Proceedings of the
47th IEEE Conference on Decision and Control
Cancun, Mexico, Dec. 9-11, 2008

Multistage Investments With Recourse: a Single-Asset Case with Transaction Costs

Ufuk Topcu, Giuseppe Calafiore, and Laurent El Ghaoui

***Abstract*— We consider a financial decision problem involving dynamic investment decisions on a single risky instrument over multiple and discrete time periods. Investment returns are assumed stochastic and possibly dependent over time, and proportional transaction costs are considered in the model. In this setting, the investor's goal is to determine investment policies that maximize the net profit while maintaining the associated risk under control. We propose approximations of the ensuing stochastic multistage optimization problem that are based on affine recourse strategies and that lead to efficiently solvable second order cone or semidefinite programs.**

I. Introduction

In this paper we study a multi-period financial decision problem involving trading in a single financial instrument over a finite decision horizon. The goal is to determine an investment policy that maximizes the net profit over the investment horizon, taking into account the associated risk and cost of transaction. Returns of the investment over successive periods are modeled as possibly time-correlated random variables, with known expectations and covariances, [1]. The problem is here tackled using a multi-period investment strategy with recourse, aimed at adjusting the current position based on the information about previous positions and associated returns. Due to randomness of the return process, the approach with recourse leads to a multi-stage decision problem with stochastic objective and constraints.

The mainstream approach for dealing with multi-period problems with recourse is currently provided by stochastic programming, see e.g. [2], [3], [4] and the many references therein. In the specific context of financial allocation, stochastic programming methods have been proposed by many authors, see for instance [5], [6], [7] and the survey in [8]. Note however that stochastic programming formulations invariably result in computationally intractable problems (see for instance [9]), that need be approximated via exponentially growing scenario trees, see [10].

In this paper, we follow a radically different route, which avoids stochastic approximations and scenario trees. Specifically, we focus on affine recourse strategies in which the current position is determined as an affine function of finitely many previous returns. This approach leads to computationally tractable optimization problems (second order cone and semidefinite programming problems) whose solutions provide suboptimal policies, since affinely parameterized recourse functions constitute only a subclass of the set of possible recourse relations. Affine recourse has been used for two-stage portfolio allocation problems in [11], where the random returns at different times are considered independent. Here, we examine the case of dependent returns focusing on single-asset case.

We take the effect of transaction costs into account by introducing in the utility function a penalty term related to the total cost of transactions. We develop two different models for handling the risk associated to the investment. In the first model we consider a quadratic penalty in the utility function which involves the sum of squared conditional volatilities for the returns in each period.[1] In the second model we directly impose a constraint on the probability of the net profit being lower than a pre-specified threshold. The first formulation leads to a second order cone programming problem, whereas the second results in a semidefinite programming problem. Although the latter approach for risk modeling provides more flexibility (as detailed in Section V), this extra freedom comes with higher computational cost of semidefinite programming compared to second order cone programming.

II. Problem setup.

We consider an investment decision problem where an investor needs to decide how much money to allocate in a single financial instrument (asset) at each stage of a decision horizon composed of $n > 1$ periods. Let p_t denote the market value or price of this asset at time t and let u_t be the "bet" made by the investor (that is, the amount of money invested in the instrument) at time t, for $t = 1, \ldots, n$, u_0 being the given initial position. The outcome of the bet is dictated by the underlying price fluctuation, which is random: if u_t is invested at t, the value of this investment at time $t+1$ is $\frac{p_{t+1}}{p_t}u_t$, hence the *net return* of the investment in this period is $y_t u_t$, where y_t denotes the random return in period t, that is

$$y_t \doteq \frac{p_{t+1} - p_t}{p_t}, \quad t = 1, \ldots, n.$$

The returns are possibly correlated in time, hence the vector $y = [y_1 \ \cdots \ y_n]^T$ containing the returns over the forward periods is assumed to be such that

$$y \doteq y(x) = \hat{y} + Yx,$$

U. Topcu is with the Department of Mechanical Engineering, University of California, Berkeley, CA, 94720 USA. e-mail: utopcu@jagger.me.berkeley.edu.

G. Calafiore is with the Dipartimento di Automatica e Informatica, Politecnico di Torino, Italy. e-mail: giuseppe.calafiore@polito.it

Laurent El Ghaoui is with the Department of Electrical Engineering and Computer Science, University of California, Berkeley, CA, 94720, USA. e-mail: elghaoui@eecs.berkeley.edu.

[1]We incidentally remark that such risk measures are not *coherent* in the sense of [12]. However, these type of risk terms are certainly still the most widely used and accepted in practice.

where $\hat{y} \doteq \mathbf{E}(y)$ is the return expectation, $Y \in \mathbf{R}^{n\times n}$ is a lower-triangular matrix describing the correlation in time of returns, such that $\mathbf{var}(y) = YY^T$ (note that Y is diagonal if returns are assumed uncorrelated), and $x \in \mathbf{R}^n$ is a zero-mean and unit covariance vector capturing the unpredictable stochastic part of the returns.

