Stochastic Singular Control Problems with State Constraints

by

Kevin J. Ross

A dissertation submitted to the faculty of the University of North Carolina at Chapel Hill in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Department of Statistics and Operations Research.

Chapel Hill
2006

Approved by:

Amarjit Budhiraja, Advisor

Chuanshu Ji, Reader

M. Ross Leadbetter, Reader

Eric Renault, Reader

Gordon Simons, Reader

ABSTRACT

KEVIN J. ROSS: Stochastic Singular Control Problems with State Constraints.
(Under the direction of Amarjit Budhiraja.)

Singular control is an important and challenging class of problems in stochastic control theory. Such control problems can rarely be solved explicitly and thus numerical approximation schemes are necessary. In this work we develop approximation schemes for singular control problems with state constraints.

The first problem we consider arises in problems of optimal consumption and investment under transaction costs. We use Markov chain approximations to develop a convergent numerical scheme. Proof of convergence uses techniques from the theory of weak convergence. Specific features that make the analysis nontrivial include unboundedness of state and control spaces and cost function; degeneracies in the dynamics; and presence of both singular and absolutely continuous controls. We present a computational algorithm and the results of a numerical study.

Numerical schemes for singular control problems can be computationally quite intensive, and thus it is of great interest to develop less expensive schemes that exploit specific features of the underlying dynamics. To this end we investigate connections between singular control and optimal stopping problems. A key technical step in establishing such connections is proving existence of an optimal singular control. We prove such a result for a general class of singular control problems with linear dynamics and state constrained to be in a polyhedral cone. A particular example of this class of models are the so-called Brownian control problems (BCPs) and thus existence of optimal controls for BCPs follows as a consequence.

Armed with this existence result, we consider a two-dimensional singular control problem that arises from queueing networks. We prove rigorously an equivalence of this

problem with an optimal stopping problem. We exploit this connection in developing simple computational schemes for the singular control problem, and we investigate performance of the schemes in a numerical study.

ACKNOWLEDGMENTS

I owe a considerable debt of gratitude to my advisor Professor Amarjit Budhiraja. Always approachable, he has provided tremendous guidance and support over the last few years. He has been the source of valuable ideas and advice, both related to this dissertation as well as academic life in general.

I would also like to thank my committee members, Professors Chuanshu Ji, M. Ross Leadbetter, Eric Renault, and Gordon Simons, whose input and guidance have been most helpful.

Many thanks to Dr. Thomas S. Royster Jr., Mrs. Caroline H. Royster, and the Graduate School at UNC for the generous fellowship in the Royster Society of Fellows. Participation in the Society of Fellows has greatly aided my own research, allowed me to broaden my intellectual horizons, and prepared me well for interdisciplinary research in the future.

Thanks always to my family, for their unwavering support of all my endeavors, academic or otherwise.

Finally, I would like to thank my wife Amy for her patience, encouragement, and friendship throughout the writing of this dissertation.

TABLE OF CONTENTS

LIST OF FIGURES **vii**

1 Introduction **1**

2 Convergent Numerical Scheme for a Problem of Optimal Consumption and Portfolio Selection with Transaction Costs **8**

2.1 Problem Description and Motivation 8

2.2 Optimal Consumption and Portfolio Selection with Transaction Costs . . 14

2.3 An Approximating Markov Decision Problem 21

2.4 Main Convergence Result 36

2.5 Computational Methods for the MDP 59

2.6 Numerical Study 66

3 Existence of Optimal Controls for Singular Control Problems with State Constraints **68**

3.1 Introduction 68

3.2 Setting and Main Result 72

3.3 Restriction to Continuous Controls 74

3.4 Existence of an Optimal Control 80

3.5 Brownian Control Problems 91

4 Numerical Scheme for a Brownian Control Problem through Optimal Stopping **93**

4.1 Introduction 93

4.2 The Criss-Cross Network 95

4.3 The Brownian Control Problem 96

4.4 Equivalent Workload Formulation of the BCP 98

4.5 Connections between EWF and BCP 100

4.6 A Further Reduction in Parameter Regime IIb 101

4.7 Equivalence between Singular Control and Optimal Stopping Problems . 103

4.8 Proof of Theorem 4.7.1 105

4.9 Numerical Study 115

Bibliography **124**

LIST OF FIGURES

2.1 State space and directions of control 15

2.2 The discrete state space $\mathbb{S}_{\ell}^{h+}$ 22

2.3 Numerically computed optimal control 67

4.1 The criss-cross network 96

4.2 Comparison of approximately optimal controls 117

4.3 Divergence of numerical algorithm for singular control MDP 119

4.4 Convergence of algorithm for optimal stopping MDP 120

4.5 Algorithm for singular control MDP based on optimal stopping 122

4.6 Approximate value function for singular control MDP 122

4.7 Comparison of singular control MDP algorithms 123

Chapter 1

Introduction

Stochastic Control Theory is an active area of research in applied probability with applications in diverse disciplines such as aerospace engineering, management science, economics, mathematical finance, and queuing networks. The area is quite well developed and now there are several excellent texts that are available ([19], [18], [30], [32], [52]). The basic problem can be described as follows. There is a stochastic dynamical system whose evolution can be influenced by exercising a control with a view towards achieving a desired goal. For example, if the dynamical system is described via a stochastic differential equation then a control may be in terms of modifying the drift or the diffusion coefficient. The control may be modulated continuously over time. It could be a bounded function of the current and past states of the controlled process or it may be unbounded, of the singular or impulsive type. Frequently, the desired goal in a stochastic control problem is to optimize a cost (or reward) function which may depend on both the state process and the control. From a computational point of view the main objective is to compute the minimal cost function, the so-called *value function*, and the control policies that either achieve or well approximate this optimal cost.

One of the main approaches in the computation of the value function is via the method of dynamic programming. For the diffusion control problems of interest in the current work, the nonlinear evolution equations of dynamic programming become the, so-called, Hamilton-Jacobi-Bellman (HJB) partial differential equations (PDEs). Rarely can such nonlinear PDEs be solved explicitly and thus in practice one needs to resort to numerical approximations. For diffusion control problems where the value function has suitable smoothness properties and is the unique classical solution of the corresponding HJB equation, there are well established finite difference methods that can be used for computing such approximations. However, many interesting modern applications involve diffusion control problems where due to degeneracies in the dynamics, non-smoothness of the boundary, singular nature of the control process, form of the boundary condition and other complexities, the associated value function is not necessarily smooth and the existence/uniqueness theory of the corresponding HJB equations is not well understood. Despite the fact that over the past twenty-five years there has been a rapid development in the theory of *viscosity solutions* of HJB equations for such diffusion control problems (cf. [10], [18], [49]), the PDE approach to the approximation of the value function becomes much more difficult.

Over the last thirty years, Kushner, Dupuis and co-workers (see [32] and references therein) have developed a powerful machinery for computational problems in stochastic control theory using a probabilistic approach. The main idea is to approximate the original controlled diffusion process by a suitable controlled Markov chain on a finite state space. The approximation should be defined so that the local properties of the controlled diffusion are similar to those of the approximating chain. Next, one introduces an appropriate Markov decision problem (MDP), a discrete time, discrete state analog of the continuous time control problem of interest. "Almost optimal" control policies and value functions for the MDP can be computed using classical iterative procedures such as value space and/or policy space approximations and their refinements. This in turn yields approximations of the value function and optimal control policy for the original control problem. Convergence analysis of the algorithm involves establishing convergence of the value function of the MDP to the value function of the original diffusion control problem as various approximation parameters approach suitable limits. This convergence analysis is completely probabilistic and is based on the theory of weak convergence of probability measures (e.g. [4]).

Although every Markov chain approximation corresponds to some particular finite difference approximation of the corresponding HJB equation, there are two main advantages of the above probabilistic approach to the approximation problem. First, the Markov chain approach is flexible, and it enables use of physical insights derived from dynamics of the controlled diffusion in choosing the approximating chain, or equivalently, the precise form of the finite difference approximation. It is well known that if the finite difference approximation is not chosen appropriately, it may lead to serious instabilities in the numerical procedure. Markov chain approximations allow one to naturally select the appropriate finite difference scheme for a given application. Frequently, numerically unstable schemes can be avoided by incorporating such insights. Second, Markov chain approximation schemes do not require smoothness of the cost or value functions; nor do they rely on associated HJB equations. This is a great advantage in many problems where complex features of the dynamics make the PDE theory for the associated HJB equation hard to tackle, as is true for problems in the current work.

In this work we are interested in two problems in stochastic optimal control of diffusions. The first one arises from problems in optimal consumption and portfolio selection with transaction costs, while the second is motivated by scheduling control problems for queuing networks in heavy traffic. Although, there are many important differences between the two control problems, both are problems of singular control with state constraints in domains with corners. *State constraints* refers to the requirement that the state process must remain within a certain feasibility region at all times. By *singular control* we mean that the control terms in the dynamics of the state process need not be absolutely continuous with respect to Lebesgue measure, and are only required to have paths of bounded variation.

Singular control is perhaps one of the most difficult classes of problems in stochastic control theory. We refer the reader to [5], especially the sections at the end of each chapter, for a thorough survey of the literature. The HJB equations for such problems, which are variational inequalities with gradient constraints, are notoriously difficult to work with. Finite difference approximations for some singular control problems have been studied in [26] and [50]. In particular, the problem we study in Chapter 2 is precisely the one undertaken in [50]. However, our approach to the approximation problem is quite different in that all our techniques are probabilistic. We follow the Markov chain approach of Kushner and Dupuis and in contrast to [50] we prove convergence of the approximation scheme. Markov chain approximations along with convergence proofs for a singular control problem have been studied by Kushner and Martins in [33]. Although the work in Chapter 2 borrows significantly from the ideas in [33], specific features of the dynamics and the model under consideration make the convergence analysis quite delicate. In particular, the model features that make our analysis considerably harder than that in [33] include: unboundedness of the cost function, domain and control space, mixed boundary conditions (Dirichlet-Neumann), degeneracies in the dynamics and presence of both singular and absolutely continuous control. The main convergence result of the chapter is Theorem 2.4.12. This result, and the corresponding proofs, illustrate how the Markov chain approach can be adapted to handle complex dynamical features. While the analysis is for a specific two-dimensional problem, we believe that many of the techniques introduced in this chapter can be applied more generally.

In Section 2.5 we present a computational algorithm for obtaining near optimal control policies for the control problem of Chapter 2. Theorem 2.4.12 and the analysis of Section 2.5 guarantees convergence of the numerical scheme. A similar approach can be taken to develop numerical schemes for other classes of singular control problems. One advantage of such an approach is the ability to establish convergence of the computational algorithms. However, such numerical schemes are often computationally intensive.

Thus, whenever possible one would like to take advantage of specific problem features to simplify the numerical scheme. One such simplification results from exploiting connections between singular control and so-called obstacle/optimal stopping problems (see [45]). Although numerical schemes for singular control problems are notoriously hard, there are relatively simple schemes available for optimal stopping problems.

