Tax-Aware Portfolio Construction
via Convex Optimization

Nicholas Moehle    Mykel J. Kochenderfer    Stephen Boyd
Andrew Ang
BlackRock AI Labs
August 13, 2020

Abstract

We describe an optimization-based tax-aware portfolio construction method that adds tax liability to a standard Markowitz-based portfolio construction approach that models expected return, risk, and transaction costs. Our method produces a trade list that specifies the number of shares to buy of each asset and the number of shares to sell from each tax lot held. To avoid wash sales (in which some realized capital losses are disallowed), we assume that we trade monthly, and cannot simultaneously buy and sell the same asset.

The tax-aware portfolio construction problem is not convex, but it becomes convex when we specify, for each asset, whether we buy or sell it. It can be solved using standard mixed-integer convex optimization methods at the cost of very long solve times for some problem instances. We present a custom convex relaxation of the problem that borrows curvature from the risk model. This relaxation can provide a good approximation of the true tax liability, while greatly enhancing computational tractability. This method requires the solution of only two convex optimization problems: the first determines whether we buy or sell each asset, and the second generates the final trade list. This method is therefore extremely fast even in the worst case. In our numerical experiments, which are based on a realistic tax-loss harvesting scenario, our method almost always solves the nonconvex problem to optimality, and when in does not, it produces a trade list very close to optimal. Backtests show that the performance of our method is indistinguishable from that obtained using a globally optimal solution, but with significantly reduced computational effort.
1 Introduction

Tax-aware investment strategies aim to match or beat the return of a benchmark index while simultaneously minimizing or controlling tax liability generated by the trades. Examples include tax-loss harvesting strategies, which close losing positions to generate capital losses that can be deducted from investors’ income or capital gains; tax-neutral strategies, which aim to generate no net capital gains; and capital-gains-budgeting strategies, which limit net capital gains to a given maximum level. Central to all of these strategies is tax-aware portfolio optimization, which identifies losing positions while maximizing (pre-tax) returns and minimizing active risk relative to a benchmark, all subject to other portfolio management considerations like transaction costs and turnover.

There are several features of the current tax law that make optimization for tax-managed strategies challenging. First, rather than specifying the actions for each asset as in a standard portfolio construction problem, a decision must be made to hold or sell each individual tax lot. Having multiple tax lots per asset makes the problem high dimensional. Second, the portfolio construction problem is nonconvex due to the wash sale rule defined in section 1091 of the United States tax code. This rule limits an investor’s deduction when they sell shares at a loss and, within 30 days before or after this sale, acquires an economically similar position. Wash sale rules cannot be avoided by simply trading monthly; we must also not sell shares of a given asset (at a loss) and repurchase those shares in the same trade. This restriction is encoded as a complementarity condition, which results in a nonconvex optimization problem.

Our contribution is to formulate a tax-aware portfolio construction problem that explicitly accounts for tax liabilities from long- and short-term capital gains, while maintaining the standard objectives of return, risk, and transaction costs. While this tax liability term is not convex, we develop a convex relaxation of the problem that allows practical trade lists to be computed extremely quickly. Because convex optimization problems can be solved efficiently and reliably, convex optimization is used in numerous engineering, science, finance, and economics problems that require massive scale and customization Boyd and Vandenberghe, 2004. Our proposed formulation potentially allows tax-aware solutions to be delivered in scale to hundreds of thousands of clients across multiple account offerings. Developing a convex-optimization-based tax-aware portfolio construction method enables tax-management to be customized to different benchmarks, divestments, overlays, portfolio tilts and risk levels, and other features demanded by clients.

In addition to choosing how many shares to buy or sell of each asset, our method also divides up each sale across lots. For a fixed sell amount, we find it is optimal to sell shares in the order that minimizes immediate tax liability, which we call the least tax first out (LTFO) method, which coincides with the well-known highest basis first out (HIFO) method if the same tax rate applies to all lots. Dickson, Shoven,
and Sialm (2000) and Berkin and Ye (2003) show that the accounting method for lot ordering significantly affects the losses that can be harvested. Even though it is not optimal, Atra and Pae (2013) show that the HIFO method generates substantial benefit to an investor’s total wealth.

Our convex relaxation combines all terms that are separable across a single asset, and then replaces the resulting (nonconvex) function with its convex envelope. This approach has the effect of taking (or ‘borrowing’) curvature from the transaction cost and specific risk terms, and introducing it into the tax liability term. This effect can also be interpreted as an application of the Shapley–Folkman Lemma. The overall effect is that the tax-aware portfolio construction problem can be very well approximated by a convex optimization problem, and the fidelity of this approximation is borne out in our numerical examples.

To demonstrate our method, we apply it to a tax-loss harvesting strategy. First, we show that in a realistic backtest scenario, the strategy tightly tracks the benchmark (the S&P 500) while harvesting capital losses. In this application, we compare the performance of our method with that of a standard mixed-integer quadratic programming formulation and show that our method delivers near-identical trade lists, despite being several hundred times faster.

1.1 Related work and background

Markowitz portfolio construction. Markowitz’s 1952 formulation of portfolio construction as an optimization problem involves a trade off of expected return and risk (Markowitz, 1952). This optimization problem, with a quadratic objective and linear equality constraints, has an analytical solution. Several extensions exist; for example, by including position limits or a long-only constraint (Sharpe, 1963; Markowitz, 1955; Grinold and Kahn, 1999), the resulting problem no longer has an analytical solution, but it can be formulated and efficiently solved as a quadratic program (QP) (Boyd and Vandenberghe, 2004, pp. 55–156). The formulation can also be extended to include other constraints and objective terms, including those related to trading, given a previous or initial portfolio (Pogue, 1970; Lobo, Fazel, and Boyd, 2007). The extension to include an initial portfolio allows the method to be used as a trading policy, which can be run periodically to determine what trades to make (Boyd et al., 2017). This method has become common in portfolio management.

Portfolio construction via convex optimization. These Markowitz-inspired portfolio construction problems are typically convex and can be efficiently solved (Boyd and Vandenberghe, 2004). Even complex portfolio construction problems can be specified succinctly in a high-level domain-specific language for convex optimization, such as CVXPY (Diamond and Boyd, 2016), CVX (Grant and Boyd, 2014), Convex.jl (Udell et al., 2014), and CVXR (Fu, Narasimhan, and Boyd, 2020). Furthermore, problems with thousands of assets and a risk model with dozens of factors
can be solved in well under a second using standard open-source solvers such as ECOS (Domahidi, Chu, and Boyd, 2013), OSQP (Stellato et al., 2020), SCS (O’Donoghue et al., 2016; O’Donoghue et al., 2019), or commercial solvers such as CPLEX (IBM Corporation, 2019), MOSEK (MOSEK ApS, 2019), or Gurobi (Gurobi Optimization LLC, 2020). Custom implementations of portfolio construction solvers can be far faster, with solve times measured in milliseconds.

Although solver speed is not essential if we are only interested in occasionally rebalancing a handful of portfolios, it can be useful when managing a very large number of individualized accounts. Additionally, fast solvers allow us to quickly run many backtest simulations of a trading algorithm. These simulations allow us to tune hyper-parameters, carry out what-if experiments, and compare different formulations or models on historical or synthesized data. This scale is crucial to building (and testing) a highly customized, tax-aware investment products.

