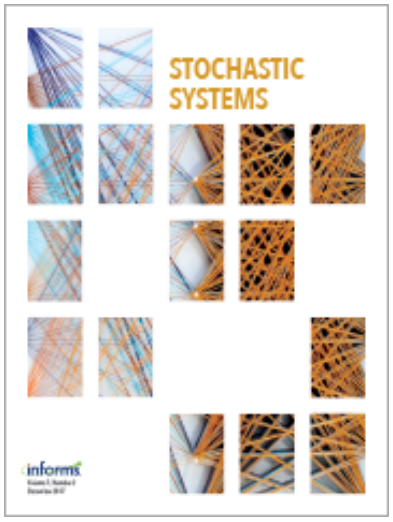


Stochastic Systems

Publication details, including instructions for authors and subscription information:
http://pubsonline.informs.org

Solving the Drift Control Problem

Melda Ormeci Matoglu, John Vande Vate, Huizhu Wang

To cite this article:
Melda Ormeci Matoglu, John Vande Vate, Huizhu Wang (2015) Solving the Drift Control Problem. Stochastic Systems 5(2):324-371. https://doi.org/10.1287/12-SSY087

Stochastic Systems
2015, Vol. 5, No. 2, 324–371
DOI: 10.1214/12-SSY087

SOLVING THE DRIFT CONTROL PROBLEM

By Melda Ormeci Matoglu[*,‡], John Vande Vate[†,‡] and Huizhu Wang[†,‡]

University of New Hampshire * *and Georgia Institute of Technology* †

We model the problem of managing capacity in a build-to-order environment as a Brownian drift control problem. We formulate a structured linear program that models a practical discretization of the problem and exploit a strong relationship between relative value functions and dual solutions to develop a functional lower bound for the continuous problem from a dual solution to the discrete problem. Refining the discretization proves a functional strong duality for the continuous problem. The linear programming formulation is so badly scaled, however, that solving it is beyond the capabilities of standard solvers. By demonstrating the equivalence between strongly feasible bases and deterministic unichain policies, we combinatorialize the pivoting process and by exploiting the relationship between dual solutions and relative value functions, develop a mechanism for solving the LP without ever computing its coefficients. Finally, we exploit the relationship between relative value functions and dual solutions to develop a scheme analogous to column generation for refining the discretization so as to drive the gap between the discrete approximation and the continuous problem to zero quickly while keeping the LP small. Computational studies show our scheme is much faster than simply solving a regular discretization of the problem both in terms of finding a policy with a low average cost and in terms of providing a lower bound on the optimal average cost.

1. Introduction. Consider the problem of managing capacity in a build-to-order environment modeled as a Brownian drift control problem with the objective of minimizing the long-term average cost. Assume the controller can, at some cost, shift the processing rate among a finite set of alternatives by, for example, adding or removing staff, increasing or reducing the number of shifts or opening or closing production lines. The controller incurs a cost for capacity per unit time and a delay cost that reflects the opportunity cost of revenue waiting to be recognized or the customer service impacts of delaying delivery of orders. Furthermore he incurs a cost per

Received November 2012.

‡This work has been supported in part by NSF grant 0825840.

AMS 2000 subject classifications: Primary 60J70; secondary 90C05; 90C46.

Keywords and phrases: Diffusion process, drift control, algorithm, strongly feasible basis.

unit to reject orders or idle resources as necessary to keep the workload of waiting orders within a prescribed range. Ormeci Matoglu and Vande Vate (2011) address this problem by introducing a practical restriction, which allows the controller to change the capacity only when the workload in the order queue is a value in a given finite set $\mathcal{S}$ (e.g., all the integer values up to some constant). They call this the $\mathcal{S}$-restricted Brownian control problem and formulate a linear programming model to solve it. They show that when the costs for changing the processing rate satisfy a triangle inequality, an optimal solution to the $\mathcal{S}$-restricted problem can be found among a special class of policies called deterministic non-overlapping control band policies. The simplicity of these policies greatly facilitates their application in industrial settings.

In this paper we address the $\mathcal{S}$-restricted problem with a new LP formulation that enjoys the important properties (e.g., pre-Leontief structure with a single scaling constraint, strong relationship between complementary dual solutions and relative value functions) of the original formulation in Ormeci Matoglu and Vande Vate (2011) and prove a cleaner and more satisfying strong duality theory for the restricted and unrestricted problems. Ormeci Matoglu and Vande Vate (2011) show how to construct a sequence of policies whose average costs converge to a lower bound on the average cost of any non-anticipating policy. They do not however state a strong duality theory for the problem because their approach never constructs a pair (γ, f) satisfying all the lower bound conditions (i.e. a feasible dual solution) for the unrestricted problem. We provide a constructive proof of the strong duality theorem by showing how to compute a sequence of policies and lower bounds (i.e. feasible dual solutions) whose values converge proving that the lower bound is tight. Theorem 3.1 proves that strong duality, namely that the infimum of the average cost over all non-anticipating policies is equal to the supremum of γ over all pairs (γ, f) satisfying the lower bound conditions for the unrestricted problem. In fact, we extend this result to the non-overlapping control band policies and show (without requiring that the transition costs satisfy a triangle inequality) that the infimum of the average cost over all non-overlapping control band policies is equal to the supremum of γ over all pairs (γ, f) satisfying the same lower bound conditions (Corollary 3.1).

The original linear programming formulation of Ormeci Matoglu and Vande Vate (2011) and indeed the new formulation we study here are extremely poorly scaled and solutions are so sensitive to small errors in the problem coefficients that completing pivots algebraically is difficult and standard solvers are not up for the task. To help bridge the gap between the theory and its practical application, we exploit the properties of the LP

formulation and the strong relationship between dual solutions and relative value functions to develop a largely combinatorial approach that allows us to solve the problem without ever computing the problem coefficients.

Our combinatorial pivoting scheme focuses on "strongly feasible bases", initially used by Cunningham (1976) and then by Orlin (1985) to solve network problems via simple combinatorial modifications to the simplex method. The main idea behind these modifications is to keep the basis "strongly feasible" via a rule for choosing the leaving basic variable at each iteration. Applying the notion of strongly feasible bases to the the drift control problem allows us to exploit the combinatorial nature of the problem and yields significant computational advantages. As a consequence, we can restrict our attention to the structure of the bases rather than the values of the primal variables and so combinatorialize much of the pivoting process, reducing the computational burden significantly. The relationship between complementary dual solutions and relative value functions allows computing the reduced costs directly without computing the dual solution (or for that matter the problem coefficients).

Using the techniques developed in Ormeci Matoglu and Vande Vate (2011) we can obtain a lower bound on the average cost of an optimal policy for the unrestricted problem from our solution to the $\mathcal{S}$-restricted problem. Refining the discretization, the obvious method for closing the gap between the solution to the $\mathcal{S}$-restricted problem and this lower bound, makes the LP large. By exploiting the relationship between complementary dual solutions and relative value functions we show how to choose the points of $\mathcal{S}$ so as to drive this gap to zero quickly. This allows us to produce an ϵ-optimal solution by adding only a relatively small number of points to $\mathcal{S}$.

Drift control problems were studied in the literature in different contexts with different cost structures and solution approaches. See, for example, (Ata, Harrison and Shepp, 2005; Avram and Karaesmen, 1996; Chernoff and Petkau, 1978; Ghosh and Weerasinghe, 2007; Liao, 1984; Perry and Bar-Lev, 1989; Rath, 1977). While (Avram and Karaesmen, 1996; Chernoff and Petkau, 1978; Perry and Bar-Lev, 1989; Rath, 1977) restrict the controller to only two drift rates, Ata, Harrison and Shepp (2005) confine the drift rate in a finite range with more general processing costs, but do not address the holding or delay costs and changeover costs and the cost of displacement at the lower boundary. (Ghosh and Weerasinghe, 2007, 2010) include a congestion cost similar to a holding cost but do not include changeover costs. Ormeci Matoglu and Vande Vate (2011) and this paper deal with the more general problem of selecting from many rates when the process incurs the cost of lost production whenever the lines are idled and the cost of rejecting

orders whenever the upper boundary is reached, reflecting both the immediate lost revenue and the potential impact on future sales to the customer.

The use of linear programming to reformulate long-term average stochastic control problems began with Manne (1960) in the context of a discrete time, finite state controlled Markov chain and now has become standard (e.g., Feldman and Valdez-Flores (1996)). In generalizing from discrete time and finite state space to continuous time and continuous state, different approaches have been developed. Infinite dimensional linear programming has been one of the tools employed (e.g., Anderson and Nash, 1989; Anderson, Nash and Perold, 1983; Hernández-Lerma and González-Hernández, 1998; Hernández-Hernández and Hernández-Lerma, 1994; Hernández-Lerma and Lasserre, 1998; Klabjan and Adelman, 2006; Taksar, 1997). By their nature these problems are in general difficult to solve and the main challenge has been to show the absence of a duality gap between the primal and the dual programs (strong duality). Strong duality results are generally developed with varying restrictions on the problems and the sets involved. Klabjan and Adelman (2006) develop infinite-dimensional linear programming theory for semi-Markov decision processes on Borel spaces with average cost criterion in which the state transitions are deterministic. Hernández-Hernández, Hernández-Lerma and Taksar (1996) employ infinite-dimensional linear programming methods to study deterministic continuous-time control problems and discrete-time Markov decision processes with discounted cost. They give conditions for solvability and strong duality. Taksar (1997) employs linear programming techniques to a singular diffusion control problem and shows that the dual program is equivalent to finding the maximal solution to a variational inequality. Another common approach used to solve the infinite dimensional problems has been discretization. Kushner and Dupuis (2001) approximate a stochastic control problem in continuous time and continuous state via a Markov Chain on a finite state discretization of the original state space and apply numerical methods to find a solution to the Markov Chain problem that optimizes an analogue of the original cost function. Kumar and Muthuraman (2004) develop a discretization scheme using finite element methods for certain singular control diffusion problems. Helmes and Stockbridge (2000) and Helmes and Stockbridge (2008) develop linear programming based approaches to solve diffusion control problems. These methods generate constraints on a finite set of moments to develop an approximate solution. In the infinite dimensional LP problem discretizing time converts the problem to an approximating LP (e.g., Buie and Abrham, 1973; Pullan, 1993). However this has some drawbacks as the LP problem becomes quite large, the solution is only approximate and the discretization may obscure

important features of the optimal solution. Another line of work involves extending the simplex method for finite dimensional linear programs to the continuous-time or infinite-dimensional problem (see Weiss, 2008).

In some sense we address the problem in both ways, but avoid some of the complications. We introduce a discretization on the state space (the $\mathcal{S}$-restricted problem), use this discretization to construct an LP and provide a mechanism to complete the simplex pivots. In doing so we restrict our attention to the structure of the bases rather than the values of the primal variables and so combinatorialize much of the pivoting process, reducing the computational burden significantly. We also obtain a lower bound on the average cost of an optimal policy for the unrestricted problem from our solution to the $\mathcal{S}$-restricted problem. Exploiting the relationship between complementary dual solutions and relative value functions, we also develop an intelligent mechanism for adding points to $\mathcal{S}$ and so variables or columns to the LP formulation (as in column generation) so as to quickly close the gap between the average cost of an optimal solution to the $\mathcal{S}$-restricted problem and a lower bound on the cost of an optimal solution to the unrestricted problem.

Thus our contributions are: (1) We formulate a structured linear program that models a practical discretization of the problem and exploit a strong relationship between relative value functions and dual solutions to develop a functional lower bound for the continuous problem from a dual solution to the discrete problem. Refining the discretization proves a functional strong duality for the continuous problem (Theorem 3.1). (2) We observe that this LP is poorly scaled and that standard algorithms and solvers fail to solve it. By demonstrating the equivalence between strongly feasible bases and deterministic unichain policies, we combinatorialize the pivoting process and by exploiting the relationship between dual solutions and relative value functions, develop a mechanism for solving the LP without ever computing its coefficients. (3) We develop a scheme analogous to column generation for refining the discretization so as to drive the gap between the discrete approximation and the continuous problem to zero quickly while keeping the LP small. (4) We present computational studies showing that our "column generation" scheme is much faster than simply solving a regular discretization of the problem.

The rest of the paper is organized as follows: In §2 we describe the Brownian drift control problem. In §3 we present our strong duality result, Theorem 3.1, and summarize the main results of this paper. In §4 we introduce a controlled random walk approximation and provide a linear programming formulation of the controlled random walk. In §5 we establish the correspondence between a relative value function and a dual solution, provide a lower

bound on the average cost of any non-anticipating policy, and prove Theorem 3.1. In §6 we introduce the notion of a strongly feasible basis. Theorem 6.1 establishes the equivalence of strongly feasible bases and deterministic unichain policies, which allows us to combinatorialize the pivoting process of the LP. In §7 we provide a mechanism to add points to the set $\mathcal{S}$ so as to reduce the gap between the $\mathcal{S}$-restricted problem and the lower bound on the optimal solution to the unrestricted problem quickly. In §8 we briefly address implementation issues and present computational studies that demonstrate that the column generation algorithm shrinks the duality gap rapidly and consistently on a range of problems that have three to eight drift rates.

2. The Brownian drift control problem. Let

$$W(T) = W(0) + \int_0^T \mu(t)dt + \int_0^T \sigma(\mu(t))dB(t), \quad T \geq 0,$$

be a diffusion process with drift $\mu(t)$ in some fixed finite set Λ for each $t \geq 0$, variance $\sigma^2(\mu) > 0$ for each $\mu \in \Lambda$ and initial level $W(0)$ on some filtered space $\{\Omega, \mathcal{F}, \mathbb{P}; \mathcal{F}_t, t \geq 0\}$. To simplify notation we consider the special case in which $\sigma^2(\mu) = \sigma^2$. The results we present apply in the more general setting in which the variance σ^2 is a function of the drift rate μ. The process $W(T)$ describes the difference between the cumulative work to have arrived and the cumulative work processed by time T, i.e. the netput process Harrison (1985). The drift rate $\{\mu(t), t \geq 0\}$, which is adapted to the Brownian motion $\{B(t) : t \geq 0\}$, is the difference between the average arrival rate and the rate $\lambda(t)$ at which work is completed. We assume the arrival process is time homogeneous with average rate μ_0 and that the controller can, at some cost, shift the processing rate among a finite set of alternatives. Further, the controller must exert the minimal instantaneous control required to keep the process within the allowed range $[0, \Theta]$. We let $A(t)$ denote the cumulative increases in work and $R(t)$ the cumulative decreases in work up to time t exerted by the controller at 0 and Θ, respectively. The resulting controlled process is

$$X(T) = X(0) + \int_0^T \mu(t)dt + \sigma B(T) + A(T) - R(T), \quad T \geq 0, \tag{2.1}$$

where $X(0) = W(0)$. $X(t)$ lives in the bounded region $[0, \Theta]$, and the controller may only adjust the drift rate by choosing from among the possible values in the finite set Λ. We assume, without loss of generality, that $W(0) \in [0, \Theta]$. To avoid tedious case analysis, we also assume that $0 \notin \Lambda$.

The special case $\mu = 0$ requires different formulas for computing the coefficients of our LP formulation and a special formula for the relative value function. All of our arguments also work for these special formulas, but the tedious case analysis would not be illuminating. See Appendix B for these coefficients and the relative value function when $\mu = 0$.

A policy defines the times at which and amounts by which we adjust the drift rate. We restrict attention to the space $\mathcal{P}$ of all non-anticipating policies $\Phi = \{(T_i, u_i) : i \geq 0\}$, where (i) $0 \leq T_0 < T_1 < T_2 < \cdots < T_i < T_{i+1}, \ldots$ is a sequence of stopping times and (ii) Each $u_i \in \Lambda$ is a random variable adapted to $\mathcal{F}_{T_i}$ indicating the rate to which we change the drift at time T_i. Under the policy $\Phi = \{(T_i, u_i) : i \geq 0\}$, the drift rate $\mu(t) = u_i$ for $T_i \leq t < T_{i+1}$.

With each policy $\{(T_i, u_i) : i \geq 0\} = \{\mu(t) : t \geq 0\}$, the associated Skorokhod problem: (a) $X(t) \in [0, \Theta]$, $t \geq 0$, (b) $A(\cdot)$, $R(\cdot)$ are nondecreasing and continuous with $A(0) = 0$, $R(0) = 0$, (c) $\int_0^T 1_{\{X(t)>0\}} dA(t) = \int_0^T 1_{\{X(t)<\Theta\}} dR(t) = 0$, $t \geq 0$, where the continuous process $\{X(t) : t \geq 0\}$ is defined by (2.1), uniquely defines $\{X, A, R\}$ (See Section 2.4 of Harrison (1985)). Note that since the drift rate controls uniquely determine the instantaneous controls exerted at the boundaries, we do not include the latter in our specification of a policy.

To change the drift from rate u to rate v, the controller must pay a fixed cost, $K(u, v) > 0$, for $u \neq v$. To simplify notation, we let $K(u, u) = 0$ for all $u \in \Lambda$.

There is a cost $c(u)$ per unit time for the capacity to process work that depends on the drift rate u and when $X(t) > 0$ there is a backlog of orders, which incurs a linear delay cost at rate h per unit per unit time. The instantaneous controls exerted at 0 (Θ) to adjust the workload up (down) incurs a unit cost of U (M).