Chapter 3 undertakes a key technical step in establishing connections between singular control and optimal stopping problems. In this chapter, we consider a broad family of multi-dimensional singular control problems with linear dynamics and state constrained to lie in a convex cone (see Section 3.2). The main result of this chapter is Theorem 3.2.3, which establishes existence of an optimal control for the singular control problem. The proof uses weak convergence arguments and a time rescaling technique. Existence of optimal controls for Brownian control problems [20], associated with a broad family of controlled stochastic networks, follows as a consequence. To the best of our knowledge, the current work is the first to address existence of an optimal control for a general multi-dimensional singular control problem with state constraints. Classical compactness arguments which are used for establishing existence of optimal controls for problems with absolutely continuous control terms (cf. [6]) do not naturally extend to singular control problems. For one-dimensional models one can typically establish existence constructively, by characterizing the optimal controlled process as a reflected diffusion (cf. [2, 3, 23]). In higher dimensions, one approach is through studying regularity of solutions of the variational inequalities associated with singular control problems and the smoothness of the corresponding free boundary. Such smoothness results are the starting points in the characterization of the optimal controlled process as a constrained diffusion with reflection at the free boundary. Except for specific models (cf. [46, 47]), this approach encounters substantial difficulties, even for linear dynamics (cf. [51]). A key difficulty is that little is known about the regularity of the free boundary in higher dimensions. Alternative approaches for establishing existence of optimal controls based on compactness arguments are developed in [40, 25, 15]. The first of these papers considers linear dynamics while the last two consider models with nonlinear coefficients. In all three papers the state space is all of $\mathbb{R}^d$, i.e. there are no state constraints. It is important to note that, in our model, although the drift and diffusion coefficients are constant, the state constraint requirement introduces a (non-standard) nonlinearity in the dynamics. While our method does not provide any characterization of the optimal control, it is quite general and should be applicable for other families of singular control problems (with or without state constraints).

The existence result of Chapter 3 (Theorem 3.2.3) is critical to establishing connections between singular control and optimal stopping problems. In Chapter 4 we investigate such connections for a two-dimensional singular control problem that arises from a scheduling control problem for the so-called criss-cross network (see Figure 4.1). This network has been studied by several authors (see [24], [9], [34], [38], and [8]). In the current work, we focus on the parameter regime IIb of [38], (Condition 4.6.1 of the current work), a regime which to date has not yielded an analytical solution. Using Theorem 3.2.3, we establish in Theorem 4.7.1 equivalence between the singular control problem and an optimal stopping problem. In Section 4.9 we exploit this connection to develop a computational scheme for the singular control problem which is much simpler than one based on a Markov chain approximation of the controlled process. The main idea is to numerically approximate the optimal stopping time and value function of the optimal stopping problem, and then use these quantities and Theorem 4.7.1 to obtain an approximation to the optimal control and value function for the singular control problem. Through several examples we illustrate how such an algorithm performs better numerically than one based on a Markov chain approximation for the original singular control problem.

Notation. We will use the following notation throughout. The set of nonnegative real numbers is denoted as $I\!R_+$. All vectors are column vectors and vector inequalities are to be interpreted componentwise. For $x \in I\!R^n$, $|x|$ denotes the Euclidean norm. For a point $x \in I\!R^n$ and a set $A \in I\!R^n$, $\text{dist}(x, A)$ will denote the distance of x from $\overline{A}$. Given a Polish space E, a function $f : [0, \infty) \to E$ is RCLL if it is right-continuous on $[0, \infty)$ and has left limits on $(0, \infty)$. We define the class of all such functions by $\mathcal{D}([0, \infty) : E)$. The subset of $\mathcal{D}([0, \infty) : E)$ consisting of all continuous functions will be denoted by $\mathcal{C}([0, \infty) : E)$. A process is RCLL if its sample paths lie in $\mathcal{D}([0, \infty) : E)$ a.s. For $T \geq 0$ and $\phi \in \mathcal{D}([0, \infty) : E)$ let $|\phi|^*_T \doteq \sup_{0 \leq t \leq T} |\phi(t)|$. The Borel σ-field for a Polish space E will be denoted by $\mathcal{B}(E)$. We will denote generic constants in $(0, \infty)$ by $c, c_1, c_2, \cdots$; their values may change from one theorem (lemma, proposition) to the next. By convention, the infimum of an empty set is ∞.

All other notation will be introduced within each chapter. Such notation will be self-contained within each chapter. To simplify notation, we may reuse variables, symbols, etc. from one chapter to the next.

Chapter 2

Convergent Numerical Scheme for a Problem of Optimal Consumption and Portfolio Selection with Transaction Costs

2.1 Problem Description and Motivation

In this chapter we consider a problem of optimal consumption and portfolio selection with proportional transaction costs that has been studied by several authors ([36, 12, 48, 50, 35]). The basic problem can be described as follows. Consider a single investor who has two instruments available for investment: a risk free asset such as a bank account which pays a fixed interest rate $r > 0$ and a risky asset, a stock, whose price evolution is modeled via a geometric Brownian motion with a mean value of return $b > r$ and constant volatility $\sigma > 0$. We assume that the investor may buy or sell stock continuously over time in not necessarily integer valued quantities. The investor is assumed to consume wealth at some time dependent rate $C(t)$ and without loss of generality we assume that the consumption is deducted from the bank account. The investor may instantaneously transfer money from the bank account to stock and vice-versa by paying a proportional transaction cost; namely, there are $\lambda \in (0,\infty)$ and $\mu \in (0,1)$ such that the investor pays λ times the amount moved from the bank account to stock as a transaction fee, and similarly, he pays μ times the amount moved from stock to the bank account as a transaction fee. All transaction fees are charged from the bank account. The basic constraint on the consumption control C and the portfolio selection control, denoted (M,N), is that the investor must be solvent at all times. More precisely, if $X(t)$ and $Y(t)$ represent the amount of investment in the bank account and the stock, respectively, at time t, then we require $(X(t),Y(t)) \in \mathbb{S}$ for all $t \geq 0$, where

$$\mathbb{S} \doteq \{(x,y) \in I\!R^2 : x + (1+\lambda)y \geq 0 \text{ and } x + (1-\mu)y \geq 0\}.$$

The goal of the investor is to maximize the expected total discounted utility of consumption, $I\!E \int_0^\infty e^{-\beta t} f(C(t))dt$, where $\beta \in (0,\infty)$ is the discount factor and the utility function $f : [0,\infty) \to [0,\infty)$ is a continuous function satisfying $f(0) = 0$. The condition $f(0) = 0$ can be relaxed if f is nondecreasing and $f(0) > -\infty$ by replacing f by $f - f(0)$.

In absence of transaction costs, Merton proved in the classical paper [41] that when the utility function is $f(c) = c^p/p, p < 1, p \neq 0$ or $f(c) = \log c$ (note that the latter utility function does not satisfy the conditions of the current chapter) the investor's optimal policy is to keep a constant proportion of total wealth in the risky asset and to consume at a rate proportional to total wealth. (For a simple and self-contained treatment see [12]). This "Merton line" target can always be achieved since transactions can be made continuously and instantaneously without affecting wealth. However, when transaction costs apply such a policy results in immediate bankruptcy. Magill and Constantinides first conjectured in [36] that there must exist a "no-transaction region" taking the form of a wedge in the wealth space. When wealth is inside this region consumption is the only control that can be exercised. Purchase or sale of stock only occurs when the wealth attempts to exit the no-transaction region. The formal arguments of [36] were put on a rigorous footing by Davis and Norman in [12] for the cases $f(c) = c^p/p$ or $f(c) = \log c$. In their work, under suitable conditions on model parameters, the free boundary problem associated with optimal consumption in the presence of proportional transaction costs is solved explicitly and C^2 regularity of the value function is established. The authors show that the (optimal) no-transaction region is a wedge, in particular, the optimal policy is to exercise the minimal amount of trading necessary to keep wealth inside the no-transaction region. Inside the region, consumption occurs at a finite rate. In [48] Shreve and Soner consider the same problem as in [12] but with conditions on the model parameters that are weaker and much more explicit. Once more, regularity properties of the value function and the associated free boundary are proved. A more general utility function which satisfies suitable smoothness, concavity and growth properties was considered in [50]. Using viscosity solution methods the authors sketch a proof for unique solvability of the associated HJB equation by the value function. A finite difference approximation scheme for approximating the value function is introduced; however, convergence of the proposed scheme for the portfolio selection problem is not proved. The authors do provide results from several numerical studies which identify near optimal control policies and the (numerical) free boundary.

In the current work, we do not impose any concavity, smoothness or growth conditions on the utility function; the key condition (Condition 2.2.1) that we require is that the value function is finite and continuous. In particular, we do not claim nor use that the value function is the unique solution of the associated HJB equation. The main goal of the study is to obtain convergent numerical approximations for the value function. The basic approach, as in [32], is to introduce a Markov Decision Problem (MDP) for an approximating, finite state, discrete time, controlled Markov chain. The main result of this chapter (Theorem 2.4.12) shows that the value function of the MDP converges to the value function of the original singular control problem as various parameters in the approximation approach their limits suitably. In Section 2.5 we use the approximating MDP to obtain computational schemes for obtaining near optimal control policies. The key result of Section 2.5 is Lemma 2.5.1, which allows us to characterize the value function and optimal control policies via solution of suitable dynamic programming equations (see Theorem 2.5.2). Finally in Section 2.6 results from a numerical study using the algorithm of Section 2.5 are described. In particular, Figure 2.3 shows the numerical no transaction region and the associated (numerical) free boundary obtained by an implementation of the algorithm.