Non-convex portfolio construction problems. Some practical constraints and objective terms are not convex. An obvious example is that asset holdings must be in integral numbers of shares. For large portfolios, this constraint is readily handled by simple heuristics, for example by ignoring it in solving the problem, and then rounding the real-valued holdings to the nearest integer values. Other more challenging constraints include limits on the number of assets in the portfolio, or a minimum nonzero trade size (Bertsimas, Darnell, and Soucy, 1999). A challenging nonconvex objective term is tax liability, the focus of this paper.

These nonconvex portfolio construction problems can be reformulated as mixed-integer convex optimization problems (which are not convex). They can be solved exactly using a variety of methods and software, such as GLPK (Makhorin, 2016), CPLEX, MOSEK, and Gurobi. Such solvers are often fast, but for some problem instances can have very long solve times, often hundreds of times more than those associated with similar convex problems. In contrast, solving convex optimization problems is always fast.

Convex approximations. An alternative to solving the nonconvex optimization problem exactly is to employ a heuristic method that finds an approximate solution, far faster than it would take to solve the problem exactly. This paper presents such heuristic methods, based on a convex approximation of the original problem. The idea that convex approximations of nonconvex problems can be used in place of global nonconvex solvers with the same practical performance has been widely noted in other areas (Diamond, Takapoui, and Boyd, 2018).

The problem we study in this paper involves the sum of many nonconvex terms that are all similar. Because the sum of a large number of nonconvex functions tends to be ‘more convex’ than the original functions, these problems are often well approximated by convex problems. This intuitive phenomenon was described by Shapley, Folkman, and Starr. Starr applied it to problems in microeconomics involving a large
number of agents (Starr, 1969). A summary of this idea is given by Bertsekas (1997), where practical algorithms for solving these problems are given, along with provable performance bounds. Other examples of such problems are sigmoidal programming (Udell and Boyd, 2013) and unit commitment (Lauer et al., 1982).

**Tax-aware investment.** Our paper is related to literature that develops optimal tax trading strategies, building on the seminal papers by Constantinides (1983; 1984). The main intuition in these papers is that investors may reduce their tax liability by deferring capital gains and realizing, or accelerating, losses. Losses can be used to offset current income or capital gains; realizing short-term losses is especially valuable because they can be deducted from highly taxed income. This intuition also applies in our setting. Most of the papers in this literature, however, apply well-known numerical solutions to solve the tax problem. For example, numerical dynamic programming has been widely applied (Dammon and Spatt, 1996; Dammon, Spatt, and Zhang, 2004; Dammon, Spatt, and Zhang, 2001). An early paper by Dybvig and Koo (1996) works with a binomial tree and formulates an optimal stopping problem. DeMiguel and Uppal (2005) formulate an optimal tax investment strategy with non-linear programming. Although Markowitz-based tax-aware portfolio construction is an old idea (Pogue, 1970), to our knowledge, ours appears to be the first contribution to the tax literature on developing a convex tax optimization problem by relaxing the original non-convex problem. We focus on the speed and reliability of convex optimization techniques with applications to managing taxable managed funds (Sialm and Zhang, 2020), the large and rapidly growing tax-loss harvesting industry (Chaudhuri, Burnham, and Lo, 2020), and security valuation in the presence of taxes (Gallmeyer and Srivastava, 2011).

### 1.2 Contributions

This paper focuses on incorporating a specific nonconvex term, the tax liability generated by the trades, into an otherwise convex portfolio construction problem. Ignoring this constraint, or using simple ad hoc rounding methods to handle it, does not work well compared to solving the problem exactly with a mixed-integer convex solver. Our contribution is to develop heuristic methods for approximately solving the tax-aware portfolio construction problem that relies on solving two convex optimization problems, making it reliably fast.

### 1.3 Outline

Section 2 describes the tax-aware portfolio construction problem, and section 3 provides the details of the tax liability generated by sales of assets. We develop two convex relaxations of this problem in section 4: (1) a basic relaxation that involves only the tax liability by itself and (2) a more sophisticated relaxation that combines
other terms such as transaction cost or risk with the tax liability. Section 5 describes a general framework for heuristics for the tax-aware portfolio construction problem. We demonstrate the performance of our methods in section 6.

2 Tax-aware portfolio construction

This section outlines our notation and describes the tax-aware portfolio optimization problem. We start by describing the trading dynamics and various objective terms.

2.1 Portfolio holdings and dynamics

We consider a universe of $n$ assets we are allowed to hold and trade. We let $h_{\text{init}} \in \mathbb{R}^n$ denote the dollar value of our pre-trade holdings of these $n$ assets. We restrict ourselves to long-only portfolios, so $h_{\text{init}} \geq 0$.

Our task is to decide how much of each of these assets to buy or sell. We represent this decision by a purchase vector $u \in \mathbb{R}^n$, denominated in dollars. If we purchase asset $i$, $u_i > 0$; if we sell asset $i$, $u_i < 0$. Our post-trade holdings are $h \in \mathbb{R}^n$, given by

$$h = h_{\text{init}} + u.$$ 

This equation ignores transaction costs, which are assumed to be small. (Following convention, we include these transaction costs in our objective function.) We require that the post-trade portfolio is also long-only, making $h \geq 0$. This constraint means we cannot sell more of any asset than we currently hold.

Cash. The cash held in the portfolio is $c_{\text{init}} \in \mathbb{R}$, which we allow to be negative. The post-trade cash balance is

$$c = c_{\text{init}} - 1^T u.$$ 

We assume the post-trade cash amount must match some desired value $c_{\text{des}}$, which translates to the constraint on $u$:

$$1^T u = c_{\text{des}} - c_{\text{init}}.$$ 

The total pre-trade portfolio value, including cash, is $1^T h_{\text{init}} + c_{\text{init}}$, which we assume is positive. While any value of $c_{\text{des}}$ is possible, a common choice is a given fraction $\eta$ of the total portfolio value,

$$c_{\text{des}} = \eta(1^T h_{\text{init}} + c_{\text{init}}).$$ (1)

The choice $\eta = 0.01$, for example, means that 1% of the total portfolio value is to be held in cash. The cash balance can be used to handle cash deposits into and withdrawals from the account by adjusting $c_{\text{init}}$ by the amount deposited or withdrawn.
2.2 Objective terms

Here we describe various objective terms and additional constraints, including the traditional ones: expected return, active risk, and transaction costs. We briefly introduce the tax liability term and provide some of its attributes, reserving a detailed description for section 3.

Risk. The risk of a managed portfolio is typically measured with respect to a benchmark portfolio, such as the S&P 500. This benchmark portfolio is described by a vector \( h_b \in \mathbb{R}^n \), scaled so that it has the same market value as our portfolio, i.e., \( 1^T h_b = 1^T h_{init} + c_{init} \).

The (active) risk is

\[
(h - h_b)^T V (h - h_b),
\]

where \( V \) is the covariance matrix of the asset returns. Our covariance matrix \( V \) has the traditional factor model form

\[
V = X \Sigma X^T + D,
\]

where \( X \in \mathbb{R}^{n \times k} \) is the factor exposure matrix, \( \Sigma \in \mathbb{R}^{k \times k} \) is the symmetric positive definite factor covariance matrix, and \( D \in \mathbb{R}^{n \times n} \) is the diagonal matrix of idiosyncratic variances with \( D_{ii} > 0 \) (Grinold and Kahn, 1999; Boyd et al., 2017). The risk can be decomposed into two components, the systematic risk

\[
(h - h_b)^T X \Sigma X^T (h - h_b), \tag{2}
\]

and the specific risk

\[
(h - h_b)^T D (h - h_b) = \sum_{i=1}^{n} D_{ii} (h_i - h_{b,i})^2. \tag{3}
\]

It is common to express active risk in terms of its square root, which has units of dollars. We note for future use that the specific risk (3) is separable, i.e., a sum of terms each associated with one asset.