We consider the *Average Cost Brownian Control Problem*, which is to find a non-anticipating policy that minimizes the long run average cost:

$$\mathrm{AC}(\Phi) = \limsup_{T\to\infty} \frac{1}{T}\mathbb{E}\Big[\int_0^T c(\mu(t)) + hX(t)dt + UA(T) + MR(T) + \sum_{i=1}^{N(T)} K(u_{i-1}, u_i)\Big],$$

where $\mu(t)$ denotes the drift rate at time t and, for each $T \geq 0, N(T) = \sup\{n \geq 0 : T_n \leq T\}$ denotes the number of changes in the drift rate by time T.

A *control band* $\psi = (u, s, S)$ is defined by a rate $u \in \Lambda$ and an interval (s, S). Given $\mu(t) = u$ and $X(t) \in (s, S)$, a policy implementing the control

band $\psi = (u, s, S)$ maintains the drift rate u until X first reaches $\{s, S\}$. If $s < 0$, a policy implementing the control band (u, s, S) maintains the drift rate u until $X(t)$ first reaches S and relies on instantaneous controls at 0 to keep the process non-negative. Similarly, if $S > \Theta$, a policy implementing the control band (u, s, S) maintains the drift rate u until $X(t)$ first reaches s and relies on instantaneous controls at Θ to keep the process below that bound.

We say that two bands (u, s, S) and (u', s', S') *overlap* if $u = u'$ and the intervals (s, S) and (s', S') overlap, i.e., $(s, S) \cap (s', S') \neq \emptyset$. Otherwise, the bands are *non-overlapping*. We say that a collection of bands is non-overlapping if no two of its members overlap.

A *control band policy* is defined by a collection of control bands $\Psi = \{\psi_i : i \in \mathcal{I}\}$ such that each point $x \in [0, \Theta]$ is contained in some band ψ_i together with a rule for switching from one control band to the next. Such a policy maintains the drift rate of the current control band (u, s, S) until $X(t)$ first reaches s or S at which point it changes to a new band by changing the drift rate as dictated by the switching rule.

We say the switching rule for a band (u, s, S) is *deterministic* if the rule associates with each endpoint s and S a unique band to switch to. A *deterministic control band policy* is a control band policy in which the switching rule from each band is deterministic. Deterministic policies are simple to describe and implement.

We consider the $\mathcal{S}$-restricted Brownian control problem introduced in Ormeci Matoglu and Vande Vate (2011) in which the controller may only change the drift rate when $X(t)$ is in a given finite set $\mathcal{S} = \{s_i : i = 1, 2, \ldots, n\}$, where $0 = s_0 < s_1 < s_2 \ldots < s_n < s_{n+1} = \Theta$. We define the subset $\mathcal{P}(\mathcal{S}) \subset \mathcal{P}$ to be those non-anticipating policies $\{(T_i, u_i) : i \geq 0\}$ in which each T_i is a hitting time for some subset of $\mathcal{S}$.

Ormeci Matoglu and Vande Vate (2011) formulate a linear program LP(Λ, $\mathcal{S}$) to find a minimum average cost control band policy in $\mathcal{P}(\mathcal{S})$ and prove that there is an optimal solution to LP$(\Lambda, \mathcal{S})$ that corresponds to a deterministic non-overlapping control band policy. They show that an optimal solution to LP$(\Lambda, \mathcal{S})$ is an optimal policy for the $\mathcal{S}$-restricted Brownian control problem and so, the $\mathcal{S}$-restricted Brownian control problem admits an optimal policy that is a deterministic non-overlapping control band policy under the assumption that the transition costs $K(u, v)$ satisfy a triangle inequality.

Note that when considering a function f defined on a subset of $\mathbb{R} \times \Lambda$ we treat f as a family of functions $\{f(\cdot, u) : u \in \Lambda\}$ each defined on the corresponding subset of $\mathbb{R}$ and so, for example, use f' and f'' to represent derivatives with respect to the first argument. For a function f we define the following smoothness conditions.

Smoothness conditions: *We say that a function f satisfies the smoothness conditions if f (i) can be written as the difference of two convex functions, (ii) is differentiable at all but a finite set of points and (iii) has bounded first and second derivatives.*

Ormeci Matoglu and Vande Vate (2011) prove:

PROPOSITION 2.1. *Suppose that for each $u \in \Lambda$, $f(\cdot, u) : [0, \Theta] \to \mathbb{R}$ satisfies the smoothness conditions and that γ is a scalar. Further, suppose that for each $u \in \Lambda$ the function $f(\cdot, u)$ and the scalar γ satisfy:*

$$\frac{\sigma^2}{2} f''(a,u) + u f'(a,u) + c(u) + ha \geq \gamma \text{ for almost all } a \in [0,\Theta], \tag{2.2}$$

$$f(a,v) - f(a,u) \geq -K(u,v) \text{ for all } a \in [0,\Theta] \text{ and } v \in \Lambda, \tag{2.3}$$

$$f'(0,u) \geq -U, \tag{2.4}$$

$$f'(\Theta,u) \leq M, \tag{2.5}$$

$$\rho(a,u) \equiv f'(a+,u) - f'(a-,u) \geq 0 \text{ for all } a \in (0,\Theta). \tag{2.6}$$

Then $\gamma \leq \mathrm{AC}(\Phi)$ for each policy $\Phi \in \mathcal{P}$ and each initial state.

A pair (γ, f) is said to satisfy *the lower bound conditions* if $f(\cdot, u)$ satisfies the smoothness conditions and γ and $f(\cdot, u)$ satisfy (2.2)–(2.6) for each $u \in \Lambda$.

3. Main results. One of the main contributions of this paper is Theorem 3.1, which extends Proposition 2.1 to prove that the lower bound is tight. Ormeci Matoglu and Vande Vate (2011) show how to construct a sequence of policies whose average costs converge to a lower bound on the average cost of any non-anticipating policy. That paper does not however state a strong duality theory for the problem because the approach never constructs a pair (γ, f) satisfying all the lower bound conditions. We introduce a modest modification of this original approach that allows us to prove Theorem 3.1 and along the way, provides a stronger result about the optimality of non-overlapping policies.

THEOREM 3.1. *The infimum of $\mathrm{AC}(\Phi)$ over all non-anticipating policies $\Phi \in \mathcal{P}$ is equal to the supremum of γ over all pairs (γ, f) satisfying the lower bound conditions.*

We further generalize the result of Ormeci Matoglu and Vande Vate (2011) by proving Corollary 3.1 without requiring that the transitions costs satisfy a triangle inequality.

COROLLARY 3.1. *The infimum of* $\mathrm{AC}(\Phi)$ *over all non-overlapping control band policies* $\Phi \in \mathcal{P}$ *is equal to the supremum of* γ *over all pairs* (γ, f) *satisfying the lower bound conditions.*

Theorem 3.1 relies on formulating a structured linear program (LP) that models a practical discretization of the problem (originating from the $\mathcal{S}$-restricted Brownian control problem) and exploiting a strong relationship between relative value functions and dual solutions to develop a functional lower bound for the continuous problem from a dual solution to the discrete problem. Refining the discretization (e.g., selecting states uniformly from the continuous state space) proves a functional strong duality for the continuous problem, thus proving Theorem 3.1.

The second result of this paper is developing a symbolic pivoting methodology that makes it possible to solve the LP without calculating its coefficients. We observe that this LP is so poorly scaled that standard algorithms and solvers fail to solve it. In §6, we demonstrate the equivalence between strongly feasible bases and deterministic unichain policies, and combinatorialize the pivoting process. We exploit the relationship between dual solutions and relative value functions to develop a mechanism for solving the LP without ever computing its coefficients. This symbolic pivoting algorithm (Lemma 6.4) allows us to solve the LP despite the poor scaling.

Third, we develop a scheme analogous to column generation for refining the discretization so as to drive the gap between the discrete approximation and the continuous problem to zero quickly while keeping the LP small. In Lemma 7.3 we describe how to add new states to the discretization for this purpose.

Finally, we provide a computational study that demonstrates that the developed "column generation" approach is much faster than solving a regular discretization of the problem.

4. A controlled random walk approximation and its LP formulation. We define a controlled random walk on the lattice $\mathcal{S} \times \Lambda$ that approximates the controlled diffusion process. We often refer to a state (s_j, u) simply by (j, u). Upon entering a state (j, u), the controller may instantaneously move the process to a new state (j, v) with $v \in \Lambda$ at the cost $K(u, v)$. If the controller does not change the drift rate, the system wanders from state to state, changing from state (j, u), $1 < j < n$ to state $(j + 1, u)$ with probability

$$p_+(j,u) = \frac{e^{\frac{2u}{\sigma^2}(s_{j-1}-s_j)} - 1}{e^{\frac{2u}{\sigma^2}(s_{j-1}-s_{j+1})} - 1} = \frac{e^{-\frac{2u}{\sigma^2}s_j} - e^{-\frac{2u}{\sigma^2}s_{j-1}}}{e^{-\frac{2u}{\sigma^2}s_{j+1}} - e^{-\frac{2u}{\sigma^2}s_{j-1}}},$$

and to state $(j-1,u)$ with probability $p_-(j,u) = 1-p_+(j,u)$. If the controller does not change the drift rate, the process moves from state $(1,u)$ to state $(2,u)$ and from state (n,u) to state $(n-1,u)$ with probability 1.

We let $\mathrm{ET}[j,u]$ denote the expected sojourn time in state (j,u) if the policy does not change the drift rate, where $\mathrm{ET}[j,u]$ is defined through the solution to the basic adjoint equations provided as in Lemma 1 of Ormeci Matoglu and Vande Vate (2011). We provide this Lemma with our notation and the explicit expression $\mathrm{ET}[j,u]$ in the Appendix. To see the relationship to the diffusion process, observe that $p_+(j,u)$ is the probability that starting at s_j, $X(t)$ reaches s_{j+1} before reaching s_{j-1}, and $\mathrm{ET}[j,u]$ is the expected time before $X(t)$, starting at s_j hits $\{s_{j-1}, s_{j+1}\}$ when the drift rate is u. The term $p_-(j,u) = 1 - p_+(j,u)$ is the probability that starting at s_j, $X(t)$ reaches s_{j-1} before reaching s_{j+1}.

To reflect the costs of maintaining the drift rates, each time the process drifts out of state (j,u) the controller is charged $c(u)\mathrm{ET}[j,u]$. To reflect the costs of instantaneous controls required to keep the process in $[0,\Theta]$, each time the process drifts out of state $(1,u)$ the controller is charged $U\mathrm{EA}[u]\mathrm{ET}[1,u]$, and each time the process drifts out of state (n,u) the controller is charged $M\mathrm{ER}[u]\mathrm{ET}[n,u]$. To reflect the costs of holding inventory, we also charge the controller $h\mathrm{ET}[j,u]\mathrm{EX}[j,u]$ each time the system drifts out of state (j,u). $\mathrm{EA}[u]$, $\mathrm{ER}[u]$ and $\mathrm{EX}[j,u]$ are defined as in Lemma 1 of Ormeci Matoglu and Vande Vate (2011), and the explicit expressions are provided in the Appendix.

Next we formulate a linear program that finds an optimal policy for the controlled random walk defined earlier in this section. For each $j = 1,2,\ldots,n$ and $u,v \in \Lambda$, we let $y(j,u,v)$ denote the rate or transitions per unit time at which the controller changes the drift rate from u to v when the random walk is in state (j,u). If the controller does not change to a new drift rate in state (j,u), the system transitions to state $(j-1,u)$ (when $j>1$) or $(j+1,u)$, (when $j<n$). In this case we let $r(j,u)$ denote the rate at which the system transitions out of state (j,u).

We obtain the following constraints on y and r. First, each arrival to state (j,u) must be accompanied by a departure from the state. The rate of departures from state (j,u) is simply $r(j,u) + \sum_{v\in\Lambda} y(j,u,v)$, namely the rate at which the process transitions out of state (j,u) plus the rate at which the controller changes the drift rate from state (j,u). The rate of arrivals to state (j,u), on the other hand, is simply

$$p_+(j-1,u)r(j-1,u) + p_-(j+1,u)r(j+1,u) + \sum_{v\in\Lambda} y(j,v,u),$$

the sum of the rates at which the system transitions into state (j,u) from

neighboring states and the rate at which the the controller moves the system to state (j,u) by changing the drift rates. Combining these expressions we obtain the *Flow Conservation Constraint* for each $j = 1, 2, \ldots, n$ and $u \in \Lambda$:

$$\begin{aligned} r(j,u) - p_+(j-1,u)r(j-1,u) - p_-(j+1,u)r(j+1,u) \\ + \sum_{v\in\Lambda}(y(j,u,v) - y(j,v,u)) = 0, \end{aligned} \tag{4.1}$$

where we adopt the convention that $p_+(0,u) = p_-(n+1,u) = 0$ and $p_+(1,u) = p_-(n,u) = 1$. It is convenient to scale the solutions so that

$$\sum_{u\in\Lambda}\sum_{j=1}^{n} r(j,u)\mathrm{ET}[j,u] = 1. \tag{4.2}$$

In this way $r(j,u)\mathrm{ET}[j,u]$ is the fraction of time the system spends in state (j,u).

We next present the linear program formulation, $\mathrm{S}LP(\Lambda,\mathcal{S})$, of the controlled random walk approximation of the $\mathcal{S}$-restricted problem:

$$\begin{aligned} &\min \sum_{u\in\Lambda}\sum_{j=1}^{n} r(j,u)\left(\mathrm{ET}[j,u]\mathrm{EX}[j,u]h + \mathrm{ET}[j,u]c(u)\right) \\ &+ \sum_{u\in\Lambda}\left(U\mathrm{EA}[u]\mathrm{ET}[1,u]r(1,u) + M\mathrm{ER}[u]\mathrm{ET}[n,u]r(n,u)\right) \\ &\qquad + \sum_{u,v\in\Lambda}\sum_{j=1}^{n} K(u,v)y(j,u,v) \end{aligned} \tag{4.3}$$

s.t. Flow Conservation Constraint for each $j = 1, 2, \ldots, n$ and $u \in \Lambda$:

$$\begin{aligned} r(j,u) - p_+(j-1,u)r(j-1,u) - p_-(j+1,u)r(j+1,u) \\ + \sum_{v\in\Lambda}(y(j,u,v) - y(j,v,u)) = 0 \end{aligned} \tag{4.4}$$

s.t. Scale Constraint:

$$\sum_{u\in\Lambda}\sum_{j=1}^{n} r(j,u)\mathrm{ET}[j,u] = 1. \tag{4.5}$$

(4.6) non-negativity: $r, y \geq 0$

Then using the dual variables $\alpha(j,u)$ for $j=1,\ldots,n$ and $u\in\Lambda$ for flow conservation constraints and γ for the scale constraint the dual of this problem is

$$\max \gamma \tag{4.7}$$

s.t. Dual Constraint for $r(1,u)$ for each $u\in\Lambda$:

$$\begin{aligned}&\alpha(1,u)-p_+(1,u)\alpha(2,u)+\gamma\mathrm{ET}[1,u]\\&\le h\mathrm{EX}[1,u]\mathrm{ET}[1,u]+c(u)\mathrm{ET}[1,u]+U\mathrm{EA}[u]\mathrm{ET}[1,u]\end{aligned} \tag{4.8}$$

s.t. Dual Constraint for $r(j,u)$ for each $j=2,3,\ldots,n-1$ and $u\in\Lambda$:

$$\begin{aligned}&\gamma\mathrm{ET}[j,u]+\alpha(j,u)-p_+(j,u)\alpha(j+1,u)-p_-(j,u)\alpha(j-1,u)\\&\le h\mathrm{ET}[j,u]\mathrm{EX}[j,u]+c(u)\mathrm{ET}[j,u]\end{aligned} \tag{4.9}$$

s.t. Dual Constraint for $r(n,u)$ for $u\in\Lambda$:

$$\begin{aligned}&\alpha(n,u)-p_-(n,u)\alpha(n-1,u)+\gamma\mathrm{ET}[n,u]\\&\le h\mathrm{ET}[n,u]\mathrm{EX}[n,u]+c(u)\mathrm{ET}[n,u]+M\mathrm{ER}[u]\mathrm{ET}[n,u]\end{aligned} \tag{4.10}$$

s.t. Dual Constraint for $y(j,u,v)$ for each $j=1,2,\ldots,n$ and $u\in\Lambda$:

$$\alpha(j,u)-\alpha(j,v)\le K(u,v) \tag{4.11}$$

Our primal model is similar to the semi-Markov decision process (SMDP) formulation of Tijms (2003) where the variables are a scaled version of the traditional MDP formulation. In the SMDP model the variables represent the number of visits to each state per unit time rather than the probabilities of being in each state. Just as in the original formulation of Ormeci Matoglu and Vande Vate (2011) each variable has exactly one positive coefficient in the Flow Conservation Constraints of the primal formulation $\mathrm{S}LP(\Lambda,\mathcal{S})$ and the coefficients for each variable on these constraints sum to 0. Thus, the Flow Conservation Constraints form a pre-Leontief system (see Veinott (1968)) and so, together with the Scale Constraint (4.5), $\mathrm{S}LP(\Lambda,\mathcal{S})$ is a scaled, pre-Leontief linear program like the original formulation of Ormeci Matoglu and Vande Vate (2011).

