) = 0$. We use CPLEX (version 11.0) to solve the mixed-integer linear programming formulation ($P_{r1}$). For $\alpha = 2$, we applied our heuristic 20 times on each problem and report the average results. For all artificial problems, we set $\epsilon = 10^{-10}$, $\text{iter}_{\text{best}} = 30$. Detailed computational results are listed in Tables 1 and 2 ($T = 290$). The meaning of each column in the following two tables is stated below.

Index: the name of the asset market;
$K$: the cardinality of the tracking portfolio;
$TR$: the tracking error defined in (7);
$PTR$: defined as ($TR$)/(standard deviation of $\{r_I(t), t = 0, 1, \cdots, T\}$);
Time: the CPU time used (in seconds).

We see from Table 1 that our approximate formulation ($P_{r1}$) makes high quality of approximation. The accuracy can reach up to $o(10^{-17})$. Meanwhile, the approximate formulation is easy to handle. In all cases, it took CPLEX less than 7 seconds to get an optimal portfolio. From Table 2, we can see that our heuristic performed quite well for the artificial index tracking problems. The average accuracy can reach to $o(10^{-9}) \sim o(10^{-10})$ and the maximum cpu time used is less than 16 seconds.

5.2. Test with real indexes. The numerical results in the last subsection has shown that the formulations ($P_{r1}$) and ($P_{r2}$) and the proposed hybrid solution method worked well in sample. In practice, we wish a portfolio perform well out of sample too. In order to investigate both the in sample and out of sample performance of our approach, we used the data from time 0 to time 145 as the training data, and test the tracking performance from time 146 to time 290.

We first tested the formulation ($P_{r1}$). We solved it by CPLEX as well as our heuristic procedure. The maximum cpu time costed by CPLEX is set to be 7200 seconds. Detailed results were shown in Table 3. We then test formulation ($P_{r2}$). For each case, we run the heuristic procedure 20 times and reported the average results in Table 4. We considered the cases $\gamma = 0.0025, 0.005, 0.0075, 0.01$, while the
Table 2. Results for artificial index ($\alpha = 2$)

| Index   | $K$ | TR       | PTR      | Time(s) |
|---------|-----|----------|----------|---------|
| Hang Seng | 5   | 9.8813E-10 | 0.00%   | 2.61    |
|         | 8   | 4.0080E-10 | 0.00%   | 2.63    |
|         | 10  | 6.3005E-10 | 0.00%   | 11.50   |
| DAX     | 5   | 2.0305E-09 | 0.00%   | 5.34    |
|         | 8   | 1.3536E-09 | 0.00%   | 4.34    |
|         | 10  | 1.4177E-09 | 0.00%   | 15.06   |
| FTSE    | 5   | 5.5373E-10 | 0.00%   | 4.36    |
|         | 8   | 5.4406E-10 | 0.00%   | 4.20    |
|         | 10  | 3.9203E-10 | 0.00%   | 15.94   |
| S&P     | 5   | 1.7699E-09 | 0.00%   | 7.05    |
|         | 8   | 7.4322E-10 | 0.00%   | 4.09    |
|         | 10  | 1.1031E-09 | 0.00%   | 13.06   |
| Nikkei  | 5   | 1.0979E-09 | 0.00%   | 6.61    |
|         | 8   | 1.1576E-09 | 0.00%   | 7.05    |
|         | 10  | 4.6438E-10 | 0.00%   | 15.64   |

case $\gamma = 0.01$ was not considered in formulation ($P_{r1}$). This is because the cpu time used by CPLEX increases dramatically as the value of $\gamma$ grows. While we solved ($P_{r1}$) with $\gamma = 0.01$ by CPLEX, we found that in most cases, we could not get an optimal solution within 7200 seconds.

The results in Table 3 show that both in sample and out of sample tracking errors are decreasing with respect to $\gamma$. Meanwhile, lower in sample tracking errors generally yield lower out of sample tracking errors. This is in accordance with the expectations that “Increasing the transaction cost limit $\gamma$ will reduce tracking error” and “Reducing in-sample tracking error reduces out of sample tracking error” ([3]).