Expected return. Suppose we have a forecast of the return of the \( n \) assets, expressed as a vector \( \alpha \in \mathbb{R}^n \), where \( \alpha_i \) is the expected return of asset \( i \). The expected active return of portfolio \( h \) is then \( \alpha^T (h - h_b) \), which is measured in dollars. Because the constant offset \( \alpha^T h_b \) is immaterial for optimization, we write the expected return as simply \( \alpha^T h \).

Transaction costs. The transaction cost follows a simple bid-ask spread model:

\[
\kappa^T |u|,
\]
where $\kappa \in \mathbb{R}_+^n$ is one-half the bid-ask spread, and $|u|$ is the element-wise absolute value of $u$. For simplicity, we neglect the standard price impact term; this omission is reasonable if we assume our trades are small relative to the total market volume over the trading period. Due to the individualized nature of tax-management, this is often reasonable.

**Tax liability.** We let $L : \mathbb{R}^n \rightarrow \mathbb{R}$ denote the tax liability function, where $L(u)$ is the immediate tax liability incurred by the trades $u$ due to realizing capital gains. We will describe $L(u)$, which derives from the history of previous transactions in the assets, in detail in section 3; for now, we simply note some of its attributes. First, it is separable across the assets, i.e., it has the form

$$L(u) = \sum_{i=1}^{n} L_i(u_i),$$

where $L_i(u_i)$ is the tax liability for asset $i$ incurred by trading. There is no immediate tax liability when buying an asset, making $L_i(u_i) = 0$ for $u_i \geq 0$. For $u_i < 0$, i.e., selling the asset, $L_i(u_i)$ is a convex piecewise linear function. While $L_i$ is convex for $u_i < 0$, it is not convex over the whole interval, which includes buying ($u_i > 0$) and selling ($u_i < 0$). The total tax liability function $L(u)$ is not convex, but it becomes convex if we restrict the sign of $u_i$, i.e., we specify whether we are buying or selling each asset.

**Constraints.** We have already mentioned several constraints, for example that $h \geq 0$ (the portfolio is long-only) and $1^T u = c_{\text{des}} - c_{\text{init}}$ (the post-trade cash matches a desired value). We also allow for additional convex constraints on the trade list and post-trade holdings. We represent these as $u \in \mathcal{U}$ and $h \in \mathcal{H}$. These could include, for example, limits on the holdings of a particular asset, or limits on the exposure of our portfolio to a certain factor. For concreteness, we assume $\mathcal{H}$ and $\mathcal{U}$ are polyhedral, i.e., described by a finite set of linear equality and inequality constraints, although our proposed method also applies more generally.

### 2.3 Tax-aware portfolio construction

**Tax-aware utility function.** We assemble our objective terms into a single utility function of $u$ and $h$, which is a weighted combination of the terms listed above,

$$U(h, u) = \alpha^T h - \gamma_{\text{risk}}(h - h_b)^T V(h - h_b) - \gamma_{\text{tc}} \kappa^T |u| - \gamma_{\text{tax}} L(u).$$

(4)

where $\gamma_{\text{risk}}$, $\gamma_{\text{tc}}$, and $\gamma_{\text{tax}}$ are nonnegative trade-off parameters. The first two terms constitute the traditional risk-adjusted (expected) return used in Markowitz portfolio construction. The third term is transaction cost, a widely used addition to the traditional Markowitz utility, with the parameter $\gamma_{\text{tc}}$ used to control turnover. The last term takes into account the tax liability consequences of the trades.
Tax-aware portfolio construction problem. Our problem is to maximize utility subject to constraints, i.e.,

\[
\begin{align*}
\text{maximize} & \quad \alpha^T h - \gamma_{\text{risk}}(h - h_b)^T V (h - h_b) - \gamma_{\text{tc}} \kappa^T |u| - \gamma_{\text{tax}} L(u) \\
\text{subject to} & \quad h = h_{\text{init}} + u, \quad 1^T u = c_{\text{des}} - c_{\text{init}} \\
& \quad u \in U, \quad h \in H,
\end{align*}
\]

with decision variables u and h. The problem data are \(\alpha, h_b, V, \kappa, h_{\text{init}}, c_{\text{des}}, c_{\text{init}}, \) the function \(L\) (described in section 3), the constraint sets \(U\) and \(H\), and the trade-off parameters \(\gamma_{\text{risk}}, \gamma_{\text{tc}}, \) and \(\gamma_{\text{tax}}\). We refer to the problem (5) as the tax-aware Markowitz problem, or TAM problem, and we denote its optimal value as \(U^*\).

Non-convexity. The constraints in the TAM problem (5) are convex, as are all terms in the objective with the exception of the tax liability term. Unfortunately, that term renders the TAM problem (5) nonconvex, which makes it difficult to solve (exactly) in general. We note, however, that the problem becomes convex when we specify the sign of the trade list \(u\), i.e., if we specify for each asset whether we are to sell it \((u_i \leq 0)\) or buy it \((u_i \geq 0)\).

The TAM problem can be formulated as a mixed-integer quadratic program (MIQP), which can be solved using various methods. It is well known that in practice, these methods can often solve problems reasonably quickly, but in many other cases, the solution times can be extremely long. The main contribution of this paper is a method for approximately solving the TAM problem, which involves solving only two convex optimization problems. As a result, our method is always very fast and never involves the very long solution time that can be observed with MIQP solvers. As we will see in section 6, for realistic instances of the TAM problem, our method delivers near identical performance as a globally optimal solution.

3 Tax liability

This section describes the tax liability function.

3.1 Tax lots and capital gains

Tax lots. For each asset, the pre-trade holdings are composed of zero or more tax lots. Each tax lot has several attributes associated with it: its quantity of shares, acquisition date, and cost basis (the price per share at which the shares were acquired).

We let \(q_{ij}\) denote the quantity of shares in the \(j\)th lot of asset \(i\). The total number of shares of asset \(i\) held is \(\sum_j q_{ij}\). We have \(h_{\text{init},i} = p_i \sum_j q_{ij}\), where \(p_i\) is the current price of asset \(i\). We let \(b_{ij}\) denote the cost basis (in dollars per share) of the \(j\)th lot of asset \(i\).
Selling shares. When shares of asset $i$ are sold, i.e., we have $u_i < 0$, we must specify which tax lots from which to take the shares. Let $s_{ij}$ denote the dollar value of shares sold from the $j$th lot of asset $i$, with $0 \leq s_{ij} \leq q_{ij}p_i$, where $q_{ij}p_i$ is the dollar value of the $j$th lot of asset $i$. The total dollar value of asset $i$ sold is then $\sum_j s_{ij}$, which must be equal to $-u_i$.

When we sell $s_{ij}$ dollars from lot $j$ of asset $i$, we incur a capital gain, which is the difference of our proceeds and our cost basis for those shares, i.e., $(1 - b_{ij}/p_i)s_{ij}$. We refer to this quantity as the gain; when it is negative, we refer to it as the loss.

Long-term and short-term gains. A tax lot is long term if the acquisition date is more than one year before the trade date, and the lot is short term otherwise. Gains from long-term and short-term lots are taxed at two different positive rates, $\rho_{lt}$ and $\rho_{st}$, respectively, with $\rho_{lt} \leq \rho_{st}$. The tax liability for selling dollar value $s_{ij}$ from the $j$th lot of asset $i$ is $\rho_{lt}(1 - b_{ij}/p_i)s_{ij}$ if lot $j$ is long term, and $\rho_{st}(1 - b_{ij}/p_i)s_{ij}$ if lot $j$ is short term.