In a linear program with m linearly independent constraints, a set of m linearly independent columns is called a basis and the variables associated with those columns are referred to as the corresponding basic variables, any remaining variables are referred to as the non-basic variables. Each basis uniquely defines a basic solution that satisfies the constraints with each non-basic variable fixed at a boundary value (either its lower bound or its

upper bound if it has one). If the resulting values for the basic variables also satisfy any bounds on their values the basis is called feasible and the solution is called a basic feasible solution. In our primal linear program, each variable is constrained to be non-negative and so a basic solution is feasible if all the basic variables are non-negative. In the dual of this linear program, the dual variables are unconstrained – only the slack variables representing the reduced costs of the primal variables are constrained to be non-negative. A variable is said to represent the Flow Conservation constraint on which it has its positive coefficient. We say that a Flow Conservation constraint is active under a basic feasible solution of $\mathrm{S}LP(\Lambda, \mathcal{S})$ if it is represented by a positive basic variable. To ensure feasibility, it is clear that each positive basic variable must have all its non-zero coefficients on active rows. Thus, Lemma 2 of Ormeci Matoglu and Vande Vate (2011) reduces to:

LEMMA 4.1. *Each basic feasible solution* (r^*, y^*) *of* $\mathrm{S}LP(\Lambda, \mathcal{S})$ *has exactly one positive basic variable representing each active row.*

We exploit this property in §6 to speed and simplify the pivots.

In §5 we use the dual of our LP to derive a relative value function that satisfies the lower bound conditions. We show that nonnegative reduced costs of the $y(j, u, v)$ and $r(j, u)$ variables imply two of the lower bound conditions are satisfied. Furthermore, our LP model conforms naturally with the discretization of the policy space, which in §7 will be instrumental to identifying attractive points to add to the discretization. Thus, instead of a naive uniformly spaced discretization in §7 we propose a formal column generation approach that provides both a series of increasingly good policies and a series of lower bounds and guarantees that the average costs of the policies and the values of the lower bounds converge to the average cost of an optimal policy.

5. Relative value functions. In this section we exploit the strong relation between complementary dual solutions and relative value functions established in the Technical Lemma (Lemma 3) of Ormeci Matoglu and Vande Vate (2011). We restate a version of that result here as it applies to the dual problem (4.7)–(4.11). Given a control band $\Phi = (u, s, S)$ and a starting a point $z \in (s, S)$, let $\mathrm{ET}[z, \Phi]$ denote the expected time for the process $X(t)$ to hit s or S when it is in control band Φ starting at a point z and maintaining drift rate u and let $\mathrm{EX}[z, \Phi]$ denote the average value of $X(t)$ over this time. Let $p(z, \Phi, S)$ denote the probability that starting at z and maintaining the drift rate u, $X(t)$ first reaches S. Then $p(z, \Phi, s) = 1 - p(z, \Phi, S)$. Note that the notation we use for sojourn times, expected value of X and transition probabilities in $SLP(\Lambda, \mathcal{S})$ is a special case of the

above notation where a control band $\Phi = (u, s, S)$ always follows the form (u, s_{j-1}, s_{j+1}) so that $\mathrm{ET}[j, u]$ is equal to $\mathrm{ET}[z, \Phi]$ where $s_j = z$, $s_{j-1} = s$ and $s_{j+1} = S$. Similarly $p_+(z, u) = p(z, \Phi, S)$ and $p_-(z, u) = p(z, \Phi, s)$.

LEMMA 5.1 (Technical Lemma of Ormeci Matoglu and Vande Vate (2011)). *Consider a control band $\Phi = (u, s, S)$. A continuous function $f : [s, S] \to \mathbb{R}$ satisfies*

$$\begin{aligned}&\mathrm{ET}[z, \Phi]\gamma + f(z) - p(z, \Phi, S)f(S) - p(z, \Phi, s)f(s)\\ &= (hEX[z, \Phi] + c(u) + UEA[u]1_{\{s=0\}} + MER[u]1_{\{S=\Theta\}})\mathrm{ET}[z, \Phi]\end{aligned}$$

where

$$p(z, \Phi, S) = \frac{e^{-\frac{2u}{\sigma^2}z} - e^{-\frac{2u}{\sigma^2}s}}{e^{-\frac{2u}{\sigma^2}S} - e^{-\frac{2u}{\sigma^2}s}} = 1 - p(z, \Phi, s)$$

if and only if

$$f(z) = -\frac{h}{2u}z^2 + \left(\frac{\gamma}{u} + \frac{h\sigma^2}{2u^2} - \frac{c(u)}{u}\right)z + Ce^{-\frac{2u}{\sigma^2}z} + F$$

for some constants C and F.

One immediate consequence of Lemma 5.1 is that we can extend a dual solution (γ, α) from the points of $\mathcal{S}$ to the entire range $[0, \Theta]$ in a natural way. In fact, we exploit the strong relationship between complementary dual solutions and relative value functions to solve the $\mathrm{SLP}(\Lambda, \mathcal{S})$ without ever computing the values of $p_+(j, u)$, $p_-(j, u)$, $\mathrm{ET}[j, u]$, $\mathrm{EA}[u]$, $\mathrm{ER}[u]$ or $\mathrm{EX}[j, u]$. This relationship allows us to rely on the relative value functions to determine the entering variables (i.e. a nonbasic variable with negative reduced cost, so that it's inclusion in the basis results with same or lower average cost) when solving $SLP(\Lambda, \mathcal{S})$ and to identify points to add to $\mathcal{S}$ so as to reduce the gap between the optimal solution to the $\mathcal{S}$-restricted problem and the lower bound for the unrestricted problem.

In a manner similar to Ormeci Matoglu and Vande Vate (2011), given an optimal solution (γ, α) to the dual problem (4.7)–(4.11), we construct continuous functions $f(\cdot, u)$ such that (γ, f) satisfies all but condition (2.3) of the lower bound conditions.

LEMMA 5.2. *Given an optimal solution (γ, α) to the problem (4.7)–(4.11), for each $u \in \Lambda$ define*

$$f(a, u) = C(j, u)e^{-\frac{2u}{\sigma^2}a} + F(j, u) - \frac{h}{2u}a^2 + \left(\frac{\gamma}{u} + \frac{h\sigma^2}{2u^2} - \frac{c(u)}{u}\right)a \tag{5.1}$$

for $s_j \le a \le s_{j+1}$, where $C(j,u)$ and $F(j,u)$ are constants satisfying:

$$C(0,u) = \frac{\sigma^2}{2u}\left(U + \frac{\gamma}{u} + \frac{h\sigma^2}{2u^2} - \frac{c(u)}{u}\right) \tag{5.2}$$

$$C(n,u) = \frac{\sigma^2}{2u}e^{\frac{2u}{\sigma^2}\Theta}\left(-M - \frac{h}{u}\Theta + \frac{\gamma}{u} + \frac{h\sigma^2}{2u^2} - \frac{c(u)}{u}\right) \tag{5.3}$$

and for $j = 1, 2, 3, \ldots, n$

$$-\frac{h}{2u}s_j^2 + \left(\frac{\gamma}{u} + \frac{h\sigma^2}{2u^2} - \frac{c(u)}{u}\right)s_j + C(j-1,u)e^{-\frac{2u}{\sigma^2}s_j} + F(j-1,u) = \alpha(j,u) \tag{5.4}$$

and

$$-\frac{h}{2u}s_j^2 + \left(\frac{\gamma}{u} + \frac{h\sigma^2}{2u^2} - \frac{c(u)}{u}\right)s_j + C(j,u)e^{-\frac{2u}{\sigma^2}s_j} + F(j,u) = \alpha(j,u). \tag{5.5}$$

Then (γ, f) satisfies all but condition (2.3) of the lower bound conditions.

PROOF. Since the system (5.2)–(5.5) is non-singular, it uniquely defines C and F. The fact that $C(0,u)$ satisfies (5.2) and $C(n,u)$ satisfies (5.3) ensures that $f(\cdot,u)$ satisfies (2.4) and (2.5) with equality.

It is easy to see that for each $u \in \Lambda$, $f(a,u)$ satisfies (2.2) with equality for each $a \in (0,\Theta) \setminus \mathcal{S}$. Next we define

$$f_i(a,u) = C(i,u)e^{-\frac{2u}{\sigma^2}a} + F(i,u) - \frac{h}{2u}a^2 + \left(\frac{\gamma}{u} + \frac{h\sigma^2}{2u^2} - \frac{c(u)}{u}\right)a.$$

Since α satisfies (4.9) substituting (5.4) and (5.5) we obtain

$$\begin{aligned}&\mathrm{ET}[j,u]\gamma + f_j(j,u) - p_+(j,u)f_{i-1}(s_{j+1},u) - p_-(j,u)f(s_{j-1},u)\\ &= (h\mathrm{EX}[j,u] + c(u))\mathrm{ET}[j,u].\end{aligned}$$

and by the Technical Lemma we have

$$\begin{aligned}&\mathrm{ET}[j,u]\gamma + f_{j-1}(s_j,u) - p_+(j,u)f_{j-1}(s_{j+1},u) - p_-(j,u)f_{j-1}(s_{j-1},u)\\ &= (h\mathrm{EX}[j,u] + c(u))\mathrm{ET}[j,u].\end{aligned}$$

Subtracting these expressions we obtain

$$p_+(j,u)\left(f_{j-1}(s_{j+1},u) - f_j(s_{j+1},u)\right) \tag{5.6}$$

$$= p_+(j,u)\left([C(j-1,u)-C(j,u)]e^{-\frac{2u}{\sigma^2}s_{j+1}}+F(j-1,u)-F(j,u)\right)\le 0.$$

By (5.4) and (5.5)

$$C(j,u)e^{-\frac{2u}{\sigma^2}s_j}+F(j,u)=C(j-1,u)e^{-\frac{2u}{\sigma^2}s_j}+F(j-1,u)$$

which yields $F(j-1,u)-F(j,u)=(C(j,u)-C(j-1,u))\,e^{-\frac{2u}{\sigma^2}s_j}$. Substituting this in (5.6) we obtain that (4.9) is equivalent to

$$p_+(j,u)\,[C(j,u)-C(j-1,u)]\left(e^{-\frac{2u}{\sigma^2}s_j}-e^{-\frac{2u}{\sigma^2}s_{j+1}}\right)\le 0.$$

Since $(e^{-\frac{2u}{\sigma^2}s_j}-e^{-\frac{2u}{\sigma^2}s_{j+1}})$ has the same sign as $\frac{2u}{\sigma^2}e^{-\frac{2u}{\sigma^2}s_j}$, this implies

$$(C(j,u)-C(j-1,u))\,\frac{2u}{\sigma^2}e^{-\frac{2u}{\sigma^2}s_j}\le 0,\qquad \text{and so}$$

$$\rho(s_j,u)=f'(s_j+,u)-f'(s_j-,u)=-\frac{2u}{\sigma^2}e^{-\frac{2u}{\sigma^2}s_j}\,(C(j,u)-C(j-1,u))\ge 0$$

proving that $f(\cdot,u)$ satisfies (2.6) for $a\in(s_1,s_n)$.

Since α satisfies (4.8) ((4.10)), one can similarly show that $\rho(a,u)\ge 0$ for $a\in(0,s_1]$ ($\rho(a,u)\ge 0$ for $a\in[s_n,\Theta)$).

Note that (5.4)–(5.5) ensure that for each $u\in\Lambda$, the function $f(\cdot,u)$ is continuous. It remains to show that for each $u\in\Lambda$, $f(\cdot,u)$ can be written as the difference of two convex functions. This argument is identical to the one used in the proof of Theorem 2 of Ormeci Matoglu and Vande Vate (2011) (see Ormeci Matoglu and Vande Vate (2010) for details). □

An interpretation of the relative value function f for a policy Φ is that it represents the total expected cost so that $f(x,u)-f(x,v)$ represents the difference in total expected costs over an infinitely long period of time by starting with drift rate u rather than with drift rate v. Lemma 5.1 stipulates the form of the relative value function. Equations (5.2) and (5.3) ensure the choice of $C(0,u)$ and $C(n,u)$ satisfy (2.4) and (2.5) with equality at the boundaries 0 and Θ, f' at the boundaries is equal to the rate cost is incurred at the boundaries as instantaneous boundary control is exerted. Equations (5.4) and (5.5) ensure that the relative value functions are continuous and coincide with α at the points of $\mathcal{S}$.

Observation 1. Note that if $r(j,u)$ is basic in an optimal solution to $SLP(\Lambda,\mathcal{S})$, then (γ,α) satisfies the corresponding dual constraint (4.9) with equality and so $C(j,u)=C(j-1,u)$ and hence $F(j,u)=F(j-1,u)$.

Similarly, if $y(j,u,v)$ is basic, then $f(s_j,u) = f(s_j,v) + K(u,v)$. In fact, we can use the relative value functions f to determine reduced costs: The reduced cost of $y(j,u,v)$ is simply

$$K(u,v) - \alpha(j,u) + \alpha(j,v) = K(u,v) + f(s_j,v) - f(s_j,u)$$

and the reduced cost of $r(j,u)$ is

$$\begin{aligned} & h\mathrm{ET}[j,u]\mathrm{EX}[j,u] + c(u)\mathrm{ET}[j,u] - (\gamma \mathrm{ET}[j,u] + \alpha(j,u) \\ & \quad - p_+(j,u)\alpha(j+1,u) - p_-(j,u)\alpha(j-1,u)) \\ = \;& p_+(j,u)\left(e^{-\frac{2u}{\sigma^2}s_{j+1}} - e^{-\frac{2u}{\sigma^2}s_j}\right)(C(j,u) - C(j-1,u)) \end{aligned}$$

which has the same sign as $\rho(s_j,u)$.

Observation 2. Observe that since α satisfies (4.11), f satisfies (2.3) for each $a \in \mathcal{S}$ and $u,v \in \Lambda$. However, f need not satisfy (2.3) between the points of $\mathcal{S}$. In fact, the Technical Lemma (Lemma 5.1) allows us to extend the dual solution α beyond the points in $\mathcal{S}$ and hence extend Observation 1 to describe the reduced cost of new primal variables $y(a,u,v)$ and $r(a,u)$ for each point $a \in (0,\Theta) \setminus \mathcal{S}$. In particular, the reduced cost of $y(a,u,v)$ is simply $K(u,v) + f(a,v) - f(a,u)$ and the reduced cost of $r(a,u)$ has the same sign as $\rho(a,u) = f'(a+,u) - f'(a-,u)$.

Observation 3. Observation 1 suggests the fundamental insight that knowing only which primal variables are basic, i.e., the structure of the policy, we can compute the average cost γ and the relative value functions f and so the (signs of) the reduced costs directly without computing the dual solution α. In particular, we can replace the conditions (5.4)–(5.5) with the conditions:

$$C(j-1,u)e^{-\frac{2u}{\sigma^2}s_j} + F(j-1,u) = C(j,u)e^{-\frac{2u}{\sigma^2}s_j} + F(j,u) \tag{5.7}$$

for each $j = 1,2,\ldots,n$ ensuring $f(\cdot,u)$ is continuous for each $u \in \Lambda$,

$$C(j-1,u) = C(j,u) \text{ and } F(j-1,u) = F(j,u) \text{ when } r(j,u) \text{ is basic} \tag{5.8}$$

ensuring the dual solution satisfies the complementary slackness conditions with respect to r,

$$f(s_j,u) = f(s_j,v) + K(u,v) \text{ when } y(j,u,v) \text{ is basic} \tag{5.9}$$

ensuring the dual solution satisfies the complementary slackness conditions with respect to y and $F(0,u^*)=0$ *for some arbitrary distinguished drift rate* u^* since adding a constant to each F has no impact on the constraints (5.7)–(5.9). The unique solution to this system gives γ, the average cost of the policy, and the relative value functions f, which allow us to compute reduced costs. In §6 we exploit this observation to develop a solution procedure that does not require the values of parameters such as $p_+(j,u)$, $\mathrm{ET}[j,u]$, $\mathrm{EX}[j,u]$, $\mathrm{EA}[u]$ and $\mathrm{ER}[u]$.

Before continuing with the proof of Theorem 3.1 we prove Lemma 5.3, a restatement Proposition 1 of Ormeci Matoglu and Vande Vate (2011) to facilitate our proof of Theorem 3.1.

LEMMA 5.3. *Suppose that for each* $u\in\Lambda$, $f(\cdot,u):[0,\Theta]\to\mathbb{R}$ *satisfies the smoothness conditions. Further, suppose that for each* $u\in\Lambda$ *the function* $f(\cdot,u)$ *satisfies* (2.2) *for some scalar* γ_ϵ, *and* (2.4)–(2.6),

$$f(a,v)-f(a,u)\geq -K(u,v)+\epsilon \text{ for } a\in[0,\Theta] \text{ and } v\in\Lambda$$

for some scalar ϵ. *Then*

$$\gamma_\epsilon+\lim_{T\to\infty}\frac{N(T)}{T}\epsilon\leq \mathrm{AC}(\Phi) \tag{5.10}$$

for each policy $\Phi\in\mathcal{P}$ *and each initial state* (a,u).