The only paper (to the best of our knowledge) that carries out a complete convergence analysis for a numerical scheme for a singular control problem is [33]. Although the current chapter crucially uses several ideas developed in [33], there are key differences in the model that make our analysis substantially delicate. First, the above paper considers a queuing problem with "finite buffers" which essentially means that the state space and control space are bounded. In the current study we first have to suitably approximate the original unbounded model by one where the consumption control and the state space are bounded. This two stage approximation procedure is carried out in Propositions 2.2.2 and 2.2.4. This is the only place where the assumption on the continuity of the value function (Condition 2.2.1) is used. Next, in contrast to [33], in addition to singular control terms, we also have an absolutely continuous control term (consumption control) that appears in a nonlinear fashion in the cost (reward) criterion through the utility function f. This requires us to introduce the relaxed formulation for the stochastic control problem in order to carry out the convergence analysis. Lemma 2.4.1 ensures that the relaxed formulation does not change the value function of the control problem. The next substantial difficulty in our analysis is the state constraint feature of the dynamics. Although, in [33] the state is constrained to be in a bounded polyhedral region, the state constraints can be easily handled by introducing the, so-called, "Skorohod map". However, in the problem considered in this chapter, the directions of control do not point inward into the state space (see Figure 2.1) and therefore do not allow for a similar Skorohod reduction. Nonetheless, one useful feature of the dynamics (see (2.1)) is that once the state of the system reaches the boundary of $\mathbb{S}$, the only admissible control corresponds to moving the state process instantaneously to the origin and keeping it there at all times. This observation allows us to convert an infinite horizon cost to an exit time criterion (see equations (2.2)-(2.4)). This reformulation makes some aspects of the convergence analysis simpler, however, degeneracies in the state dynamics make treatment of convergence properties of exit times quite subtle. To see the basic difficulty consider the following simple example. Let ξ_n be a sequence of positive reals such that $\xi_n \to 0$ as $n \to \infty$. Let x_n be the solution of the ODE $\dot{x} = x$ with initial condition ξ_n and x the solution of the same ODE with 0 initial condition. Clearly $x_n \to x$ uniformly on compacts, however if $\tau_n \doteq \inf\{t | x_n(t) = 0\}$ and $\tau \doteq \inf\{t | x(t) = 0\}$, then clearly $\tau_n \not\to \tau$. In other words, convergence of processes in general need not imply the convergence of the corresponding exit times. The issue is especially problematic when, as is the case for the controlled dynamics considered in this chapter, the diffusion coefficients in the state dynamics are not uniformly non-degenerate. This is another key difference between the current model and the problem studied in [33].

One of the major obstacles in proving the convergence of the value function of a sequence of approximating discrete MDPs to the value of the original singular control problem is proving the tightness of the sequence of singular control terms in the Skorohod $D[0, \infty)$ space. A powerful technique for bypassing this tightness issue, based on suitable stretching of time scale was introduced in [33]. Although such time transformation ideas go back to the work of of Meyer and Zheng [42] (see also Kurtz [31]), the papers [33, 37] were the first to use such ideas in stochastic control problems. A similar technique was also recently used in [7]. A key ingredient to this technique is the uniform moment estimate obtained in Lemma 2.3.5. In [33] such a moment estimate follows easily from the form of the cost function where a strictly positive proportional cost is incurred for exercising the singular control. In the current problem there is no direct contribution to the (cost) reward function from the singular control term and as a result, the proof of this uniform estimate becomes more involved. Roughly speaking, the main idea of the proof is that a controller cannot make too much use of a singular control without pushing the process to the boundary of the domain.

The chapter is organized as follows. In Section 2.2 we give a precise formulation of the control problem of interest. We also present here two propositions (Propositions 2.2.2 and 2.2.4) which allow approximation of the original control problem by one with a bounded state space and bounded consumption actions. Section 2.3 introduces the discrete MDP that approximates the original singular control problem. It also introduces the continuous time interpolations and the time transformation that are key to the convergence analysis. In Section 2.4 we present the main convergence result that establishes the convergence of the value function of the MDP to that of the original singular control problem. Section 2.5 is devoted to computational methods for the MDP. A key result here is Lemma 2.5.1 which allows, via Theorems 2.5.2 and 2.5.3, iterative methods for computation of the value function and optimal control policies for the MDP. In problems with only absolutely continuous controls, estimates of the form in Lemma 2.5.1 are straightforward consequences of a contraction property that follows from the strictly positive discount factor in the cost (cf. Chapter 6 of [32]). However for singular control problems, due to the instantaneous nature of the control, such contraction estimates are typically unavailable. Here, once again, we use the special feature of the dynamics, which says that too much use of the singular control will rapidly bring the process to the boundary, in obtaining such an estimate. Finally, in Section 2.6 we present results from a numerical study of the algorithm.

Notation. The following notation will be used in this chapter. For a RCLL path $\{\xi(t)\}$, the jump at $t > 0$ will be denoted by $\delta\xi(t)$. As a convention we take $\delta\xi(0) \doteq \xi(0)$. For a sequence of random variables $\{\xi_n\}_{n\geq 0}$, we will use the notation $\delta\xi_n$ for the increment $\xi_{n+1} - \xi_n$.

2.2 Optimal Consumption and Portfolio Selection with Transaction Costs

We begin with a precise mathematical formulation of the optimal consumption-investment problem described in the previous section. Let $(\Omega, \mathcal{F}, I\!P)$ be a probability space on which is given a filtration $\{\mathcal{F}_t\}_{t\geq 0}$ satisfying the usual hypothesis. Let W be a real-valued $\{\mathcal{F}_t\}$-Brownian motion. We will denote the probability system $(\Omega, \mathcal{F}, I\!P, \{\mathcal{F}_t\}, W)$ by Φ. The Wiener process represents the source of uncertainty of the risky asset. The state process, which represents the wealth of the investor, is a controlled Markov process $Z \equiv (X, Y)$ given on the above probability system via the equations:

$$
\begin{aligned}
dX(t) &= (rX(t) - C(t))dt - (1+\lambda)dM(t) + (1-\mu)dN(t), \\
dY(t) &= bY(t)dt + \sigma Y(t)dW(t) + dM(t) - dN(t),
\end{aligned}
\tag{2.1}
$$

with initial condition $X(0-) = x, Y(0-) = y$, where $z \doteq (x, y) \in \mathbb{S}$. Here C is an $\{\mathcal{F}_t\}$-progressively measurable process such that for all $t \in [0, \infty)$, $C(t) \geq 0$ a.s. and $I\!E \int_0^t e^{-rs} C(s)ds < \infty$. Also, M and N are $\{\mathcal{F}_t\}$-adapted, non-decreasing, RCLL processes satisfying $M(0) \geq 0$ and $N(0) \geq 0$ a.s. The processes X and Y represent the amounts invested in the bond and the stock, respectively; $M(t), N(t)$ denote the cumulative purchases and sales of stock, respectively, over $[0, t]$. The process C represents the consumption of the investor. The processes C, M, and N are the control processes. Since M and N are not required to be absolutely continuous (with respect to the Lebesgue measure), they are referred to as singular controls. Denote by $\mathcal{A}(\Phi, z) \equiv \mathcal{A}(z)$ the set of "admissible controls", i.e. all $U \equiv (C, M, N)$ of the form described above. Let $\partial\mathbb{S}$ denote

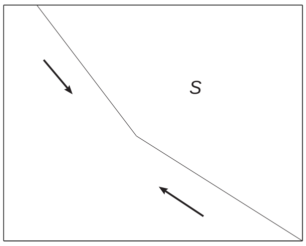

Figure 2.1: State space and singular control directions.

the boundary of $\mathbb{S}$. From the dynamical description of Z it follows that if $z \in \partial\mathbb{S}$ then the only control that keeps the investor solvent takes Z to the origin instantly and keeps it there at all times (see Figure 2.1).

Recall the utility function f of Section 2.1. Since $f(0) = 0$, one can reformulate the state constraint control problem on an infinite time horizon described in Section 2.1 to an exit time control problem, as follows. For $z \in \mathbb{S}$ and $U \in \mathcal{A}(z)$, let $\tau \equiv \tau(z, U)$ be defined as

$$\tau \doteq \inf\{t \in [0, \infty) : Z(t) \notin \mathbb{S}^o\}, \tag{2.2}$$

where Z is the controlled process corresponding to initial condition z and control U. Define the cost, $J(z, U)$, for using the control U by

$$J(z, U) \doteq I\!\!E \int_{[0,\tau)} e^{-\beta t} f(C(t))dt. \tag{2.3}$$

The value function of the control problem is then given by

$$V(z) = \sup_{\Phi} \sup_{U \in \mathcal{A}(z)} J(z, U), \tag{2.4}$$

where the outside supremum is over all probability systems Φ. The following will be a standing assumption in this chapter.

Condition 2.2.1 *For all $z \in \mathbb{S}$, $V(z) < \infty$ and $V : \mathbb{S} \to I\!\!R_+$ is a continuous map.*

We refer the reader to [28, 48, 50] for some sufficient conditions for the above assumption to hold.

State and Control Space Truncation. In order to develop numerical methods for computing $V(z)$, we will need to first approximate the control problem by an analogous control problem with a bounded state space and control set. We now present the convergence result which says that the value function of the "truncated control problem" converges to V as the truncation parameters approach their limits. We begin by considering the control space truncation.

For $p \in (0, \infty)$, let $\mathcal{A}_p(\Phi, z) \equiv \mathcal{A}_p(z)$ be the subset of $\mathcal{A}(z)$ consisting of $U = (C, M, N)$ which satisfy $0 \le C(t) \le p$, for all $t \ge 0$, a.s. Define $V_p(z)$ by replacing $\mathcal{A}(z)$ with $\mathcal{A}_p(z)$ in (2.4). The following is the first convergence result.

Proposition 2.2.2 *V_p converges to V, uniformly on compact subsets of $\mathbb{S}$, as $p \to \infty$.*

Proof. We first establish pointwise convergence, i.e. $V_p(z) \to V(z)$ as $p \to \infty$. Since $V_p(z) \le V(z)$, it suffices to show that, for all $z \in \mathbb{S}$,

$$\liminf_{p \to \infty} V_p(z) \ge V(z).$$

Fix $\epsilon > 0$ and choose an "ϵ-optimal control", i.e. $U_\epsilon \in \mathcal{A}(z)$ such that $V(z) - \epsilon < J(z, U_\epsilon)$. Suppose τ_ϵ is the associated exit time from $\mathbb{S}^o$. Define a control $\tilde{U}_p \equiv (\tilde{C}_p, \tilde{M}_p, \tilde{N}_p)$ by $\tilde{C}_p(t) \doteq C_\epsilon(t) \wedge p$, $\tilde{M}_p(t) \doteq M_\epsilon(t)$, $\tilde{N}_p(t) \doteq N_\epsilon(t)$, $t \ge 0$. It follows from the fact that $\tilde{C}_p \le C_\epsilon$ and standard comparison results for solutions of stochastic differential equations (cf. Proposition 5.2.18 of [29]) that the wealth process under control $\tilde{U}_p$ is never less than the wealth process under control U_ϵ. In particular, denoting by τ_p the exit time from $\mathbb{S}^o$ by the controlled process corresponding to the control $\tilde{U}_p$, we have $\tau_p \ge \tau_\epsilon$. Combining this with the observations that $\tilde{C}_p(t) \uparrow C_\epsilon(t)$ as $p \to \infty$ a.s. for all $t \ge 0$ and f is continuous, we have from Fatou's lemma

$$\liminf_{p\to\infty} J(z,\tilde{U}_p) \geq \liminf_{p\to\infty} \mathbb{E}\int_{[0,\tau_\epsilon)} e^{-\beta t} f(\tilde{C}_p(t))dt \ \geq \ \mathbb{E}\int_{[0,\tau_\epsilon)} e^{-\beta t} f(C_\epsilon(t))dt \ \geq V(z)-\epsilon.$$