The total tax liability from selling all assets is

$$\sum_{i,j} \rho_{ij}(1 - b_{ij}/p_i)s_{ij} = \sum_{i,j} T_{ij}s_{ij},$$

where the tax rates $\rho_{ij}$ are given by:

$$\rho_{ij} = \begin{cases} 
\rho_{lt} & \text{lot } j \text{ of asset } i \text{ is long term} \\
\rho_{st} & \text{lot } j \text{ of asset } i \text{ is short term}.
\end{cases}$$

We refer to $T_{ij} = \rho_{ij}(1 - b_{ij}/p_i)$ as the tax rate for lot $j$ of asset $i$. This is the dollar tax liability generated per dollar sold of the lot. It is positive if the lot basis exceeds the current asset price, i.e., the lot is held at a gain, and is negative if the lot is held at a loss.

3.2 Tax liability function

Suppose that for asset $i$ we have $u_i < 0$, i.e., we are selling $-u_i$ dollars of asset $i$, which translates to $-u_i/p_i$ shares. We can solve the problem of allocating the sale across lots in order to minimize the tax liability incurred. We define

$$L_i(u_i) = \min_{s_{ij}} \left\{ \sum_j T_{ij}s_{ij} \mid \sum_j s_{ij} = -u_i, \ 0 \leq s_{ij} \leq q_{ij}p_i \right\},$$

which is the smallest tax gain achievable to carry out this sale. We define $L_i(u_i) = +\infty$ for $-u_i < p_i \sum_j q_{ij}$, i.e., if we ask to sell more shares of the asset than we hold. We also define $L_i(u_i) = 0$ for $u_i \geq 0$, i.e., we are buying shares instead of selling. These properties hold because purchasing additional shares incurs no immediate tax liability.
Figure 1: Tax liability functions $L_i$ for two assets. The solid black curve is for an asset with four lots with two held at a loss and two held at a gain. The dashed red curve is for an asset with two lots, both held at a gain.

Least-tax-first-out lot policy. For a given value of $u_i < 0$, it is easy to determine optimal values of $s_{ij}$; it is a convex optimization problem with an analytical solution. We simply sort the values of $T_{ij}$ from least (most negative) to greatest, breaking ties arbitrarily. Then we sell shares from lots in this order. For example, we start by selling shares from the lot with the smallest (or most negative) value of $T_{ij}$, which is the term-adjusted tax liability rate. If we need to sell more shares than that, we go to the lot with second smallest value, and so on. This greedy approach is optimal, i.e., it minimizes the tax liability when selling $-u_i$ dollars of asset $i$. We refer to this approach of choosing lots from which to sell shares as the least tax first out (LTFO) method. Note that it takes into account whether the lots are long term or short term. If all lots are short term, this scheme coincides with the well-known highest basis first out (HIFO) method.

The tax liability function $L_i$ is continuous and piecewise affine, and also has the following convexity properties. If none of the lots are at a loss, $L_i$ is convex and nonnegative. If at least one lot is held at a loss, then $L_i$ takes on negative values and is not convex. However, when the domain of $L_i$ is restricted to either $u_i \leq 0$ or $u_i \geq 0$, the resulting function is convex.

Figure 1 shows two different tax liability functions. The dashed red curve shows the tax liability function for an asset which we hold in two lots, both at a gain (i.e., with current price greater than basis). The solid black curve shows the tax liability function of a different asset, which we hold in four lots, two at a loss (i.e., the basis is greater than the current price), and two at a gain. Each linear segment corresponds to a tax lot, with the slope given by the tax rate of the lot, and width given by the total value of the lot.
4 Convex relaxations

The first step in developing our convex-optimization-based heuristic for approximately solving the TAM problem is to form a convex relaxation or approximation of the problem. We start with a very simple one and then describe one that is more sophisticated.

4.1 Convex envelope of a function

We review a standard concept, the convex envelope of a function $f : \mathbb{R} \rightarrow \mathbb{R}$, denoted $f^{**}$. It is defined as

$$f^{**}(x) = \inf \{ \theta f(v) + (1 - \theta)f(w) \mid \theta \in [0, 1], \ x = \theta v + (1 - \theta)w \}. \quad (6)$$

The infimum is over $\theta, v,$ and $w$. The convex envelope function $f^{**}$ is convex, and it satisfies $f^{**}(x) \leq f(x)$ for all $x$, i.e., it is a global underestimator of $f$. If $f$ is convex, then $f^{**}$ is equal to $f$.

The convex envelope can be defined several other equivalent ways. For example, $f^{**}$ is the greatest convex function that is a global underestimator of $f$. It is also the (Fenchel) conjugate of the conjugate of $f$, i.e., $(f^{**})^*$, where the superscript $*$ is the traditional notation for the conjugate function. (This explains why we denote the convex envelope of $f$ as $f^{**}$.) An example is shown in figure 2.

If $f(x)$ is convex when restricted to $x \leq 0$ and also when restricted to $x \geq 0$, we can require that $v \geq 0$ and $w \leq 0$ in (6), i.e., we can define the convex envelope as

$$f^{**}(x) = \inf \{ \theta f(v) + (1 - \theta)f(w) \mid \theta \in [0, 1], \ x = \theta v + (1 - \theta)w, \ v \geq 0, \ w \leq 0 \}. \quad (7)$$

4.2 Basic convex relaxation

We now focus on the tax liability function $L_i(u_i)$. The convex envelope $L_i^{**}$ is easily found in terms of the point that minimizes it. Let $u_{\min}$ be the value that minimizes $L_i(u_i)$. In fact, $u_{\min}$ is the total value of all lots held at a loss, and $L_i(u_{\min})$ is the
associated minimum capital gain (i.e., maximum possible realized loss). Then we have
\[ L_i^{**}(u_i) = \begin{cases} L_i(u) & u_i \leq u_{\text{min}} \\ L_i(u_{\text{min}}) & u_i > u_{\text{min}}. \end{cases} \]

This function is shown in figure 3.

**Interpretation as ignoring wash sale rule.** The convex envelope \( L_i^{**} \) is a poor approximation of the tax liability function \( L_i \), since it suggests that for any value \( u_i \geq u_{\text{min}} \), you can realize the maximum possible loss, obtained by liquidating all lots held at a loss. Indeed, this can be done: for \( u_i \geq u_{\text{min}} \), you sell all lots held at a loss, and then buy as much of the asset as needed to get to \( u_i \). This scheme generates an immediate capital loss of \(-L_i(u_{\text{min}})\). Unfortunately, it would also be considered a wash sale, and so the loss would be disallowed by the IRS. (The wash sale rule is designed to prevent precisely this.) We can interpret the convex envelope \( L_i^{**}(u_i) \) as the tax liability that could be obtained if we ignore the wash sale rule.