PROOF. The proof is analogous to that of Proposition (2.1). Consider a policy $\Phi=\{(T_i,u_i),i\geq 0\}\in\mathcal{P}$. Then by a simple adaptation of the Meyer–Ito formula (see for example Theorem 70 of Protter (2004)) we have

$$\begin{aligned}
&\mathbb{E}[f(X(T),\mu(T))]-f(X(0),\mu(0))\\
&=\mathbb{E}\Bigg[\int_0^T\left(\frac{\sigma^2}{2}f''(X(t),\mu(t))+\mu(t)f'(X(t),\mu(t))\right)dt\\
&+\int_0^T f'(X(t),\mu(t))dA(t)-\int_0^T f'(X(t),\mu(t))dR(t)\\
&+\sum_{i=1}^{N(T)}\left(f\left(X\left(T_i\right),u_i\right)-f\left(X\left(T_i-\right),u_{i-1}\right)\right.\\
&+\frac{1}{2}\sum_{u\in\Lambda}\int_0^\Theta L(T,a,u)\rho(da,u)\Bigg],
\end{aligned}$$

where L is the local time. Using (2.2), (2.4)–(2.6) and (5.10) we see that

(5.11)

$$\begin{aligned}&\mathbb{E}[f(X(T),\mu(T))]-f(X(0),\mu(0))\\&\geq \mathbb{E}\bigg[\int_0^T \gamma_\epsilon - c(\mu(t)) - hX(t)dt - UA(t) - MR(t) - \sum_{i=1}^{N(T)}(K(u_{i-1},u_i)-\epsilon)\bigg].\end{aligned}$$

Dividing both sides by T and taking the limit yields

$$\begin{aligned}&\limsup_{T\to\infty}\frac{1}{T}\mathbb{E}[f(X(T),\mu(T))-f(X(0),\mu(0))]\\ \geq\ & \gamma_\epsilon - \limsup_{T\to\infty}\frac{1}{T}\mathbb{E}\bigg[\int_0^T c(\mu(t))+hX(t)dt+UA(t)+MR(t)\\ &+\sum_{i=1}^{N(T)}(K(u_{i-1},u_i)-\epsilon)\bigg]\\ =\ & \gamma_\epsilon + \lim_{T\to\infty}(N(T)/T)\epsilon - \mathrm{AC}(\Phi).\end{aligned}$$

Since $f(X(t),\mu(t))$ is bounded

$$\limsup_{T\to\infty}\frac{\mathbb{E}[f(X(T),\mu(T))]-f(X(0),\mu(0))}{T}=0$$

and so,

$$\mathrm{AC}(\Phi)\geq \gamma_\epsilon + \limsup_{T\to\infty}(N(T)/T)\epsilon. \tag{5.12}$$

□

COROLLARY 5.1. *Let $(\gamma_\epsilon,\alpha_\epsilon)$ be an optimal solution to (4.7)–(4.11) where $K(u,v)$ is replaced with $K(u,v)-\epsilon$ in (4.11) and let Φ_ϵ be the corresponding policy. Then*

$$\mathrm{AC}(\Phi_\epsilon)=\gamma_\epsilon+\lim_{T\to\infty}\frac{N(T)}{T}\epsilon=\gamma_\epsilon+\sum_{s\in\mathcal{S},u,v\in\Lambda}y(s,u,v)\epsilon,$$

where $y(s,u,v)$ is the rate of transition under policy Φ_ϵ.

PROOF. Construct $(\gamma_\epsilon,f_\epsilon)$ from $(\gamma_\epsilon,\alpha_\epsilon)$ as defined in (5.1)–(5.5). By Observation 1, (5.12) are satisfied with equality in this case. □

Note that a pair (γ, f) obtained from a dual solution (γ, α) will satisfy (2.3) at the points in $\mathcal{S}$ but not necessarily for the whole range $[0, \Theta]$. The extent to which it violates the constraint (i.e. ϵ) can be controlled by the level of discretization. It can be made arbitrarily small by choosing a fine enough discretization. The proof of Theorem 3.1 relies on this.

Since $\mathrm{EX}[j,u]$, $\mathrm{EA}[u]$, $\mathrm{ER}[u]$ and $c(u)$ are all bounded, we can derive a lower bound $\underline{AC}$ on the average cost of any policy ignoring the transition costs. And since $K^* = \min\{K(u,v), u \neq v\} > 0$, we can use the the average cost of any policy in $\Phi \in \mathcal{P}(\mathcal{S})$ together with the lower bound $\underline{AC}$ to bound $\lim_{T\to\infty} \frac{N(T)}{T} = \sum_{s\in\mathcal{S},u,v\in\Lambda} y(s,u,v)$. In particular, an optimal policy $\Phi^* \in \mathcal{P}(\mathcal{S})$ must satisfy $\underline{AC} + K^* \lim_{T\to\infty} N(T)/T \leq \mathrm{AC}(\Phi^*) \leq \mathrm{AC}(\Phi)$. So $\lim_{T\to\infty} N(T)/T \leq (\mathrm{AC}(\Phi) - \underline{AC})/K^* \equiv H$.

We are now ready to prove Theorem 3.1. Our approach is to construct, for each $\epsilon > 0$ sufficiently small, a non-overlapping control band policy $\Phi_\epsilon \in \mathcal{P}$ and a pair $(\gamma_\epsilon, f_\epsilon)$ satisfying the lower bound conditions, such that $\gamma_\epsilon \leq \mathrm{AC}(\Phi_\epsilon) \leq \gamma_\epsilon + \epsilon H$. Taking $\epsilon \to 0$ proves the result.

Proof of Theorem 3.1. Let d be the supremum of γ over all pairs (γ, f) satisfying the lower bound conditions and let p be the infimum of $\mathrm{AC}(\Phi)$ over all policies $\Phi \in \mathcal{P}$. Then $\gamma_\epsilon \leq d \leq p \leq \mathrm{AC}(\Phi_\epsilon) \leq \gamma_\epsilon + \epsilon H$, where the first inequality follows because $(\gamma_\epsilon, f_\epsilon)$ satisfies the lower bound conditions, the second inequality follows from weak duality, i.e. Proposition 2.1, and the third inequality follows from the fact that $\Phi_\epsilon \in \mathcal{P}$. Taking $\epsilon \to 0$ proves $p = d$.

It remains to construct a pair $(\gamma_\epsilon, f_\epsilon)$ satisfying the lower bound conditions and a policy Φ_ϵ satisfying $\gamma_\epsilon \leq AC(\Phi_\epsilon) \leq \gamma_\epsilon + \epsilon H$. By Theorem 3 of Ormeci Matoglu and Vande Vate (2011) if we choose $\mathcal{S}$ so that $\delta \geq s_{i+1} - s_i$, where $(\max\{\overline{f}'(u), 0\} - \min\{\underline{f}'(v), 0\})\delta \leq \epsilon$ for each pair of rates u and v in Λ and

$$\overline{f}'(u) = \begin{cases} e^{\frac{2u}{\sigma^2}\Theta}\left(\frac{h\Theta}{u} + M + U\right) - \frac{c(u)}{u} & \text{if } u > 0 \\ -\frac{h\Theta}{u} + \frac{h\sigma^2}{2u^2} + M & \text{if } u < 0 \end{cases}$$

$$\underline{f}'(u) = \begin{cases} -\frac{h\Theta}{u} - U & \text{if } u > 0 \\ (\frac{h\sigma^2}{2u^2} - \frac{c(u)}{u} + U)(1 - e^{-\frac{2u}{\sigma^2}\Theta}) - U & \text{if } u < 0 \end{cases}$$

and solve the LP with $K(u,v)$ replaced by $K(u,v) - \epsilon$, then the relative value function f_ϵ constructed as in (5.1)–(5.5) from an optimal solution $(\gamma_\epsilon, \alpha)$ to the dual satisfies (2.2)–(2.6). Note that, f_ϵ satisfies (5.10) for $a \in \mathcal{S}$, by the choice of discretization we ensure f_ϵ satisfies (2.3) for all $a \in [0, \Theta]$. A policy Φ_ϵ corresponding to an optimal solution to the primal is in $\mathcal{P}(\mathcal{S}) \subseteq \mathcal{P}$ and satisfies $\gamma_\epsilon \leq \mathrm{AC}(\Phi_\epsilon) \leq \gamma_\epsilon + \epsilon H$ by Lemma 5.3 and Corollary 5.1. □

Observe that in the proof of Theorem 3.1, for each $\epsilon > 0$, Φ_ϵ is a non-overlapping control band policy. Hence, we can restrict $\mathcal{P}$ to only the non-overlapping control band policies proving Corollary 3.1.

6. Strongly feasible bases and deterministic unichain policies. $SLP(\Lambda, \mathcal{S})$ is so poorly scaled that completing pivots is algebraically difficult and beyond the capability of standard solvers. Furthermore, as we drive the points of $\mathcal{S}$ close together, the problem grows quite large. To help resolve this issue, we follow a classic path forged by Cunningham (1976), Orlin (1985) and others who applied the notion of strongly feasible bases to network flow problems, a special case of pre-Leontief systems, but with the added complication of bounds on the variables. Strongly feasible bases extend the unique representation property of Lemma 4.1 to inactive rows and provide a combinatorial characterization of feasible bases, i.e. rather than relying on algebraic computations involving the coefficients of the constraint matrix we can simply check the connectedness properties of a simple directed graph to determine linear independence. This approach has significant computational advantages and in fact confines all the computational challenges to the problem of computing reduced costs. In this section we define strongly feasible bases, prove lemmas about their properties and characterize them in a number of increasingly combinatorial ways in Theorem 6.1 and Lemma 6.4. The primal pivoting procedure in Lemma 6.4 together with the method for computing γ and the relative value functions f, in Observation 3, allow us to solve SLP$(\Lambda, \mathcal{S})$ without computing the parameters $p_+(j,u)$, ET$[j,u]$, EX$[j,u]$, EA$[u]$ and ER$[u]$.

Our formulation S$LP(\Lambda, \mathcal{S})$ is a scaled pre-Leontief linear program of the form:

$$\text{Minimize } cx \tag{6.1}$$

subject to the constraints

$$Ax = 0 \tag{6.2}$$

$$wx = 1 \tag{6.3}$$

$$x \geq 0 \tag{6.4}$$

where c and w are row vectors corresponding to the coefficients of (4.3) and (4.5), and $A = (a_{ij})$ is an $m \times n$ pre-Leontief matrix representing the Flow Conservation Constraints (4.4) and satisfying:

Assumption 1: Each column has a unique positive entry and that entry is $+1$,

Assumption 2: The coefficients of each column sum to 0, i.e., $\mathbf{1}A = 0$, where $\mathbf{1}$ is the row vector of all ones

The fact in Assumption 1 that the unique positive entry in each column is +1 only simplifies the arguments slightly. All the results of this section are true if the unique positive entry on each column is at most 1. Note that Assumption 2 ensures that the constraints $Ax = 0$ are linearly dependent and so we may eliminate any one of these constraints, henceforth called the *root*, without affecting the feasible solutions.

6.1. *Strongly feasible bases.* We say that a feasible basis B for (6.2)–(6.4) is *strongly feasible* if for some row i^*, row i^* is represented by a column of B and for each row $i = 1, 2, \ldots m$, B is a feasible basis for

$$\begin{bmatrix} A \\ w \end{bmatrix} x = e_{m+1} + \epsilon(e_i - e_{i^*}) \qquad x \geq 0 \tag{6.5}$$

for some $\epsilon > 0$, where e_i represents the unit column vector with its one in row i.

LEMMA 6.1. *Each strongly feasible basis for* (6.2)–(6.4) *has exactly one basic variable representing each row of A.*

PROOF. Consider a strongly feasible basis B for the system (6.2)–(6.4). By definition B includes a variable representing the row i^*. Clearly, to be a feasible basis for (6.5) for each $i = 1, 2, \ldots, m$, $i \neq i^*$, B must include at least one variable representing each of the other rows of A. Since A includes a linearly dependent row, a basis can have at most one variable for each row of A and so a strongly feasible basis must have exactly one variable representing each row of A. □

We say that the m-vector b is a *perturbation* if $b_i > 0$ for $i = 1, 2, \ldots, m$, $i \neq i^*$, for some index $1 \leq i^* \leq m$ and $\sum_{i=1}^{m} b_i = 0$. We call i^* the *root* of the perturbation b.

LEMMA 6.2. *A feasible basis B for* (6.2)–(6.4) *is strongly feasible if and only if B is feasible for*

$$\begin{bmatrix} A \\ w \end{bmatrix} x = \begin{pmatrix} b \\ 1 \end{pmatrix} \qquad x \geq 0 \tag{6.6}$$

for some perturbation b and includes a column representing the root of the perturbation.

PROOF. Suppose B is strongly feasible. Then B is a feasible basis for (6.2)–(6.4), row i^* is represented in B and for each $i = 1, 2, \ldots, m$, there is $\epsilon_i > 0$ such that the basic solution x^i to (6.5) is feasible. Clearly $x = \frac{1}{m-1}\sum_{i \neq i^*} x^i \geq 0$ is the corresponding basic solution for (6.6) where $b_i = \frac{1}{m-1}\epsilon_i > 0$ for $i \neq i^*$ and $b_{i^*} = -\sum_{i \neq i^*} b_i$.

Now suppose B is a feasible basis for (6.2)–(6.4) and for (6.6) for some perturbation b and includes a column representing the root i^* of b. We argue that B is a strongly feasible basis. Let x_B be the components of x corresponding to the basis B. We assume without loss of generality that x_1 is a basic variable representing row i^* and re-write the system

$$Bx_B = \begin{pmatrix} b \\ 1 \end{pmatrix} \qquad \text{as} \qquad \begin{bmatrix} a & \tilde{B} \\ 1 & \tilde{r} \\ w_1 & \tilde{w} \end{bmatrix} \begin{pmatrix} x_1 \\ \tilde{x} \end{pmatrix} = \begin{pmatrix} \tilde{b} \\ b_{i^*} \\ 1 \end{pmatrix}, \text{ where}$$

$\tilde{b} > 0$ and $a \leq 0$. Let $x_B^* = (x_1^*, \tilde{x}^*) \geq 0$ be the unique solution to this system.

Note that $\tilde{B}$ is a square pre-Leontief matrix and $\tilde{x}^* \geq 0$ is a feasible solution to the system $\tilde{B}x = \tilde{b} - x_1^* a \geq \tilde{b} > 0$. It follows (See Veinott (1968)) that $\tilde{B}$ is Leontief and so has non-negative inverse.

For an arbitrary choice of $i \neq i^*$, consider the system

$$\begin{bmatrix} a & \tilde{B} \\ 1 & \tilde{r} \\ w_1 & \tilde{w} \end{bmatrix} \begin{pmatrix} x_1 \\ \tilde{x} \end{pmatrix} = e_{m+1} + \epsilon(e_i - e_{i^*}).$$

Since each column of A sums to 0, the system

$$\begin{bmatrix} a & \tilde{B} \\ 1 & \tilde{r} \end{bmatrix} \begin{pmatrix} x_1 \\ \tilde{x} \end{pmatrix} = \epsilon(e_i - e_{i^*})$$

is equivalent to the system $ax_1 + \tilde{B}\tilde{x} = \epsilon e_i$ and so, for each choice of $x_1 \geq 0$ and $\epsilon \geq 0$, $\tilde{x} = \tilde{B}^{-1}(\epsilon e_i - a x_1) \geq 0$.

In particular, since B is a feasible basis for the unperturbed problem (6.2)–(6.4) corresponding to $\epsilon = 0$, there is a value of $x_1^* > 0$ such that $w_1 x_1^* + \tilde{w}\tilde{x} = (w_1 - \tilde{w}\tilde{B}^{-1}a)x_1^* = 1$ and so $(w_1 - \tilde{w}\tilde{B}^{-1}a) = \frac{1}{x_1^*} > 0$. It follows then that for each $i \neq i^*$ we can choose $\epsilon_i > 0$ and $x_1 > 0$ such that $wx_1 + \tilde{w}\tilde{x} = (w_1 - \tilde{w}\tilde{B}^{-1}a)x_1 + \tilde{w}\tilde{B}^{-1}e_i\epsilon_i = 1$, proving that B is strongly feasible. □

Combining the observations of Lemma 4.1 and Lemma 6.1 leads to Lemma 6.3.

LEMMA 6.3. *Suppose B is a strongly feasible basis for (6.2)–(6.4) and let $x_B = (x^+, x^0)$ be the corresponding basic variables with positive basic variables x^+ and degenerate basic variables x^0. Then the system Bx_B can be written as*

$$\begin{bmatrix} 0 & R \\ M & N \\ w^+ & w^0 \end{bmatrix} \begin{bmatrix} x^+ \\ x^0 \end{bmatrix} \tag{6.7}$$

where $N \leq 0$ and R is Leontief and so has non-negative inverse.

PROOF. By Lemma 6.1, B has exactly one variable representing each row of A and since the non-zero coefficients of the positive basic variables must all fall on active rows, the system can be written as in (6.7). Observe that R is square, each row of R has exactly one positive element, and since B is strongly feasible, the rows of R are non-trivial. Hence, R is Leontief, and so R is non-singular and $R^{-1} \geq 0$. □

6.2. *Deterministic unichain policies.* A stationary deterministic policy for the problem (6.2)–(6.4) is defined by a set of columns B with exactly one column representing each row of A. The column representing row i defines the action to invoke in state i. We use these notions of a stationary deterministic policy and the corresponding set of columns interchangeably.

Given a set B of columns of (6.2)–(6.4) we define $G(B)$ to be the directed graph with a node for each row of A, i.e. for each state, and a directed edge (i, j) for each ordered pair (i, j) of rows of A such that some column of B representing row i has a negative coefficient on row j. A directed graph G is said to be *strongly connected* if there is a directed path between each pair of nodes. Otherwise, G contains at least two maximal strongly connected subgraphs called strongly connected components. If we contract each strongly connected component to a single node, the resulting graph is acyclic. So at least one strongly connected component has no out-going edges. We refer to the set of nodes of a strongly connected component with no out-going edges as a *recurrent class*. A recurrent class is non-trivial if it contains more than one node. A graph G may have one or more recurrent classes. The remaining nodes form the *transient class*, $\mathcal{T}$.

