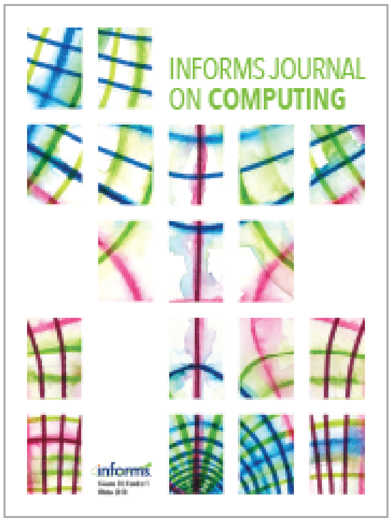

INFORMS Journal on Computing

Linear Programming and the Control of Diffusion Processes

Andrew Ahn, Martin Haugh

INFORMS Journal on Computing
Vol. 27, No. 4, Fall 2015, pp. 646–657
ISSN 1091-9856 (print) | ISSN 1526-5528 (online)

http://dx.doi.org/10.1287/ijoc.2015.0651

Linear Programming and the Control of Diffusion Processes

Andrew Ahn, Martin Haugh
Department of IE and OR, Columbia University, New York, New York 10027
{aja2133@columbia.edu, mh2078@columbia.edu}

Recent work by Han and Van Roy [Han J, Van Roy B (2011) Control of diffusions via linear programming. Infanger G, ed. *Stochastic Programming: The State of the Art, in Honor of George B. Dantzig* (Springer, New York), 329–354] introduced a linear programming technique to compute good suboptimal solutions to high-dimensional control problems in a diffusion-based setting. Their problem formulation worked with finite horizon problems where the horizon, T, is an exponentially distributed random variable. We extend their approach to finite horizon problems with a fixed horizon T. We also apply these techniques to dynamic portfolio optimization problems and then simulate the resulting policies to obtain lower bounds on the optimal value functions. We also use these policies in conjunction with convex duality methods designed for portfolio optimization problems to construct upper bounds on the optimal value functions. In our numerical experiments we find that the primal and dual bounds are very close, and so we conclude, for these problems at least, that the linear programming approach performs very well.

Keywords: linear programming; portfolio optimization; duality
History: Accepted by Winfried K. Grassmann, Area Editor for Computational Probability and Analysis; received June 2014; revised February 2015; accepted March 2015. Published online November 13, 2015.

1. Introduction

Because of the so-called curse of dimensionality, solving high-dimensional control problems is a notoriously difficult problem. It is not surprising, then, that suboptimal control has been an active area of research for many years. Moreover, the advent of ever-increasing computational power has seen many developments in the related area of *approximate dynamic programming* (ADP), particularly for discrete-time control problems. (See Bertsekas 2005 for a comprehensive introduction to classical suboptimal control techniques. Bertsekas 2012 also contains an excellent treatment of approximate dynamic programming.) Linear programming (LP) methods have played an important role in the development of several ADP techniques, beginning with Schweitzer and Seidmann (1985) and continuing with the important contributions of de Farias and Van Roy (2003), de Farias and Van Roy (2004), among others.

Recently, Han and Van Roy (2011) proposed an LP-based approach for the approximate solution of the Hamilton-Jacobi-Bellman (HJB) equation that arises from continuous-time control problems. Their approach applies to diffusion problems with an exponentially distributed horizon, T, and their numerical results were promising, with the LP-based policy outperforming other base-case policies. In this paper we extend their approach to continuous-time control problems with a fixed horizon, T. We apply these techniques to dynamic portfolio optimization problems and then simulate the resulting policies to obtain primal, i.e., lower, bounds on the optimal value functions. We also use these policies in conjunction with the convex duality methodology of Haugh et al. (2006a) (hereafter, HKW) to construct dual, i.e., upper, bounds on the optimal value functions. (See also Haugh and Jain 2011 and, more recently, Bick et al. 2012, who also use this dual approach.) By comparing the resulting primal and dual bounds, we can easily assess the quality of the suboptimal policy produced by the LP approach. In our numerical experiments we find that the primal and dual bounds are very close, and so we can conclude that, for these problems at least, the LP approach performs very well indeed.

The remainder of this paper is organized as follows. In §2 we formulate the continuous-time portfolio optimization problem and also discuss here the exponentially distributed and fixed horizon versions of the problem. In §3 we review the approach of Han and Van Roy (2011) for approximately solving the HJB equation when the horizon is an exponentially distributed random variable. We extend their methodology to the fixed horizon case in §4, and our numerical results are presented in §5. We conclude in §6. The online appendices (available as supplemental material at http://dx.doi.org/10.1287/ijoc.2015.0651) contain additional details, including an overview of the aforementioned dual approach of HKW.

2. The Portfolio Optimization Problem Formulation

In formulating the dynamic portfolio optimization problems that we will consider throughout the paper, we will follow the formulation of HKW. There are N risky stocks and an instantaneously risk-free bond in a market. We do note, however, that the LP-based approach to solving the HJB equation applies to control problems in diffusion settings more generally than the portfolio optimization problem that we consider here. The vector of stock prices is denoted by $P_t = (P_{1t}, \dots, P_{Nt})^\top$, and the instantaneously risk-free rate of return on the bond is denoted by r_t. Without loss of generality, we assume the stocks pay no dividends. Assets return dynamics depend on the M-dimensional vector of state variables, $Z_t = (Z_{1t}, \dots, Z_{Mt})^\top$, taking values in a state space S, so that

$$r_t = r(Z_t), \tag{1a}$$

$$\frac{dP_t}{P_t} = \mu_P(Z_t)\,dt + \Sigma_P(Z_t)\,dB_t, \tag{1b}$$

$$dZ_t = \mu_Z(Z_t)\,dt + \Sigma_Z(Z_t)\,dB_t, \tag{1c}$$

where $B_t = (B_{1t}, \dots, B_{Nt})^\top$ is an N-dimensional standard Brownian motion; $\mu_Z(Z_t)$ and $\mu_P(Z_t)$ are M- and N-dimensional drift vectors; and $\Sigma_Z(Z_t)$, $\Sigma_P(Z_t)$ are $M \times N$ and $N \times N$ diffusion matrices of the state variable and security prices, respectively. We assume that the diffusion matrix, $\Sigma_P(Z_t)$, of the asset return process is nondegenerate for each Z_t so that $x^\top \Sigma_P(Z_t)\Sigma_P(Z_t)^\top x \geq \epsilon \|x\|^2$ for all x and some $\epsilon > 0$. We can then define a process, η_t, according to

$$\eta_t(Z_t) := \Sigma_P(Z_t)^{-1}(\mu_P(Z_t) - r(Z_t)\cdot \mathbf{1}),$$

where $\mathbf{1} = (1, \dots, 1)^\top$. In a market without portfolio constraints, η_t corresponds to the market-price-of-risk process (see, e.g., Duffie 1996, §6.G). We make the standard assumption that the process η_t is square integrable so that

$$\mathbb{E}_0\left[\int_0^T \|\eta_t\|^2\,dt\right] < \infty.$$

(We use $\mathbb{E}_t[\cdot]$ to denote an expectation conditional on time t information throughout the paper.) Under this opportunity set, our portfolio consists of positions in the N stocks and the risk-free bond. We also assume that continuous rebalancing of the portfolio is permitted and that $\theta_t(Z_t) := (\theta_{1t}(Z_t), \dots, \theta_{Nt}(Z_t))^\top$ is the vector of risky security weights in the portfolio at time t. To rule out arbitrage, we require the portfolio strategy to satisfy a square integrability condition, namely, that $\int_0^T \|\theta\|^2\,dt < \infty$ almost surely. The value of the portfolio, W_t, associated with θ_t then changes according to the stochastic differential equation (SDE):

$$\frac{dW_t}{W_t} = [r_t + \theta_t^\top \lambda_t]\,dt + \theta_t^\top \Sigma_{P_t}\,dB_t, \tag{2}$$

where $\lambda_t := \mu_{P_t} - r_t \cdot \mathbf{1}$. (For ease of exposition, we will use r, μ_P, μ_Z, etc. (or r_t, μ_{P_t}, μ_{Z_t}, etc.), in place of $r(Z_t)$, $\mu_P(Z_t)$, $\mu_Z(Z_t)$, etc., throughout the paper.) We also assume that the portfolio is constrained so that

$$\theta_t(Z_t) \in \mathbf{K} \tag{3}$$

for all t and where $\mathbf{K}$ is some fixed convex set containing zero.