**Basic relaxed TAM problem.** We can now obtain a relaxation of the TAM problem, by replacing \( L_i \) with \( L_i^{**} \), to obtain

\[
\begin{aligned}
\text{maximize} & \quad \alpha^T h - \gamma_{\text{risk}}(h - h_b)^T V(h - h_b) - \gamma_{\text{tc}} \kappa^T |u| - \gamma_{\text{tax}} L^{**}(u) \\
\text{subject to} & \quad h = h_{\text{init}} + u, \quad 1^T u = c_{\text{des}} - c_{\text{init}}, \\
& \quad u \in U, \quad h \in H,
\end{aligned}
\]

with decision variables \( u \) and \( h \). Here \( L^{**}(u) = \sum_{i=1}^n L_i^{**}(u_i) \) is the convex envelope of the tax liability function. This is a convex problem, in fact a quadratic program (QP), and can be solved very efficiently. Our interpretation above, however, suggests that the relaxation might not be that useful, since it essentially pretends that you
can simultaneously sell lots at a loss and buy them back without triggering a wash sale. The next section develops a variation on this relaxation that is more accurate and useful.

Since the objective in the relaxation (8) is an upper bound on the objective in the original TAM problem, its optimal value $U_{br}^\star$ is an upper bound on the optimal value $U^\star$ of the TAM problem. (The subscript ‘br’ stands for ‘basic relaxation’.)

### 4.3 Convex relaxation with borrowed curvature

This section provides a more sophisticated convex relaxation of the TAM problem (5). We first eliminate the post-trade holdings variable $h$ to express the problem in terms of the trade list $u$, with an objective that is the sum of a separable function and one that is not separable. This form is:

$$\text{maximize} \quad -f_0(u) - \sum_{i=1}^n f_i(u_i)$$
subject to $u \in \tilde{U}$

with variable $u \in \mathbb{R}^n$, where the constraint set is

$$\tilde{U} = \{u \in U \mid h_{init} + u \in H, \ 1^T u = c_{init} - c_{des}\}.$$

The constraint set $\tilde{U}$ includes the original constraint $u \in U$ as well as the holdings constraint $h \in H$ and the post-trade cash constraint, and is convex. The non-separable part of the objective function is

$$f_0(u) = \gamma_{\text{risk}}(h_{init} - h_{b} + u)^T \Sigma X^T (h_{init} - h_{b} + u),$$

which is the systematic component of risk (2), and is convex. The separable part corresponding to asset $i$ is

$$f_i(u_i) = -\alpha_i u_i + \gamma_{\text{risk}} D_{ii}(h_{init,i} - h_{b,i} + u_i)^2 + \gamma_{tc} \kappa_i |u_i| + \gamma_{\text{tax}} L_i(u_i),$$

which includes contributions from the expected return, specific risk (3), transaction cost, and tax liability.

The functions $f_i$ are piecewise quadratic and nonconvex in general, but are convex when $u_i \leq 0$ or $u_i \geq 0$. The problem (9), which is equivalent to the original TAM problem (5), is not convex because the functions $f_i$ are not convex. However, if we fix the sign of each $u_i$, the problem (9) becomes convex, and therefore easy to solve. (In fact it suffices to fix the sign of $u_i$ for each asset where we hold at least one lot at a loss; the other $f_i$ are convex.)

**Relaxed TAM problem with borrowed curvature.** We can now obtain another convex relaxation of the TAM problem, by replacing $f_i$ with $f_i^{**}$, to obtain

$$\text{maximize} \quad -f_0(u) - \sum_{i=1}^n f_i^{**}(u_i)$$
subject to $u \in \tilde{U}$. 

(11)
This is a convex problem, which can be formulated as a second-order cone problem (SOCP), as shown in appendix A (Boyd and Vandenberghe, 2004, §4.4.2). The convex envelopes \( f_i^{**} \) are convex and also piecewise quadratic. Figure 4 plots \( f_i \) and \( f_i^{**} \).

The objective of the relaxation (11) is an upper bound on the objective of the original TAM problem. It follows that its optimal objective value \( U_{\text{relax}}^* \) is an upper bound on the optimal value of the original TAM problem. In fact, we have

\[
U^* \leq U_{\text{relax}}^* \leq U_{\text{br}}^*,
\]

i.e., the upper bound found by this relaxation is tighter than the upper bound found from the basic relaxation.

In fact, the gap \( U_{\text{relax}}^* - U^* \) can be bounded in terms of \( k \), the number of factors in the risk model, and the distances between the separable functions \( f_i \) and their convex envelopes \( f_i^{**} \). This is an application of the Shapley–Folkman Lemma (Bertsekas, 1982; Udell and Boyd, 2016).

**TAM problem with approximate tax liability.** By re-introducing the post-trade holdings variable, the relaxed problem (11) can be written as

\[
\begin{align*}
\text{maximize} & \quad \alpha^T h - \gamma_{\text{risk}}(h - h_b)^T V(h - h_b) - \gamma_{\text{te}} \kappa^T |u| - \gamma_{\text{tax}} \hat{L}(u) \\
\text{subject to} & \quad h = h_{\text{init}} + u, \quad \mathbf{1}^T u = c_{\text{des}} - c_{\text{init}} \\
& \quad u \in U, \quad h \in H.
\end{align*}
\]

with decision variables \( u \) and \( h \). This is the TAM problem with the tax liability functions \( L \) replaced by approximate tax liability function \( \hat{L} \), defined as

\[
\hat{L}(u) = \sum_{i=1}^{n} \hat{L}_i(u_i),
\]

where \( \hat{L}_i = L_i + (f_i^{**} - f_i)/\gamma_{\text{tax}} \). Problem (12) can be solved exactly using convex optimization, even though the functions \( \hat{L}_i \) are not convex. This is possible because the nonconvex function \( \hat{L} \) borrows curvature from the other separable objective terms, resulting in an objective function that is concave.

This interpretation recalls the basic relaxation (8), in which \( L \) is replaced by \( L^{**} \). However, because \( \hat{L} \) is typically a much better approximation of \( L \) than \( L^{**} \), this yields a much tighter relaxation. An example is shown in figure 4.

## 5 Approximate solution methods

In this section we describe heuristic solution methods for approximately solving the TAM problem. The methods involve solving two convex optimization problems, and work in two stages.
Figure 4: Left. $f_i(u_i)$ (solid black line) and its convex envelope $f_{i^*}(u_i)$ (blue dashed line). Right. Nonconvex tax liability function $L_i(u_i)$ (solid black line), its convex envelope $L_{i^*}(u_i)$ (dashed green line), and the approximation used in our sophisticated relaxation $\hat{L}_i(u_i)$ (dashed blue line). Even though $\hat{L}_i(u_i)$ is nonconvex, we can still solve the problem globally and efficiently.

1. **Guess the vector $s$ of signs of an optimal $u$.** This is done by solving a relaxation of the TAM problem (which in addition provides an upper bound on $U^*$).

2. **Solve TAM with these sign constraints.** Add the sign constraints $s_i u_i \geq 0$, $i = 1, \ldots, n$ to the TAM problem and solve. With these constraints, the TAM problem is a convex QP and can be efficiently solved.

There are several choices for step 1, which we describe below. We also note that we only need to specify the sign of $u_i$ for assets in which we hold at least one lot at a loss.

**Methods for guessing the sign.** For step 1, we can solve either the basic relaxation (8) or the more sophisticated relaxation (11). There are several choices for guessing the signs of $u_i$ from the solution of one of these relaxations. The most obvious method is to use the sign of the solution of the relaxation, i.e., $s = \text{sign}(u_{br}^*)$ or $s = \text{sign}(u_{relax}^*)$.