One basic goal of the investor is clearly to accumulate positive returns over time, that is to maximize the expectation of $\sum_t y_t u_t$. However, a sensible investment the problem is really a multi-criterion one, since the investor also needs to take into some account the cost of transactions and the exposure to risk. To this end, we introduce the following three terms:

$$\mathcal{P} \doteq \sum_{t=1}^{n} y_t u_t = y^T(x)u \quad \text{:cumulative profit} \tag{1}$$

$$\mathcal{R} \doteq \sum_{t=1}^{n} \sigma_t^2 u_t^2 = u^T \Sigma u \quad \text{:risk term} \tag{2}$$

$$\mathcal{C} \doteq \sum_{t=1}^{n} c|u_t - g_t u_{t-1}| + c|g_n u_n| \tag{3}$$

$$\text{:transaction cost term,}$$

where σ_t^2 is the variance of y_t, $\Sigma \doteq \mathbf{diag}(\sigma_1^2, \dots, \sigma_n^2)$, $g_t \doteq (1 + \hat{y}_t)$, and $c \geq 0$ is the unit transaction cost. Note that $\mathcal{P}$ represents the cumulative return of the investments, $\mathcal{R}$ accounts for risk (intended as level of deviation from expectation, or volatility) in each period with a term proportional to the variance of the investment return conditional to the investment level being u_t, and $\mathcal{C}$ is a term accounting for cost of transactions. Specifically, this term represents the total cost of transactions, conditional to the returns being fixed to their expectation $\hat{y}$.

We next introduce a dynamic decision setup in which the vector of investment decisions $u = [u_1 \cdots u_n]^T$ is a generic strictly causal function $u(x)$ of the return innovations. That is, we are concerned with the following two stochastic optimal control problems, where the first problem is

$$P_1 : \quad J^*_{cl1} := \max_{u\in\mathcal{U}} \mathbf{E}_x \{\mathcal{P} - \mathcal{C} - \mathcal{R}\}, \tag{4}$$

being $\mathcal{U}$ is the class of causal functions such that the corresponding expected objective exists, and the second problem is

$$P_2 : \quad J^*_{cl2} := \max_{u\in\mathcal{U}} \mu - \mathbf{E}_x \{\mathcal{C}\} \tag{5}$$
$$\text{subject to:} \quad \sup_{x\sim(0,I)} \mathbf{Prob}\{x : \mathcal{P} < \mu\} \leq \epsilon,$$

where $x \sim (\hat{x}, X)$ means that x has first and second centered moments equal to $\hat{x}$ and X, $\epsilon \in (0, 1)$ is a pre-specified constant, and the supremum is taken over all distributions having the specified moments.

Notice that in the case of problem P_1 risk is handled by means of the penalty term $\mathcal{R}$ in the utility function. In problem P_2, risk is instead accounted for directly as a worst-case downside probability, and we seek to maximize an upper bound μ on the net profit that is violated with a probability at most ϵ. In a single-period setting and when the probability distribution of returns is known, this latter approach to risk control is known in the literature as the Value-at-Risk (V@R) approach, see, e.g., [13], [14]. Here we do not assume that the exact probability distribution of the returns is known. Instead, we consider a class of all possible distributions having the specified first and second moments, and enforce the downside probability constraint robustly over this class. In this sense, we name the condition in problem (5) a worst-case Value-at-Risk constraint.

We remark further that considering a general set $\mathcal{U}$ of causal recourse functions in problems (4), (5) makes them extremely hard to solve in practice. Indeed, even the computation of the objective may not be tractable for a general set $\mathcal{U}$ of causal functions. On the other extreme, a conventional approximation providing lower bounds on the objectives is to restrict $\mathcal{U}$ to the class of "open-loop" strategies, that is, control functions that are actually independent of the returns. This latter approach, however, might be too coarse and fails to capture the dynamic nature of the decision problem at hand.

In the following sections we shall consider tractable approximations to the above problems that use a restricted class of recourse functions (namely, affine recourse functions) and that lead to efficiently computable convex programs. In particular, in Section III we introduce an affine class of recourse policies and develop suitable approximations for the expectation of the transaction cost term $\mathcal{C}$. With these premises, we next show in Section IV that problem P_1 can be approximately reformulated and efficiently solved as a second order cone program (SOCP), [15]. Section V is devoted to the solution of problem P_2, and Section VI provides a simple numerical example. Conclusions are finally drawn in Section VII. Some of the technical proofs are reported in the Appendix.

III. Affine recourse strategy and transactions penalty

In our approach we postulate that the position u_t for $t > 1$ is an affine, strictly causal function of the returns' innovations (we need strict causality since a decision u_t shall not depend on returns $y_t, y_{t+1}, \dots$ that have not yet been observed at time t), specifically:

$$u_t(x) = \hat{u}_t + U_{t,1}x_1 + \ldots + U_{t,t-1}x_{t-1}, \quad t = 2, \dots, n,$$

where $\hat{u}_t, U_{t,i}$ are new variables. More compactly, we assume that u is the affine, strictly causal function

$$u(x) = \hat{u} + Ux, \tag{6}$$

where $\hat{u} \in \mathbf{R}^n$ and $U \in \mathbf{R}^{n\times n}$ is strictly lower triangular (that is, U is lower triangular with zeros on the main diagonal). We hence consider the following lower bound on problem (4):