Structures in $G(B)$ translate in a natural, though not necessarily unique way to sets of columns in B. For example, a simple path $p = (i_1, i_2, i_3, \ldots, i_t)$ corresponds to a set $B' = [B_{j_1}, B_{j_2}, \ldots, B_{j_{t-1}}]$ of columns of B, where column B_{j_k} represents row i_k and has a negative coefficient on row i_{k+1}. The translation is not necessarily unique as B may include several different sets

of columns matching these requirements. For each choice of B' meeting these requirements, however, $G(B')$ contains the path p.

The stationary deterministic policy B is said to be *unichain* if $G(B)$ contains exactly one recurrent class $\mathcal{R}(B)$ and *multi-chain* otherwise. We let $\mathcal{T}(B)$ denote the transient class of a policy B.

We characterize strongly feasible bases in a number of increasingly combinatorial ways.

THEOREM 6.1. *The following are equivalent:*

i. B *is a strongly feasible basis for* (6.2)–(6.4).
ii. B *is a basis for* (6.2)–(6.4) *with a unique basic variable representing each row of* A.
iii. The variables of B *define a deterministic unichain policy.*

PROOF. We already proved that i. implies ii. in Lemma 6.1.

ii. implies iii.: Since B has a unique basic variable representing each row of A, the corresponding policy is deterministic. By Assumption 2, $G(B)$ contains at least one recurrent class. If $G(B)$ contains exactly one recurrent class, then the policy is also unichain. To see that $G(B)$ cannot contain more than one recurrent class, observe that for each recurrent class C, the corresponding Markov chain with transition matrix P, i.e. P_{ij} is the probability of moving from state i to state j, has a unique stable distribution, i.e. a unique probability vector π such that $P^T\pi = \pi$ and $\pi_i > 0$ for each $i \in C$ (Puterman, 2005).

We argue that for each recurrent class C, $\sum_{i\in C} w_i\pi_i > 0$ and hence z, where $z_j = \pi_j / \sum_{i\in C} w_i\pi_i \geq 0$, satisfies $wz = 1$ and since $P^T z = z$, we see that it also satisfies $Bz = e_{m+1}$.

To see that $\sum_{i\in C} w_i\pi_i > 0$, observe that $\sum_{i\in C} w_i\pi_i = 0$ only if every transition in the recurrent class C has $w_i = 0$ and so the columns of B representing rows in C must be linearly dependent, contradicting the fact that B is a basis.

We now argue that $G(B)$ contains exactly one recurrent class. For if $G(B)$ contained more than one recurrent class, then with each such class C_i there would be a distinct solution z^i satisfying $Bz^i = e_{m+1}$, contradicting the fact that B is a basis.

iii. implies i.: Let P be the transition matrix corresponding to the unichain policy defined by B. Since the policy is unichain there is a unique $\pi \geq 0$ such that $P^T\pi = \pi$ and $\sum_{i=1}^m \pi_i = 1$. Further, for each recurrent state i, $\pi_i > 0$.

Choose a recurrent state i^* and re-write the system $P^T\pi = \pi$ as

$$\begin{bmatrix} a & \tilde{B} \\ 1 & \tilde{r} \end{bmatrix} \begin{pmatrix} \pi_{i^*} \\ \tilde{\pi} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \tag{6.8}$$

where $a \leq 0$ (since π_i^* represents row i^*) and $\tilde{B}$ is a square pre-Leontief matrix.

We first argue that $\tilde{B}$ is Leontief and $\tilde{B}^{-1} \geq 0$. To see this, consider the system in which we change the transition probabilities out of state i^* so that every other state is reachable in a single transition. In particular, observe that the system

$$\begin{bmatrix} a' & \tilde{B} \\ 1 & \tilde{r} \end{bmatrix} \begin{pmatrix} \pi_{i^*} \\ \tilde{\pi} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \tag{6.9}$$

where each entry of a' is $-\frac{1}{m-1}$, corresponds to a finite state Markov process that is not only unichain, but in which each state is recurrent. Thus, there is a unique $\pi' \geq 0$ such that

$$\begin{bmatrix} a' & \tilde{B} \\ 1 & \tilde{r} \end{bmatrix} \begin{pmatrix} \pi'_{i^*} \\ \tilde{\pi}' \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

and $\sum_{i=1}^m \pi'_i = 1$. In fact, since each state of this process is recurrent $\pi'_i > 0$ for each i. Note that since $a'\pi'_{i^*} + \tilde{B}\tilde{\pi}' = 0$, $\tilde{\pi}' \geq 0$ is a solution to $\tilde{B}\tilde{\pi}' = -a'\pi'_{i^*} > 0$ and so again invoking the results of Veinott (1968), we see that $\tilde{B}$ is Leontief and $\tilde{B}^{-1} \geq 0$.

We next argue, returning to the system (6.8), that $w_{i^*} - \tilde{w}\tilde{B}^{-1}a > 0$, where $\tilde{w}$ represents the components of w corresponding to columns of B excluding the column representing row i^*. To see this, observe that since $a\pi_{i^*} + \tilde{B}\tilde{\pi} = 0$, $\tilde{\pi} = -\tilde{B}^{-1}a\pi_{i^*}$. So $w\pi = w_{i^*}\pi_{i^*} + \tilde{w}\tilde{\pi} = w_{i^*}\pi_{i^*} - \tilde{w}\tilde{B}^{-1}a\pi_{i^*} = (w_{i^*} - \tilde{w}\tilde{B}^{-1}a)\pi_{i^*}$. Since $w \geq 0$ and $\pi \geq 0$, $w\pi \geq 0$. So either $w\pi > 0$ and $w_{i^*} - \tilde{w}\tilde{B}^{-1}a > 0$ as desired, or $w\pi = 0$. But this can only happen if $\pi_i > 0$ exclusively for columns corrresponding to y variables, which contradicts the statement that the policy is unichain. To see this, note that if $G(B)$ includes a recurrent class defined only by y variables indicating changes in the drift rate, then $G(B)$ is not connected and each connected component of $G(B)$ must contain a recurrent class, contradicting the assumption that the policy is unichain. So, $w_{i^*} - \tilde{w}\tilde{B}^{-1}a > 0$ and since $\pi_{i^*} > 0$, $w\pi > 0$.

We use this fact to argue that B is a feasible basis for (6.2)–(6.4). To see this, observe that $v = \pi/w\pi \geq 0$ is a solution to $Bv = e_{m+1}$. We next argue that v is the unique solution. To see this, suppose there is another solution v' (not necessarily non-negative) such that $Bv' = e_{m+1}$. Then, letting

$v' = (v'_{i^*}, \tilde{v}')$, we see that $av'_{i^*} + \tilde{B}\tilde{v}' = 0$ and so $\tilde{v}' = -\tilde{B}^{-1}av'_{i^*} = \tilde{v}\frac{v'_{i^*}}{v_{i^*}}$. It follows that $v' = \frac{v'_{i^*}}{v_{i^*}}v$ and since $wv = wv' = 1$, $v = v'$. Thus, B is a feasible basis for (6.2)–(6.4).

Finally, we argue that B is a strongly feasbile basis. Consider a state $i \neq i^*$. For each value of $x_{i^*} > 0$ and scalar $\epsilon_i > 0$, letting $\tilde{x} = \tilde{B}^{-1}(\epsilon_i e_i - ax_{i^*})$ yields a vector $x = (x_{i^*}, \tilde{x}) \geq 0$ satisfying $ax_{i^*} + \tilde{B}\tilde{x} = \epsilon_i e_i$. Further, since $w_{i^*} - \tilde{w}\tilde{B}^{-1}a > 0$ we can choose values for $x_{i^*} > 0$ and $\epsilon_i > 0$ such that $wx = w_{i^*}x_{i^*} + \tilde{w}\tilde{x} = w_{i^*}x_{i^*} + \tilde{w}\tilde{B}^{-1}(\epsilon_i e_i - ax_{i^*}) = \left(w_{i^*} - \tilde{w}\tilde{B}a\right)x_{i^*} + \tilde{w}\tilde{B}^{-1}e_i\epsilon_i = 1$. For these values of x_{i^*} and ϵ_i, $x = (x_{i^*}, \tilde{x}) \geq 0$ is a feasible solution to the perturbed problem $Bx = e_{m+1} + \epsilon_i(e_1 - e_{i^*})$. Since this is true for each $i \neq i^*$, it follows that B is strongly feasible. □

Next we look at the pivoting process restricted to strongly feasible bases. In Lemma 6.4 we identify several ways a pivot can progress. We simplify the pivoting process by developing combinatorial mechanisms for recognizing these different ways, for identifying a leaving variable and for constructing a new strongly feasible solution (if necessary).

LEMMA 6.4. *Given a strongly feasible basis B for (6.2)–(6.4) and an entering variable x_e, assume without loss of generality that x_e represents row j and that x_1 is the current basic variable representing row j.*

Case 1: *x_1 is a positive basic variable*

- *The pivot is non-degenerate, x_1 is a candidate to leave the basis and*
- *If we choose x_1 to leave the basis, the resulting basis is strongly feasible.*

Case 2: *x_1 is a degenerate basic variable either **(A)***

- *The pivot is degenerate, x_1 is a candidate to leave the basis and*
- *If we choose x_1 to leave the basis, the resulting basis is strongly feasible,*

*or **(B)***

- *The pivot is non-degenerate, x_1 is not a candidate to leave the basis and*
- *The resulting basis is not strongly feasible, but a new strongly feasible basis representing the new basis can be constructed easily.*

PROOF. Let a be the column of the entering variable and re-write the system

$$Bd = -a \qquad \text{as} \qquad \begin{bmatrix} 0 & R \\ M & N \\ w^+ & w^0 \end{bmatrix} \begin{bmatrix} d^+ \\ d^0 \end{bmatrix} = - \begin{bmatrix} a_R \\ a_M \\ w_e \end{bmatrix},$$

where d^+ corresponds to the rows for positive basic variables, d^0 corresponds to the rows of degenerate basic variables, $N \leq 0$, and R is Leontief.

We first consider Case 1 in which x_1 is a positive basic variable. In this case, row j will have its positive coefficient in M. The entering variable x_e represents the same row as x_1, hence a_M has a positive element and since A is pre-leontief $a_R \leq 0$. Since R is Leontief, $Rd^0 = -a_R$ has the unique solution $d^0 = -R^{-1}a_R \geq 0$ and so the degenerate variables will not leave and the pivot is non-degenerate. If x_1 is not a candidate to leave the basis, then after the pivot, both x_1 and x_e will be positive basic variables representing row j in the resulting basic feasible solution, contradicting Lemma 4.1. Thus, we have showed that if x_1 is positive the pivot is non-degenerate.

We next consider Case 2 in which $x_1 = 0$ and so j is an inactive row and can not have its positive coefficient in M. Since $N \leq 0$ row j must have its positive coefficient in R, and so a_R contains a positive coefficient. We re-write the system $Rd^0 = -a_R$ as

$$Rd^0 = \begin{bmatrix} 1 & k \\ \ell & \tilde{R} \end{bmatrix} \begin{pmatrix} d_1 \\ \tilde{d} \end{pmatrix} = \begin{pmatrix} -1 \\ -\tilde{a}_R \end{pmatrix} = -a_R, \tag{6.10}$$

where $\tilde{a}_R \leq 0$, $\ell \leq 0$. Since R is Leontief there is a column vector $x = (x_1, \tilde{x}) \geq 0$ such that $Rx > 0$. Then $\ell x_1 + \tilde{R}\tilde{x} > 0$. Since $\ell \leq 0$, $\tilde{R}\tilde{x} > 0$ and so $\tilde{R}$ is Leontief. Let $(d_1^*, \tilde{d}^*)$ be the unique solution to this system. There are three cases to consider:

Case 2$-(A)$**:** $d_1^* < 0$. In this case, the pivot is degenerate, x_1 is a candidate to leave the basis and if we choose x_1 to leave, the resulting basis has a unique representative for each row of A and so is strongly feasible.

Case 2$-(B)$**:** $d_1^* = 0$. In this case $\tilde{d} = -\tilde{R}^{-1}(\tilde{a}_R + \ell) \geq 0$ and so the pivot is non- degenerate. Unfortunately, x_1 is not a candidate to leave, and the resulting basis is not strongly feasible. All is not lost, however. Since the pivot is non-degenerate the new basic feasible solution is strictly better than its predecessor. We simply construct a strongly feasible basis representing the new basic feasible solution and continue.

Case 2$-(C)$**:** $d_1^* > 0$. In this case, since $\tilde{R}$ is Leontief, $\tilde{d} = -\tilde{R}^{-1}(\tilde{a}_R + \ell) \geq 0$, so the pivot is non-degenerate and after the pivot, both x_1 and x_e are positive

basic variables representing row j, contradicting Lemma 4.1. So this case cannot arise. □

In order to make this approach "combinatorial", we require combinatorial mechanisms for recognizing Case 2−(B), for identifying a leaving variable and for constructing a new strongly feasible basis in this case.

A strongly feasible basis corresponds to a deterministic unichain policy B and so can be recognized by the fact that $G(B)$ has a single recurrent class. This observation provides a simple combinatorial mechanism for recognizing Case 2−(B): Let B' be the deterministic policy obtained by replacing x_1, the current basic variable representing row j, with x_e, the entering variable. If $G(B')$ has more than one recurrent class, x_1 was not a candidate to leave the basis and so we have Case 2−(B).

We argue that $G(B')$ has either one recurrent class (corresponding to Case 2−(A)) or two recurrent classes (corresponding to Case 2−(B)). To see this, note that $G(B)$ and $G(B')$ only differ in the edges out of node j and so any new recurrent class in $G(B')$ must contain node j. Example 1 illustrates that Case 2−(B) can indeed arise.

EXAMPLE 6.1. Let $\Lambda = \{u_1, u_2\}$ and $\mathcal{S} = \{s_1, s_2\}$. Start with the strongly feasible basis that changes the drift rate from u_2 to u_1 at s_1 and maintains the drift rate otherwise. This unichain policy has the unique recurrent class $\{(1, u_1), (2, u_1)\}$. Suppose $r(1, u_2)$ is to enter the basis. If we choose $y(1, u_2, u_1)$ to leave the basis, the new policy maintains the drift rate everywhere and has two recurrent classes: $\{(1, u_1), (2, u_1)\}$ and $\{(1, u_2), (2, u_2)\}$. To resolve this, note that the leaving variable must be either $r(1, u_1) > 0$ or $r(2, u_1) > 0$ and so the pivot must be non-degenerate. That means that the average cost of the policy with the new recurrent class $\{(1, u_2), (2, u_2)\}$ is lower, so we construct a strongly feasible basis with this recurrent class. The unichain policy that changes the drift rate from u_1 to u_2 at s_1 is one example we could choose.

Consider Case 2−(B) in which $G(B')$ also contains a second recurrent class $\mathcal{R}'$ with $j \in \mathcal{R}'$. In this case, the leaving variable must represent a state in the original recurrent class $\mathcal{R}(B)$ and the basic variables in the new basic feasible solution x' must include the basic columns, other than x_1, representing the states of $\mathcal{R}'$. Let $p = (i_1, i_2, \ldots, i_t)$ be a shortest path in $G(A)$ from a node in $\mathcal{R}(B)$ to a node in $\mathcal{R}'$ and let Y be a set of columns representing this path. The policy B^* obtained by replacing the representatives of $\{i_1, i_2, \ldots, i_{t-1}\}$ in B' with the corresponding columns of Y is a unichain policy with the unique recurrent class $\mathcal{R}'$. To see this observe that

$\mathcal{R}'$ is a recurrent class of B^* and any other recurrent class of B^* must include at least one of the nodes of $\{i_1, i_2, \ldots, i_{t-1}\}$. Since $G(B^*)$ contains a path from each of these nodes to $\mathcal{R}'$ and no path from $\mathcal{R}'$ to any of these nodes, it follows that $\mathcal{R}'$ is the unique recurrent class in $G(B^*)$. Thus, B^* is a strongly feasible basis representing the new basic feasible solution.

Combining the primal pivoting procedure outlined in Lemma 6.4 with the method outlined in Observation 3 for computing γ and the relative value functions f allows us to solve $\mathrm{SLP}(\Lambda, \mathcal{S})$ without computing the parameters $p_+(j,u)$, $\mathrm{ET}[j,u]$, $\mathrm{EX}[j,u]$, $\mathrm{EA}[u]$ and $\mathrm{ER}[u]$.

In the next section we show how to keep the LP's we must solve small. In particular, we close the gap between the average cost of the $\mathcal{S}$-restricted Brownian control problem and the lower bound on the average cost of the unrestricted problem by intelligently adding points to $\mathcal{S}$.

7. Column generation. By selecting the points in $\mathcal{S}$ sufficiently close to each other we can obtain solutions arbitrarily close to a lower bound on the average cost of the unrestricted problem. However, adding many equally spaced points increases the size of the LP and the computational effort required to solve it. It is natural then to ask whether we can instead judiciously add points where they will have greatest impact, focusing more points in places where the preformance of the policy is particularly sensitive. We employ the words "Column Generation" as a convenient shorthand for this process of intelligently adding points to $\mathcal{S}$ and so variables or columns to the LP formulation to close the gap between the average cost of an optimal solution to the $\mathcal{S}$-restricted problem, which we henceforth denote by $\mathrm{AC}_{\mathcal{S}}$, and a lower bound on the cost of an optimal solution to the unrestricted problem, which we denote by $\underline{\gamma}$. Thus, Column Generation involves two issues: **Upper Bounds:** Adding points to $\mathcal{S}$ that reduce $\mathrm{AC}_{\mathcal{S}}$, the average cost of an optimal policy for the $\mathcal{S}$-restricted problem, **Lower Bounds:** Developing a better lower bound $\underline{\gamma}$ on the average cost of an optimal policy for the unrestricted problem.