A less obvious method is a random choice of the signs with probabilities taken from the solution of the relaxation. For each $i$, we obtain the values $\theta_i$ in the convex envelope definition (7) for $L_i$ (in the basic relaxation) or $f_i$ (in the more sophisticated relaxation). We then set $s_i = 1$ with probability $\theta_i$, and $s_i = -1$ with probability $1 - \theta_i$. (This is done independently for each $i$.) Thus we use the values in the convex envelope as probabilities on whether we buy or sell each asset. This method can be used to generate multiple candidate sign vectors, and we can compare the objectives after step 2 and use the one with the largest objective.
In many numerical experiments we found that the method that performs best is to solve the relaxation (11), and then use the randomized method to guess a set of signs. (This is despite the fact that the simple rounding method is guaranteed to produce feasible sign constraints for the TAM problem, and the randomized method is not.) We have also found that generating multiple sets of candidate signs does not substantially improve the results. This method requires two convex optimization solves: one to solve the relaxation (an SOCP), and one to solve the original TAM problem with the sign of \( u \) fixed (a QP).

6 Numerical examples

We demonstrate these methods by simulating a tax-loss harvesting strategy, in which we solve the TAM problem once a month to generate the trade list. First, we show a six-year backtest of such a strategy. Then, we use this backtest (and others like it) to generate realistic instances of the TAM problem, which we use to evaluate the methods of section 5.

6.1 Benchmark and data

All of our simulations use the S&P 500 as the benchmark, with data over the period 2002 to 2019. Our universe includes all assets that were in the S&P 500 at any point over that time interval, which gives \( n = 998 \). We included a constraint that we only purchase shares of current S&P 500 constituents. This prevents us from purchasing assets that, at the time of the simulated trade, have never been in the benchmark. It also means we don’t increase our holdings of former S&P 500 constituents (but we also do not require them to be immediately sold). If any asset is delisted, we liquidate the asset immediately, incurring the associated tax liability.

We take \( \alpha = 0 \), i.e., we do not have any views on the active returns, so our goal is to simply track the benchmark portfolio while minimizing tax liability. Our risk model parameters \( \Sigma, X, \) and \( D \) are from the Barra US Equity model (Menchero, Orr, and Wang, 2011), which uses \( k = 72 \) factors. Our cash target \( c_{\text{des}} \) is given by (1) with \( \eta = 0.005 \), i.e., we hold 50 basis points in cash after each trade. We use tax rates \( \rho_{\text{lt}} = 0.238 \) and \( \rho_{\text{st}} = 0.408 \), which reflect the current highest marginal tax rates in the United States for long-term and short-term capital gains, respectively. We used the conservative value \( \kappa_i = 0.0005 \) for all transaction costs, i.e., the bid-ask spread is 10 basis points for all assets. The parameter \( \gamma_{\text{risk}} \) was scaled with the account value, so that \( \gamma_{\text{risk}} = \tilde{\gamma}_{\text{risk}}(1^T h_{\text{init}} + c_{\text{init}}) \), with \( \tilde{\gamma}_{\text{risk}} = 200 \). The other trade-off parameters were chosen as

\[
\gamma_{tc} = 1, \quad \gamma_{\text{tax}} = 1.
\]
6.2 Backtests

Our data consists of 204 months, over a 17 year period from August 2002 through August 2019. We use this data to carry out 12 different, staggered six-year-long backtests. The first one starts in August 2002 and ends in July 2008; the last one starts in August 2013 and ends in July 2019. In these backtests, monthly trading means we trade on the first business day more than 31st calendar days after the last trade. For each trade, the initial cash amount $c_{\text{init}}$ is adjusted for the realized transaction cost $\kappa^T|u|$ of the last trade, as well as cash inflows due to dividends and other corporate actions. In the backtests, we round the trade lists to an integral number of shares. Each backtest starts with a portfolio of $1$M in cash.

Each month, the trade list is determined by solving the TAM problem using one of three methods:

- **Basic heuristic.** We use the basic relaxation (8), combined with the simple rounding method.

- **Sophisticated heuristic.** We use the more sophisticated relaxation (11), combined with the randomized rounding method.

- **Mixed-integer method.** We use the mixed-integer mode of CPLEX (version 12.9) to solve the TAM problem directly, with a time limit of 300 seconds.

For both heuristics, we used CPLEX (as a QP/SOCP solver) to solve the convex relaxations and to generate the final trade list.

**Example backtest.** Figure 5 shows the results of one of our backtests, initiated in August 2013. The top plot shows the active risk, and the bottom plot shows the cumulative tax liability, which is the net realized gain, accounting for long- and short-term tax rates. (This quantity is negative, meaning we are realizing a net loss). Here we use the conventional definition of active risk, which is the square root of the definition given in section 2.

These results show that a tax-aware trading scheme can indeed track a benchmark while simultaneously realizing capital losses. It is interesting to note that losses are harvested even during bull markets. The rate of tax-loss harvesting decreases with the life of the fund, since more lots are held as a gain. We note that all three methods achieve good performance, with the mixed-integer method and the sophisticated heuristic modestly outperforming the basic heuristic. Backtests initiated at other months have similar results.

Finally, in figure 6, we show the cumulative time required to carry out this backtest using all three methods. The time limit of 300 seconds for the mixed integer method was used as the solve time for the 17 instances in which CPLEX timed out (out of 59 total); increasing the limit would of course increase the cumulative solve time. Note the very different vertical scales for the two plots, which illustrates that both
Figure 5: The active risk (top) and cumulative tax liability (bottom) of a backtest for all three solution methods.

heuristics are substantially faster than the mixed integer method. The solve time of the mixed integer method is more irregular than the heuristics, i.e., sometimes it solves quickly, but sometimes it takes much longer. (These are seen as the large jumps in cumulative solve time in the top plot.) For this back test the heuristics are about 300× faster than the mixed-integer solver, which itself failed to solve more than one quarter of the problems globally.

6.3 Detailed comparison of solution methods

This section compares the performance and solve time of the three methods. The data we use is taken from the backtests described above. We have 12 backtests, each six years long, giving a total of 744 instances of the TAM problem. (We exclude the initial trade, in which the account holds only cash.) To make the utility (4) comparable across the problem instances, we divide it by the account value $1^T h_{init} + c_{init}$. (This number is the monthly after-tax expected return adjusted for risk and transaction costs and is measured in percent or basis points.)

Comparison of heuristics to their bounds. For the 744 instances of the TAM problem, we compute the utility achieved by solving the basic heuristic and the sophisticated heuristic, denoted $U_{br, round}$ and $U_{relax, round}$. We compare these with the bounds $U_{br}^*$ and $U_{relax}^*$, obtained by solving the two relaxations (8) and (11). In figure 7, we show histograms of the suboptimality gaps $U_{br} - U_{br, round}$ and $U_{relax} - U_{relax, round}$. 

These differences must be nonnegative; when the differences are equal to zero, this means that the heuristic has solved the problem and produced a certificate of optimality.

The basic heuristic is optimal in roughly one quarter of cases, but it is often suboptimal, sometimes by more than 50 bp. (The greatest suboptimality gap, which is not depicted, is 147 bp.) On the other hand, the sophisticated heuristic produces provably optimal trades for the vast majority of problem instances. In all other problem instances, the sophisticated heuristic is never suboptimal by more than a few basis points.