$$J^*_{ar1} \doteq \max_{\hat{u},U} \mathbf{E}_x \{\mathcal{P} - \mathcal{C} - \mathcal{R}\}, \tag{7}$$

where the new decision variables are the vector $\hat{u}$ and the matrix U. Plugging the recourse policy (6) into (1) and (2) and taking expectation, we obtain that

$$\begin{aligned}\mathbf{E}_x \mathcal{P} &= \mathbf{E}_x\, y^T(x)u(x) = \hat{y}^T\hat{u} + \mathbf{Tr}\, Y^T U \\ \mathbf{E}_x \mathcal{R} &= \mathbf{E}_x\, u^T(x)\Sigma u(x) = \mathbf{Tr}\,(\hat{u}\hat{u}^T + UU^T)\Sigma.\end{aligned}$$

The expectation of the transactions cost penalty term $\mathcal{C}$, instead, cannot be determined exactly in closed form. Hence, we next consider computable upper and lower bounds for $\mathbf{E}_x \mathcal{C}$, that result to be useful for approximating the solution of (7).

A. Approximations for the transaction cost penalty term

The transaction penalty term $\mathcal{C}$ under the control policy (6) takes the form

$$\begin{aligned}\mathcal{C} &= \sum_{t=1}^{n} |c(U_t^T - g_t U_{t-1}^T)x + c(\hat{u}_t - g_t\hat{u}_{t-1})| \\ &\qquad + |cg_n U_n^T x + cg_n\hat{u}_n| \\ &= \|Ax+b\|_1,\end{aligned}$$

where U_t^T denotes the t-th row in U, and where we set for $t = 1,\ldots,n$

$$\begin{aligned}&a_t^T \doteq c(U_t^T - g_tU_{t-1}^T), \qquad b_t \doteq c(\hat{u}_t - g_t\hat{u}_{t-1}),\\ &a_{n+1}^T \doteq cg_nU_n^T, \qquad b_{n+1} \doteq cg_n\hat{u}_n,\\ &A \doteq \begin{bmatrix} a_1^T \\ \vdots \\ a_{n+1}^T\end{bmatrix} = A(U), \quad b \doteq \begin{bmatrix} b_1 \\ \vdots \\ b_{n+1}\end{bmatrix} = b(\hat{u}).\end{aligned}$$

1) A lower bound on the expected transaction penalty: A lower bound on $\mathbf{E}_x \mathcal{C}$ can be derived from Jensen's inequality, [15], which states that for any random variable X and convex function $f(\cdot)$ it holds that $f(\mathbf{E}(X)) \leq \mathbf{E}(f(X))$. In particular, considering $f(\cdot) = |\cdot|$, we have $|\mathbf{E}X| \leq \mathbf{E}|X|$, thus

$$\begin{aligned}\mathbf{E}_x\, \mathcal{C}(x) = \mathbf{E}_x \|Ax+b\|_1 &\geq \sum_{t=1}^{n+1} |\mathbf{E}_x\,(a_t^T x + b_t)| \\ &= \sum_{t=1}^{n+1} |b_t| = \|b(\hat{u})\|_1 \doteq \mathcal{C}_{lb}(\hat{u}),\end{aligned}$$

from which we obtain that

$$\mathbf{E}_x\{\mathcal{P} - \mathcal{C} - \mathcal{R}\} \leq J_{ub}(\hat{u}, U),$$

where

$$\begin{aligned}J_{ub}(\hat{u},U) &\doteq \mathbf{E}_x\mathcal{P} - \mathbf{E}_x\mathcal{R} - \mathcal{C}_{lb}(\hat{u}) \\ &= (\hat{y}^T\hat{u} + \mathbf{Tr}\,Y^TU) \\ &\qquad -(\mathbf{Tr}\,(\hat{u}\hat{u}^T + UU^T)\Sigma) - \|b(\hat{u})\|_1.\end{aligned}$$

2) An upper bound on the expected transaction penalty: Applying again Jensen's inequality to the random variable $|X|$, with $f(\cdot) = (\cdot)^2$, we obtain $\mathbf{E}|X| \leq \sqrt{\mathbf{E}X^2}$, hence

$$\begin{aligned}\mathbf{E}_x\,\mathcal{C} = \mathbf{E}_x\|Ax+b\|_1 &\leq \sum_{t=1}^{n+1} \sqrt{\mathbf{E}_x\,(a_t^Tx + b_t)^2} \\ &= \sum_{t=1}^{n+1}\sqrt{a_t^Ta_t + b_t^2} \\ &= \sum_{t=1}^{n+1}\left\|\begin{bmatrix} a_t \\ b_t\end{bmatrix}\right\|_2 \doteq \mathcal{C}_{ub}(\hat{u},U), \quad (8)\end{aligned}$$

from which it follows that

$$\mathbf{E}_x\{\mathcal{P} - \mathcal{R} - \mathcal{C}\} \geq J_{lb}(\hat{u},U),$$

with

$$\begin{aligned}J_{lb}(\hat{u},U) &\doteq \mathbf{E}_x\mathcal{P} - \mathbf{E}_x\mathcal{R} - \mathcal{C}_{ub}(\hat{u},U) \\ &= (\hat{y}^T\hat{u} + \mathbf{Tr}\,Y^TU) - (\mathbf{Tr}\,(\hat{u}\hat{u}^T + UU^T)\Sigma) \\ &\qquad - \sum_{t=1}^{n+1}\left\|\begin{bmatrix} a_t \\ b_t\end{bmatrix}\right\|_2.\end{aligned}$$

The following lemma shows that the bound (8) has a relative accuracy of at least $2/\pi$. A proof for this lemma is given in the Appendix.