Since $\epsilon > 0$ is arbitrary, the pointwise convergence of V_p to V follows. Next we show that for each p, V_p is lower semicontinuous (l.s.c.) Fix $z \in \mathbb{S}$ and let $\mathbb{S} \ni z_n \to z$ as $n \to \infty$. To prove that V_p is l.s.c. it suffices to show that

$$\liminf_{n\to\infty} V_p(z_n) \geq V_p(z). \tag{2.5}$$

Fix $\epsilon > 0$ and let $U_\epsilon = (C_\epsilon, M_\epsilon, N_\epsilon) \in \mathcal{A}_p(z)$ be an ϵ-optimal control, i.e. $V_p(z) - \epsilon < J(z, U_\epsilon)$. Let Z_ϵ be the controlled process according to U_ϵ and define τ_ϵ via (2.2) with Z replaced by Z_ϵ. Define $U_n \equiv (C_n, M_n, N_n)$ as $C_n \doteq C_\epsilon$, $M_n(t) \doteq M_\epsilon(t)1_{t<\tau_\epsilon} + M_n^* 1_{t\geq\tau_\epsilon}$, $N_n(t) \doteq N_\epsilon(t)1_{t<\tau_\epsilon} + N_n^* 1_{t\geq\tau_\epsilon}$, where $M_n^*, N_n^* \geq 0$ are chosen so that the controlled process Z_n corresponding to U_n and initial condition z_n satisfies $Z_n(\tau_\epsilon) \notin \mathbb{S}^o$. (Note, clearly $U_n \in \mathcal{A}_p(z_n)$.) This insures that $\tau_n \doteq \inf\{t : Z_n(t) \notin \mathbb{S}^o\}$ is at most τ_ϵ. Note that on the set $\{\tau_\epsilon = \infty\}$, we have $U_n(t) = U_\epsilon(t)$ for all $t \geq 0$. We claim that on the set $\{\tau_\epsilon < \infty\}$ we have $\liminf_{n\to\infty} \tau_n \geq \tau_\epsilon$ a.s., which implies $\tau_n \to \tau_\epsilon$ a.s. as $n \to \infty$ on the set $\{\tau_\epsilon < \infty\}$. To see the claim, suppose that $\liminf \tau_n < \tau_\epsilon - \delta$ for some $\delta > 0$. Then there exists $N_0 \geq 1$ such that $\tau_n < \tau_\epsilon - \delta/2$ for all $n \geq N_0$. Also, from the choice of the control U_n we see that, for all $\delta > 0$ and $L \in (0, \infty)$, $\sup_{0\leq t\leq(\tau_\epsilon-\delta/2)\wedge L} |Z_n(t) - Z(t)| \to 0$, in probability, as $n \to \infty$. Combining this with the fact that $Z_n(\tau_n) \notin \mathbb{S}^o$ we have that $Z(t) \notin \mathbb{S}^o$ for some $t \leq \tau_\epsilon - \delta/2$. However, this contradicts the definition of τ_ϵ. Thus we have shown $\tau_n \to \tau_\epsilon$ a.s. on the set $\{\tau_\epsilon < \infty\}$.

Next, recalling the choice of U_ϵ and that $C_\epsilon(t) = C_n(t)$ for all $t \geq 0$ on the set

$\{\tau_\epsilon = \infty\}$, we have

$$
\begin{aligned}
V_p(z) - V_p(z_n) &\leq J(z, U_\epsilon) - J(z_n, U_n) + \epsilon \\
&= I\!E\Big[1_{\{\tau_\epsilon<\infty\}} \int_{[\tau_n,\tau_\epsilon)} e^{-\beta t}(f(C_\epsilon(t)) - f(C_n(t)))dt\Big] + \epsilon \\
&\leq f_*(p) I\!E\Big[1_{\{\tau_\epsilon<\infty\}} \int_{[\tau_n,\tau_\epsilon)} e^{-\beta t} dt\Big] + \epsilon,
\end{aligned}
$$

where $f_*(p) \doteq \sup_{0\leq c\leq p} f(c) < \infty$. Since $\tau_n \to \tau_\epsilon$ a.s. on the set $\{\tau_\epsilon < \infty\}$, the first term on the right in the last line above approaches 0 as $n \to \infty$. The inequality (2.5) now follows from the above display on taking $n \to \infty$ and then $\epsilon \to 0$. Finally note that for each z, $V(z) - V_p(z) \downarrow 0$. The result now follows from Dini's Theorem (cf. Theorem M8 [4]). ∎

Next, we consider the truncation of the state space. The reduction will be achieved by replacing the original dynamical system given by (2.1) with one which evolves exactly as before in the interior of some compact domain but is instantaneously reflected back when the controlled process is about to exit the domain. The reflection mechanism is made precise via the notion of a Skorohod map. We begin with the following definition. Fix $\ell \in (0, \infty)$.

Definition 2.2.3 *Let $\phi \in \mathcal{D} \doteq D([0,\infty) : I\!R^2)$ be such that $\phi(0) \in (-\infty, \ell] \times (-\infty, \ell]$. We will denote the space of all such ϕ by $\mathcal{D}_0$. We say a pair $(\psi, \eta) \in \mathcal{D} \times \mathcal{D}$ solves the Skorohod problem (SP) for ϕ in $(-\infty, \ell]\times(-\infty, \ell]$, with normal reflection, if the following hold: (i) $\psi(0) = \phi(0)$. (ii) $\psi(t) = \phi(t) - \eta(t)$, $t \in (0, \infty)$. (iii) $\psi(t) \in (-\infty, \ell] \times (-\infty, \ell]$ for all $t \geq 0$. (iv) $\eta(\cdot)$ is componentwise nondecreasing. (v) $\eta_i(t) = \int_{(0,t]} 1_{\{\psi_i(t)=\ell\}} d\eta_i(t)$, $i = 1, 2$, where $\eta(t) = (\eta_1(t), \eta_2(t))'$, $\psi(t) = (\psi_1(t), \psi_2(t))'$.*

It is well known (cf. [16], [22]) that for every $\phi \in \mathcal{D}_0$, there is a unique solution (ψ, η) to the above SP. We will write $\psi = \Gamma(\phi)$ and refer to the map $\Gamma : \mathcal{D}_0 \to \mathcal{D}_0$ as the Skorohod map. The following Lipschitz property (cf. [16]) is quite useful in various estimates.

There exists $\kappa \in (0,\infty)$, independent of ℓ, such that, for all $\phi_1, \phi_2 \in \mathcal{D}_0$,

$$|\Gamma(\phi_1) - \Gamma(\phi_2)|_T^* \leq \kappa|\phi_1 - \phi_2|_T^*, \; T \in (0,\infty). \tag{2.6}$$

We will now introduce the modified constrained dynamics of the controlled Markov process. Set $\mathbb{S}_\ell \doteq \mathbb{S} \cap (-\infty,\ell] \times (-\infty,\ell]$. Let $Z_\ell \equiv (X_\ell, Y_\ell)$ solve the following system of equations:

$$\begin{aligned} dX_\ell(t) &= (rX_\ell(t) - C(t))dt - (1+\lambda)dM(t) + (1-\mu)dN(t) - dR_1(t), \\ dY_\ell(t) &= bY_\ell(t)dt + \sigma Y_\ell(t)dW(t) + dM(t) - dN(t) - dR_2(t), \end{aligned} \tag{2.7}$$

where $Z_\ell(0-) = z$, $U \equiv (C, M, N) \in \mathcal{A}_p(z)$, $z = (x,y) \in \mathbb{S}_\ell$ and $R = (R_1, R_2)'$ is a componentwise nondecreasing, RCLL, $\mathcal{F}_t$-adapted process satisfying

$$\int_0^\infty 1_{\{X_\ell(t)<\ell\}} dR_1(t) = 0, \quad \int_0^\infty 1_{\{Y_\ell(t)<\ell\}} dR_2(t) = 0. \tag{2.8}$$

The unique solvability of (2.7) and (2.8) follows from the Lipschitz continuity property (2.6) of the Skorohod map and the usual Picard iteration method. Define τ_ℓ and $J_\ell(z,U)$ as in (2.2) and (2.3) with Z replaced by Z_ℓ in (2.2) and τ replaced by τ_ℓ in (2.3). Define $V_{\ell,p}$ as

$$V_{\ell,p}(z) = \sup_{\Phi} \sup_{U \in \mathcal{A}_p(\Phi,z)} J_\ell(z,U). \tag{2.9}$$

The following is the second convergence result of this section.

Proposition 2.2.4 *For all $p \in (0,\infty)$, $V_{\ell,p}$ converges to V_p, uniformly on compact subsets of $\mathbb{S}$, as $\ell \to \infty$.*

Proof. Let $Z \equiv (X,Y)$ be as in (2.1) and τ as in (2.2), with $C \equiv 0$. It is easy to check that for each $T \in (0,\infty)$ and compact subset $\mathbb{S}_0 \subset \mathbb{S}$, there exists $\Lambda \equiv \Lambda(T) \in (0,\infty)$

such that

$$\sup_{\Phi} \sup_{(M,N)} \sup_{z\in\mathbb{S}_0} I\!\!E \sup_{0\le t\le T\wedge\tau} (X^+(t)+Y^+(t)) \le \Lambda,$$

where the supremum is taken over all $\{\mathcal{F}(t)\}$-adapted, nondecreasing, RCLL processes M and N such that $M(0) \ge 0$, $N(0) \ge 0$; and over all systems Φ. Thus in particular we have that

$$\sup_{\ell} \sup_{\Phi} \sup_{U\in\mathcal{A}_p(\Phi,z)} \sup_{z\in\mathbb{S}_0} I\!\!E \sup_{0\le t\le T\wedge\tau_\ell} (X_\ell^+(t)+Y_\ell^+(t)) \le \Lambda, \tag{2.10}$$

where $Z_\ell \equiv (X_\ell, Y_\ell)$ are as defined in (2.7), and τ_ℓ is as introduced below (2.8).

Fix $\delta > 0$. Let $z \in \mathbb{S}_0$ and $\epsilon > 0$ be arbitrary. Let Φ and $U \in \mathcal{A}_p(z,\Phi)$ be such that $V_{\ell,p}(z) \le J_\ell(z,U) + \epsilon$. Choose $T \in (0,\infty)$ such that $f_*(p)e^{-\beta T}/T < \epsilon$. Then $V_{\ell,p}(z) \le I\!\!E \int_0^{T\wedge\tau_\ell} e^{-\beta t} f(C(t))dt + 2\epsilon$.