Comparison of heuristics to mixed-integer method. We now compare the heuristics to the mixed-integer method. We denote by $U_{\text{mip}}$ the utility obtained by the mixed-integer method, which we compare to $U_{\text{br, round}}$ and $U_{\text{relax, round}}$, the utility obtained by our two heuristic methods. Note that if the CPLEX solves the mixed-integer problem within the 300 second time limit, we have $U_{\text{mip}} = U^*$. Figure 8 shows histograms of $U_{\text{mip}} - U_{\text{br, round}}$ and $U_{\text{mip}} - U_{\text{relax, round}}$. (For problem instances where $U_{\text{mip}} = U^*$, these values are nonnegative, but can be negative otherwise.) We observe that the utility obtained by the basic relaxation often matches the mixed-integer method, but is occasionally much larger, by as much as 2 percent. The utility obtained by the sophisticated relaxation almost always matches the mixed-integer method; indeed, the two methods never differ by more than 2 basis points, despite
the heuristic being much faster (as we will examine next).

We now compare the sophisticated heuristic and the mixed-integer method in more detail. The mixed-integer method times out (and therefore, does not necessarily globally solve the problem) in 179 of the 744 cases. Among the 565 instances that the mixed-integer method solves (within 300 seconds), in 549 instances the sophisticated method solves the problem to within numerical precision, which is a utility of 0.05 basis points. It is never more than 0.3 basis points suboptimal in the other 16 cases. For the 179 cases in which the mixed integer method fails to solve the problem, we observe that the sophisticated heuristic achieves a normalized utility within 0.05 bp of the mixed integer method in 89 cases, and outperforms it by more than 0.05 basis points in 68 cases (by up to 2 bp). In only 22 of the 179 instances did the sophisticated heuristic underperform the mixed integer method by more than 0.05 basis points, and never by more than 2 basis points.

**Solve times.** Figure 9 shows the solve times for the sophisticated heuristic and the mixed integer method on a scatter plot. Note that all of the points are below the dashed black line, which indicates that the heuristic method was faster in all cases. A significant proportion (179 out of 744) of the problem instances took 300 seconds using the mixed-integer method, which was the maximum time allowed. The solve times of the basic heuristic were similar to the sophisticated heuristic, and are not shown.
Figure 8: The differences in (scaled) utility achieved by the two heuristics compared to the utility achieved by the mixed integer method.

Figure 9: The solve times of the 720 problem instances using the relax-and-round heuristic and the mixed-integer solution, with each problem instance shown as a single dot. The dashed red line shows the maximum allowed time of the mixed-integer solver.
7 Conclusion

We formulate tax-aware portfolio construction as a nonconvex optimization problem, and we present a heuristic for this problem based on convex optimization. This method is reliably fast: for problems with several hundred assets and several dozen factors, it takes less than a second. We compare our heuristic against the standard, mixed-integer quadratic programming formulation, solved using CPLEX, on realistic problem instances. When the mixed-integer method is limited to five minute solve times, we find that our heuristic outperforms it more often than not, despite being several hundred times faster. This speed is not necessary for monthly (or even daily) trading, but is useful for backtesting and Monte Carlo simulation, possibly over hundreds of thousands of individualized accounts. Our method also produces a bound on the optimal value. For realistic data, the bound usually tight enough that it certifies that the heuristic solved the problem globally. In future work, we will extend this method to other nonconvex terms that are often present in practical portfolio optimization problems.

Acknowledgements. We would like to thank Emmanuel Candès for useful discussions and feedback. We would also like to thank Eric Kisslinger for identifying an important error in an early version of the software.
References

Atra, R. and Y. Pae (2013). “Likely Benefits from HIFO Accounting”. In: Midwest Finance Association Annual Meeting.

Berkin, A. and J. Ye (2003). “Tax Management, Loss Harvesting, and HIFO Accounting”. In: Financial Analysts Journal 59.4, pp. 91–102.

Bertsekas, D. (1982). Constrained optimization and Lagrange multiplier methods. Academic Press.

Bertsekas, D. (1997). “Nonlinear programming”. In: Journal of the Operational Research Society 48.3, pp. 334–334.

Bertsimas, D., C. Darnell, and R. Soucy (1999). “Portfolio construction through mixed-integer programming at Grantham, Mayo, Van Otterloo and Company”. In: Interfaces 29.1, pp. 49–66.

Boyd, S. and L. Vandenberghe (2004). Convex optimization. Cambridge University Press.

Boyd, S., E. Busseti, S. Diamond, R. Kahn, P. Nystrup, and J. Speth (2017). “Multi-Period Trading via Convex Optimization”. In: Foundations and Trends in Optimization 3.1, pp. 1–76.

Chaudhuri, S., T. Burnham, and A. Lo (2020). “An Empirical Evaluation of Tax-Loss-Harvesting Alpha”. In: Financial Analysts Journal, p. 1.

Constantinides, G. (1983). “Capital Market Equilibrium with Personal Tax”. In: Econometrica 51.3, p. 611.

Constantinides, G. (1984). “Optimal stock trading with personal taxes”. In: Journal of Financial Economics 13.1, pp. 65–89.

Dammon, R. and C. Spatt (1996). “The Optimal Trading and Pricing of Securities with Asymmetric Capital Gains Taxes and Transaction Costs”. In: Review of Financial Studies 9.3, pp. 921–952.

Dammon, R., C. Spatt, and H. Zhang (2001). “Optimal Consumption and Investment with Capital Gains Taxes”. In: Review of Financial Studies 14.3, p. 5.

Dammon, R., C. Spatt, and H. Zhang (2004). “Optimal Asset Location and Allocation with Taxable and Tax-Deferred Investing”. In: Journal of Finance 59.3, pp. 999–1037.

DeMiguel, V. and R. Uppal (2005). “Portfolio Investment with the Exact Tax Basis via Nonlinear Programming”. In: Management Science 51.2, pp. 277–290.

Diamond, S. and S. Boyd (2016). “CVXPY: A Python-embedded modeling language for convex optimization”. In: Journal of Machine Learning Research 17.83, pp. 1–5.

Diamond, S., R. Takapoui, and S. Boyd (2018). “A general system for heuristic minimization of convex functions over non-convex sets”. In: Optimization Methods and Software 33.1, pp. 165–193.

Dickson, J., J. Shoven, and C. Sialm (2000). “Tax Externalities of Equity Mutual Funds”. In: National Tax Journal 53.3, pp. 607–628.
Domahidi, A., E. Chu, and S. Boyd (2013). “ECOS: An SOCP solver for embedded systems”. In: European Control Conference, pp. 3071–3076.

Dybvig, P. and H. Koo (1996). Investment with Taxes. Tech. rep. Washington University in Saint Louis.

Fu, A., B. Narasimhan, and S. Boyd (2020). “CVXR: An R package for disciplined convex optimization”. To appear, Journal of Statistical Software.

Gallmeyer, M. and S. Srivastava (2011). “Arbitrage and the tax code”. In: Mathematics and Financial Economics 4.3, pp. 183–221.

Grant, M. and S. Boyd (2008). “Graph implementations for nonsmooth convex programs”. In: Recent Advances in Learning and Control. Ed. by V. Blondel, S. Boyd, and H. Kimura. Lecture Notes in Control and Information Sciences. Springer-Verlag Limited, pp. 95–110.

Grant, M. and S. Boyd (2014). CVX: Matlab Software for Disciplined Convex Programming, version 2.1. http://cvxr.com/cvx.

Grinold, R. and R. Kahn (1999). Active portfolio management. second. McGraw-Hill.

Gurobi Optimization LLC (2020). Gurobi Optimizer Reference Manual. http://www.gurobi.com.

IBM Corporation (2019). CPLEX. https://www.ibm.com/support/knowledgecenter/SSSA5P_12.9.0/ilog.odms.studio.help/Optimization_Studio/topics/COS_home.html. Version 12.9.