Lemma 1: Let $A \in \mathbf{R}^{n\times m}$, $b \in \mathbf{R}^n$, and define

$$\psi := \sup_{x\sim(0,I)} \mathbf{E}_x\|Ax+b\|_1, \quad \overline{\psi} := \sum_{i=1}^{n}\left\|\begin{pmatrix} a_i \\ b_i\end{pmatrix}\right\|_2,$$

where a_i^T (resp. b_i) are the rows (resp. elements) of A (resp. b). Then,

$$\frac{2}{\pi}\overline{\psi} \leq \psi \leq \overline{\psi}. \quad (9)$$

$\triangleleft$

IV. Efficiently computable bounds on J_{ar1}^*

In this section we focus on the computation of upper and lower bounds for problem P_1 in (7), based on the developments of the previous section. We have the following proposition.

Proposition 1: The following inequality holds for the optimal value $\mathcal{J}_{ar1}^*$ of problem P_1 under affine recourse rule:

$$J_{lb1}^* \leq \mathcal{J}_{ar1}^* \leq J_{ub1}^*,$$

where J_{ub1}^* is the optimal value of the convex quadratic programming problem:

$$\begin{aligned}J_{ub1}^*: \quad \max_{\hat{u},U,z\geq 0} \quad & \hat{y}^T\hat{u} - \mathbf{Tr}\,\hat{u}\hat{u}^T\Sigma - \textstyle\sum_{i=1}^{n+1} z_i \\ & + \mathbf{Tr}\,(Y^TU - UU^T\Sigma) \\ \text{subject to:} \quad & U \text{ strictly lower triangular}, \hat{u}_0 = u_0, \\ & -z_i \leq b_i(\hat{u}) \leq z_i, \quad i = 1,\ldots,n+1.\end{aligned}$$

and J_{lb1}^* is the optimal value of the following SOCP:

$$\begin{aligned}J_{lb1}^*: \quad \max_{\hat{u},U,\mu_1,\mu_2} \quad & (\hat{y}^T\hat{u} + \mathbf{Tr}\,Y^TU) - \mu_1 - \mu_2 \quad (10) \\ \text{subject to:} \quad & U \text{ strictly lower triangular}, \hat{u}_0 = u_0, \\ & \mathbf{Tr}\,(\hat{u}\hat{u}^T + UU^T)\Sigma \leq \mu_1, \\ & \textstyle\sum_{t=1}^{n+1}\left\|\begin{bmatrix} a_t \\ b_t\end{bmatrix}\right\|_2 \leq \mu_2.\end{aligned}$$

$\triangleleft$

Remark 1: Notice that, since $\mathcal{C}_{lb}$ only depends on $\hat{u}$ (and not on the recourse parameter U), the upper bound problem (10) results to be separable in the variables $\hat{u}$ and U. This means that the nominal investments $\hat{u}$ are found as the solution of the reduced problem

$$\max_{\hat{u},z\geq 0} \quad \hat{y}^T\hat{u} - \mathbf{Tr}\,\hat{u}\hat{u}^T\Sigma - \sum_{i=1}^{n+1} z_i \qquad (11)$$
$$\text{subject to:} \quad \hat{u}_0 = u_0,$$
$$-z_i \leq b_i(\hat{u}) \leq z_i, \quad i = 1,\ldots,n+1,$$

which simply corresponds to the "open-loop" problem one would encounter if no recourse action was applied. Once the optimal $\hat{u}$ is computed by solving (11), the optimal recourse parameter U for problem (10) is determined as the strictly lower triangular matrix that maximizes $h(U) = \mathbf{Tr}(UY^T + UU^T\Sigma)$. To this end, note that if we write $Y = Y_d + Y_{slt}$, where Y_d is the diagonal part of Y and Y_{slt} is its strictly lower triangular part, then $h(U) = \mathbf{Tr}(UY_d + UY_{slt}^T + UU^T\Sigma) = \mathbf{Tr}(UY_{slt}^T + UU^T\Sigma)$, since U is strictly lower triangular. Now, taking the matrix derivative of h with respect to U we have that the maximum is achieved when $Y_{slt} - 2\Sigma U = 0$, that is for

$$U = \frac{1}{2}\Sigma^{-1}Y_{slt}.$$

V. Profit maximization with downside risk constraint

In this section we consider the computation of upper and lower bounds for the solution of problem P_2 under affine recourse. Specifically, we shall consider the problem:

$$P_2: \quad J_{ar2}^* := \max_{\hat{u},U} \ \mu - \mathbf{E}_x\{\mathcal{C}\} \qquad (12)$$
$$\text{subject to:} \quad U \text{ is strictly lower triangular, } \hat{u}_0 = u_0$$
$$\sup_{x\sim(0,I)} \mathbf{Prob}\{x : \mathcal{P} < \mu\} \leq \epsilon \quad (13)$$
$$|\hat{u}_t| \leq u_{\max}, \ t = 1,\ldots,n,$$

where $u_{\max}$ is some given upper bound on the maximum nominal exposure in the investment, which is introduced for insuring boundedness of the solution set. The following key result, whose proof is reported in the Appendix, provides a key step for showing that the condition in (13) can be expressed explicitly via a convex semidefinite constraint.