Choose $\ell_0 \equiv \ell_0(\delta)$ such that $\ell_0 > (\Lambda f_*(p))/(\delta\beta)$. Define

$$A_{\ell_0} \doteq \{\omega : \sup_{0\le t\le T\wedge\tau_\ell} (X_\ell^+(t)+Y_\ell^+(t)) > \ell_0\}.$$

Then

$$I\!\!E \int_0^{T\wedge\tau_\ell} e^{-\beta t} f(C(t))dt = I\!\!E[1_{A_{\ell_0}} \int_0^{T\wedge\tau_\ell} e^{-\beta t} f(C(t))dt] + I\!\!E[1_{A_{\ell_0}^c} \int_0^{T\wedge\tau_\ell} e^{-\beta t} f(C(t))dt]. \tag{2.11}$$

It follows from Markov's inequality and (2.10) that $I\!\!P[A_{\ell_0}] \le \Lambda/\ell_0$. Thus the first integral on the right side of (2.11) is bounded by $(f_*(p)/\beta) I\!\!P[A_{\ell_0}] \le \delta$. Next, for $\ell \ge \ell_0$, on the set $A_{\ell_0}^c$, $Z_\ell(\cdot\wedge T\wedge\tau_\ell) = Z(\cdot\wedge T\wedge\tau_\ell)$. In particular, $T\wedge\tau \ge T\wedge\tau_\ell$. Thus

$$I\!\!E[1_{A_{\ell_0}^c} \int_0^{T\wedge\tau_\ell} e^{-\beta t} f(C(t))dt] \le I\!\!E[\int_0^{T\wedge\tau} e^{-\beta t} f(C(t))dt] \le V_p(z).$$

Combining the above bounds, we have $V_{\ell,p}(z) \le V_p(z) + \delta + 2\epsilon$. Since $\epsilon > 0$ is arbitrary, we have that, for all $\ell \ge \ell_0$ and $z \in \mathbb{S}_0$, $V_{\ell,p}(z) \le V_p(z) + \delta$. It is easily seen that the roles of V_p and $V_{\ell,p}$ can be interchanged in the above argument. Thus we have that, given

$\delta > 0$, there exists an ℓ_0 such that $|V_{\ell,p}(z) - V_p(z)| \le \delta$ if $\ell > \ell_0$, for all $z \in \mathbb{S}_0$. Since $\mathbb{S}_0$ is an arbitrary compact subset of $\mathbb{S}$, the result follows. ■

Corollary 2.2.5 *For all $z \in \mathbb{S}$, $\lim_{p\to\infty} \lim_{\ell\to\infty} V_{\ell,p}(z) = V(z)$.*

2.3 An Approximating Markov Decision Problem

In this section we will present the Markov decision problem whose value function approximates $V_{\ell,p}$. Since throughout this section ℓ, p will be fixed, we will drop them from the notations: $V_{\ell,p}, \tau_\ell, J_\ell, \mathcal{A}_p(z)$ and $Z_\ell \equiv (X_\ell, Y_\ell)$. We will introduce a discrete time, discrete state controlled Markov chain to approximate the continuous time process given by (2.7).

Fix $h > 0$ and define the two-dimensional h-grid, $L^h \doteq \{(jh, kh) : -\infty < j, k < +\infty\}$. The symbol h denotes the approximation parameter and as h approaches 0, a suitable interpolation of the controlled Markov chain, to be introduced below, "approaches" a controlled diffusion process of the form in (2.7). We will assume for simplicity that ℓ is an integer multiple of h.

A natural definition of the state space for the approximating chain is $\mathbb{S}_\ell^h \doteq \mathbb{S}_\ell \cap L^h$. However, due to reflection terms in the dynamics of the controlled process, it is convenient to consider a slightly "enlarged" state space, namely, $\mathbb{S}_\ell^{h+} \doteq \mathbb{S}_{\ell+h} \cap L^h$. The "solvency boundary" of the space $\mathbb{S}_\ell^{h+}$ is defined as

$$\partial^h \doteq \{(x, y) \in \mathbb{S}_\ell^{h+} : x + (1+\lambda)y \le h(1+\lambda) \text{ or } x + (1-\mu)y \le h\}.$$

The points $(x, y) \in \mathbb{S}_\ell^{h+}$ for which $x = \ell + h$ or $y = \ell + h$ form the reflecting boundary, ∂_{R}^h.

Let $\{Z_n^h, n = 0, 1, 2, \ldots\}$ be a discrete time controlled Markov chain with state space $\mathbb{S}_\ell^{h+}$, with $Z_n^h = (X_n^h, Y_n^h)$. The transition probabilities will be defined so that the chain's

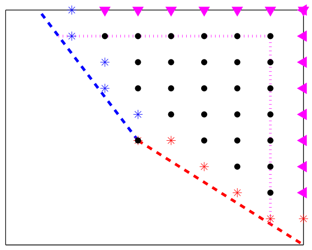

Figure 2.2: The discrete state space $\mathbb{S}_\ell^{h+}$.

evolution law well approximates the local behavior of the controlled diffusion (2.7). For each n, the increments of the chain δZ_n^h will approximate exactly one of the following dynamical descriptions:

- "Controlled diffusion step": $(rX_t - C_t, bY_t)'dt + (0,\sigma)'dW_t$.
- "Purchase control step": $(-(1+\lambda), 1)'dM_t$.
- "Sales control step": $(1-\mu, -1)'dN_t$.
- "Reflection step": dR_t.

Each of these steps is described precisely in what follows. We also introduce a family of "interpolation intervals" $\{\Delta^h, h > 0\}$ used in defining the approximating cost function and in the convergence arguments. For each pair $(z, c) \in \mathbb{S}_\ell^{h+} \times [0, p]$ we first define a family $\tilde{\Delta}^h(z, c)$. For the controlled diffusion steps, if the state of the chain is z and the exercised consumption control is c, Δ^h will be taken to be $\tilde{\Delta}^h(z, c)$; whereas for singular control steps and reflection steps Δ^h will be taken to be 0. This reflects the fact that for the controlled diffusion (2.7), reflection and singular control terms can change the state instantaneously. Suitable conditions on $\tilde{\Delta}^h(z, c)$ in order to obtain convergence of the continuous time interpolated processes to corresponding controlled diffusions are introduced below.

Controlled Diffusion Steps and Local Consistency. By a controlled diffusion step we mean that the Markov chain evolves according to a transition law which is "locally consistent" in the sense of [32], with a (controlled) diffusion given as:

$$d\tilde{X}(t) = (r\tilde{X}(t) - C(t))dt, \; d\tilde{Y}(t) = b\tilde{Y}(t)dt + \sigma\tilde{Y}(t)dW(t).$$

Formally, given $h > 0$, we choose for each $c \in [0,p]$ and $z \in \mathbb{S}_{\ell}^{h+} \setminus \partial^h$ a probability measure $q_h^{(0)}(z,c,d\tilde{z})$ on L^h along with an interpolation interval $\tilde{\Delta}^h(z,c)$ which satisfy the following local consistency conditions for some $\rho > 0$:

$$m_0(z,c) \doteq \int_{L^h} (\tilde{z} - z) q_h^{(0)}(z,c,d\tilde{z}) = \begin{pmatrix} rx - c \\ by \end{pmatrix} \tilde{\Delta}^h(z,c) + O(h^\rho \tilde{\Delta}^h(z,c)), \tag{2.12}$$

$$\begin{aligned} \sigma_0(z,c) &\doteq \int_{L^h} (\tilde{z} - z - m_0(z,c))(\tilde{z} - z - m_0(z,c))' q_h^{(0)}(z,c,d\tilde{z}) \\ &= \begin{pmatrix} 0 & 0 \\ 0 & |\sigma y|^2 \end{pmatrix} \tilde{\Delta}^h(z,c) + O(h^\rho \tilde{\Delta}^h(z,c)). \end{aligned} \tag{2.13}$$

In the above displays $\tilde{z} = (\tilde{x}, \tilde{y})$ and throughout, by the symbol $O(k)$ we will mean an expression which is bounded above by $\alpha|k|$ where α is a constant depending only on the coefficients of the model and the truncation parameters ℓ, p. In addition we assume that there exists $\zeta \in (0,\infty)$ such that $q_h^{(0)}(z,c,B_{\zeta h}(z)) = 1$ for all $c \in [0,p]$ and $h > 0$, where $B_{\zeta h}(z)$ is a ball of radius ζh centered at z. The interpolation intervals are required to satisfy:

$$\tilde{\Delta}_*^h \doteq \sup_{z,c} \tilde{\Delta}^h(z,c) \to 0 \text{ as } h \to 0, \;\; \inf_{z,c} \tilde{\Delta}^h(z,c) > 0 \text{ for each } h > 0, \tag{2.14}$$

where the sup and inf in the above displays are taken over all $(z,c) \in \mathbb{S}_{\ell}^{h+} \times [0,p]$. For the sake of specificity we make the following choice for $q_h^{(0)}$. Let $Q(x,y) \equiv Q^h(x,y) \doteq$

$hr|x| + hp + hb|y| + \sigma^2 y^2$. Define for all $(x,y) \in \mathbb{S}_\ell^{h+} \setminus \partial^h$:

$$
\begin{aligned}
q_h^{(0)}((x,y),c,(x+h,y)) &\doteq \frac{hrx^+}{Q(x,y)} \quad , \quad q_h^{(0)}((x,y),c,(x-h,y)) \doteq \frac{hrx^- + hc}{Q(x,y)}, \\
q_h^{(0)}((x,y),c,(x,y+h)) &\doteq \frac{hby^+ + \frac{1}{2}\sigma^2 y^2}{Q(x,y)} \quad , \quad q_h^{(0)}((x,y),c,(x,y-h)) \doteq \frac{hby^- + \frac{1}{2}\sigma^2 y^2}{Q(x,y)}, \\
q_h^{(0)}((x,y),c,(x,y)) &\doteq \frac{h(p-c)}{Q(x,y)}, \\
\tilde{\Delta}^h(z,c) &\doteq \frac{h^2}{Q(x,y)}.
\end{aligned}
\tag{2.15}
$$

It is easy to check that $q_h^{(0)}, \tilde{\Delta}^h$ defined above satisfy (2.12), (2.13) and (2.14).

Singular Control Steps. The singular control terms in the controlled diffusion are the nondecreasing RCLL processes M and N. The process M pushes the state process in the direction $v_1 = (-(1+\lambda), 1)'$, whereas N pushes the state process in the direction $v_2 = ((1-\mu), -1)'$. For the approximating chain we will assume that at most one among the sales control and purchase control are exercised at any given time instant and the magnitude of the corresponding displacement is $O(h)$. In order to capture the "singular" behavior of the limit diffusion — namely the feature that the state process can instantaneously be displaced by large amounts — we will take the interpolation interval for all singular control steps in the approximating chain to be 0.