The portfolio optimization problem is to choose a self-financing trading strategy that maximizes the expected utility of terminal wealth. The horizon, T, is assumed to be finite, but it may be either random or deterministic, depending on the specific formulation under consideration. The utility function $u(W)$ is assumed to be strictly increasing, concave, and smooth. Moreover, it is assumed to satisfy the Inada conditions at zero and infinity so that $\lim_{W\to 0} u'(W) = \infty$ and $\lim_{W\to\infty} u'(W) = 0$. In this paper, we will use the constant relative risk aversion (CRRA) utility function so that

$$u(W) := \frac{W^{1-\gamma}}{1-\gamma}$$

with $\gamma > 1$. (We note that log utility is obtained in the limit as γ decreases to 1.)

2.1. When the Horizon T Is Fixed

When the problem has a fixed horizon, T, the investor's portfolio optimization problem at time t is to solve for

$$J^*(w, z, t) = \sup_{\{\theta_s:\, s\in[t,T]\}} \mathbb{E}_t[u(W_T)] \quad \text{subject to (1), (2), and (3),} \tag{4}$$

where w and z are the wealth and state vector values at time t. A well-known implication of CRRA utility is that J^* is separable in w and (z, t) so that we can write $J^*(w, z, t) = u(w)V^*(z, t)$. The optimal strategy is therefore independent of the wealth process, w_t.

To write the HJB equation for this problem, we first define the HJB operator:

$$\begin{aligned} H_\theta V(z,t) := {} & (1-\gamma)V(z,t)(\theta^\top\lambda + r) + V_z(z,t)^\top \mu_Z(z) \\ & + \tfrac{1}{2}\gamma(\gamma-1)V(z,t)\theta^\top \Sigma_P \Sigma_P^\top \theta \\ & + (1-\gamma)V_z(z,t)^\top \Sigma_Z \Sigma_P^\top \theta \\ & + \tfrac{1}{2}\mathrm{tr}[V_{zz}(z,t)\Sigma_Z\Sigma_Z^\top] + V_t(z,t). \end{aligned} \tag{5}$$

Note that V_z is the M-dimensional gradient of V with respect to the state variable z. Similarly, V_{zz} is the $M \times M$ Hessian matrix of V with respect to z. The HJB equation is then given by

$$0 = HV(z, t) := \inf_{\theta \in \mathbf{K}} H_\theta V(z, t), \tag{6}$$

and we note that $V^*(z, t) = J^*(w, z, t)/u(w)$ is a solution to this equation.

2.2. When the Horizon T Is Exponentially Distributed

We now assume the horizon T is an exponentially distributed random variable with mean τ. Moreover, T is assumed to be independent of all other random sources. In this case the investor's problem is identical to (4) but now with the understanding that the expectation must also be taken with respect to T. By first taking expectation with respect to T, it is easy to see the problem may also be formulated as

$$J^*(w, z) = \sup_{\{\theta_s\}} \mathbb{E}\left[\int_{t=0}^{\infty} e^{-t/\tau} u(W_t)\, dt\right]$$
$$\text{subject to (1), (2), and (3).} \tag{7}$$

Although t is no longer a state variable, J^* is still separable, so we can again write $J^*(w, z) = u(w)V^*(z)$. The HJB operator for this problem is then defined as

$$\begin{aligned} H_\theta V(z) := {} & (1-\gamma)V(z)(\theta^\top \lambda + r) + V_z(z)^\top \mu_Z(z) \\ & + \tfrac{1}{2}\gamma(\gamma - 1)V(z)\theta^\top \Sigma_P \Sigma_P^\top \theta \\ & + (1-\gamma)V_z(z)^\top \Sigma_Z \Sigma_P^\top \theta + \tfrac{1}{2}\mathrm{tr}\big[V_{zz}(z)\Sigma_Z\Sigma_Z^\top\big] \\ & - V(z)/\tau + 1, \end{aligned} \tag{8}$$

and the HJB equation is given by

$$0 = HV(z) = \inf_{\theta \in \mathbf{K}} H_\theta V(z). \tag{9}$$

We note that $V^*(z) = J^*(w, z)/u(w)$ is a solution to this equation.

3. Review of Han and Van Roy's LP Approach

In this section we review Han and Van Roy's (2011) LP approach for approximately solving (7) when the horizon T is exponentially distributed. In a standard argument they show that the optimal solution, V^*, to the HJB equation (9) is also the unique optimum of the following static optimization problem:

$$\begin{aligned} \max_{V(z)} \quad & \int V(z)\rho(dz) \\ \text{subject to} \quad & H_\theta V(z) \ge 0, \quad \forall\, \theta \in K,\ z \in S, \\ & V \in C^2, \end{aligned} \tag{$\mathcal{P}_1$}$$

where ρ is a prespecified positive measure for the integral. Although the objective and constraints in ($\mathcal{P}_1$) are linear, the problem is still very challenging to solve because there are uncountably many decision variables and constraints; indeed, there is one constraint for every (θ, z) pair. We therefore solve an approximation to ($\mathcal{P}_1$), and this approximation is obtained via the following steps:

1. We first choose a suitable set of basis functions $\{\phi_1(z), \ldots, \phi_k(z)\}$ with the goal of finding a linear combination, $\sum_{j=1}^k r_j \phi_j(z)$, that we will use to approximate $V^*(z)$. The original problem then reduces to the problem of solving for k decision variables, $r_1, r_2, \ldots, r_k$. The algorithm is initialized with a predetermined weight vector, $r^{(0)} = (r_1^{(0)}, \ldots, r_k^{(0)})$.
2. We generate a finite sample set $z_1, \ldots, z_Q$ and approximate the integral in ($\mathcal{P}_1$) by a corresponding finite sum of Q terms. Although any positive measure, ρ, can be used in theory, the performance of the algorithm depends on how the samples are generated. Han and Van Roy define

$$\rho(dz) := \frac{1}{\tau}\mathbb{E}\left[\int_{t=0}^{\infty} e^{-t/\tau} \mathbf{1}_{\{Z_t \in [z, z+dz]\}}\, dt\right]$$

and then generate $z_1, \ldots, z_Q$ by simulating (approximately) from this measure. In particular, they first simulate the horizon $T \sim \mathrm{Exp}(1/\tau)$ and then simulate a discrete-time approximation to the dynamics of the state variables (1c). The value of the state vector at the simulated time T is taken as one of our Q samples.
3. For each sample, z_j, we choose a single corresponding θ_j as follows: given an approximation, $\Phi r := \phi_1 r_1 + \cdots + \phi_k r_k$, to V^*, we myopically choose a greedy action with respect to Φr. That is, we select $\theta_j \in \arg\min_{\theta \in \mathbf{K}} H_\theta[(\Phi r)(z_j)]$ for each j.
4. Given a weight vector r, we find a new weight vector r' by solving an approximation of ($\mathcal{P}_1$) (see phase 2 of the algorithm below); this approximation is a linear program that we obtain from steps 1–3.