Lemma 2: Let x have zero mean and unit covariance and $q(x) \doteq [x^T\,1]Q[x^T\,1]^T$ be a quadratic form, with $Q = Q^T \in \mathbf{R}^{(n+1)\times(n+1)}$. Define

$$\kappa(Q) := \sup_{x\sim(0,I)} \mathbf{Prob}\{x : q(x) < 0\}.$$

Then,

$$\kappa(Q) = \min_{\tau\geq 0} \mathbf{Tr}\,(J - \tau Q)_+,$$

where $J = ee^T$ with $e = [0\ldots 0\,1]^T \in \mathbf{R}^{n+1}$ and for a symmetric matrix A, A_+ denotes the matrix obtained by projecting A onto the positive semidefinite cone.[2]

[2] A_+ can be computed from A by simply replacing all negative eigenvalues, if any, by zeros.

Now, note that in the affine recourse case we have

$$\begin{aligned}\mathcal{P} &= z^TG(\hat{u},U)z \\ &\doteq z^T\begin{bmatrix} \frac{1}{2}(Y^TU + U^TY) & \frac{1}{2}(U^T\hat{y} + Y^T\hat{u}) \\ \frac{1}{2}(\hat{y}^TU + \hat{u}^TY) & \hat{y}^T\hat{u}\end{bmatrix}z,\end{aligned}$$

with $z^T \doteq [x^T\,1]$, whence $\mathcal{P} - \mu = z^T(G(\hat{u},U) - \mu J)z$. Then, from Lemma 2 it follows that

$$\begin{gathered}\sup_{x\sim(0,I)} \mathbf{Prob}\{x : \mathcal{P} < \mu\} \leq \epsilon \\ \Leftrightarrow \min_{\tau\geq 0} \mathbf{Tr}\,(J - \tau G(\hat{u},U) + \tau\mu J)_+ \leq \epsilon \\ \Leftrightarrow \min_{\tau\geq 0} \tau\mathbf{Tr}\left(\tau^{-1}J - G(\hat{u},U) + \mu J\right)_+ \leq \epsilon \\ [\text{letting } \lambda = \tau^{-1}] \\ \Leftrightarrow \min_{\lambda\geq 0} \lambda^{-1}\mathbf{Tr}\,((\lambda+\mu)J - G(\hat{u},U))_+ \leq \epsilon \\ \Leftrightarrow \exists\lambda > 0 : \lambda^{-1}\mathbf{Tr}\,((\lambda+\mu)J - G(\hat{u},U))_+ \leq \epsilon \\ \Leftrightarrow \exists\lambda > 0 : \mathbf{Tr}\,((\lambda+\mu)J - G(\hat{u},U))_+ \leq \lambda\epsilon.\end{gathered}$$

Further, notice that for any symmetric matrix A

$$\begin{aligned}\mathbf{Tr}A_+ \leq \eta \quad &\Leftrightarrow \quad \min_{M\succeq 0, M\succeq A} \mathbf{Tr}M \leq \eta \\ &\Leftrightarrow \quad \exists M \succeq 0, M \succeq A : \mathbf{Tr}M \leq \eta,\end{aligned}$$

therefore continuing the previous chain of equivalences, we have that

$$\begin{gathered}\mathbf{Tr}\,((\lambda+\mu)J - G(\hat{u},U))_+ \leq \lambda\epsilon \\ \Leftrightarrow \exists M \succeq 0, M \succeq (\lambda+\mu)J - G(\hat{u},U) : \mathbf{Tr}M \leq \lambda\epsilon,\end{gathered}$$

which is a convex semidefinite condition in $M, \lambda, \mu, \hat{u}, U$. Finally, considering the upper and lower bounds on the transaction penalty term developed in Section III, we can easily prove the result in the next proposition.