In order to obtain weak convergence of the interpolated chain to the controlled diffusion, we need to ensure that the control directions match asymptotically those for the physical problem. More precisely, given $h > 0$ we define for each $z \in \mathbb{S}_\ell^{h+}$ two probability measures $q_h^{(i)}(z, d\tilde{z}), i = 1, 2$, on L^h as follows. For states $(x,y) \in \mathbb{S}_\ell^{h+} \setminus \partial^h$:

$$q_h^{(1)}((x,y),(x-h,y)) = \lambda/(\lambda+1) \quad , \quad q_h^{(1)}((x,y),(x-h,y+h)) = 1/(\lambda+1); \tag{2.16}$$

$$q_h^{(2)}((x,y),(x,y-h)) = \mu \quad , \quad q_h^{(2)}((x,y),(x+h,y-h)) = 1-\mu. \tag{2.17}$$

It is easy to check that $q_h^{(1)}$ and $q_h^{(2)}$ introduced above satisfy the following consistency conditions:

$$m_i(z) \doteq \int_{L^h} (\tilde{z} - z) q_h^{(i)}(z, d\tilde{z}) = hv_i, \tag{2.18}$$

$$\sigma_i(z) \doteq \int_{L^h} (\tilde{z} - z - m_i(z))(\tilde{z} - z - m_i(z))' q_h^{(i)}(z, d\tilde{z}) = O(h^2). \tag{2.19}$$

Reflection Steps. We will define a transition kernel that with probability 1 moves a state in ∂_{R}^h to some state in $\mathbb{S}_\ell^h$. Once more, since reflection in the diffusion control problem occurs instantaneously, we take the interpolation interval at reflection steps to be 0. Since the directions of reflection in the diffusion control problem are normal, a natural choice of the transition kernel for reflection step is as follows for $z \in \partial_{\mathrm{R}}^h$:

$$q_h^{(3)}((\ell+h, y), (\ell, y)) = 1, \;\; q_h^{(3)}((x, \ell+h), (x, \ell)) = 1, \;\; q_h^{(3)}((\ell+h, \ell+h), (\ell, \ell)) = 1. \tag{2.20}$$

For $z \notin \partial_{\mathrm{R}}$, $q_h^{(3)}(z, \cdot)$ can be defined arbitrarily. It will be seen from the definition of admissible controls given below that for such states the definition of $q_h^{(3)}$ is immaterial.

The Controlled Markov Chain. As described above, the control at each step is first specified by the choice of an action: controlled diffusion, singular control, or reflection. Therefore, we define a sequence of control actions $\{I_n^h, n = 0, 1, 2, \ldots\}$ with $I_n^h = 0, 1, 2, 3$ if the nth step in the chain is a controlled diffusion step, purchase control step, sales control step, or a reflection step, respectively. In the case of a controlled diffusion step, the magnitude of the consumption control must also be specified. Consequently, the space of controls is given by $\mathcal{U} \doteq \{0, 1, 2, 3\} \times [0, p]$.

The probability measures associated with each of the control actions will now be combined into a single probability measure for use in defining the controlled Markov chain. For each $z \in \mathbb{S}_\ell^{h+} \setminus \partial^h, u \in \mathcal{U}$ $(u = (i, c))$, we define a probability measure

$p_h(z,u,d\tilde{z})$ on L^h by:

$$p_h(z,u,d\tilde{z}) = q_h^{(0)}(z,c,d\tilde{z})\, 1_{\{i=0\}} + q_h^{(i)}(z,d\tilde{z})\, 1_{\{i\in\{1,2,3\}\}}. \tag{2.21}$$

The definition of the transition function for $z \in \partial^h$ is not important since in the analysis of the control problem the chain will be stopped the first time it hits ∂^h. For sake of specificity we set $p_h(z,u,z) = 1$ for all $z \in \partial^h$ and $u \in \mathcal{U}$.

We are now ready to specify the controlled Markov chains. Given a sequence $U^h = \{U_n^h, n = 0,1,2,\ldots\}$ (where $U_n^h = (I_n^h, C_n^h)$) of $\mathcal{U}$-valued random variables we construct a controlled Markov chain $\{Z_n^h, n = 0,1,2,\ldots\}$ with initial condition $z_h = (x_h, y_h) \in \mathbb{S}_\ell^{h+}$ and state space $\mathbb{S}_\ell^{h+}$, as follows:

$$Z_0^h = z_h, \;\; I\!P[Z_{n+1}^h \in E|\mathcal{F}_n^h] = p_h(Z_n^h, U_n^h, E), \;\; n \geq 0, \;\; E \in \mathcal{B}(\mathbb{S}_\ell^{h+}), \tag{2.22}$$

where $\mathcal{F}_n^h = \sigma\{Z_0^h, \ldots, Z_n^h, U_0^h, \ldots, U_n^h\}$. The following definition of admissible controls ensures that $Z_n^h \in \mathbb{S}_\ell^{h+}$ for all n and so the definition in (2.22) is meaningful.

Definition 2.3.1 *The control sequence $U^h = \{U_n^h, n = 0,1,2,\ldots\}$ is said to be admissible for the initial condition z_h and $\{Z_n^h\}$ ($\{Z_n^h, U_n^h\}$) is called the corresponding controlled Markov chain (resp. controlled pair) if:*

1. *U_n^h is $\sigma\{Z_0^h, \cdots, Z_n^h, U_0^h, \cdots, U_{n-1}^h\}$-adapted.*
2. *$I\!P[I_n^h = 3|Z_n^h \in \mathbb{S}_\ell^h] = 0$ and $I\!P[I_n^h = 3|Z_n^h \in \partial_R^h \setminus \partial^h] = 1$ for all n.*
3. *Condition (2.22) holds.*

The class of all admissible control sequences for initial state z_h will be denoted by $\mathcal{A}^h(z_h)$.

We also define for each $z \in \mathbb{S}_\ell^{h+}$ and $u = (i, c) \in \mathcal{U}$ the interpolation intervals

$$\Delta^h(z, u) = \tilde{\Delta}^h(z, c)\, 1_{\{i=0\}}. \tag{2.23}$$

For an admissible pair $\{Z_n^h, U_n^h\}$, we denote the associated sequence of interpolation intervals $\Delta^h(Z_n^h, U_n^h)$ by $\{\Delta_n^h, n = 0, 1, 2, \ldots\}$. Define, $t_0^h \doteq 0$ and $t_n^h \doteq \sum_{i=0}^{n-1} \Delta_i^h$ for $n \geq 1$.

Markov Decision Problem (MDP) for the Chain. Given an admissible pair $\{Z_n^h, U_n^h\}$ let $\eta_h \doteq \inf\{n : Z_n^h \in \partial^h\}$. The cost function for the controlled Markov chain is defined as:

$$J^h(z_h, U^h) = \mathbb{E} \sum_{n=0}^{\eta_h - 1} e^{-\beta t_n^h} f(C_n^h) \Big(\frac{1 - e^{-\beta \Delta_n^h}}{\beta} \Big). \tag{2.24}$$

Note that we have used the factor $(1 - e^{-\beta \Delta_n^h})/\beta$ rather than the more intuitive (and asymptotically equivalent) Δ_n^h. This somewhat simplifies the convergence proofs without affecting the limiting results. The value function of the MDP is defined as:

$$V^h(z_h) = \sup_{U^h \in \mathcal{A}^h(z_h)} J^h(z_h, U^h). \tag{2.25}$$

Continuous Time Interpolation. One of the main goals of the study is to show that the value function of the MDP defined in (2.25) converges, as $h \to 0$, to the value function of the limit diffusion control problem. This convergence result allows the computation of near optimal policies for the diffusion control problem introduced below (2.6) by numerically solving the above MDP. We next introduce the continuous time interpolation and time rescaling techniques that will be used in the proof of our main convergence result.

The continuous time interpolations of various processes will be constructed to be piecewise constant on the time intervals $[t_n^h, t_{n+1}^h)$, $n \geq 0$. For use in this construction we define $n^h(t) \doteq \max\{n : t_n^h \le t\}$, $t \ge 0$. Note that $n^h(t)$ is an $\{\mathcal{F}_n^h\}$-stopping time. Setting $\mathcal{F}^h(t) \doteq \mathcal{F}^h_{n^h(t)}$ we obtain a continuous time filtration $\{\mathcal{F}^h(t), t \ge 0\}$. Define $U^h(t) \doteq U^h_{n^h(t)}, t \ge 0$. Also, define the continuous time processes associated with the controlled diffusion steps as follows. First let $B_0^h = 0$ and $S_0^h = 0$ and define for $n \ge 1$,

$$B_n^h \doteq \sum_{k=0}^{n\wedge\eta_h - 1} I\!E[\delta Z_k^h | \mathcal{F}_k^h]\, 1_{\{I_k^h = 0\}}, \qquad S_n^h \doteq \sum_{k=0}^{n\wedge\eta_h - 1} \left(\delta Z_k^h - I\!E[\delta Z_k^h|\mathcal{F}_k^h]\right) 1_{\{I_k^h=0\}}. \tag{2.26}$$

Define the continuous time process B^h by setting $B^h(0) \doteq 0$ and $B^h(t) \doteq B^h_{n^h(t)}$ for $t > 0$. The process S^h is defined in a similar manner. We define the interpolations associated with the purchase control and sales control as follows. Let $M_0^h = 0, N_0^h = 0$, $E_{i,0}^h = 0$, $i = 1, 2$ and define for $n \ge 1$:

$$M_n^h \doteq \sum_{k=0}^{n\wedge\eta_h-1} h\, 1_{\{I_k^h=1\}}, \quad N_n^h \doteq \sum_{k=0}^{n\wedge\eta_h-1} h\, 1_{\{I_k^h=2\}}, \quad E_{i,n}^h \doteq \sum_{k=0}^{n\wedge\eta_h-1} (\delta Z_k^h - hv_i)\, 1_{\{I_k^h=i\}}.$$

The continuous time processes M^h and N^h are defined as $M^h(0) \doteq 0, N^h(0) \doteq 0$ and $M^h(t) \doteq M^h_{n^h(t)}, N^h(t) \doteq N^h_{n^h(t)}$ for $t \ge 0$. The processes E_1^h and E_2^h are defined analogously. The continuous time process associated with reflection is defined as follows. If $n^h(t) = 0$ define $R^h(t) = 0$; otherwise let

$$R^h(t) \doteq -\sum_{k=0}^{n^h(t)-1} \delta Z_k^h\, 1_{\{I_k^h=3\}}. \tag{2.27}$$

We define the continuous time interpolation Z^h of the controlled Markov chain Z_n^h introduced in Definition 2.3.1 by $Z^h(t) \doteq Z^h_{n^h(t)}$, $t \ge 0$. The following representation for $Z^h(t)$ is easily verified:

$$Z^h(t) = z_h + B^h(t) + S^h(t) + v_1 M^h(t) + v_2 N^h(t) + E_1^h(t) + E_2^h(t) - R^h(t),\ t \ge 0. \tag{2.28}$$

Also, it follows from condition (2.12) that on the set $\{I_n^h = 0, \eta_h > n\}$,

$$\mathbb{E}[\delta Z_n^h | \mathcal{F}_n^h] = \begin{pmatrix} rX_n^h - C_n^h \\ bY_n^h \end{pmatrix} \Delta^h(Z_n^h, 0, C_n^h) + O(h^\rho \Delta^h(Z_n^h, 0, C_n^h)) \text{ a.s.}$$

This fact, together with the piecewise constant nature of the processes, yields

$$B^h(t) = \int_0^{t\wedge\tau^h} \begin{pmatrix} rX^h(s) - C^h(s) \\ bY^h(s) \end{pmatrix} ds + \delta_1^h(t), \tag{2.29}$$

where $\tau^h \doteq t^h_{\eta_h}$ and δ_1^h is an $\{\mathcal{F}^h(t)\}$-adapted process which, in view of (2.14), satisfies for all $t \geq 0$ and $m \geq 1$

$$\sup_{0\leq s\leq t} \mathbb{E}|\delta_1^h(s)|^m \to 0 \text{ as } h \to 0.$$

A similar calculation gives the following representation of the cost function (2.24):

$$J^h(z_h, U^h) = \mathbb{E}\int_{[0,\tau^h]} e^{-\beta t} f(C^h(t))dt. \tag{2.30}$$

Time Rescaling. A common approach for proving the convergence of V^h to V as $h \to 0$ is to begin by showing that the collection $\{(Z^h(\cdot), \tau^h), h \geq 0\}$ is tight and then characterize the subsequential weak limits suitably. However, for problems with singular controls, showing the tightness of the above family becomes problematic since, in general, the processes $\{(M^h(\cdot), N^h(\cdot)), h \geq 0\}$ may fail to be tight. A powerful method for handling this tightness issue was introduced by Kushner and Martins [37]. The basic idea is to suitably stretch out the time scale so that the various processes involved in the convergence analysis, in the new time scale, are tight; carry out the weak convergence analysis with the rescaled processes; and finally revert back to the original time scale to argue the convergence of V^h to V.