We also see from the table that the cpu time consumed by the CPLEX increases dramatically with respect to the size of the problem and the value of the transaction cost limit $\gamma$. For Hang Seng index ($n = 31$) with $\gamma = 0.0025$, it took less than 1.0 seconds to get an optimal tracking portfolio. While, for the Nikkei index ($n = 225$) with $\gamma = 0.0075$, it took more than 2 hours to get an optimal solution. On the other hand, our heuristic seems less sensitive with respective to $n$ and $\gamma$. In all cases, it took less than 15 seconds to get a tracking portfolio. In addition, for some of the problems, eg. DAX with $\gamma = 0.0075$, S&P with $\gamma = 0.0025 \sim 0.0075$, our hybrid method could get a portfolio with smaller tracking error than that of CPLEX. This would be not surprising because we used the original tracking error (7) to guide our exploration, while the portfolio obtained by CPLEX was an optimal portfolio with respect to our approximate tracking error measurement (10).

From Table 4, we can see that the our hybrid method could get a portfolio with average tracking error $o(10^{-5})$. The results also in accordance with the two expectations mentioned above. To see the effectiveness of the formulation and the algorithm, we draw the tracking portfolio in Figure 1 (fixed the market price and the portfolio price to be 100 at $T = 145$). From Figure 1, we can see that our portfolio tracked the index pretty well both in sample and out of sample.

Next we compare our method with some existing heuristic methods. [3] and [9] also considered index tracking problems used heuristics to get the solutions. The heuristic in [3] was based on evolutionary while the heuristic in [9] was based on
Figure 2. In sample versus out of sample performance ($\alpha = 2$, $K = 10$, $\gamma = 0.1$).

a threshold accepting rule. As we have shown in Section 4, our hybrid heuristic is a combination of the quadratic programming and the simulated annealing. Table 5 lists the performance of these three heuristics. As the definition of return in [3], [9] and ours are different, it is hard to say whose tracking portfolio is of the least tracking error. So, in the table, we did not list the exact value of the tracking errors, instead, we list the order of them. The results in Table 5 show that all the three method could get optimal portfolios with similar accuracy orders. Since [9] did not test their heuristic for market index, we leave the corresponding cells blank. Average cpu time costed were reported in Table 5, where “s” stands for seconds and “m” stands for minutes. As we are testing the methods under different operational
A HYBRID APPROACH FOR INDEX TRACKING WITH CONSTRAINTS 925

Table 3. Results for market index ($\alpha = 1$, $K = 10$, $T = 145$)

| Index  | $\gamma$ | Our method | CPLEX |
|--------|----------|------------|-------|
|        |          | TR in sample | TR out of sample | Time | TR in sample | TR out of sample | Time |
| Hang Seng | 0.0025 | 8.7926E-03 | 9.8732E-03 | 2.41 | 8.7926E-03 | 9.8732E-03 | 0.50 |
|         | 0.0050 | 7.0231E-03 | 7.7252E-03 | 3.79 | 7.0231E-03 | 7.7252E-03 | 0.44 |
| DAX     | 0.0025 | 9.4665E-03 | 1.5762E-02 | 3.46 | 9.4665E-03 | 1.5762E-02 | 1.56 |
|         | 0.0050 | 7.3926E-03 | 1.3190E-02 | 5.01 | 7.3926E-03 | 1.3190E-02 | 2.48 |
| FTSE    | 0.0025 | 8.3966E-03 | 8.6472E-03 | 3.92 | 8.2697E-03 | 8.6523E-03 | 1.36 |
|         | 0.0050 | 6.8252E-03 | 8.4108E-03 | 5.59 | 6.7041E-03 | 8.6305E-03 | 14.17 |
| S&P     | 0.0025 | 7.3333E-03 | 7.5609E-03 | 3.73 | 7.3985E-03 | 6.6952E-03 | 3.53 |
|         | 0.0050 | 6.0642E-03 | 7.2365E-03 | 8.41 | 6.1468E-03 | 5.9437E-03 | 86.92 |
| Nikkei  | 0.0025 | 5.4048E-03 | 8.1038E-03 | 11.45 | 5.4960E-03 | 7.0842E-03 | 148.83 |
|         | 0.0050 | 4.6010E-03 | 6.6864E-03 | 8.18 | 4.5304E-03 | 7.7663E-03 | 4440.17 |