Proposition 2: The following inequality holds for the optimal value $\mathcal{J}_{ar2}^*$ of problem P_2 in (12) under affine recourse rule:

$$J_{lb2}^* \leq \mathcal{J}_{ar2}^* \leq J_{ub2}^*,$$

where J_{ub2}^* is the optimal value of the convex semidefinite programming problem (SDP):

$$J_{ub2}^*: \quad \max_{\hat{u},U,M\succeq 0,\mu,\lambda\geq 0,z\geq 0} \mu - \sum_{i=1}^{n+1} z_i \qquad (14)$$
$$\begin{aligned}&\text{subject to:} \\ &U \text{ strictly lower triangular, } \hat{u}_0 = u_0, \\ &-z_i \leq b_i(\hat{u}) \leq z_i, \quad i = 1,\ldots,n+1, \\ &|\hat{u}_t| \leq u_{\max}, \ t = 1,\ldots,n, \\ &\mathbf{Tr}\,M \leq \lambda\epsilon \\ &M \succeq (\lambda+\mu)J - G(\hat{u},U),\end{aligned}$$

and J_{lb2}^* is the optimal value of the SDP

$$J_{lb2}^*: \quad \max_{\hat{u},U,\mu,\lambda\geq 0,M\succeq 0,\mu_2} \mu - \mu_2 \qquad (15)$$
$$\begin{aligned}&\text{subject to:} \\ &U \text{ strictly lower triangular, } \hat{u}_0 = u_0, \\ &\sum_{t=1}^{n+1} \left\| \begin{bmatrix} a_t \\ b_t \end{bmatrix} \right\|_2 \leq \mu_2, \\ &|\hat{u}_t| \leq u_{\max}, \ t = 1,\ldots,n, \\ &\mathbf{Tr}\,M \leq \lambda\epsilon \\ &M \succeq (\lambda+\mu)J - G(\hat{u},U).\end{aligned}$$

$\triangleleft$

VI. EXAMPLE

We present an example comparing the results from the solution of the problem in (10) to the open-loop solution using daily Euro price in Dollars. Expected returns $\hat{y}$ and covariance factor Y have been estimated by fitting a 10-th order AR model to historical data. We stress the fact that this example has the sole purpose of testing the numerical functionality of the proposed techniques. In particular we are not claiming that an AR model is the fittest model to use for describing returns of liquid assets such as exchange rates. We solved the open-loop problem and the problem in (10) over a forward decision horizon consisting of 10 periods. The unit cost of transaction has been assumed to be $c = 0.1\%$. We simulated the three computed strategies on future unseen data for 20 periods . The results are compared in Figure 1, which shows the net profit (profit minus transaction cost) of two investing strategies following the open loop policy and the policy from (10).

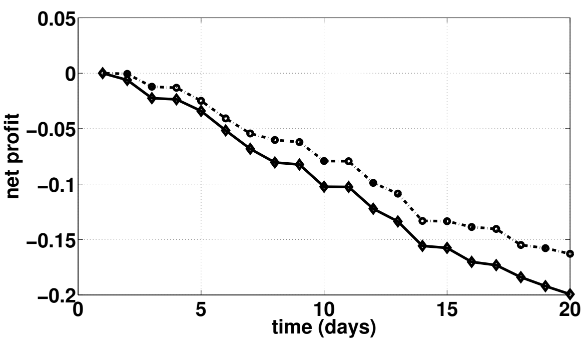

Fig. 1. Net profit computed using actual return vs. time for the open loop strategy (solid curve) and the policy from problem (10) (dashed-dot curve).

VII. CONCLUSIONS

In this paper we considered a dynamic investment strategy on a single risky asset, based on affine recourse policies. The proposed approach takes into account transaction costs and the possible statistical dependence of returns over different periods. This formulation admits relaxations that can be solved efficiently by means of second order cone or semidefinite optimization solvers. Some preliminary numerical experiments on real financial data confirm that the proposed technique can significantly outperform open loop strategies, while being computationally more attractive than a full stochastic programming formulation. Current research is being directed towards extensions of this model to the case of portfolios with many assets.

APPENDIX

Preliminaries. The following theorem from [16] will be used in the proof of Lemma 1.

Theorem 1: Let $B^0, B^1, \ldots, B^L \in \mathbf{R}^{n\times n}$ be symmetric and $B^1, \ldots, B^L$ be of rank 2. Let the problem P_ρ be defined as:

$$\mathrm{P}_\rho : \quad \text{Is } B^0 + \sum_{l=1}^{L} u_l B^l \succeq 0,\ \forall u :\ \|u\|_\infty \leq \rho \ ? \qquad (16)$$

and the problem P_{relax} be defined as:
P_{relax}: Do there exist symmetric matrices $X_1, \ldots, X_L \in \mathbf{R}^{n\times n}$ satisfying

$$\begin{array}{l} X_l \succeq \rho B^l,\ l = 1, \ldots, L, \\ \sum_{l=1}^{L} X_l \succeq B^0 ? \end{array}$$

Then, the following statements hold:

1) If P_{relax} is feasible, then P_ρ is feasible.
2) If P_{relax} is not feasible, then $\mathrm{P}_{\frac{\pi}{2}\rho}$ is not feasible. $\lhd$

Theorem 1 is adapted from a more general theorem (Theorem 2.1) in [16].