We now introduce the time rescaling that will be used in our study. The rescaled time increments, $\{\hat{\Delta}_n^h, n = 0, 1, 2, \ldots\}$, are defined as $\hat{\Delta}_n^h \doteq \Delta_n^h 1_{\{I_n^h=0\}} + h\, 1_{\{I_n^h \in \{1,2\}\}}$. Define $\hat{t}_0^h \doteq 0$ and $\hat{t}_n^h \doteq \sum_{i=0}^{n-1} \hat{\Delta}_i^h$ for $n \geq 1$.

Definition 2.3.2 *The rescaled time process $\hat{T}^h(t)$ is the unique continuous nondecreasing process satisfying: (1) $\hat{T}^h(0) = 0$; (2) the derivative of $\hat{T}^h(t)$ is 1 for $t \in (\hat{t}_n^h, \hat{t}_{n+1}^h)$ if $I_n^h = 0$; (3) the derivative of $\hat{T}^h(t)$ for $t \in (\hat{t}_n^h, \hat{t}_{n+1}^h)$ is 0 if $I_n^h = 1, 2, 3$.*

It is easy to check that $\hat{T}^h(\hat{t}_n^h) = t_n^h$ and that $\hat{T}^h(\hat{t}_{n+1}^h) - \hat{T}^h(\hat{t}_n^h) = \Delta_n^h$. Let $\hat{n}^h(t) \doteq \max\{n : \hat{t}_n^h \leq t\}$, $t \geq 0$. Using the observation that every reflection step must be followed by either a singular control step or a diffusion control step, it follows that $\hat{n}^h(t)$ is a bounded $\{\mathcal{F}_n^h\}$-stopping time, with bound

$$\hat{n}^h(t) \leq 2\Big(\frac{t}{h} + \frac{t}{\inf_{z,c} \hat{\Delta}^h(z, 0, c)}\Big) < \infty. \tag{2.31}$$

Define the continuous time filtration $\{\hat{\mathcal{F}}^h(t), t \geq 0\}$ by setting $\hat{\mathcal{F}}^h(t) \doteq \mathcal{F}_{\hat{n}^h(t)}$.

The rescaled processes (denoted with a ˆ) are defined in a manner similar to the processes defined below (2.26) with appropriate adjustments to the time variable. For example, we define $\hat{B}^h(0) = 0$ and $\hat{B}^h(t) \doteq B_{\hat{n}^h(t)}^h$ if $\hat{n}^h(t) > 0$. We define the processes $\hat{U}^h(t)$, $\hat{S}^h(t)$, $\hat{M}^h(t)$, $\hat{N}^h(t)$, $\hat{E}_1^h(t)$, $\hat{E}_2^h$, $\hat{R}^h(t)$, $\hat{Z}^h(t)$ analogously (that is, by replacing $n^h(t)$ with $\hat{n}^h(t)$ in the definitions below (2.26)). Then we have the following rescaled version of (2.28)

$$\hat{Z}^h(t) = z_h + \hat{B}^h(t) + \hat{S}^h(t) + v_1 \hat{M}^h(t) + v_2 \hat{N}^h(t) + \hat{E}_1^h(t) + \hat{E}_2^h - \hat{R}^h(t). \tag{2.32}$$

Remark 2.3.3 *From the definition of $\hat{T}^h(t)$ if follows that $\hat{n}^h(t) = n^h(\hat{T}^h(t))$. This equality yields a straightforward relationship between the original interpolated processes and the rescaled processes. For example, $\hat{B}^h(t) = B^h(\hat{T}^h(t))$. Similar equations hold between $U^h(t)$, $S^h(t)$, $M^h(t)$, $N^h(t)$, $E_1^h(t)$, E_2^h, $R^h(t)$, $Z^h(t)$ and their corresponding rescaled versions.*

Using the fact that $\hat{T}^h(\hat{t}^h_{n+1}) - \hat{T}^h(\hat{t}^h_n) = \Delta^h_n$ which is 0 for singular control and reflection steps, a calculation similar to that which produced (2.29) yields

$$\hat{B}^h(t) = \int_0^{t\wedge\hat{\tau}^h_1} \begin{pmatrix} r\hat{X}^h(s) - \hat{C}^h(s) \\ b\hat{Y}^h(s) \end{pmatrix} d\hat{T}^h(s) + \hat{\delta}^h_1(t), \tag{2.33}$$

where $\hat{\tau}^h_1 \doteq \inf\{t : \hat{Z}^h(t) \in \partial^h\}$ and $\hat{\delta}^h_1$ is an $\{\hat{\mathcal{F}}^h(t)\}$-adapted process satisfying for all $m \geq 1$,

$$I\!\!E \sup_{0\leq s\leq t} |\hat{\delta}^h_1(s)|^m \to 0 \text{ as } h \to 0. \tag{2.34}$$

We now state several lemmas related to the time rescaling. The following "change of variables" formula (cf. Theorem IV.3.45 [43]) will be used several times in our analysis.

Lemma 2.3.4 *Let $\hat{G} : [0,\infty) \to [0,\infty)$ be a continuous and nondecreasing function. Suppose that $\hat{G}(t) \to \infty$ as $t \to \infty$. Define the inverse $G : [0,\infty) \to [0,\infty)$ as $G(t) = \inf\{s : \hat{G}(s) > t\}$. Then for all bounded and measurable functions $g : [0,\infty) \to [0,\infty)$,*

$$\int_{[0,G(t)]} g(s) d\hat{G}(s) = \int_{[0,t]} g(G(s))ds. \tag{2.35}$$

The following lemma is at the heart of the time transformation idea. It ensures that the weak limits of $\hat{T}^h(t)$ increase to ∞ as $t \to \infty$ and thus makes the reverting back to the original time scale, in the limit, possible (see Theorem 2.4.6).

Lemma 2.3.5 *Let $\{U^h_n, n = 0, 1, 2, \ldots\}_{h>0}$ be a family of admissible control sequences. Then for all $t \geq 0$*

$$\sup_h I\!\!E|M^h(t) + N^h(t)| < \infty. \tag{2.36}$$

Proof. Without loss of generality, assume $h \in (0, 1)$. Define

$$Y^h_i(t) \doteq \sum_{k=0}^{n^h(t)\wedge\eta_h-1} \delta Z^h_k \, 1_{\{I^h_k=i\}} \ , n^h_i(t) \doteq \sum_{k=0}^{n^h(t)\wedge\eta_h-1} 1_{\{I^h_k=i\}}, \ i = 1, 2.$$

Writing $Y_i^h \equiv (Y_{i,1}^h, Y_{i,2}^h)'$, it follows from (2.16) and (2.17) that,

$$\mathbb{E}Y_{1,2}^h(t) = h\frac{1}{1+\lambda}\mathbb{E}[n_1^h(t)], \quad \mathbb{E}Y_{2,1}^h(t) = h(1-\mu)\mathbb{E}[n_2^h(t)].$$

A straightforward calculation shows $|B^h(t)| \leq c_1(1+t)$ and $\mathbb{E}|S^h(t)| \leq c_2(1+t)$, where the constants c_1, c_2 are independent of h and t. From (2.16), (2.17) we see that $hn_1^h(t) = M^h(t) = Y_{1,1}^h(t)$ and $hn_2^h(t) = N^h(t) = Y_{2,2}^h(t)$. Thus from (2.28) there is $\tilde{c}_1 \in (0,\infty)$ such that

$$hn_1^h(t) \leq \tilde{c}_1(1+t) + |S_1^h(t)| + Y_{2,1}^h(t), \; hn_2^h(t) \leq \tilde{c}_1(1+t) + |S_2^h(t)| + Y_{1,2}^h(t). \qquad (2.37)$$

Combining the above inequalities we have, for some $c_3 \in (0,\infty)$, $h\mathbb{E}[n_1^h(t)] \leq c_3(1+t) + h(1-\mu)\mathbb{E}[n_2^h(t)]$ and $h\mathbb{E}[n_2^h(t)] \leq c_3(1+t) + h\mathbb{E}[n_1^h(t)]/(1+\lambda)$. It follows that $h\mathbb{E}[n_1^h(t)]$ and $h\mathbb{E}[n_2^h(t)]$ are "close" to each other. More precisely, there exist constants $\alpha \geq 1$, $c_4 > 0$, $L_0 > 0$ such that for $L \geq L_0$

$$h(\mathbb{E}[n_1^h(t)] \vee \mathbb{E}[n_2^h(t)]) > L \Rightarrow h(\mathbb{E}[n_1^h(t)] \wedge \mathbb{E}[n_2^h(t)]) > \alpha L - c_4.$$

In particular, we have $\sup_h h\mathbb{E}[n_1^h(t)] = \infty$ if and only if $\sup_h h\mathbb{E}[n_2^h(t)] = \infty$. Now suppose $\sup_h h\mathbb{E}[n_1^h(t)] = \infty$ and $\sup_h h\mathbb{E}[n_2^h(t)] = \infty$. By Cramer's theorem (see Theorem 2.1.24 [13]), for all $\delta > 0$ there exists a constant $c(\delta) \in (0,\infty)$ such that for all $k_0 \in \mathbb{N}_0$ and $h > 0$

$$\max\Big\{\mathbb{P}[|Y_{2,1}^h - h(1-\mu)n_2^h(t)| > \delta h n_2^h(t), n_2^h(t) = k_0], \\ \mathbb{P}[|Y_{1,2}^h - h(1/(1+\lambda))n_1^h(t)| > \delta h n_1^h(t), n_1^h(t) = k_0]\Big\} \leq c(\delta)e^{-k_0 c(\delta)}.$$