Steps 3 and 4 are repeated to obtain a sequence of weight vectors, $\{r^{(0)}, r^{(1)}, r^{(2)}, \ldots\}$. We are now ready to define the *adaptive constraint selection algorithm* of Han and Van Roy.

Adaptive Constraint Selection Algorithm I

for $i = 1$ *to* ∞ *do*

for $j = 1$ *to* Q *do*

$$\theta_j \in \arg\min_{\theta \in \mathbf{K}} H_\theta[(\Phi r^{(i-1)})(z_j)] \qquad \text{(phase 1)}$$

end for

$$r^{(i)} \in \arg\max_{r \in \mathbb{R}^k} \sum_{j=1}^{Q} (\Phi r)(z_j) \qquad \text{(phase 2)}$$

$$\text{subject to } H_{\theta_j}[(\Phi r)(z_j)] \ge 0, \quad \forall\, j = 1, \ldots, Q$$

end for

This algorithm does not necessarily generate an optimal solution to ($\mathcal{P}_1$), and there is no theoretical guarantee that the sequence $r^{(i)}$ will converge. (Han and Van Roy 2011 did not specify how they handled nonconvergence, but we suspect they simply used a different starting point, $r^{(0)}$, in that event.) However, if it does converge, then Han and Van Roy show it must converge to an optimal solution of the following problem, which is an approximation of ($\mathcal{P}_1$):

$$\max_r \sum_{j=1}^{Q} (\Phi r)(z_j) \quad \text{subject to} \quad H[(\Phi r)(z_j)] \geq 0, \quad \forall j = 1, \ldots, Q. \tag{10}$$

We now discuss in further detail the steps required to execute phases 1 and 2.

Phase 1

We can expand the objective of phase 1 using (8). If we then remove terms that do not depend on θ and eliminate the common factor, $\gamma - 1$, then the problem of phase 1 can be expressed as

$$\theta_j \in \arg\min_{\theta \in \mathbf{K}} \Big\{ \tfrac{1}{2}\theta^\top [\gamma(\Phi r^{(i-1)})(z_j)\Sigma_P \Sigma_P^\top]\theta - [(\Phi r^{(i-1)})(z_j)\lambda(z_j) + \Sigma_P \Sigma_Z^\top (\Phi r^{(i-1)})_z(z_j)]^\top \theta \Big\}. \tag{11}$$

If $(\Phi r^{(i-1)})(z_j) > 0$, then (11) is a convex quadratic program and therefore easy to solve. Otherwise, the objective in (11) may be unbounded if $\mathbf{K}$ is not compact, and some other heuristic approach for computing θ_j would be required.

Phase 2

If we expand the constraints of the LP in phase 2 using (8), then we obtain the following LP:

$$\max_{r \in \mathbb{R}^k} c^\top r \quad \text{subject to} \quad Ar \geq -\mathbf{1}, \tag{12}$$

where

$$\begin{aligned} A_{ij} := {} & (1-\gamma)\phi_j(z_i)[\theta_i^\top \lambda(z_i) + r(z_i)] + [(\phi_j)_z(z_i)]^\top \mu_Z(z_i) \\ & + \tfrac{1}{2}\gamma(\gamma-1)\phi_j(z_i)\theta_i^\top \Sigma_P \Sigma_P^\top \theta_i \\ & + (1-\gamma)[(\phi_j)_z(z_i)]^\top \Sigma_Z \Sigma_P^\top \theta_i \\ & + \tfrac{1}{2}\mathrm{tr}[(\phi_j)_{zz}(z_i)\Sigma_Z \Sigma_Z^\top] - \phi_j(z_i)/\tau, \end{aligned}$$

and $c_i := \phi_i(z_1) + \phi_i(z_2) + \cdots + \phi_i(z_Q)$. This linear program has a k-dimensional decision vector and Q linear constraints.

4. Extending the LP Approach to the Case of a Fixed Horizon T

The LP approach of the previous section applies to problems with an exponentially distributed horizon T, but we would also like to apply it to the case of a fixed horizon. As we shall see, this extension is not immediate and requires some work because time t is also a state variable in this case. We show in §A of the online supplement that under some technical conditions, the solution to the HJB equation (6), V^*, is also the unique solution to the following optimization problem:

$$\begin{aligned} \max_V \quad & \int V(z,t)\,\rho(dz,dt) \\ \text{subject to} \quad & H_\theta V(z,t) \geq 0, \quad \forall \theta \in K,\ z \in S,\ t \in [0,T], \\ & V(z,T) \leq 1, \quad \forall z \text{ (boundary condition)}, \\ & V \in C^2, \end{aligned} \tag{$\mathcal{P}_2$}$$

where ρ is again some prespecified measure. In contrast to problem ($\mathcal{P}_1$), the boundary condition $V(z,T) \leq 1$ is required in this case.

The extension of the adaptive constraint selection algorithm seems straightforward: we choose basis functions $\{\phi_1(z,t), \ldots, \phi_k(z,t)\}$ to approximate $V^*(z,t)$ and generate $\{(z_1,t_1),(z_2,t_2),\ldots,(z_Q,t_Q)\}$ as a representative sample of (z,t). As a simple heuristic for generating this sample, we first generate the t_j's as IID $\sim U[0,T]$, and then, for each t_j, we set $z_j := Z_{t_j}$, where Z_{t_j} is obtained by simulating a discrete-time approximation to the market state dynamics (1c) and then terminating at time t_j. The adaptive constraint selection algorithm in this case is as follows.

Adaptive Constraint Selection Algorithm II (For a fixed horizon T)

for $i = 1$ *to* ∞ *do*

for $j = 1$ *to* Q *do*

$$\theta_j \in \arg\min_{\theta \in \mathbf{K}} H_\theta[(\Phi r^{(i-1)})(z_j, t_j)] \qquad \text{(phase 1)}$$

end for

$$r^{(i)} \in \arg\max_{r \in \mathbb{R}^k} \sum_{j=1}^{Q} (\Phi r)(z_j, t_j) \qquad \text{(phase 2)}$$

$$\text{subject to } H_{\theta_j}[(\Phi r)(z_j,t_j)] \geq 0, \quad \forall j = 1, \ldots, Q, \qquad (\Phi r)(z,T) \leq 1, \quad \forall z$$

end for

We note the boundary condition $(\Phi r)(z,T) \leq 1$ in phase 2 applies to all possible z rather than just the sampled z_j's. We could, of course, impose this boundary constraint on just a finite subset of z values, but we will see later that it is straightforward to impose the general constraint through an appropriate choice of basis functions. For example, we could ensure that each nonconstant basis function has a common factor $T-t$. As a result, the only contribution to the left-hand side of the constraint $(\Phi r)(z,T)\le 1$ will come from a constant basis function, say, $\phi_1\equiv 1$. We can then impose the boundary condition by adding the linear constraint $r_1\le 1$ to the LP phase 2.

We now discuss the objectives and constraints in phases 1 and 2 in further detail and, in particular, why this algorithm is problematic to implement in its current form.

Phase 1

If we substitute (5) into the objective function of phase 1, drop all terms that do not depend on θ, and eliminate the common factor, $\gamma-1$, then the problem of phase 1 may be written as

$$\theta_j\in\arg\min_{\theta\in\mathbf{K}}\Big\{\tfrac{1}{2}\theta^\top[\gamma(\Phi r^{(i-1)})(z_j,t_j)\Sigma_P\Sigma_P^\top]\theta - [(\Phi r^{(i-1)})(z_j,t_j)\lambda(z_j) + \Sigma_P\Sigma_Z^\top(\Phi r^{(i-1)})_z(z_j,t_j)]^\top\theta\Big\}. \quad (13)$$

If $(\Phi r^{(i-1)})(z_j,t_j)>0$, then the phase 1 problem is a convex quadratic program and therefore easy to solve. Otherwise, depending on $\mathbf{K}$, (13) may be unbounded. In this case we simply take θ_j to be the myopic portfolio described in §B of the online supplement. We note here, however, that in the numerical experiments of §5.1, we rarely encountered negative values of $(\Phi r^{(i-1)})(z_j,t_j)$. (We considered problems with various combinations of risk-aversion parameter, γ, and horizon T in those numerical experiments. In the worst case among all such problems, we observed negative values of $(\Phi r^{(i-1)})(z_j,t_j)$ only 1.46% of the time. In most of these problems we never encountered negative values of $(\Phi r^{(i-1)})(z_j,t_j)$.)