Lemma 5.3 provides a potentially powerful tool for developing a lower bound $\underline{\gamma}$ from a solution $\mathrm{AC}_{\mathcal{S}}$ to the $\mathcal{S}$-restricted problem: take $\underline{\gamma} = \mathrm{AC}_{\mathcal{S}} - \epsilon H$, where $\{f(\cdot, u) : u \in \Lambda\}$ are the relative value functions of an optimal policy for the $\mathcal{S}$-restricted problem and $\epsilon \geq \max_{a\in[0,\Theta],u,v\in\Lambda}\{f(a,u) - f(a,v) - K(u,v)\}$. One difficulty with this approach is that unless consecutive points of $\mathcal{S}$ are close together, ϵ is generally large and so yields a rather weak lower bound on the average cost of an optimal policy for the unrestricted problem. Further, successively adding to $\mathcal{S}$ the points at which $f(a,u) - f(a,v) - K(u,v)$ achieves a maximum value generally forces us to

add numerous points to $\mathcal{S}$ that have little effect on either the upper bound $\mathrm{AC}_{\mathcal{S}}$ or the lower bound $\underline{\gamma}$.

Lemma 7.1 offers alternative strategies for reducing the gap between the optimal solution to the $\mathcal{S}$-restricted problem and the lower bound on the optimal solution to the unrestricted problem.

LEMMA 7.1. *Suppose that for each $u \in \Lambda$, $f(\cdot, u) : [0, \Theta] \to \mathbb{R}$ satisfies the smoothness conditions and (2.3)–(2.5),*

$$\begin{aligned} \frac{\sigma^2}{2} f''(a,u) + u f'(a,u) + c(u) + ha &\geq \gamma(1-\delta_1) \text{ for almost all } a \in [0,\Theta] \\ &\text{and} \\ \rho(a,u) \equiv f'(a+,u) - f'(a-,u) &\geq -\gamma\delta_2 \text{ for all } a \in (0,\Theta). \end{aligned}$$

for some scalars $\gamma, \delta_1, \delta_2$. Then $\gamma(1-\delta_1-\delta_2) \leq \mathrm{AC}(\Phi)$ for each policy $\Phi \in \mathcal{P}$ and each initial state (a, u).

Lemma 7.1 suggests approaching the problem of finding a lower bound through relaxations of (2.2) measured in terms of δ_1 and of (2.6) measured in terms of δ_2. These relaxations prove more useful in identifying appropriate points to add to $\mathcal{S}$ and, as we show in Lemma 7.2, it is relatively easy to construct functions f that satisfy (2.3)–(2.5) from the relative value functions defined in §5.

LEMMA 7.2. *Given an optimal dual solution (α, γ) for the $\mathcal{S}$-restricted problem, define*

$$\tilde{f}(a,u) = \min_{v \in \Lambda} \{ f(a,v) + K(u,v) \}, \tag{7.1}$$

where the functions $f(\cdot, u)$ are defined by (5.1)–(5.5). Then the functions $\tilde{f}(\cdot, u)$ satisfy the smoothness conditions and (2.3)–(2.5).

PROOF. By definition, the functions $\tilde{f}(\cdot, u)$, $u \in \Lambda$ satisfy (2.3). To see this let $\tilde{f}(a,u) = \min_{v\in\Lambda}\{f(a,v) + K(u,v)\} = f(a,w) + K(u,w)$. Note that $\tilde{f}(a,z) = \min_{v\in\Lambda}\{f(a,v) + K(z,v)\} \leq f(a,z) + K(z,z) = f(a,z)$. Then $\tilde{f}(a,u) - \tilde{f}(a,z) = f(a,w) + K(u,w) - \tilde{f}(a,z) \leq f(a,z) + K(u,z) - \tilde{f}(a,z) \leq f(a,z) + K(u,z) - f(a,z) = K(u,z)$. Further, since each function $f(\cdot, u)$ satisfies (2.4) and (2.5), clearly each function $\tilde{f}(\cdot, u)$ does as well.

Since each $\tilde{f}(\cdot, u)$ is defined as the minimum of continuous functions, these functions are continuous. It remains to show that $\tilde{f}(\cdot, u)$ satisfies the smoothness conditions. Since each of the functions $f(\cdot, u)$ can be written as the difference of two convex function their minimum can also be written as

the difference of two convex functions (see p. 375 of Floudas and Pardalos (2001)). Thus, it is enough to show that for each pair of drift rates v and v' the functions $f(a,v)+K(u,v)$ and $f(a,v')+K(u,v')$ cross at finitely many points. Let $g(a,u,v,v') = f(a,v)+K(u,v)-(f(a,v')+K(u,v'))$. Observe that g''' has no roots, hence g'' can have at most one root and g' can have at most 2 roots and g can have at most 3 roots for $a \in [s_i, s_{i+1}]$. Since $|\mathcal{S}|$ is finite, the functions cross finitely many times for each pair of drift rates. □

COROLLARY 7.1. *The functions* $v(a,u) = \mathit{argmin}_{v\in\Lambda}\{f(a,v)+K(u,v)\}$ *for each* $u \in \Lambda$ *are piecewise constant with finitely many pieces.*

We can use the functions $\tilde{f}$ to construct an alternative lower bound on the unrestricted problem: find δ_1 and δ_2 such that for each $u \in \Lambda$

$$\Gamma\tilde{f}(a,u) \equiv \frac{\sigma^2}{2}\tilde{f}''(a,u) + u\tilde{f}'(a,u) + c(u) + ha - \gamma \geq -\gamma\delta_1 \text{ for almost all } a \in [0,\Theta], \tag{7.2}$$

$$\tilde{\rho}(a,u) \equiv \tilde{f}'(a+,u) - \tilde{f}'(a-,u) \geq -\gamma\delta_2 \text{ for all } a \in [0,\Theta]. \tag{7.3}$$

We show that these two criteria can be used to identify $a^* \notin \mathcal{S}$ such that the reduced cost $\mathrm{rc}(a^*,u,v)$ of $y(a^*,u,v)$ is strictly negative and that adding such a^* to $\mathcal{S}$ decreases $\mathrm{AC}_{\mathcal{S}}$, thus reducing the gap between $\mathrm{AC}_{\mathcal{S}}$ and $\underline{\gamma}$.

LEMMA 7.3. *If* (7.2) *or* (7.3) *holds with equality for some* $\delta_1 > 0$ *or* $\delta_2 > 0$ *for some* $s_i \leq a < s_{i+1}$ *then there exists a point* $a^* \in (s_{i-1}, s_{i+1})$ *and a variable* $y(a^*,u,v)$ *with negative reduced cost.*

PROOF. We handle (7.2) and (7.3) separately. We first look at the case where $\tilde{\rho}(z,u) < 0$ for some $z \in [s_i, s_{i+1})$.

Case 1: $\tilde{\rho}(z,u) < 0$. First observe that if $v(a,u) = v$ for all a in some interval (b,d), then $\tilde{\rho}(a,u) = f'(a+,v) - f'(a-,v) \geq 0$ for all $a \in (b,d)$. So, $\tilde{\rho}(a,u)$ can only be negative when a is in the finite set $Z(u) = \{a \in [0,\Theta] : v(a+,u) \neq v(a-,u)\}$. Consider a point $z \in Z(u)$ such that $\tilde{\rho}(z,u) < 0$ and suppose

$$\tilde{f}(a,u) = \begin{cases} f(a,v)+K(u,v) & a \in [b,z] \\ f(a,v')+K(u,v') & a \in [z,d] \end{cases}$$

where $b < z < d$ and $v = v(z-,u) \neq v(z+,u) = v'$. Then by definition (7.1) (i) $\tilde{\rho}(z,u) = f'(z+,v') - f'(z-,v) < 0$, (ii) $\tilde{f}(z,u) = f(z,v)+K(u,v) \leq f(z,u)$ and so the reduced cost of $y(z,u,v)$ is non-positive, (iii) $\tilde{f}(z,u) = f(z,v')+K(u,v') \leq f(z,u)$ and so the reduced cost of $y(z,u,v')$ is non-positive.

First consider the case $z \notin \mathcal{S}$. Then, $f'(z, v') = f'(z+, v') < f'(z-, v) = f'(z, v)$ and so either $f'(z, u) > f'(z, v')$ or $f'(z, u) < f'(z, v)$ or both. If $f'(z, u) > f'(z, v')$, the reduced cost $\mathrm{rc}(a, u, v') \equiv f(a, v') + K(u, v') - f(a, u)$ of $y(a, u, v')$ is a non-positive and decreasing function of a on the interval $[z, a^*]$, where $a^* > z$ is the first point at which $f'(a^*, u) = f'(a^*, v')$. Observe that since the reduced cost of $y(s_{i+1}, u, v')$ is non-negative, it follows that a^* exists and satisfies $z < a^* < s_{i+1}$. Next observe that $\mathrm{rc}(a^*, u, v') < 0$ is a local minimum of $\mathrm{rc}(\cdot, u, v')$ and so we add the point a^* to $\mathcal{S}$.

On the other hand, if $f'(z, u) < f'(z, v)$, then with analogous arguments one can show that there exists a $s_i < a^* < z$ where $\mathrm{rc}(a^*, u, v) < 0$ is a local minimum of $\mathrm{rc}(\cdot, u, v)$ and so we add the point a^* to $\mathcal{S}$.

For the case with $z = s_i \in \mathcal{S}$ with analogous analysis one can show that there is a point a^* with $\mathrm{rc}(a^*, u, v) < 0$, so we add the point a^* to $\mathcal{S}$.

Case 2: $\Gamma\tilde{f}(z, u) < 0$**.** If $v(z, u) = v$, $\tilde{f}(z, u) = f(z, v) + K(u, v)$, and so

$$\Gamma\tilde{f}(z, u) = \frac{\sigma^2}{2}\tilde{f}''(z, u) + u\tilde{f}'(z, u) + c(u) + hz - \gamma = -f'(z, v)(v - u) - c(v) + c(u).$$

Observe that between consecutive points of $\mathcal{S} \cup Z(u)$, $\Gamma\tilde{f}(a, u) = -f(a, v)(v - u) - c(v) + c(u)$ for some $v \in \Lambda$. Thus, the extrema of $\Gamma\tilde{f}(a, u)$ occur only at points of $\mathcal{S} \cup Z(u)$ or at points $z \in Z'(u) = \{a \in (0, \Theta) : f''(a, v(a, u)) = 0\}$. So we need only consider (1) $\Gamma\tilde{f}(z-, u) \equiv \lim_{a \uparrow z} \Gamma\tilde{f}(a, u)$ for $z \in \mathcal{S} \cup Z(u)$, (2) $\Gamma\tilde{f}(z+, u) \equiv \lim_{a \downarrow z} \Gamma\tilde{f}(a, u)$ for $z \in \mathcal{S} \cup Z(u)$ and (3) $\Gamma\tilde{f}(z, u)$ for $z \in Z'(u) \equiv \{a \in (0, \Theta) : f''(a, v(a, u)) = 0\}$. We only consider the case in which $\Gamma\tilde{f}(z-, u) < 0$ for $z \in Z(u) \setminus \mathcal{S}$. The other cases are similar.

Consider the case in which $\Gamma\tilde{f}(z-, u) < 0$ for some point $z \in Z(u) \setminus \mathcal{S}$ so that $s_i < z < s_{i+1}$. Then by (7.1) (i) $\tilde{f}(z, u) = f(z, v) + K(u, v) \le f(z, u)$, (ii) $\tilde{f}(z, u) = f(z, v') + K(u, v') \le f(z, u)$, (iii) $f'(z, v) \ge f'(z, v')$.

If $f'(z, u) > f'(z, v')$ then the arguments of Case 1 imply that there is $z < a^* < s_{i+1}$ such that $f'(a^*, u) = f'(a^*, v')$ and the reduced cost $\mathrm{rc}(a^*, u, v')$ of $y(a^*, u, v')$ is strictly negative. Similarly if $f'(z, u) < f'(z, v)$ then there is $z > a^* > s_i$ such that $f'(a^*, u) = f'(a^*, v)$ and the reduced cost $\mathrm{rc}(a^*, u, v)$ of $y(a^*, u, v)$ is strictly negative. We add a^* to $\mathcal{S}$. Thus, we need only consider the case in which $f'(z, u) = f'(z, v) = f'(z, v')$. Since $\Gamma\tilde{f}(z-, u) < 0$,

$$\begin{aligned}&\frac{\sigma^2}{2}f''(z-, v) + uf'(z-, v) + c(u) + hz\\ &\quad < \gamma = \frac{\sigma^2}{2}f''(z-, u) + uf'(z-, u) + c(u) + hz\end{aligned}$$

and so $f''(z-,v) < f''(z-,u)$. It follows that there is a point $s_i < a^* < z$ such that $f'(a^*,v) = f'(a^*,u)$ and $r(a,u,v) = f(a,v) + K(u,v) - f(a,u)$, the reduced cost of $y(a,u,v)$, is a non-positive increasing function on (a^*,z). Thus $y(a^*,u,v)$ has negative reduced cost and we choose a^* to add to $\mathcal{S}$. □

8. Implementation issues and computational experience. In this section we briefly discuss implementation issues and present computational results. In our implementation we used several tools to achieve computational savings, specifically we scaled C and F in the relative value function to overcome numerical issues, applied the approach described in Observation 3 and made use of a "restricted basis" as discussed in Ormeci Matoglu and Vande Vate (2011) to compute $\text{AC}_{\mathcal{S}}$.

In an effort to avoid numerical scaling issues in our computations, we define

$$\tilde{C}(i,\mu) = C(i,\mu)\left(e^{-\frac{2\mu}{\sigma^2}s_j} - e^{-\frac{2\mu}{\sigma^2}s_i}\right),$$

where $C(i,\mu)$ is the C constant of the band (μ, s_i, s_j) of the policy. If s_j is the first (lowest value) point at which we transition from drift rate μ, the band is $(\mu, 0, s_j)$ and we define

$$\tilde{C}(0,\mu) = C(0,\mu)\left(e^{-\frac{2\mu}{\sigma^2}s_j} - 1\right).$$

Similarly, if s_i is the last (highest value) point at which we transition from drift rate μ, the band is (μ, s_i, Θ) and we define

$$\tilde{C}(i,\mu) = C(i,\mu)\left(e^{-\frac{2\mu}{\sigma^2}\Theta} - e^{-\frac{2\mu}{\sigma^2}s_i}\right).$$

We refer to control bands belonging to a recurrent class as *recurrent bands*, and those belonging to a transient class as *non-recurrent bands*. Working with $\tilde{C}$ in place of C works well for the recurrent bands, but consider for example a non-recurrent band $(u, 0, s_1)$ where $-\frac{2u}{\sigma^2}s_1$ is large, e.g., > 50. In this case, the values of $C(0,u)$ and $F(0,u)$ are defined by:

$$C(0,u) = \frac{\sigma^2}{2u}\left(U + \frac{\gamma}{u} + \frac{h\sigma^2}{2u^2} - \frac{c(u)}{u}\right), \text{ and}$$
$$C(0,u)e^{-\frac{2u}{\sigma^2}s_1} + F(0,u) = \alpha(s_1,u)$$

and so, since its coefficient $e^{-\frac{2u}{\sigma^2}s_1}$ in the second constraint is so large, $C(0,u)$ must be represented with over 20 decimal places of precision to ensure the (near) continuity of the relative value function $f(\cdot,u)$ at s_1. We know of no

other resolution for this issue than to carry the computations out to the required level of precision.

A second tool we use to solve the drift control problem involves restricted bases. The positive basic variables of the LP form a basis of the linear program restricted to the active constraints and to the variables whose coefficients appear only on these constraints. We refer to this as a *restricted basis*. As we calculate the average cost of the $\mathcal{S}$-restricted problem after adding a point a^* to $\mathcal{S}$ we we use only the restricted basis. Thus we solve a linear system of equalities and obtain $\mathrm{AC}_{\mathcal{S}}$ and the C and F parameters associated with the recurrent bands of the policy. In later steps, we compute the remaining C and F parameters defining the relative value function on the transient bands. This approach speeds computations in the case of degenerate pivots as they only affect the relative value function on transient bands.

As pointed out in Observations 1 and 2 once these parameters, and hence the relative value functions, are obtained we can identify candidates to enter via simple tests using the relative value function.

We should also point out here the tools of linear programming that we did not employ. Our goal was not to reinvent linear programming, but to apply the basic tools at the required level of precision. For example, we employed the very straightforward approach of pricing out every non-basic column and choosing the one with the most negative reduced cost to enter the basis. We also did not factor the reduced "basis" when computing the coefficients C and F defining the relative value functions and instead solved the linear system "from scratch" each time. Finally, we did not exploit the computational savings available with degenerate pivots.

Another point worth mentioning is the relative importance of variability. Recall that the variance appears in our equations in the exponential term $2\mu/\sigma^2$. This exponential term is one of the main sources of numerical issues. The Brownian model makes sense when variability is high and is less relevant when it is low. Our method works well in the former case and may run into numerical issues in the latter case. Put crudely: The Brownian model focuses on the impacts of variability. If variability is important, our model helps understand how important it is and how best to address it. If it is not important, our model is probably not appropriate and is likely to struggle with numerical issues. In the next section we present results of computational studies.