Table 4. Results for market index ($\alpha = 2$, $K = 10$, $T = 145$)

| Index  | $\gamma$ | In sample | Out of sample | Time(s) |
|--------|----------|------------|---------------|---------|
|        |          | TR | PTR | TR | PTR | Time |
| Hang Seng | 0.0025 | 9.2399E-04 | 2.78% | 1.0566E-03 | 3.18% | 14.28 |
|         | 0.0050 | 7.4254E-04 | 2.33% | 8.9386E-04 | 2.69% | 11.38 |
| DAX     | 0.0025 | 1.0157E-03 | 4.99% | 1.6760E-03 | 8.23% | 14.03 |
|         | 0.0050 | 8.1005E-04 | 3.98% | 1.5505E-03 | 7.61% | 23.59 |
| FTSE    | 0.0025 | 8.9498E-04 | 5.11% | 9.3617E-04 | 5.35% | 39.03 |
|         | 0.0050 | 7.2091E-04 | 4.12% | 8.2545E-04 | 4.72% | 19.52 |
| S&P     | 0.0025 | 8.9498E-04 | 5.11% | 9.3617E-04 | 5.35% | 39.03 |
|         | 0.0050 | 7.2091E-04 | 4.12% | 8.2545E-04 | 4.72% | 19.52 |
| Nikkei  | 0.0025 | 6.1801E-04 | 2.16% | 7.5369E-04 | 2.63% | 14.28 |
|         | 0.0050 | 5.1771E-04 | 1.81% | 6.6038E-04 | 2.31% | 57.59 |

platform and using different environment, we could not say which method is least expensive. However, we think that our method is competitive with others.

6. Conclusions. We formulated the index tracking problem with realistic constraints into a constrained optimization problem. We considered two cases where the mean absolute error and the mean square error were used to measure the tracking performance. The formulations are non-convex optimization problems. We then proposed a mixed-integer linear program (for mean absolute error measurement) and a mixed-integer quadratic program (for mean square error measurement)
Table 5. Comparison of computational results ($\alpha = 2$, $K = 10$, $T = 145$, $\gamma = 0.01$)

| Method          | Artificial index TR | Time | Market index TR (In sample) | Time |
|-----------------|---------------------|------|----------------------------|------|
| [3]             | $o(10^{-4}) \sim o(10^{-5})$ | 6.7 m | $o(10^{-3})$ $o(10^{-5}) \sim o(10^{-4})$ | 15.2 m |
| [9]             | $o(10^{-3}) \sim o(10^{-5})$ | 7.6 s | $o(10^{-4})$ $o(10^{-3}) \sim o(10^{-4})$ | 34.3 s |
| Our method      | $o(10^{-9}) \sim o(10^{-10})$ | 14.2 s | $o(10^{-4})$ $o(10^{-3}) \sim o(10^{-4})$ | 34.3 s |

to approximate them. Under reasonable conditions, we proved that the approximate formulation could provide a good approximation to the original non-convex tracking problem. To solve the approximation problem, we proposed a hybrid algorithm by combining numerical method for convex programs and simulate annealing approach for combinatorial optimization problems. Our experiments show that the approximate formulations works well and the proposed hybrid method is efficient.

REFERENCES

[1] E. Aarts and J. Korst, Selected topics in simulated annealing, in Essays and Surveys in Metaheuristics (eds. C. C. Ribeiro and P. Hansen) (Angra dos Reis, 1999), Oper. Res./Comput. Sci. Interfaces Ser., 15, Kluwer Academic Publishers, Boston, MA, 2002, 1–37.