Phase 2

If we substitute (5) into the constraints of phase 2, then the problem of phase 2 can be formulated as

$$\begin{aligned}\max_{r\in\mathbb{R}^k}\ & c^\top r\\ \text{subject to}\ & Ar\ge\mathbf{0}, \\ & (\Phi r)(z,T)\le 1,\quad \forall z,\end{aligned}\quad (14)$$

where

$$\begin{aligned}A_{ij} &= (1-\gamma)\phi_j(z_i,t_i)[\theta_i^\top\lambda(z_i)+r(z_i)] + (\phi_j)_z(z_i,t_i)^\top\\ &\quad\cdot\mu_Z(z_i)+\tfrac{1}{2}\gamma(\gamma-1)\phi_j(z_i,t_i)\theta_i^\top\Sigma_P\Sigma_P^\top\theta_i\\ &\quad+(1-\gamma)[(\phi_j)_z(z_i,t_i)]^\top\Sigma_Z\Sigma_P^\top\theta_i\\ &\quad+\tfrac{1}{2}\mathrm{tr}[(\phi_j)_{zz}(z_i,t_i)\Sigma_Z\Sigma_Z^\top]+(\phi_j)_t(z_i,t_i),\end{aligned}$$

and $c_i=\phi_i(z_1,t_1)+\phi_i(z_2,t_2)+\cdots+\phi_i(z_Q,t_Q)$. Ignoring the boundary conditions, $(\Phi r)(\cdot,T)\le 1$ (which we can handle through the choice of basis functions as previously discussed), we see that phase 2 is an LP with a k-dimensional decision vector and Q linear constraints.

It turns out that phase 2 here is very problematic. In particular, the constraint $Ar\ge\mathbf{0}$ of (14) is much more difficult to handle than the constraint $Ar\ge-\mathbf{1}$ of (12) that occurs in the exponentially distributed horizon case. This latter set of constraints is always satisfied by all points in some ball around the zero vector $\mathbf{0}$. This is not true in the fixed horizon case where the corresponding constraints are $Ar\ge\mathbf{0}$. Since r is k-dimensional, each of the Q constraints in $Ar\ge\mathbf{0}$ defines a k-dimensional closed half-space containing $\mathbf{0}$ on its boundary. Any feasible point must therefore lie in the intersection of these Q half-spaces. Moreover, since Q is typically much larger than k, the intersection is generally just a single point—namely, the origin $\{\mathbf{0}\}$. This makes the problem (14) trivial to solve, but the solution is hardly desirable.

One ad hoc approach for resolving this issue would be to relax the constraint $Ar\ge\mathbf{0}$ to $Ar\ge-\epsilon\cdot\mathbf{1}$, where ϵ is some small positive number. For example, in our initial numerical experiment of §5.1, the value $\epsilon=1$ appeared to yield the best results among several different values of ϵ. We did not, however, have a sensible rule for choosing an appropriate value of ϵ in advance. Moreover, in §4.1, we propose an alternative problem formulation that yields superior results.

4.1. An Alternative Formulation

We propose here a new problem formulation based on the certainty equivalent return, r_{ce}, which is defined as the certain annualized rate of return that makes the investor indifferent between accepting it and following his optimal trading strategy. It is therefore given implicitly via

$$u(w)V^*(z,t)=u(we^{r_{ce}(z,t)(T-t)}),$$

which implies

$$\ln(V^*(z,t))=r_{ce}(z,t)(T-t)(1-\gamma). \quad (15)$$

Because $r_{ce}(z,t)$ is generally "less nonlinear" than the value function (especially when γ is large), it makes some sense to approximate the log-value function rather than the value function itself. We assume basis functions of the form

$$\{\phi_1(z,t),\ldots,\phi_k(z,t)\} = \{(T-t)\tilde{\phi}_1(z,t),\ldots,(T-t)\tilde{\phi}_k(z,t)\} \quad (16)$$

and will use a linear combination of them to approximate $\ln(V^*(z,t))$. We also note that any such linear combination of these functions will automatically satisfy the boundary condition $e^{(\Phi r)(z,T)} \leq 1$ so that this constraint does not need to be explicitly imposed in our new adaptive constraint selection algorithm.

Adaptive Constraint Selection Algorithm III (For a fixed horizon T)
for $i=1$ *to* ∞ *do*
  for $j=1$ *to* Q *do*

$$\theta_j \in \arg\min_{\theta\in\mathbf{K}} H_\theta[e^{(\Phi r^{(i-1)})(z_j,t_j)}] \qquad \text{(phase 1)}$$

  end for

$$r^{(i)} \in \arg\max_{r\in\mathbb{R}^k} \sum_{j=1}^{Q} e^{(\Phi r)(z_j,t_j)} \qquad \text{(phase 2)}$$

$$\text{subject to } H_{\theta_j}[e^{(\Phi r)(z_j,t_j)}] \geq 0, \quad \forall\, j=1,\ldots,Q$$

end for

We provide further details on the steps required for phases 1 and 2 below.

Phase 1
If we substitute (5) into the objective function of phase 1, drop terms that do not depend on θ, and then eliminate the common factor, $e^{\Phi r^{(i-1)}}(\gamma-1)$, then the problem of phase 1 may be reduced to

$$\theta_j \in \arg\min_{\theta\in\mathbf{K}} \Big\{ \tfrac{1}{2}\theta^\top[\gamma\Sigma_P\Sigma_P^\top]\theta - [\lambda(z_j) + \Sigma_P\Sigma_Z^\top(\Phi r^{(i-1)})_z(z_j,t_j)]^\top\theta \Big\}, \quad (17)$$

where the subscript z in (17) denotes a gradient vector. In contrast to the previous algorithm, phase 1 is always a convex quadratic program and therefore easy to solve.

Phase 2
Similarly, if we substitute (5) into the constraints of phase 2, eliminate the common factor $e^{\Phi r}$, and rearrange, we obtain

$$\begin{aligned}(\Phi r)_z(z,t)^\top\mu_Z(z) &+ (1-\gamma)(\Phi r)_z(z,t)^\top\Sigma_Z\Sigma_P^\top\theta \\ &+ \tfrac{1}{2}\mathrm{tr}[((\Phi r)_{zz}(z,t) + (\Phi r)_z(z,t)(\Phi r)_z(z,t)^\top)\Sigma_Z\Sigma_Z^\top] \\ &+ (\Phi r)_t(z,t) \\ &\geq -(1-\gamma)(\theta^\top\lambda + r) - \tfrac{1}{2}\gamma(\gamma-1)\theta^\top\Sigma_P\Sigma_P^\top\theta \end{aligned} \quad (18)$$

for $(z,t) = (z_1,t_1),\ldots,(z_Q,t_Q)$. Note that if we used the alternative formulation based on approximating the log-value function for the case of the exponentially distributed horizon, we would not be able to eliminate the factor $e^{\Phi r}$ because of the constant term "1" that appears in the HJB operator in (8). Note also that the phase 2 objective function contains the exponential term $e^{(\Phi r)(z_j,t_j)}$ and the constraint (18) contains the term $(\Phi r)_z(z,t)(\Phi r)_z(z,t)^\top$, which is quadratic in r. Phase 2 is therefore not an LP.