Lauer, G., N. Sandell, D. Bertsekas, and T. Posbergh (1982). “Solution of large-scale optimal unit commitment problems”. In: IEEE Transactions on Power Apparatus and Systems 101.1, pp. 79–86.

Lobo, M., M. Fazel, and S. Boyd (2007). “Portfolio optimization with linear and fixed transaction costs”. In: Annals of Operations Research 152.1, pp. 341–365.

Makhorin, A. (2016). GNU Linear Programming Kit. https://www.gnu.org/software/glpk/. GNU Project.

Markowitz, H. (1952). “Portfolio Selection”. In: Journal of Finance 7.1, pp. 77–91.

Markowitz, H. (1955). The optimization of a quadratic function subject to linear constraints. Tech. rep. RAND Corporation.

Menchero, J., D. Orr, and J. Wang (2011). The Barra US equity model (USE4), methodology notes. English. MSCI.

Moehle, N. and S. Boyd (2015). “A perspective-based convex relaxation for switched-affine optimal control”. In: Systems & Control Letters 86, pp. 34–40.

MOSEK ApS (2019). The MOSEK optimization toolbox for MATLAB manual, version 9.0. http://docs.mosek.com/9.0/toolbox/index.html.

O’Donoghue, B., E. Chu, N. Parikh, and S. Boyd (2016). “Conic Optimization via Operator Splitting and Homogeneous Self-Dual Embedding”. In: Journal of Optimization Theory and Applications 169.3, pp. 1042–1068.

O’Donoghue, B., E. Chu, N. Parikh, and S. Boyd (2019). SCS: Splitting Conic Solver, version 2.1.2. https://github.com/cvxgrp/scs.
Pogue, G. (1970). “An extension of the Markowitz portfolio selection model to include variable transactions’ costs, short sales, leverage policies and taxes”. In: The Journal of Finance 25.5, pp. 1005–1027.
Sharpe, W. (1963). “A simplified model for portfolio analysis”. In: Management Science 9.2, pp. 277–293.
Sialm, C. and H. Zhang (2020). “Tax-Efficient Asset Management: Evidence from Equity Mutual Funds”. In: Journal of Finance 75.2, pp. 735–777.
Starr, R. (1969). “Quasi-equilibria in markets with non-convex preferences”. In: Econometrica: Journal of the Econometric Society, pp. 25–38.
Stellato, B., G. Banjac, P. Goulart, A. Bemporad, and S. Boyd (2020). “OSQP: An Operator Splitting Solver for Quadratic Programs”. To appear, Mathematical Programming Computation.
Udell, M. and S. Boyd (2013). “Maximizing a Sum of Sigmoids”. Unpublished. http://web.stanford.edu/~boyd/papers/max_sum_sigmoids.html.
Udell, M. and S. Boyd (2016). “Bounding duality gap for separable problems with linear constraints”. In: Computational Optimization and Applications 64.2, pp. 355–378.
Udell, M., K. Mohan, D. Zeng, J. Hong, S. Diamond, and S. Boyd (2014). “Convex optimization in Julia”. In: Workshop for High Performance Technical Computing in Dynamic Languages. IEEE, pp. 18–28.
A SOCP formulation

Here we explain how to represent the convex envelope \( f_{i}^{**} \) in a cone program, by expressing its epigraph using a cone representation, as described by Grant and Boyd (2008). The technique given here are similar to those used to represent perspectives of convex functions (Moehle and Boyd, 2015, § 2).

Consider a function \( f : \mathbb{R} \to \mathbb{R} \cup \{\infty\} \), of the form

\[
f(x) = \begin{cases} f_-(x) & x < 0 \\ f_+(x) & x \geq 0, \end{cases}
\]

where \( f_- \) and \( f_+ \) are both convex, with \( f_-(x) = +\infty \) for \( x > 0 \) and \( f_+(x) = +\infty \) for \( x < 0 \). We assume that each of these functions has a so-called cone representation. This means that \( f_- \) is the optimal value of a cone program

\[
\begin{align*}
\text{minimize} & \quad c^T z_- \\
\text{subject to} & \quad A_-(x, z_-) = b_-, \quad (x, z_-) \in \mathcal{K}_-, \\
\end{align*}
\]

with variable \( z_- \), where \( \mathcal{K}_- \) is a cone. We assume a similar representation for \( f_+ \).

Our goal is to represent the convex envelope (7) as the optimal value of a cone program. Using the cone representations of \( f_- \) and \( f_+ \), we can express \( f^{**}(x) \) as the optimal value of the problem

\[
\begin{align*}
\text{minimize} & \quad \theta c^T z_- + (1 - \theta)c^T z_+ \\
\text{subject to} & \quad A_-(v, z_-) = b_-, \quad (v, z_-) \in \mathcal{K}_- \\
& \quad A_+(w, z_+) = b_+, \quad (w, z_+) \in \mathcal{K}_+, \\
& \quad x = \theta v + (1 - \theta)w, \\
& \quad v \geq 0, \quad w \leq 0, \quad 0 \leq \theta \leq 1,
\end{align*}
\]

with variables \( \theta, z_- \), \( z_+ \), \( v \), and \( w \). The objective terms and the equality constraint involving \( x \) contain the product of two variables, and is not convex.

We will now change variables to obtain an equivalent convex problem. Define the variables

\[
\tilde{z}_- = \theta z_- , \quad \tilde{v} = \theta v, \quad \tilde{z}_+ = (1 - \theta)z_+, \quad \tilde{w} = (1 - \theta)w.
\]

We can express the problem above using these variables, and the original variable \( \theta \), as

\[
\begin{align*}
\text{minimize} & \quad c^T \tilde{z}_- + c^T \tilde{z}_+ \\
\text{subject to} & \quad A_-(\tilde{v}, \tilde{z}_-) = \theta b_-, \quad (\tilde{v}, \tilde{z}_-) \in \mathcal{K}_-, \\
& \quad A_+(\tilde{w}, \tilde{z}_+) = (1 - \theta)b_+, \quad (\tilde{w}, \tilde{z}_+) \in \mathcal{K}_+, \\
& \quad x = \tilde{v} + \tilde{w}, \\
& \quad \tilde{v} \geq 0, \quad \tilde{w} \leq 0, \quad 0 \leq \theta \leq 1,
\end{align*}
\]
with variables \( \theta, \tilde{z}_-, \tilde{z}_+, \tilde{v}, \) and \( \tilde{w}. \) This problem is jointly convex in all variables, and \( x, \) so it is a cone representation of \( f^{**}. \)

We note that for the change of variables (13) to be invertible, we must include in problem (14) the constraint that \( \tilde{z}_- \) must be 0 if \( \theta \) is 0. Because this additional constraint only restricts points on the boundary of the feasible set of problem (14), we can safely ignore it without changing the optimal value of the problem, assuming Slater’s condition holds. Similar arguments apply for \( \tilde{v}, \tilde{z}_+, \) and \( \tilde{w}. \)

For the specific case where \( f \) is piecewise affine (e.g., a tax liability function \( L_i \) for a single asset), the cone representations of \( f^- \) and \( f^+ \) are linear programs (LPs), and therefore so is the cone representation of \( f^{**}. \) This means the basic relaxation (8) is a quadratic program. Likewise, if \( f \) is piecewise quadratic (e.g., the separable cost functions \( f_i \) given in (10)), the cone representations of \( f^-, f^+, \) and \( f^{**} \) are second-order cone programs (SOCPs). This means the sophisticated relaxation (11) is an SOCP.