8.1. *Computational studies.* Our approach enables the controller to evaluate the impact of opening/closing a line. More lines means both more capacity and more flexibility to respond to variations in demand. It is relatively

easy to understand the value of more capacity. Solving the drift control problem provides a quantitative tool for analyzing the marginal value of additional flexibility. While in the analysis we do not consider the fixed investment cost of building a line (we only consider the costs of turning on an idle line and idling an existing line), the savings achieved through having the flexibility to operate an additional line augment the analysis of whether additional capacity is worth the investment. In this section we present a detailed numerical study and consider different numbers of lines to handle the variability in demand.

We present the results for our approach as applied to the numeric example in Ormeci Matoglu and Vande Vate (2010). The average arrival rate μ_0 =26,400 units per day and the standard deviation of the arrival process, σ, is 7,348.47 units per day. Each line has a processing rate of 10,560 units per day, so the set of drift rates Λ includes the appropriate members of $\{15840, 5280, -5280, -15840, -26400, -36960, -47520, -58080\}$ corresponding to one, two, three up to eight lines operating. Clearly, three lines provide sufficient capacity, but we explore the value of the added flexibility of four, six and even eight lines. Starting up a line costs approximately \$20,000 (note this is the cost of starting to run an existing line and does not include the initial investment cost to build the line) and shutting one down temporarily costs approximately \$2,000, thus $K(u, v)$ is 20,000 times the number of lines started up and 2,000 times the number of lines shut down. The delay cost, h, is \$2.88 per unit per day, $c(\mu)$, the cost of operating at drift rate μ, is $-18.18 * \mu$, Θ, the maximum queue length is 40,000 units, U, the cost of idling, is \$400 per unit, and M, the cost of rejecting orders, is \$1,000 per unit. In each case, we begin with the initial set of candidate control points $\mathcal{S} = \{10000, 20000, 30000\}$.

Figure 1 shows the value of the flexibility additional lines provide by showing the average cost of an optimal policy with different numbers of lines available to the policy. As expected, the marginal value of additional lines decreases. Changing from 6 to 8 lines reduced the average cost of an optimal policy by less than \$5/hour, so it did not make sense to consider more than 8 lines. In each case we stopped the column generation process when the policy was within \$5/hour of the lower bound.

Figure 2 illustrates the structure our near optimal control band policy for the three rate case. The policy uses three control bands (μ_i, s_i, S_i): $(\mu_1, 0, 13052)$, $(\mu_2, 5417, 17110)$, $(\mu_3, 8525, 40000)$ meaning that when the queue is small we should run just one line, when it reaches 13,052, we should start up a second line. If the queue falls to 5,417 we should shut down the second line, but if it rises to 17,110, we should start up the third line. When

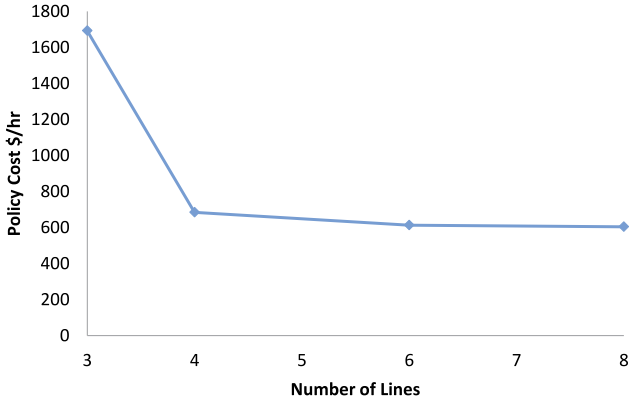


FIG 1. *Average cost with respect to number of lines.*

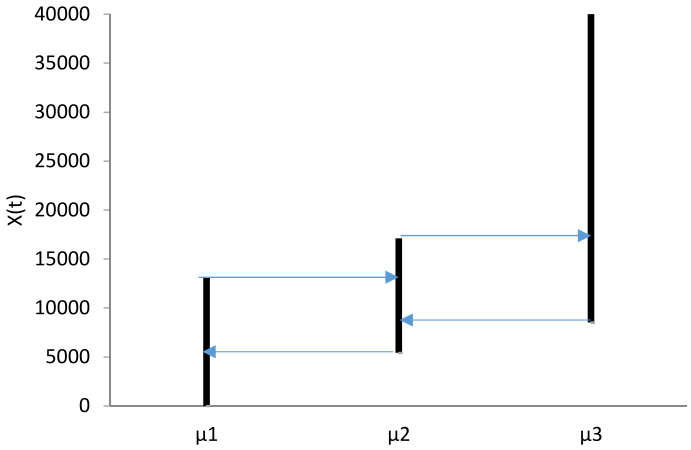


FIG 2. *Optimal policy for 3 rates problem,* $\mu_i > \mu_{i+1}$.

TABLE 1
The progress of the "Column Generation" approach with 3 drift rates

$\lvert\mathcal{S}\rvert$	Cum. Pivots	Policy $/hr	Lower Bound $/hr	Run Time (sec)
3	3	2416.48	−5068.86	2
.	.	.	.	.
8	17	1766.86	−178.97	13
.	.	.	.	.
12	23	1698.76	1258.11	26
.	.	.	.	.
22	44	1693.36	1684.46	89
23	48	1693.36	1689.60	99

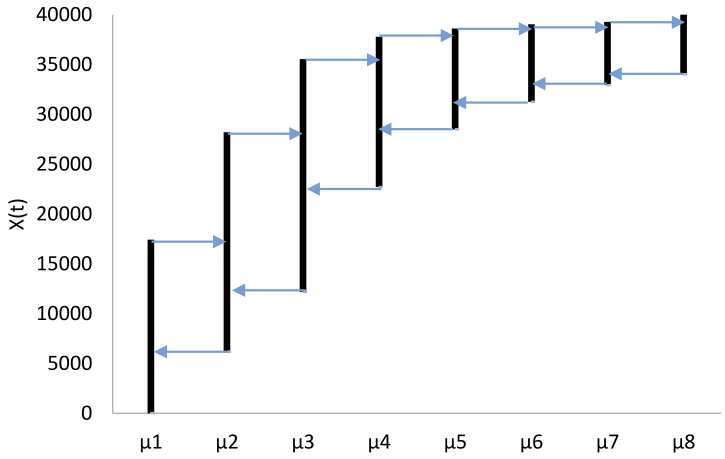


FIG 3. *Optimal policy for 8 rates problem,* $\mu_i > \mu_{i+1}$.

all three lines are running, if the queue falls to 8,525, we should shut down the third line. The average cost of this policy is \$1,693 per hour. Our lower bounds prove that the average cost of an optimal policy is not less than \$1,689 per hour. The policies for 4,6 and 8 rates exhibit a similar pattern. Figure 3 displays our near optimal control band policy for the eight rate case. The policy uses 8 control bands. It is interesting to note that the widths of the bands get smaller with each additional line, starting with the fourth line.

Table 1 summarizes the progress of the column generation approach. The "Policy \$/hr" column reports the average cost of the policy, the "Lower Bound \$/hr" column reports the lower bound on the average cost of an optimal policy, the $|\mathcal{S}|$ column reports the number of points in $\mathcal{S}$, the "Pivots" column reports the cumulative number of pivots and the "Run Time (sec)" column reports the cumulative run time. Note that the average cost of the policy and the lower bound converge quickly as we increase the number of points in $\mathcal{S}$. Each additional point we add to $\mathcal{S}$ requires only a few additional pivots.

To emphasize the importance of the column generation approach's judicious choice of points in $\mathcal{S}$, we compare this with the average cost (\$1701 per hour) and the lower bound (\$796 per hour) obtained using a grid of 23 equally spaced points. Table 2 shows the average cost, lower bound and number of pivots required as we refine the equally spaced grid. As we see in Figure 4 the policy cost and the lower bound converge much more quickly using the column generation approach. In this chart the lower bound is displayed on the left axis, and the policy cost is displayed on the right axis.

Tables 3–5 show the progress of our column generation procedure in the

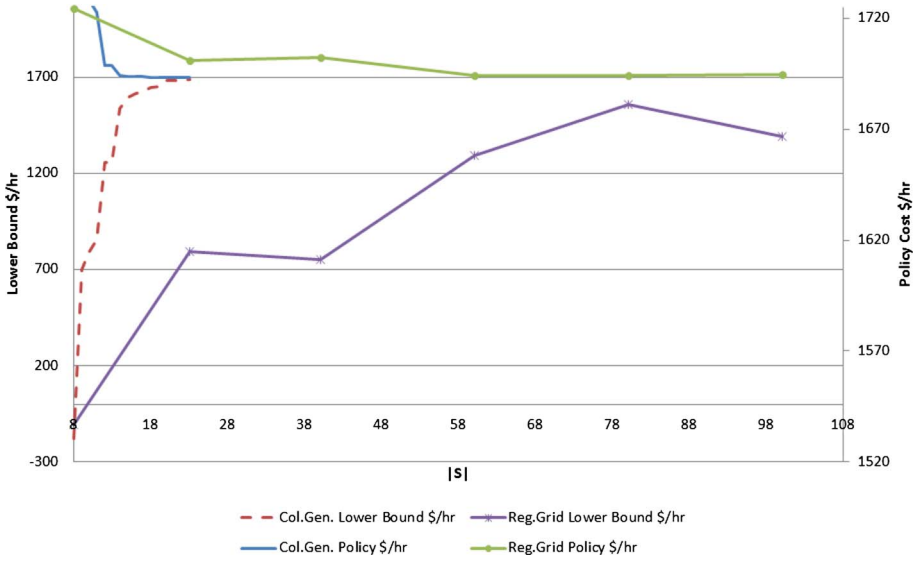


FIG 4. *Policy Cost and Lower Bound w.r.t |S| in 3 rates problem with Column Generation Approach and Regular Grid.*

TABLE 2
Optimal solution to the $\mathcal{S}$-restricted problem with $|\mathcal{S}|$ equally spaced points, 3 rates

$\|\mathcal{S}\|$	Pivots	Policy $/hr	Lower Bound $ /hr	Run Time (sec)
23	29	1701.16	795.93	13
40	53	1702.49	749.94	45
60	78	1694.18	1294.32	152
80	106	1694.23	1560.22	393
100	132	1694.50	1393.66	826

TABLE 3
The progress of the "Column Generation" approach with 4 drift rates

$\|\mathcal{S}\|$	Cum. Pivots	Policy $/hr	Lower Bound $/hr	Run Time (sec)
3	11	949.63	−2570.33	4
.	.	.	.	.
19	64	685.69	523.30	180
.	.	.	.	.
30	95	685.23	679.13	507
31	97	685.23	681.91	555

4, 6 and 8 drift rate cases and Table 6 and Table 7 outline comparable performance for the 4 and 6 drift rate cases using equally spaced grids.

As these tables illustrate, the column generation procedure requires far fewer points and far fewer pivots to find a policy of comparable average cost and produce a lower bound of comparable quality. Column generation does,

TABLE 4
The progress of the "Column Generation" approach with 6 drift rates

$\lvert\mathcal{S}\rvert$	Cum. Pivots	Policy \$/hr	Lower Bound \$/hr	Run Time(sec)
3	30	949.63	−8685.20	17
.	.	.	.	.
26	150	614.50	503.89	14162
.	.	.	.	.
60	299	613.79	610.22	20369
61	303	613.79	611.17	21552

TABLE 5
The progress of the "Column Generation" approach with 8 drift rates

$\lvert\mathcal{S}\rvert$	Cum. Pivots	Policy \$/hr	Lower Bound \$/hr	Run Time (sec)
3	40	949.63	−14588.80	49
.	.	.	.	.
39	290	605.11	436.11	14162
.	.	.	.	.
84	535	604.81	601.25	249238
85	545	604.81	602.03	262309

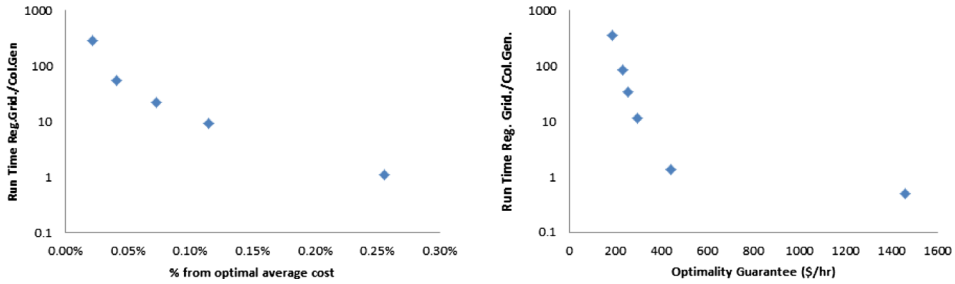
TABLE 6
Optimal solutions to the $\mathcal{S}$-restricted problem with equally spaced points, 4 Rates

$\lvert\mathcal{S}\rvert$	Cum. Pivots	Policy \$/hr	Lower Bound \$/hr	Run Time (sec)
31	147	686.98	245.89	170
60	288	686.01	390.39	1630
100	488	685.51	452.21	64900

TABLE 7
Optimal solutions to the $\mathcal{S}$-restricted problem with equally spaced points, 6 Rates

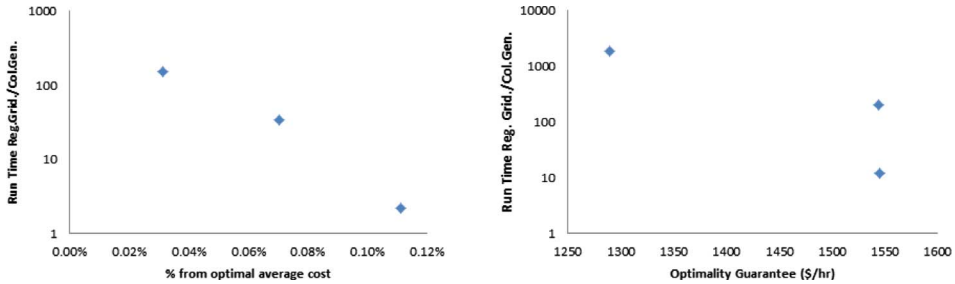
$\lvert\mathcal{S}\rvert$	Pivots	Policy \$/hr	Lower Bound \$ /hr	Run Time (sec)
39	301	614.47	−930.48	3408
79	459	614.22	−930.00	57218
159	796	613.98	−675.36	532359

however, include the extra computation of selecting the points to add to $\mathcal{S}$. To give a sense of how this added speed in terms of number of points and pivots compares with the added cost to find those points, Figure 5 shows the ratio of the running times (regular grid/column generation on a logarithmic scale) required to produce a policy for the four drift rate policy within a target percentage of the optimal average cost as the target percentage decreases. In this case, the optimal average cost is known. Normally, we do not know the optimal average cost in advance and must rely on the lower bounds produced by the procedures to gauge the quality of the policies. It is interesting to compare the ratio of the running times to produce a policy and a lower bound so that the policy is within a given percentage of the lower bound.



(a) known optimal cost (b) optimal cost unknown

FIG 5. *Ratio of run times (Regular grid/Column generation) to find a solution, 4 rates.*



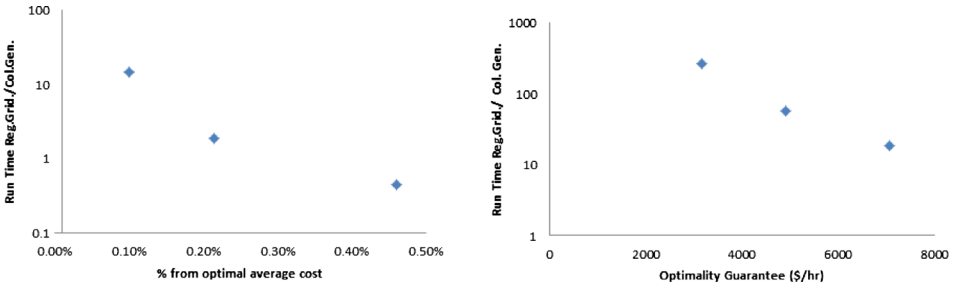
(a) known optimal cost (b) optimal cost unknown

FIG 6. *Ratio of run times (Regular grid/Column generation), 6 rates.*

This proved problematic for the fixed grid approach since in many cases, the lower bound it provided was negative. Thus, we used the absolute difference between the average cost of the policy and the lower bound as our measure of the quality of the solutions. Figure 5 shows the ratio of the running times (on a logarithmic scale) to produce a policy and a lower bound of a given quality using this measure for the four drift rate problem. Figures 6–7 outline the same comparison for the 6 and 8 drift rates problems. We observe that the time to find a policy whose average cost converges to the optimal cost is significantly lower when the column generation approach is employed.