We resolve this problem by (i) linearizing the objective function using a first-order Taylor series expansion of e^x around zero and (ii) simply dropping the terms in (18) that are quadratic in r. (We could, of course, have also used a Taylor expansion to linearize the constraints, but in our numerical experiments we obtained very good results by simply dropping the quadratic terms.) This yields the following LP for phase 2:

$$\begin{aligned}\max_{r\in\mathbb{R}^k}\ & c^\top r \\ \text{subject to }\ & Ar \geq -d,\end{aligned} \quad (19)$$

where

$$\begin{aligned}A_{ij} &:= (\phi_j)_z(z_i,t_i)^\top\mu_Z(z_i) + (1-\gamma)[(\phi_j)_z(z_i,t_i)]^\top\Sigma_Z\Sigma_P^\top\theta_i \\ &\quad + \tfrac{1}{2}\mathrm{tr}[(\phi_j)_{zz}(z_i,t_i)\Sigma_Z\Sigma_Z^\top] + (\phi_j)_t(z_i,t_i), \\ d_i &:= (1-\gamma)[\theta_i^\top\lambda(z_i) + r(z_i)] + \tfrac{1}{2}\gamma(\gamma-1)\theta_i^\top\Sigma_P\Sigma_P^\top\theta_i, \\ c_i &:= \phi_i(z_1,t_1) + \phi_i(z_2,t_2) + \cdots + \phi_i(z_Q,t_Q).\end{aligned}$$

In our numerical experiments with this algorithm, we will choose $\phi_1(z,t) = (T-t)/T$ as one of our basis functions. It is then easy to see that $A_{1j} = (T-t_j)/T > 0$ for all j, so regardless of d, we can ensure $Ar \geq -d$ holds by taking r_1 sufficiently large. We therefore do not need to relax the constraints, $Ar \geq -d$, as we needed to do with (14) in our original problem formulation for the problem with a fixed horizon T.

5. Numerical Experiments

We now illustrate the performance of the LP-based algorithms of §4. We consider several portfolio optimization problems and assume that in each of them the horizon T is fixed. We consider three different trading strategies: the strategies that are greedy with respect to the approximate value functions that are obtained from the adaptive constraint selection Algorithms II and III, as well as the well-known myopic strategy we will use as a benchmark. The myopic strategy is known to perform well under the various numerical experiments in HKW and Haugh and Jain (2011). (Indeed, when $\gamma = 1$—corresponding to log utility—the myopic strategy is known to be optimal.) Further details on the myopic strategy can be found in §B of the online supplement.

The LP-based strategies at any state (t, Z_t) are found by solving

$$\theta_t^{LP1} = \arg\min_{\theta\in\mathbf{K}} \Big\{ \tfrac{1}{2}\theta^\top[\gamma(\Phi r^*)(Z_t,t)\Sigma_P\Sigma_P^\top]\theta - [(\Phi r^*)(Z_t,t)\lambda + \Sigma_P\Sigma_Z^\top(\Phi r^*)_z(Z_t,t)]^\top\theta \Big\}, \quad (20)$$

$$\theta_t^{LP2} = \arg\min_{\theta \in \mathbf{K}} \Big\{ \tfrac{1}{2}\theta^\top [\gamma \Sigma_P \Sigma_P^\top]\theta - [\lambda + \Sigma_P \Sigma_Z^\top (\Phi r^*)_z (Z_t, t)]^\top \theta \Big\}. \quad (21)$$

We note that θ_t^{LP1} and θ_t^{LP2} are obtained from (13) and (17), i.e., phase 1 of Algorithms II and III, respectively, by replacing Φr with Φr^*, where r^* is the solution we obtain from implementing these algorithms. Similarly the myopic strategy in state (t, Z_t) is found by solving the convex quadratic program

$$\theta_t^m = \arg\min_{\theta \in \mathbf{K}} \Big\{ \tfrac{1}{2}\theta^\top [\gamma \Sigma_P \Sigma_P^\top]\theta - \lambda^\top \theta \Big\}, \quad (22)$$

where λ is the time t vector of excess returns.

We used a standard Euler scheme to generate sample paths of the security prices and state vector Z_t. (We also used stratified sampling as a variance reduction technique. In particular, we stratified upon the terminal value of the vector Brownian motion driving the price and state dynamics and then used the Brownian bridge construction to simulate the Euler scheme. See Glasserman 2004 for a discussion of Euler schemes as well as stratified sampling and the Brownian bridge construction.) At each time step all three strategies are found by solving (20), (21), and (22), respectively. By simulating many paths and averaging the utility of terminal wealth across all paths for each strategy, we can obtain estimates of the value functions associated with each of the strategies. In our numerical results, we will report these value functions as certainly equivalent (CE) annualized returns. Since these strategies are all feasible, their CE returns are therefore lower bounds on the CE return for the (in general) unknown optimal strategy. Finally we can use the dual approach of HKW (see §C of the online supplement for a review) to construct upper bounds on the optimal value. These upper bounds are also reported as CE returns.

All of the computations in the following three examples were performed using `Matlab` running on a laptop with 4GB RAM and a 2.53 GHz processor. We note that in all of these examples it took just a couple of seconds to execute phases 1 and 2 of Algorithm III—the algorithm we ultimately favor—for each (T, γ) pair and each set of trading constraints. When it comes to actually simulating/evaluating the policies, the computation times were as follows. For each (T, γ) pair and set of trading constraints in Example 1, it took approximately 5–10 minutes for precalculation of the greedy policy (with respect to the approximate value function obtained in phases 1 and 2) on a predefined grid of sample points (t, z). Note that we only did this precalculation in the case of Example 1 because the state vector is one-dimensional in that case and precalculation on a one-dimensional grid was feasible. It then took approximately 1.5 hours to simulate the one million sample paths. In the case of Examples 2 and 3, it took between one and four hours to simulate the one million sample paths for each (T, γ) pair with the specific time depending on which set of trading constraints were imposed. But we note that these sample paths are also used for the myopic policy. Moreover, if one ever wanted to use these strategies in practice, then this large simulation step would not be required. (To actually implement the computed policy in practice, we would only need to compute the strategy along the single realized path rather than one million simulated paths.)

5.1. Example 1: Three Risky Assets and a Single State Variable

Our first numerical example is from HKW, who in turn based their model on the discrete-time market model in Lynch (2001). They consider a financial market with three risky assets and a single state variable associated with a four-dimensional Brownian motion. In our framework of §2 we assumed (without loss of generality) that the volatility matrix, Σ_P, is invertible. We can enforce this here by simply assuming that the state variable is in fact a fourth risky security that we are not allowed to trade. We assume the drift term of risky assets returns is affine in the state variable, which itself follows an Ornstein–Uhlenbeck process with a long-term mean of zero. The asset return dynamics therefore satisfy

$$\begin{aligned} r_t &\equiv r, \\ \frac{dP_t}{P_t} &= (\mu_0 + Z_t \mu_1)\, dt + \Sigma_P\, dB_t, \\ dZ_t &= -kZ_t\, dt + \Sigma_Z\, dB_t, \end{aligned}$$

where $r = 0.01$, $k = 0.366$, and

$$\mu_0 = \begin{bmatrix} 0.142 \\ 0.109 \\ 0.089 \end{bmatrix}, \quad \mu_1 = \begin{bmatrix} 0.065 \\ 0.049 \\ 0.049 \end{bmatrix},$$

$$\Sigma_P = \begin{bmatrix} 0.256 & 0 & 0 & 0 \\ 0.217 & 0.054 & 0 & 0 \\ 0.207 & 0.062 & 0.062 & 0 \\ -0.741 & 0.04 & 0.034 & 0.288 \end{bmatrix},$$

$$\Sigma_Z = \begin{bmatrix} -0.741 \\ 0.04 \\ 0.034 \\ 0.288 \end{bmatrix}^\top.$$

Note that r, Σ_P, and Σ_Z are constant in this model. When performing simulations, we set the initial value, $Z_0 = 0$, of the state variable. We use a discretization time step of $dt = 1/100$ in our simulations as well as in the simulations of the later models of §§5.2

and 5.3. The horizon T is fixed at either 5 or 10 years, and the parameter γ of the CRRA utility function can be either 1.5, 3, or 5. We also consider two sets of trading constraints:

(i) The unconstrained case where the agent does not face any trading constraints (except, of course, for the fourth asset, which is really the state variable and therefore not tradable). We refer to this as the "incomplete markets" case.