Proof of Lemma 1. Let $z \doteq [x^T\ 1]^T$. Then, ψ can be written as

$$\begin{array}{rl} \psi = & \sup\ \mathbf{E}\|Ax + b\|_1 \\ \text{subject to:} & \mathbf{E}\left(zz^T\right) = I_{m+1}. \end{array} \qquad (17)$$

The Lagrangian dual of the problem in (17) can be written (using a result from Section 4.1 of [17]) as

$$\begin{array}{rl} \psi^D = & \inf_{M\succeq 0} \mathbf{Tr}\, M \\ \text{subject to:} & z^T M z \geq \|Ax + b\|_1,\ \forall x \in \mathbf{R}^m. \end{array} \qquad (18)$$

Since strong duality holds for (17) and (18) (see Section 16.4 of [17]), it follows that $\psi = \psi^D$. In order to invoke Theorem 1, we write ψ in the following form

$$\begin{array}{rcl} \psi = & & \inf_{M\succeq 0} \mathbf{Tr} M \\ & \text{subject to:} & z^T M z \geq u^T(Ax + b),\ \forall x \in \mathbf{R}^m, \\ & & \forall u \in \mathbf{R}^n : \|u\|_\infty \leq 1 \\ \text{[eliminating } x\text{]} & = & \inf_{M\succeq 0} \mathbf{Tr} M \\ & \text{subject to:} & M \succeq \begin{bmatrix} 0 & \frac{1}{2}A^T u \\ \frac{1}{2}u^T A & u^T b \end{bmatrix} \\ & & \forall u \in \mathbf{R}^n : \|u\|_\infty \leq 1 \\ & = & \inf_{M\succeq 0} \mathbf{Tr} M \\ & \text{subject to:} & M \succeq \sum_{t=1}^{n} u_t C_t, \\ & & \forall u \in \mathbf{R}^n : \|u\|_\infty \leq 1, \end{array} \qquad (19)$$

where

$$C_t = \begin{bmatrix} 0 & a_t^T/2 \\ a_t/2 & b_t \end{bmatrix}. \qquad (20)$$

Note that the constraint in (19) is in the form of the constraint in (16). Next, consider the problem

$$\begin{array}{rl} \varphi := & \inf_{M\succeq 0, X_t = X_t^T} \mathbf{Tr} M \\ \text{subject to:} & -X_t + C_t \preceq 0,\ -X_t - C_t \preceq 0, \\ & \text{for } t = 1, \ldots, n, \\ & \sum_{i=1}^{n} X_i - M = 0,\ M \succeq 0. \end{array} \qquad (21)$$

The dual of the problem in (21) can be written as

$$\begin{aligned} \varphi^D := \quad & \sup_{\Lambda_t \succeq 0, \Gamma_t \succeq 0} \sum_{t=1}^{n} \mathbf{Tr}((\Lambda_t - \Gamma_t)C_t) \\ \text{subject to:} \quad & \Lambda_t + \Gamma_t = I_{m+1}, \\ & \Gamma_t \succeq 0,\ \Lambda_t \succeq 0,\ t = 1, \dots, n. \end{aligned} \tag{22}$$

Since the problem in (21) is convex and Slater conditions are satisfied, $\varphi = \varphi^D$. Next we show that, $\varphi^D = \overline{\psi}$. To this end, diagonalizing C_t, the problem in (22) can be reduced to n decoupled scalar problems, from which it follows that $\varphi^D = \sum_{t=1}^{n} \|\text{eig}(C_t)\|_1$, where $\text{eig}(\cdot)$ denotes the vector of the eigenvalues of its argument. For C_t as in (20), we have $\|\text{eig}(C_t)\|_1 = \| (a_t,\ b_t)^T \|_2$, since the characteristic equation of C_t is $s^{n-2}(s^2 - sb_t - \frac{1}{4}\sum_{t=1}^{n} a_t^2) = 0$ and $\|\text{eig}(C_t)\|_1 = \left(b_t^2 + \sum_{t=1}^{n} a_t^2\right)^{1/2} = \| (a_t,\ b_t)^T \|_2$. Then, by the first conclusion in Theorem 1 we have $\overline{\psi} = \varphi = \varphi^D$ and $\psi \leq \overline{\psi}$. For the lower bound on ψ in (9), assume that the problem in (22) is not feasible. Then, for $M \succeq 0$, we have that

$$\left\{M :\ \mathbf{Tr}M = \varphi^D\right\} \bigcap \{M :\ X_t \succeq \pm C_t,\ \textstyle\sum_{t=1}^{n} X_t \preceq M\} = \emptyset.$$

This last emptiness statement implies, by the second conclusion in Theorem 1, that

$$\left\{M :\ \mathbf{Tr}M = \varphi^D\right\} \bigcap \{M :\ M \succeq \textstyle\sum_{t=1}^{n} u_t C_t,\ \forall u :\ |u_t| \leq \pi/2\} = \emptyset$$

and

$$\left\{\tilde{M} :\ \mathbf{Tr}\tilde{M} = \tfrac{\varphi^D}{\pi/2}\right\} \bigcap \left\{\tilde{M} :\ \tilde{M} \succeq \textstyle\sum_{t=1}^{n} \tilde{u}_t C_t,\ \forall \tilde{u} :\ |\tilde{u}_t| \leq 1\right\} = \emptyset$$

Consequently, we have $\psi \geq \frac{\varphi^D}{\pi/2} = \frac{\overline{\psi}}{\pi/2}$. □