APPENDIX A: BASIC ADJOINT RELATIONSHIPS AND CALCULATION OF SOME PERFORMANCE METRICS

LEMMA A.1. *For each state* (j, u), $u \neq 0$ *the following Basic Adjoint Relationships hold:*



(a) known optimal cost (b) optimal cost unknown

FIG 7. *Ratio of run times (Regular grid/Column generation), 8 rates.*

If $j = 1$, *then*

$$\begin{aligned} 0 &= u + EA[u] + \delta(1,2,u)(s_1 - s_2), \\ 0 &= 2EX[1,u]u + \sigma^2 + \delta(1,2,u)\left(s_1^2 - s_2^2\right), \\ 0 &= \delta(1,2,u)\left(e^{-\frac{2u}{\sigma^2}s_1} - e^{-\frac{2u}{\sigma^2}s_2}\right) - \frac{2u}{\sigma^2}EA[u], \ \textit{and} \\ ET[1,u] &= \frac{1}{\delta(1,2,u)}. \end{aligned}$$

If $1 < j < n$, *then*

$$\begin{aligned} &0 = u + \delta(j,j-1,u)(s_j - s_{j-1}) + \delta(j,j+1,u)(s_j - s_{j+1}), \\ &0 = 2EX[j,u]u + \sigma^2 + \delta(j,j-1,u)\left(s_j^2 - s_{j-1}^2\right) + \delta(j,j+1,u)\left(s_j^2 - s_{j+1}^2\right), \\ &0 = \delta(j,j-1,u)\left(e^{-\frac{2u}{\sigma^2}s_j} - e^{-\frac{2u}{\sigma^2}s_{j-1}}\right) + \delta(j,j+1,u)\left(e^{-\frac{2u}{\sigma^2}s_j} - e^{-\frac{2u}{\sigma^2}s_{j+1}}\right), \end{aligned}$$

$$\begin{aligned} p_-(j,u) &= \frac{\delta(j,j-1,u)}{\delta(j,j-1,u) + \delta(j,j+1,u)}, \\ p_+(j,u) &= \frac{\delta(j,j+1,u)}{\delta(j,j-1,u) + \delta(j,j+1,u)}, \ \textit{and} \\ ET[j,u] &= \frac{1}{\delta(j,j-1,u) + \delta(j,j+1,u)}. \end{aligned}$$

If $j = n$, *then*

$$\begin{aligned} 0 &= u - ER[u] + \delta(n,n-1,u)(s_n - s_{n-1}), \\ 0 &= 2EX[n,u]u + \sigma^2 - 2\Theta ER[u] + \delta(n,n-1,u)\left(s_n^2 - s_{n-1}^2\right), \\ 0 &= \frac{2u}{\sigma^2}e^{-\frac{2u}{\sigma^2}\Theta}ER[u] + \delta(n,n-1,u)\left(e^{-\frac{2u}{\sigma^2}s_n} - e^{-\frac{2u}{\sigma^2}s_{n-1}}\right), \ \textit{and} \end{aligned}$$

$$ET[n,u] = \frac{1}{\delta(n,n-1,u)}.$$

Here $\delta(j,k,u)$ may be interpreted as the frequency with which s_k is hit when in state (j,u). Solving for EX, EA, and ER yields the following expressions:

LEMMA A.2.

$$\mathrm{ET}[j,u] = \begin{cases} \dfrac{2u(s_2-s_1)+\left(e^{-\frac{2us_2}{\sigma^2}}-e^{-\frac{2us_1}{\sigma^2}}\right)\sigma^2}{2u^2}, & \text{if } j=1, \\ -\dfrac{2u(s_n-s_{n-1})-\left(e^{\frac{2u(\Theta-s_{n-1})}{\sigma^2}}-e^{\frac{2u(\Theta-s_n)}{\sigma^2}}\right)\sigma^2}{2u^2}, & \text{if } j=n, \\ \dfrac{\left(e^{\frac{2us_{j-1}}{\sigma^2}}(s_{j-1}-s_j)+e^{\frac{2us_{j+1}}{\sigma^2}}(s_j-s_{j+1})+e^{\frac{2u(s_{j-1}+s_{j+1}-s_j)}{\sigma^2}}(s_{j+1}-s_{j-1})\right)}{\left(e^{\frac{2us_{j-1}}{\sigma^2}}-e^{\frac{2us_{j+1}}{\sigma^2}}\right)u}, & \\ \text{otherwise.} & \end{cases}$$

$$EA[u] = \begin{cases} \dfrac{-\left(e^{-\frac{2us_2}{\sigma^2}}-e^{-\frac{2us_1}{\sigma^2}}\right)u\sigma^2}{2u(s_2-s_1)+\left(e^{-\frac{2us_2}{\sigma^2}}-e^{-\frac{2us_1}{\sigma^2}}\right)\sigma^2}, & \text{if } j=1, \\ 0, & \text{otherwise} \end{cases}$$

and

$$ER[u] = \begin{cases} \dfrac{e^{\frac{2u\Theta}{\sigma^2}}\left(-e^{-\frac{2us_{n-1}}{\sigma^2}}+e^{-\frac{2us_n}{\sigma^2}}\right)u\sigma^2}{2u(s_n-s_{n-1})-\left(e^{\frac{2u(\Theta-s_{n-1})}{\sigma^2}}-e^{\frac{2u(\Theta-s_n)}{\sigma^2}}\right)\sigma^2}, & \text{if } j=n, \\ 0, & \text{otherwise.} \end{cases}$$

$$EX[j,u] = \begin{cases} \dfrac{u\left(s_2^2-s_1^2\right)}{2u(s_2-s_1)+\left(e^{-\frac{2us_2}{\sigma^2}}-e^{-\frac{2us_1}{\sigma^2}}\right)\sigma^2}-\dfrac{\sigma^2}{2u}, & \text{if } j=1, \\ \Theta-\dfrac{\sigma^2}{2u}-\dfrac{u(s_n-s_{n-1})(2\Theta-s_{n-1}-s_n)}{2u(s_n-s_{n-1})-\left(e^{\frac{2u(\Theta-s_{n-1})}{\sigma^2}}-e^{\frac{2u(\Theta-s_n)}{\sigma^2}}\right)\sigma^2}, & \text{if } j=n, \\ \frac{1}{2}\Bigg(s_{j-1}+s_{j+1}-\dfrac{\sigma^2}{u} & \\ \quad+\dfrac{\left(e^{\frac{2us_{j-1}}{\sigma^2}}-e^{\frac{2us_{j+1}}{\sigma^2}}\right)(s_j-s_{j-1})(s_j-s_{j+1})}{e^{\frac{2us_{j-1}}{\sigma^2}}(s_j-s_{j-1})-e^{\frac{2us_{j+1}}{\sigma^2}}(s_j-s_{j+1})+e^{\frac{2u(s_{j-1}+s_{j+1}-s_j)}{\sigma^2}}(s_{j-1}-s_{j+1})}\Bigg), & \text{ow.} \end{cases}$$

APPENDIX B: BASIC ADJOINT RELATIONSHIPS, CALCULATION OF SOME PERFORMANCE METRICS AND RELATIVE VALUE FUNCTION WHEN $\mu = 0$

LEMMA B.1. *For each state $(j,0)$ the following Basic Adjoint Relationships hold:*
If $j = 1$, then

$$\begin{aligned} 0 &= EA[0] + \delta(1,2,0)(s_1 - s_2), \\ 0 &= \sigma^2 + \delta(1,2,0)\left(s_1^2 - s_2^2\right), \\ 0 &= 3\sigma^2 EX[1,0] + \delta(1,2,0)\left(s_1^3 - s_2^3\right), \text{ and} \\ ET[1,0] &= \frac{1}{\delta(1,2,0)}. \end{aligned}$$

If $1 < j < n$, then

$$\begin{aligned} 0 &= \delta(j,j-1,0)(s_j - s_{j-1}) + \delta(j,j+1,0)(s_j - s_{j+1}), \\ 0 &= \sigma^2 + \delta(j,j-1,0)\left(s_j^2 - s_{j-1}^2\right) + \delta(j,j+1,0)\left(s_j^2 - s_{j+1}^2\right), \\ 0 &= 3\sigma^2 EX[j,0] + \delta(j,j-1,0)\left(s_j^3 - s_{j-1}^3\right) + \delta(j,j+1,0)\left(s_j^3 - s_{j+1}^3\right), \\ p_-(j,0) &= \frac{\delta(j,j-1,0)}{\delta(j,j-1,0) + \delta(j,j+1,0)}, \\ p_+(j,0) &= \frac{\delta(j,j+1,0)}{\delta(j,j-1,0) + \delta(j,j+1,0)}, \text{ and} \\ ET[j,0] &= \frac{1}{\delta(j,j-1,0) + \delta(j,j+1,0)}. \end{aligned}$$

If $j = n$, then

$$\begin{aligned} 0 &= -ER[0] + \delta(n,n-1,0)(s_n - s_{n-1}), \\ 0 &= \sigma^2 - 2\Theta ER[0] + \delta(n,n-1,0)\left(s_n^2 - s_{n-1}^2\right), \\ 0 &= 3\sigma^2 EX[n,0] - 3\Theta^2 ER[0] + \delta(n,n-1,0)\left(s_n - s_{n-1}^2\right), \text{ and} \\ ET[n,0] &= \frac{1}{\delta(n,n-1,0)}. \end{aligned}$$

Solving for EX, EA, and ER yields the following expressions:

LEMMA B.2.

$$\mathrm{ET}[j,0] = \begin{cases} \frac{s_2^2 - s_1^2}{\sigma^2}, & \text{if } j = 1, \\ \frac{(2\Theta - s_n - s_{n-1})(s_n - s_n - 1)}{\sigma^2}, & \text{if } j = n, \\ \frac{(s_{j+1} - s_j)(s_j - s_{j-1})}{\sigma^2}, & \text{otherwise.} \end{cases}$$

$$EA[0] = \begin{cases} \frac{\sigma^2}{s_1+s_2}, & \text{if } j = 1, \\ 0 & \text{otherwise.} \end{cases}$$

and

$$ER[0] = \begin{cases} -\frac{\sigma^2}{s_n+s_{n-1}-2\Theta}, & \text{if } j = n, \\ 0, & \text{otherwise.} \end{cases}$$

$$EX[j,0] = \begin{cases} \frac{s_1^2+s_1s_2+s_2^2}{3(s_1+s_2)}, & \text{if } j = 1, \\ \frac{s_n^2+s_ns_{n-1}+s_{n-1}^2-2\Theta^2}{3(s_n+s_{n-1}-2\Theta)}, & \text{if } j = n, \\ \frac{s_j+s_{j-1}+s_{j+1}}{3}, & \text{otherwise.} \end{cases}$$

The relative value function, when $\mu = 0$ is defined as follows: Given an optimal solution (γ, α) to the problem (4.7)–(4.11)

$$f(a,0) = -\frac{h}{3\sigma^2}a^3 + \frac{(c(0)-\gamma)}{\sigma^2}a^2 + C(j,0)a + F(j,0),$$

for $s_j \le a \le s_{j+1}$, where $C(j,0)$ and $F(j,0)$ are constants satisfying:

$$C(0,0) = -U,$$
$$C(n,0) = \frac{2\gamma\Theta - h\Theta^2 + M\sigma^2 - 2\Theta c(0)}{\sigma^2},$$

and for $j = 1, 2, 3, \ldots, n$

$$-\frac{h}{3\sigma^2}s_j^3 + \frac{(c(0)-\gamma)}{\sigma^2}s_j^2 + C(j-1,0)s_j + F(j-1,0) = \alpha(j,u)$$
$$-\frac{h}{3\sigma^2}s_j^3 + \frac{(c(0)-\gamma)}{\sigma^2}s_j^2 + C(j,0)s_j + F(j,0) = \alpha(j,u).$$

REFERENCES

Anderson, E. J., Nash, P. and Perold, A. F. (1983). Some properties of a class of continuous linear programs. *SIAM Journal on Control and Optimization* **21** 758–765. MR0710999

Anderson, E. J. and Nash, P. (1989). *Linear Programming in Infinite-Dimensional Spaces*. John Wiley, Chichester. MR0893179

Ata, B., Harrison, J. M. and Shepp, L. A. (2005). Drift rate control of a Brownian processing system. *The Annals of Applied Probability* **15** 1145–1160. MR2134100

Avram, F. and Karaesmen, F. (1996). A method for computing Double Band Policies for switching between two diffusions. *Probability in Engineering Sciences* **10** 569–590. MR1408617

Buie, R. N. and Abrham, J. (1973). Numerical solutions to continuous linear programming problems. *Mathematical Methods of Operations Research* **17** 107–117. MR0429103

Chernoff, H. and Petkau, A. J. (1978). Optimal control of a Brownian motion. *SIAM Journal on Applied Mathematics* **34** 717–731. MR0491451

Cunningham, W. H. (1976). A network simplex method. *Mathematical Programming* **11** 105–116. MR0462532

Feldman, R. M. and Valdez-Flores, C. (1996). *Applied Probability and Stochastic Processes*. Pws Pub Co, Boston.

Floudas, C. A. and Pardalos, P. M. (2001). *Encyclopedia of Optimization* **6**. Springer.

Ghosh, A. P. and Weerasinghe, A. P. (2007). Optimal buffer size for a stochastic processing network in heavy traffic. *Queueing Systems* **55** 147–159. MR2319488

Ghosh, A. P. and Weerasinghe, A. P. (2010). Optimal buffer size and dynamic rate control for a queueing system with impatient customers in heavy traffic. *Stochastic Processes and Their Applications* **120** 2103–2141. MR2684738

Harrison, J. M. (1985). *Brownian Motion and Stochastic Flow Systems*. Wiley, New York. MR0798279

Helmes, K. and Stockbridge, R. H. (2000). Numerical comparison of controls and verification of optimality for stochastic control problems. *Journal of Optimization Theory and Applications* **106** 107–127. MR1780110

Helmes, K. and Stockbridge, R. H. (2008). Determining the optimal control of singular stochastic processes using linear programming. In *Markov Processes and Related Fields: A Festschrift in Honor of Thomas G. Kurtz*, (S. N. Ethier, J. Feng and R. H. Stockbridge, eds.). *IMS Collections* 137–153. IMS, Ohio. MR2574229

Hernández-Hernández, D. and Hernández-Lerma, O. (1994). Discounted cost Markov decision processes on Borel spaces: The linear programming formulation. *J. Math. Anal. Appl* **183** 335–351. MR1274144

Hernández-Hernández, D., Hernández-Lerma, O. and Taksar, M. (1996). The linear programming approach to deterministic optimal control problems. *Applicationes Mathematicae* **24** 17–33. MR1404981

Hernández-Lerma, O. and González-Hernández, J. (1998). Infinite linear programming and multichain markov control processes in uncountable spaces. *SIAM Journal on Control and Optimization* **36** 313–335. MR1616582

Hernández-Lerma, O. and Lasserre, J. B. (1998). Approximation schemes for infinite linear programs. *SIAM Journal on Optimization* **8** 973–998. MR1641282

Klabjan, D. and Adelman, D. (2006). Existence of optimal policies for semi-Markov decision processes using duality for infinite linear programming. *SIAM Journal on Control and Optimization* **44** 2104–2122. MR2248177

Kumar, S. and Muthuraman, K. (2004). A numerical method for solving singular stochastic control problems. *Operations Research* **52** 563–582. MR2075794

Kushner, H. J. and Dupuis, P. (2001). *Numerical Methods for Stochastic Control Problems in Continuous Time*, second ed. *Applications of Mathematics: Stochastic Modelling and Applied Probability* **24**. Springer-Verlag, New York. MR1800098 (2001g:93002)

Liao, Y.-C. (1984). Switching and impulsive control of reflected diffusion. *Applied Mathematical Optimization* **11** 153–159. MR0743924

Manne, A. S. (1960). Linear programming and sequential decisions. *Management Science* **6** 259–267. MR0129022

Orlin, J. B. (1985). On the simplex algorithm for networks andgeneralized networks. *Mathematical Programming Studies* **24** 166–178. MR0820998

Ormeci Matoglu, M. and Vande Vate, J. (2010). A Technical note on drift control with changeover cost. www.isye.gatech.edu/~jvandeva/DriftControlNote.pdf.

Ormeci Matoglu, M. and Vande Vate, J. (2011). Drift contol with changeover costs. *Operations Research* **59** 427–439. MR2830168

Perry, D. and Bar-Lev, S. K. (1989). A control of a Brownian storage system with two switchover drifts. *Stochastic Analysis and Applications* **7** 103–115. MR0977253

Protter, P. (2004). *Stochastic Integration and Differential Equations*. Springer-Verlag Berlin. MR2020294

Pullan, M. C. (1993). An algorithm for a class of continuous linear programs. *SIAM Journal on Control and Optimization* **31** 1558–1577. MR1242216

Puterman, M. L. (2005). *Markov Decision Processes: Discrete Stochastic Dynamic Programming*. Wiley, New York. MR1270015

Rath, J. H. (1977). The optimal policy for a controlled Brownian motion process. *SIAM Journal on Applied Mathematics* **32** 115–125. MR0423599

Taksar, M. I. (1997). Infinite-dimensional linear programming approach to singular stochastic control. *SIAM Journal of Control and Optimization* **35** 604–625. MR1436641

Tijms, H. C. (2004). *A First Course in Stochastic Models*. John Wiley & Sons, Ltd. MR2190630

Veinott, A. F. (1968). Extreme points of Leontief substitution systems. *Linear Algebra and Applications* **1** 181–194. MR0232590

Weiss, G. (2008). A simplex based algorithm to solve separated continuous linear programs. *Mathematical Programming* **115** 151–198. MR2403756

Melda Ormeci Matoglu
10 Garrison Ave. Durham
NH, USA
E-mail: melda.ormecimatoglu@unh.edu

John Vande Vate
Huizhu Wang
756 Ferst Dr. Atlanta
GA 30332, USA
E-mail: jvandeva@isye.gatech.edu
hwang8@gatech.edu