(ii) There are no short sales on all the risk securities as well as a no-borrowing constraint. We refer to this as the "incomplete markets + no short sales and no borrowing" case.

In each of our numerical experiments (here and elsewhere in the paper), we use $Q = 10{,}000$ sample points in our two LP-based algorithms. When we use the adaptive constraint selection Algorithm II, we use

$$\mathbb{B} = \{1\} \cup \left\{ P_i(z) \cdot \left(\frac{T-t}{T}\right)^j \,\middle|\, 0 \le i \le 5,\ 1 \le j \le 5 \right\}$$

as our basis functions where $P_i(\cdot)$ is the Chebyshev polynomial of the first kind of degree i. These polynomials up to degree $i = 5$ are

$$\begin{aligned}
P_0(z) &= 1,\\
P_1(z) &= z,\\
P_2(z) &= 2z^2 - 1,\\
P_3(z) &= 4z^3 - 3z,\\
P_4(z) &= 8z^4 - 8z^2 + 1,\\
P_5(z) &= 16z^5 - 20z^3 + 5z.
\end{aligned}$$

Note that except for the first one, all of our basis functions contain a factor of $T - t$. As stated earlier, this allows us to easily impose the constraint $(\Phi r)(z, T) \le 1$. In phase 2 of Algorithm II, we set $\epsilon = 1$ to relax the constraints in the linear program (14). When using Algorithm III associated with our alternative formulation, we use

$$\mathbb{B} = \left\{ P_i(z) \cdot \left(\frac{T-t}{T}\right)^j \,\middle|\, 0 \le i \le 5,\ 1 \le j \le 5 \right\}$$

as our set of basis functions.

It is perhaps worth mentioning at this point that ADP methods are often extremely sensitive to the choice of basis functions and the number of simulated sample paths Q. In our experiments, we found that the use of Chebyshev polynomials resulted in a very stable performance in that convergence of the $r^{(i)}$ sequence was rarely an issue. This was not necessarily the case when we experimented with other sets of basis functions. This, however, is a criticism of ADP methods in general rather than our (and Han and Van Roy's 2011) approach in particular. Moreover, this weakness of ADP methods can be partly addressed through the use of the duality approach to construct dual bounds on the optimal value function. In particular, if we find that the computed duality gap, i.e., the difference between the lower and upper bounds, is too wide, then this suggests that the ADP algorithm is failing to find a sufficiently good solution. We could then seek to improve it possibly by changing the set of basis functions or increasing the number of basis functions that we use, etc.

Tables 1 and 2 present the results of Algorithms II and III, respectively. We observe the two trading strategies driven by the LP approach perform better than the myopic strategy even when γ is close to 1. In the incomplete markets case, it is actually possible to compute the optimal solution by solving a system of ordinary differential equations. This optimal solution is reported in the row labeled "V^u." If we compare the performances of the LP strategies to the optimal strategy, we see they are generally very close to each other, although their performances do deteriorate somewhat with T and γ.

In comparing Tables 1 and 2 more closely, we also note that Algorithm III is clearly superior to Algorithm II and that this is especially noticeable when $(T, \gamma) = (10, 5)$. This seems to suggest that the error due to the linearization in phase 2 of Algorithm III is quite small. We also noticed similar behavior in our other numerical experiments, and for this reason, we will only report results from Algorithm III henceforth.

5.2. Example 2: A Zero-Premium Long-Term Bond and Three State Variables

The second model we consider is taken from Haugh et al. (2006b), who in turn based their model and parameters on Wachter and Sangvinatsos (2005). In this model there is only one risky asset, which is a long-term bond maturing at time T. The bond has no risk premium and there is a three-dimensional state variable and three-dimensional Brownian motion. In contrast to the previous model (and the duality development in §C of the online supplement), we do not explicitly define artificial assets so that the number of risky assets equals the dimension of the Brownian motion (in which case Σ_P will be invertible). Instead, we directly set the risk premium of the risky bond as well as the market price of risk process, η_t, to be zero. With these choices, it is clear that $\Sigma_P \eta = \lambda$ will be satisfied. Therefore, if necessary we could explicitly define artificial asset price dynamics so that our choice of η, i.e., zero in this example,

Table 1 **Results of Model 1 Using Algorithm II Strategy**

	$T=5$			$T=10$		
Function	$\gamma=1.5$	$\gamma=3$	$\gamma=5$	$\gamma=1.5$	$\gamma=3$	$\gamma=5$
			Incomplete markets			
LB^{LP}	16.79	10.25	6.98	17.73	11.32	7.80
	(16.77, 16.80)	(10.23, 10.28)	(6.96, 7.00)	(17.71, 17.74)	(11.30, 11.35)	(7.78, 7.83)
UB^{LP}	16.81	10.37	7.25	17.77	11.53	8.29
	(16.78, 16.84)	(10.26, 10.47)	(7.10, 7.40)	(17.74, 17.79)	(11.38, 11.69)	(7.93, 8.66)
LB^{m}	16.64	9.87	6.61	17.45	10.58	7.09
	(16.62, 16.66)	(9.86, 9.89)	(6.59, 6.62)	(17.44, 17.47)	(10.56, 10.59)	(7.08, 7.10)
UB^{m}	16.83	10.43	7.31	17.81	11.65	8.37
	(16.81, 16.86)	(10.32, 10.54)	(7.16, 7.47)	(17.79, 17.83)	(11.48, 11.82)	(7.94, 8.81)
V^{u}	16.79	10.32	7.06	17.76	11.55	8.12
			Incomplete markets + No short sales and no borrowing			
LB^{LP}	10.16	7.82	5.63	10.38	8.48	6.36
	(10.16, 10.17)	(7.81, 7.83)	(5.61, 5.65)	(10.37, 10.38)	(8.47, 8.49)	(6.35, 6.38)
UB^{LP}	10.17	7.87	5.78	10.39	8.59	6.63
	(10.17, 10.18)	(7.84, 7.91)	(5.67, 5.89)	(10.38, 10.39)	(8.56, 8.62)	(6.43, 6.84)
LB^{m}	10.16	7.63	5.34	10.37	8.17	5.80
	(10.15, 10.16)	(7.63, 7.64)	(5.33, 5.35)	(10.36, 10.37)	(8.16, 8.18)	(5.79, 5.80)
UB^{m}	10.21	7.98	5.85	10.46	8.85	6.80
	(10.21, 10.22)	(7.94, 8.02)	(5.74, 5.97)	(10.45, 10.46)	(8.80, 8.90)	(6.56, 7.05)

Notes. Rows marked "LB^{LP}" and "LB^{m}" report estimates of the CE annualized percentage returns r_{CE} from the strategy determined by Algorithm II and the myopic strategy, respectively. Approximate 95% confidence intervals are reported in parentheses. Estimates are based on one million simulated paths. The row "V^{u}" reports the optimal value function for the problem. Rows marked "UB^{LP}" and UB^{m}" report estimates of the upper bound on the true value function computed using these strategies.

would be the unique market price of risk process in the unconstrained market. Clearly, then, we do not need to explicitly define artificial asset price dynamics in order to apply the dual methodology. (See also the final paragraph of §C of the online supplement, where we explain why the choice of artificial asset price dynamics does not impact the dual bound in our numerical examples.)