Proof of Lemma 2. Let $z \doteq [x^T\ 1]^T$. Then, $\kappa(Q)$ can be written as

$$\begin{aligned} \kappa(Q) = \quad & \sup_{p(\cdot)} \mathbf{Prob}\left\{x : q(x) < 0\right\} \\ \text{subject to:} \quad & \textstyle\int_{\mathbf{R}^n} zz^T p(x) dx = I_{n+1}, \end{aligned} \tag{23}$$

where maximization is over all probability distributions p with the first and second moments equal to 0 and I_n, respectively. Let $\mathbf{1}_q(\cdot) : \mathbf{R}^n \to \{0, 1\}$ be the indicator function of the set $\{x \in \mathbf{R}^n :\ q(x) < 0\}$, i.e., $\mathbf{1}_q(x) = 1$ if $q(x) < 0$ and $\mathbf{1}_q(x) = 0$, otherwise. Since strong duality, by Theorem 8 Section 16.4 in [17], holds for the problem in (23), dualizing with respect to the constraint in (23), we can equivalently write $\kappa(Q)$ as

$$\inf_{M} \sup_{p(\cdot)} \int \mathbf{1}_q(x)p(x)dx + \mathbf{Tr}\left(M(I - \int zz^T p(x)dx)\right)$$

$$= \inf_{M} \left\{\mathbf{Tr}M + \sup_{p(\cdot)} \int \left(\mathbf{1}_q(x) - z^T M z\right) p(x)dx\right\}$$

[$\mathbf{1}_q(x) - z^T M z \leq 0$ to get the supremum inside the curly brackets definite]

$$= \inf_{M} \mathbf{Tr}M \text{ subject to: } \mathbf{1}_q(x) \leq z^T M z\ \forall x \in \mathbf{R}^n$$

[using the definition of $\mathbf{1}_q$]

$$= \inf_{M \succeq 0} \mathbf{Tr}M \quad \text{subject to: } z^T M z \geq 1\ \forall x \in \{x \in \mathbf{R}^n :\ q(x) < 0\}$$

[using the S-procedure]

$$= \inf_{M \succeq 0, \tau \geq 0} \mathbf{Tr}M \text{ subject to: } M - J + \tau Q \succeq 0$$

$$= \inf_{\tau \geq 0} \mathbf{Tr}(J - \tau Q)_+,$$

where $M \in \mathbf{R}^{(n+1)\times(n+1)}$ and all integrals are over $\mathbf{R}^n$. □

REFERENCES

[1] D. Luenberger, *Investment Science*. Oxford University Press, 1997.
[2] J. Birge and F. Louveaux, *Introduction to Stochastic Programming*. Springer Verlag, 1997.
[3] N. Gulpinar, B. Rustem, and R. Settergren, "Multistage stochastic programming in computational finance," in *Decision Making, Economics and Finance: Optimization Models*. Kluwer Academic Publishers, 2002, pp. 33–47.
[4] A. Shapiro and A. Ruszczynski, *Lectures on Stochastic Programming*. Available on line, 2007. [Online]. Available: http://www2.isye.gatech.edu/ ashapiro/publications.html
[5] G. Dantzig and G. Infanger, "Multi-stage stochastic linear programs for portfolio optimization," *Annals of Operations Research*, vol. 45, pp. 59–76, 1993.
[6] J. M. Mulvey and B. Shetty, "Financial planning via multi-stage stochastic optimization," *Comput. Oper. Res.*, vol. 31, no. 1, pp. 1–20, 2004.
[7] N. Hibiki, "Multi-period stochastic optimization models for dynamic asset allocation," *Journal of Banking and Finance*, vol. 30, pp. 365–390, 2006.
[8] M. Steinbach, "Markowitz revisited: mean-variance models in financial portfolio analysis," vol. 43, no. 1, pp. 31–85, 2001.
[9] A. Shapiro, "On complexity of multistage stochastic programs," *Operation Research Letters*, vol. 34, pp. 1–8, 2006.
[10] G. Pflug, "Scenario tree generation for multiperiod financial optimization by optimal discretization," *Mathematical Programming*, vol. 89, pp. 251–271, 2001.
[11] G. Calafiore and M. Campi, "On two-stage portfolio allocation problems with affine recourse," in *Proc. Conference on Decision and Control*, Seville, Spain, 2005, pp. 8042–8047.
[12] P. Artzner, F. Delbaen, J. Eber, and D. Heath, "Coherent measures of risk," *Mathematical Finance*, vol. 9, no. 3, pp. 203–228, 1999.
[13] P. Jorion, *Value at Risk: The New Benchmark for Managing Financial Risk*. McGraw-Hill Trade, 2001.
[14] A. McNeil, F. Rüdiger, and P. Embrechts, *Quantitative Risk Management: Concepts Techniques and Tools*. Princeton University Press, 2005.
[15] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge University Press, Cambridge University Press.
[16] A. Ben-Tal and A. Nemirovski, "On tractable approximations of uncertain linear matrix inequalities affected by interval uncertainty," *SIAM Journal on Optimization*, vol. 12, no. 3, pp. 811–833, 2002.
[17] D. Bertsimas and J. Sethuraman, "Moment problems and semidefinite programming," in *Handbook of Semidefinite Programming*, H. Wolkovitz, R. Saigal, and L. Vandenberghe, Eds. Kluwer, 2000, pp. 469–509.