[2] J. E. Beasley, OR-Library: Distributing test problems by electronic mail, Journal of the Operational Research Society, 41 (1990), 1069–1072.

[3] J. E. Beasley, N. Meade and T.-J. Chang, An evolutionary heuristic for the index tracking problem, European Journal of Operation Research, 148 (2003), 621–643.

[4] S. Browne, Beating a moving target: Optimal portfolio strategies for outperforming a stochastic benchmark, Finance and Stochastics, 3 (1999), 275–294.

[5] T.-J. Chang, N. Meade and J. E. Beasley, Heuristics for cardinality constrained portfolio optimisation, Computers and Operations Research, 27 (2000), 1271–1302.

[6] N. A. Canakgoz and J. E. Beasley, Mixed-integer programming approaches for index tracking and enhanced indexation, European Journal of Operational Research, 196 (2009), 384–399.

[7] E. Çınlar, Introduction to Stochastic Processes, Prentice-Hall, Inc., Englewood Cliffs, New Jersey, 1975.

[8] R. Flethcher, A general quadratic programming algorithm, J. Inst. Math. Appl., 7 (1971), 76–91.

[9] M. Gill and E. Kûllezi, Threshold Accepting for Index Tracking, Computing in Economics and Finance Series, 72, Society for Computational Economics, 2001.

[10] J. H. Holland, Adaptation in Natural and Artificial Systems. An Introductory Analysis with Applications to Biology, Control, and Artificial Intelligence, University of Michigan Press, Ann Arbor, Mich., 1975.

[11] R. Horst, A general class of branch-and-bound methods in global optimization with some new approaches for concave minimization, Journal of Optimization Theory and Applications, 51 (1986), 271–291.

[12] L. Ingber, Simulated annealing: Practice versus theory, Mathematical and Computer Modelling, 18 (1993), 29–57.

[13] S. Kirkpatrick, C. D. Gelatt, Jr. and M. P. Vecchi, Optimization by simulated annealing, Science, 220 (1983), 671–680.

[14] P. J. Laarhoven and E. H. Aarts, Simulated Annealing: Theory and Applications, Mathematics and its Applications, 37, Springer, 1997.

[15] H. Markowitz, Mean-Variance Analysis in Portfolio Choice and Capi7al Markets, Basil Blackwell, Oxford, 1987.

[16] R. Moral-Escudero, R. Ruiz-Turrubiano and A. Suárez, Selection of optimal investment portfolio with cardinality constraints, in Proceedings of the IEEE Congress on Evolutionary Computation, 2006, 2382–2388.
A HYBRID APPROACH FOR INDEX TRACKING WITH CONSTRAINTS

[17] I. H. Osman and J. P. Kelly, eds., Meta-Heuristics: Theory & Applications, Papers from the 1995 International Conference (MIC) held in Breckenridge, Colorado, July 22–26, 1995, Kluwer Academic Publishers, Boston, MA, 1996.

[18] A. F. Perold, C. D. Gelatt and M. P. Vecchi, Dynamic strategies for asset allocation, Financial Analysis Journal, 44 (1988), 17–27.

[19] R. Ruiz-Torrubiano and A. Suárez, A hybrid optimization approach to index tracking, Annals of Operation Research, 166 (2009), 57–71.

[20] J. Shapcott, Index Tracking: Genetic Algorithms for Investment Portfolio Selection, Technical Report EPCC-SS92-24, Edinburgh, Parallel Computing Center, 1992.

[21] C. M. S. Sutcliffe, Stock Index Futures: Theories and International Evidence, 2nd edition, International Thompson Business Press, 1997.

Received March 2012; 1st revision March 2013; 2nd revision August 2013.

E-mail address: yjli@amss.ac.cn
E-mail address: xgyang@iss.ac.cn
E-mail address: sszhu@fudan.edu.cn
E-mail address: dhli@scnu.edu.cn