Table 2 **Results of Model 1 Using Algorithm III Strategy**

	$T=5$			$T=10$		
Function	$\gamma=1.5$	$\gamma=3$	$\gamma=5$	$\gamma=1.5$	$\gamma=3$	$\gamma=5$
			Incomplete markets			
LB^{LP}	16.79	10.32	7.05	17.77	11.50	8.07
	(16.77, 16.81)	(10.29, 10.35)	(7.00, 7.09)	(17.75, 17.78)	(11.44, 11.55)	(8.01, 8.14)
UB^{LP}	16.79	10.34	7.06	17.77	11.54	8.29
	(16.76, 16.82)	(10.26, 10.43)	(6.87, 7.25)	(17.75, 17.80)	(11.42, 11.67)	(8.04, 8.55)
LB^{m}	16.63	9.86	6.59	17.46	10.57	7.09
	(16.61, 16.64)	(9.84, 9.87)	(6.58, 6.61)	(17.45, 17.47)	(10.56, 10.58)	(7.08, 7.10)
UB^{m}	16.82	10.41	7.12	17.82	11.72	8.50
	(16.79, 16.84)	(10.33, 10.50)	(6.93, 7.32)	(17.80, 17.85)	(11.59, 11.84)	(8.24, 8.77)
V^{u}	16.79	10.32	7.06	17.76	11.55	8.12
			Incomplete markets + No short sales and no borrowing			
LB^{LP}	10.16	7.83	5.68	10.38	8.52	6.55
	(10.15, 10.16)	(7.82, 7.84)	(5.66, 5.71)	(10.38, 10.38)	(8.51, 8.53)	(6.52, 6.58)
UB^{LP}	10.16	7.83	5.67	10.38	8.53	6.63
	(10.15, 10.16)	(7.80, 7.86)	(5.55, 5.80)	(10.38, 10.39)	(8.50, 8.55)	(6.50, 6.77)
LB^{m}	10.15	7.63	5.33	10.37	8.17	5.80
	(10.15, 10.16)	(7.62, 7.64)	(5.32, 5.34)	(10.36, 10.37)	(8.16, 8.18)	(5.79, 5.80)
UB^{m}	10.20	7.96	5.77	10.46	8.84	6.98
	(10.20, 10.21)	(7.91, 8.00)	(5.62, 5.91)	(10.45, 10.47)	(8.79, 8.89)	(6.81, 7.15)

Notes. Rows marked "LB^{LP}" and "LB^{m}" report estimates of the CE annualized percentage returns r_{CE} from the strategy determined by Algorithm III and the myopic strategy, respectively. Approximate 95% confidence intervals are reported in parentheses. Estimates are based on one million simulated paths. The row "V^{u}" reports the optimal value function for the problem. Rows marked "UB^{LP}" and "UB^{m}" report estimates of the upper bound on the true value function computed using these strategies.

Assume that the state variable follows a three-dimensional Ornstein–Uhlenbeck process reverting to zero vector. More precisely, the asset return dynamics satisfy the following SDEs:

$$r_t = \delta_0 + \delta_1 Z_t,$$
$$\frac{dP_t}{P_t} = r_t\,dt + \Sigma_P\,dB_t,$$
$$dZ_t = -KZ_t\,dt + \Sigma_Z\,dB_t,$$

where

$$K = \begin{bmatrix} 0.576 & 0 & 0 \\ 0 & 3.343 & 0 \\ -0.421 & 0 & 0.083 \end{bmatrix}, \quad \Sigma_Z = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix},$$

$$\delta_0 = 0.056, \quad \delta_1 = \begin{bmatrix} 0.018 \\ 0.007 \\ 0.010 \end{bmatrix}^\top,$$

and $\Sigma_P = -\delta_1 K^{-1}(I - e^{-(T-t)K})\Sigma_Z$, which forces the bond price to equal the face value at maturity. Note that the diffusion vector, Σ_P, of the asset return is time dependent in this model.

The initial state variable is $Z_0 = 0$, the time to maturity is $T = 5$ years, and the risk-aversion coefficient is $\gamma = 15$, which reflects a high degree of risk aversion. This is intentional because we can guess in this case that the policy of holding all of the portfolio in the long-term bond should be very close to optimal. In this particular model, then, we can consider the buy-and-hold policy, which invests all in the long-term bond, as another benchmark. Note that the myopic strategy in this case will simply invest everything in the cash account since the risk premium on the long-term bond is zero. In this model (and model 3 below), we consider two sets of trading constraints. In the first, the investor does not face any trading constraints and simply has an incomplete markets problem. In the second case, the investor faces a no-borrowing constraint in addition to an incomplete market.

In applying Algorithm III, we choose

$$\mathbb{B} = \{P_i(z_1)\cdot P_j(z_2)\cdot P_k(z_3)\cdot (T-t/T)^l \;|\; 0 \le i+j+k \le 3,\ 1 \le l \le 3\}$$

as our set of basis functions where once again P_i denotes the Chebyshev polynomial of degree i.

Table 3 displays the results of this experiment. As expected, because of the high value of γ, we observe that the performance of the buy-and-hold policy on the long-term bond is much better than that of the myopic policy. Surprisingly, however, the LP approach performs even better and produces lower and upper bounds that, to two decimal places at least, are identical.

Table 3 Results of Model 2 Using Algorithm III Strategy

Function	Incomplete markets	Incomplete markets + No borrowing
LB^{LP}	5.52	5.52
	(5.52, 5.52)	(5.52, 5.52)
UB^{LP}	5.52	5.52
	(5.52, 5.52)	(5.52, 5.52)
LB^{m}	4.41	4.41
	(4.40, 4.41)	(4.41, 4.42)
UB^{m}	5.52	5.52
	(5.52, 5.52)	(5.52, 5.52)
LB^{LT}	5.51	5.51
	(5.51, 5.51)	(5.51, 5.51)
UB^{LT}	5.53	5.58
	(5.52, 5.53)	(5.57, 5.59)

Notes. Rows "LB^{LP}," "LB^{m}," and "LB^{LT}" report estimated CE annualized percentage returns r_{CE} from the strategy determined by Algorithm III, the myopic strategy, and the buy-and-hold strategy on the long-term bond, respectively. Approximate 95% confidence intervals are reported in parentheses. Estimates are based on one million simulated paths. The rows marked "UB^{LP}," "UB^{m}," and "UB^{LT}" report estimates of the upper bound on the true value function computed using these strategies.

5.3. Example 3: Two Risky Bonds and a Stock Index with Four State Variables

In our final model, which is again taken from Haugh et al. (2006b), there are three risky assets: two bonds with maturities 3 years and 10 years, respectively, and a stock index. There is a four-dimensional state variable and a five-dimensional Brownian motion. As was the case with Example 2, we explicitly define the market price of risk process η_t instead of defining additional artificial risky assets that the investor will not be permitted to trade. (We note from the dynamics in (23b) and (23c) that our choice of η satisfies $\Sigma_P \eta = \lambda$. The same argument that we

Table 4 Parameters for Model 3 Defining the Instantaneous Risk-Free Rate, Risk Premium, and State Variable Processes in (23a), (23b), and (23d), Respectively

Parameter	Value				
K	0.576	0	0	0	
	0	3.343	0	0	
	−0.421	0	0.083	0	
	0	0	0	0.080	
Σ_Z	1.0000	0	0	0	0
	0	1.0000	0	0	0
	0	0	1.0000	0	0
	0	0	0	0.1600	0.3664
δ_0	0.056				
δ_1	0.018	0.007	0.010	0	
$\lambda_0^\top$	−0.5630	−0.2450	−0.2190	0.4400	0
$\lambda_1^\top$	0	0	0.5370	0.1110	0
	1.7540	−1.8150	0.3760	0.3050	0
	0	0	−0.0820	−0.0170	0
	0	0	0	0.0700	0

Table 5 **Results of Model 3 Using Algorithm III Strategy**

	Incomplete markets			Incomplete markets + No borrowing		
Function	$\gamma=1.5$	$\gamma=3$	$\gamma=5$	$\gamma=1.5$	$\gamma=3$	$\gamma=5$
LB^{LP}	59.47	35.82	24.65	32.26	20.59	15.45
	(59.39, 59.54)	(35.67, 35.98)	(24.20, 25.14)	(32.12, 32.40)	(20.46, 20.72)	(15.31, 15.59)
UB^{LP}	59.49	35.97	25.91	32.31	20.77	15.79
	(59.38, 59.61)	(35.53, 36.41)	(24.97, 26.89)	(32.11, 32.52)	(20.40, 21.15)	(15.15, 16.44)
LB^{m}	59.06	34.59	23.61	32.15	20.23	15.09
	(59.00, 59.13)	(34.52, 34.66)	(23.55, 23.67)	(32.03, 32.28)	(20.13, 20.32)	(15.02, 15.17)
UB^{m}	60.01	37.54	27.40	32.55	21.18	16.34
	(59.90, 60.12)	(37.17, 37.92)	(26.39, 28.44)	(32.34, 32.75)	(20.74, 21.64)	(15.45, 17.27)

Notes. Rows "LB^{LP}" and "LB^{m}" report estimated CE annualized percentage returns r_{CE} from the strategy determined by Algorithm III and the myopic strategy, respectively. These estimates are based on one million simulated paths for the incomplete market problem and 100,000 paths for the no-borrowing problem. Approximate 95% confidence intervals are reported in parentheses. Rows "UB^{LP}" and "UB^{m}" report the estimates of the corresponding upper bounds on the true value function.

provided in the case of Example 2 would then apply here. In particular, there is no need to explicitly define artificial asset price dynamics in order to complete the unconstrained market.) We assume the state vector follows a four-dimensional mean-reverting Ornstein–Uhlenbeck process with mean zero vector. The asset return and the risk premium dynamics satisfy the SDEs:

$$r_t = \delta_0 + \delta_1 Z_t, \tag{23a}$$

$$\eta_t = \lambda_0 + \lambda_1 Z_t, \tag{23b}$$

$$\frac{dP_t}{P_t} = (r_t \cdot \mathbf{1} + \Sigma_P \eta_t)\,dt + \Sigma_P\,dB_t, \tag{23c}$$

$$dZ_t = -KZ_t\,dt + \Sigma_Z\,dB_t, \tag{23d}$$

where $Q = K + \Sigma_Z \lambda_1$ and

$$\Sigma_P = \begin{bmatrix} & -\delta_1 Q^{-1}(I - e^{-3\cdot Q})\Sigma_Z & \\ & -\delta_1 Q^{-1}(I - e^{-10\cdot Q})\Sigma_Z & \\ -0.0126 \quad 0.0057 & -0.0295 \quad 0.143 & 0 \end{bmatrix}.$$

The particular forms of the first and second rows of Σ_P imply that we use dynamic rollover strategies for the 3-year and 10-year bonds so that the duration of the bonds should be maintained at 3 and 10 years by continuous reinvestment. The other parameter values are K, Σ_Z, δ_0, δ_1, λ_0, and λ_1, reported in Table 4. The initial state vector is $Z_0 = 0$, the horizon is $T = 5$ years, and the constant relative risk-aversion coefficient γ is set to 1.5, 3, and 5.

When we use the adaptive constraint selection Algorithm III, we use

$$\mathbb{B} = \big\{P_i(z_1)P_j(z_2)P_k(z_3)P_l(z_4)(T - t/T)^m \;\big|\; 0 \le i+j+k+l \le 2,\; 1 \le m \le 10\big\}$$

as our set of basis functions.

Table 5 displays the numerical results for this model. The results are consistent with our earlier examples in that, regardless of the market trading constraints, the LP strategy outperforms the myopic strategy. The gap between the two trading strategies is more visible here than in model 1, for example. When $\gamma = 5$ in the incomplete markets case, the duality gap of $25.91 - 24.65 = 1.26\%$ suggests that the LP-based strategy is still reasonably far from the optimal strategy. As stated earlier, we suspect that we could improve the LP-based strategy via a more careful linearization of the constraints in phase 2 of Algorithm III.

6. Conclusions

We have extended the linear programming approach of Han and Van Roy (2011) to compute good suboptimal solutions for high-dimensional control problems in a diffusion-based setting with fixed time horizons. In considering numerical examples drawn from portfolio optimization, we were able to show that our suboptimal solutions are indeed very good by using them to construct tight lower and upper bounds on the optimal value functions for these problems. These results suggest that the LP approach is a very promising one for tackling high-dimensional control problems.

There are several possible directions for future research. First, it would be interesting to extend the methodology to jump diffusions and other more general settings. There is also scope for additional theoretical work in order to better understand the properties of these LP-based algorithms. Given some of the necessary ad hoc steps of the LP approach in this paper and the original work of Han and Van Roy (2011), this may be particularly challenging.

Supplemental Material

Supplemental material to this paper is available at http://dx.doi.org/10.1287/ijoc.2015.0651.

References

Bertsekas DP (2005) *Dynamic Programming and Optimal Control*, 3rd ed., Vol. I (Athena Scientific, Nashua, NH).

Bertsekas DP (2012) *Dynamic Programming and Optimal Control*, 4th ed., Vol. II (Athena Scientific, Nashua, NH).

Bick B, Kraft H, Munk C (2012) Solving constrained consumption investment problems by simulation of artificial market strategies. *Management Sci.* 59:485–503.

de Farias DP, Van Roy B (2003) The linear programming approach to approximate dynamic programming. *Oper. Res.* 51:850–865.

de Farias DP, Van Roy B (2004) On constraint sampling in the linear programming approach to approximate dynamic programming. *Math. Oper. Res.* 29:462–478.

Duffie D (1996) *Dynamic Asset Pricing Theory* (Princeton University Press, Princeton, NJ).

Glasserman P (2004) *Monte Carlo Methods in Financial Engineering* (Springer-Verlag, New York).

Han J, Van Roy B (2011) Control of diffusions via linear programming. Infanger G, ed. *Stochastic Programming: The State of the Art, in Honor of George B. Dantzig* (Springer, New York), 329–354.

Haugh MB, Jain A (2011) The dual approach to portfolio evaluation: A comparison of the static, myopic and generalized buy-and-hold strategies. *Quant. Finance* 11:81–99.

Haugh MB, Kogan L, Wang J (2006a) Evaluating portfolio strategies: A duality approach. *Oper. Res.* 54:405–418.

Haugh MB, Kogan L, Wu Z (2006b) Portfolio optimization with position constraints: An approximate dynamic programming approach. Working paper, Columbia University, New York.

Lynch A (2001) Portfolio choice and equity characteristics: Characterizing the hedging demands induced by return predictability. *J. Financial Econom.* 62:67–130.

Schweitzer P, Seidmann A (1985) Generalized polynomial approximations in Markovian decision processes. *J. Math. Anal. Appl.* 110:568–582.

Wachter J, Sangvinatsos A (2005) Does the failure of the expectations hypothesis matter for long-term investors? *J. Finance* 60:179–230.