Choose δ such that $\mu + \delta < 1$ and $1/(1+\lambda) - \delta > 0$ (which is possible since $\mu \in (0,1)$ and $\lambda \in (0,\infty)$). Define $\alpha_1 = 1 - (1-\mu-\delta)(1/(1+\lambda)-\delta) < 1$ and $\theta = \alpha_1/4$. Fix

$\epsilon \in (0,1)$ and choose K large enough so that

$$\frac{c(\delta)}{1-e^{-c(\delta)}}e^{-c(\delta)(K+1)} < \frac{\epsilon}{8} \quad \text{and} \quad \frac{c_2(1+t)}{\theta K - \tilde{c}_1(1+t)} < \frac{\epsilon}{8}. \tag{2.38}$$

Since, by assumption $\sup_h h\mathbb{E}[n_1^h(t)] = \sup_h h\mathbb{E}[n_2^h(t)] = \infty$, there exists $h' \le 1$ such that

$$\mathbb{P}[n_1^{h'}(t) > \frac{K}{h'}, n_2^{h'}(t) > \frac{K}{h'}] > \epsilon. \tag{2.39}$$

Then for all $t \ge 0$,

$$\begin{aligned}
&\mathbb{P}[|Y_{2,1}^{h'}(t) - h'(1-\mu)n_2^{h'}(t)| > \delta h' n_2^{h'}(t), n_2^{h'}(t) > \frac{K}{h'}] \\
= &\sum_{j=[K/h']+1}^{\infty} \mathbb{P}[|Y_{2,1}^{h'}(t) - h'(1-\mu)n_2^{h'}(t)| > \delta h' n_2^{h'}(t), n_2^{h'}(t) = j] \\
\le &\sum_{j=[K/h']+1}^{\infty} c(\delta)e^{-c(\delta)j} = \frac{c(\delta)}{1-e^{-c(\delta)}}e^{-c(\delta)([\frac{K}{h'}]+1)} < \frac{\epsilon}{8},
\end{aligned}$$

where the last inequality follows from the choice of K in (2.38). Similarly, $\mathbb{P}[|Y_{1,2}^{h'}(t) - \frac{h'}{1+\lambda}n_1^{h'}(t)| > \delta h' n_1^{h'}(t), n_1^{h'}(t) > \frac{K}{h'}] < \frac{\epsilon}{8}$. Hence, in view of (2.39) we have

$$\begin{aligned}
&\mathbb{P}[|Y_{2,1}^{h'}(t) - h'(1-\mu)n_2^{h'}(t)| \le \delta h' n_2^{h'}(t), \\
&\quad |Y_{1,2}^{h'}(t) - \frac{h'}{1+\lambda}n_1^{h'}(t)| \le \delta h' n_1^{h'}(t), \min\{n_1^{h'}(t), n_2^{h'}(t)\} > \frac{K}{h'}] > \frac{\epsilon}{2}.
\end{aligned}$$

Let E denote the event in the equation above. From (2.37) and (2.38)

$$\begin{aligned}
\mathbb{P}[h' n_1^{h'}(t) - Y_{2,1}^{h'}(t) \ge \theta K] &\le \mathbb{P}[|S_1^{h'}(t)| \ge \theta K - \tilde{c}_1(1+t)] \\
&\le \frac{\mathbb{E}|S_1^{h'}(t)|}{\theta K - \tilde{c}_1(1+t)} \\
&\le \frac{c_2(1+t)}{\theta K - \tilde{c}_1(1+t)} < \frac{\epsilon}{8}.
\end{aligned}$$

Similarly, $I\!P[h'n_2^{h'}(t) - Y_{1,2}^{h'}(t) \geq \theta K] < \frac{\epsilon}{8}$. Thus

$$
\begin{aligned}
\frac{\epsilon}{2} < I\!P[E] &\leq I\!P[E, h'n_1^{h'}(t) - Y_{2,1}^{h'}(t) < \theta K, h'n_2^{h'}(t) - Y_{1,2}^{h'}(t) < \theta K] \\
&+ I\!P[h'n_1^{h'}(t) - Y_{2,1}^{h'}(t) \geq \theta K] + I\!P[h'n_2^{h'}(t) - Y_{1,2}^{h'}(t) \geq \theta K] \\
&\leq I\!P[\tilde{E}] + \frac{\epsilon}{8} + \frac{\epsilon}{8},
\end{aligned}
$$

where $\tilde{E}$ is the event in the first term on the right side above. It follows that $I\!P[\tilde{E}] > \epsilon/4$ and thus $\tilde{E}$ is nonempty. Now for any $\omega \in \tilde{E}$ we have from the definition of $\tilde{E}$ that

$$
h'n_1^{h'}(t) - h'(1-\mu-\delta)n_2^{h'}(t) < \theta K, \;\; h'n_2^{h'}(t) - h'(1/(1+\lambda)-\delta)n_1^{h'}(t) < \theta K.
$$

A straightforward calculation using these inequalities shows that for such ω

$$
h'n_1^{h'}(t) \leq \frac{2\theta}{1-(1-\mu-\delta)(1/(1+\lambda)-\delta)} K = \frac{K}{2}.
$$

However, this contradicts the fact that $h'n_1^{h'}(t) > K$ on $\tilde{E}$. Thus we must have that both $\sup_h hI\!E[n_1^h(t)]$ and $\sup_h hI\!E[n_2^h(t)]$ are finite. The result now follows on recalling that $M^h(t) = hn_1^h(t)$ and $N^h(t) = hn_2^h(t)$. ■

An important consequence of the above lemma is the following.

Lemma 2.3.6 *There exists an $h_0 \in (0,\infty)$ such that for all $h < h_0$, $\hat{T}^h(t) \to \infty$ with probability 1 as $t \to \infty$.*

Proof. Since $\tilde{\Delta}_*^h \to 0$ as $h \to 0$, we can find an h_0 such that $\tilde{\Delta}_*^h < 1$ for all $h < h_0$. We will argue via contradiction. Suppose $h < h_0$ and $I\!P[\sup_{t\geq 0} \hat{T}^h(t) < \infty] > 0$. Then there exist $\epsilon > 0$ and $T_0 > 0$ such that

$$
I\!P[\sup_{t\geq 0} \hat{T}^h(t) < T_0 - 1] > \epsilon. \tag{2.40}
$$

Using Lemma 2.3.5 we can find a K large enough so that

$$\mathbb{P}[M^h(T_0) \geq K] \leq \frac{\mathbb{E}M^h(T_0)}{K} < \frac{\epsilon}{4}, \quad \mathbb{P}[N^h(T_0) \geq K] \leq \frac{\mathbb{E}N^h(T_0)}{K} < \frac{\epsilon}{4}.$$

We will now show that

$$\mathbb{P}[\hat{T}^h(T_0 + 2K) < T_0 - 1] \leq \frac{\epsilon}{2}. \tag{2.41}$$

This will lead to a contradiction in view of (2.40) and hence prove the lemma. Note that

$$\begin{aligned} & \mathbb{P}[\hat{T}^h(T_0 + 2K) < T_0 - 1] \\ \leq\ & \mathbb{P}[\hat{T}^h(T_0 + M^h(T_0) + N^h(T_0)) < T_0 - 1, M^h(T_0) < K, N^h(T_0) < K] \\ +\ & \mathbb{P}[M^h(T_0) \geq K] + \mathbb{P}[N^h(T_0) \geq K] \\ \leq\ & \mathbb{P}[\hat{T}^h(T_0 + M^h(T_0) + N^h(T_0)) < T_0 - 1] + \frac{\epsilon}{4} + \frac{\epsilon}{4}. \end{aligned} \tag{2.42}$$

Furthermore, for each fixed t, $t + M^h(t) + N^h(t) \geq \sum_{k=0}^{n^h(t)-1}(\Delta_k^h 1_{\{I_k^h=0\}} + h\, 1_{\{I_k^h=1,2\}})$. Since $\hat{T}^h$ is nondecreasing and $\hat{T}^h(\hat{t}_n^h) = t_n^h$,

$$\begin{aligned} \hat{T}^h(t + M^h(t) + N^h(t)) &\geq \hat{T}^h\Big(\sum_{k=0}^{n^h(t)-1} (\Delta_k^h 1_{\{I_k^h=0\}} + h\, 1_{\{I_k^h=1,2\}})\Big) \\ &= \hat{T}^h(\hat{t}^h_{n^h(t)}) = t^h_{n^h(t)} = \sum_{k=0}^{n^h(t)-1} \Delta_k^h 1_{\{I_k^h=0\}} \geq t - \tilde{\Delta}_*^h. \end{aligned}$$

The last inequality above is a consequence of the inequalities: $\sum_{k=0}^{n^h(t)-1} \Delta_k^h 1_{\{I_k^h=0\}} \leq t \leq \sum_{k=0}^{n^h(t)} \Delta_k^h 1_{\{I_k^h=0\}}$. Recalling that $\tilde{\Delta}_*^h < 1$ we see that $\hat{T}^h(t + M^h(t) + N^h(t)) \geq t - 1$ for all $t \geq 0$. Using this inequality in (2.42) proves (2.41) and hence the result. ■

Let $T^h(t) \doteq \inf\{s : \hat{T}^h(s) > t\}$. Observe that $\hat{T}^h(T^h(t)) = t$ and that, due to Lemma 2.3.6, $T^h(t) < \infty$ almost surely for all $t \geq 0$. Define $\hat{\tau}^h \doteq T^h(\tau^h)$.

Lemma 2.3.7 *For $z_h \in \mathbb{S}_\ell^{h+}$ and $\{U_n^h\} \in \mathcal{A}^h(z_h)$,*

$$J^h(z_h, U^h) = I\!\!E \int_{[0,\hat{\tau}_1^h]} e^{-\beta \hat{T}^h(t)} f(\hat{C}^h(t)) d\hat{T}^h(t).$$

Proof. Note that

$$\hat{\tau}_1^h = \inf\{t : \hat{Z}^h(t) \in \partial^h\} \;=\; \inf\{t : Z^h(\hat{T}^h(t)) \in \partial^h\} \;=\; \inf\{t : \hat{T}^h(t) \geq \tau^h\}.$$

If $\tau^h = \infty$ then clearly $\hat{\tau}_1^h = \infty$. Suppose $\tau^h < \infty$. Then the above display shows that $\hat{\tau}_1^h = T^h(\tau^h-)$. Also, clearly $\hat{T}^h$ is constant over the interval $(T^h(\tau^h-),\, T^h(\tau^h)]$. The result now follows from (2.30) and Lemma 2.3.4. ■

2.4 Main Convergence Result

In this section we show that $V^h(z_h)$ converges to $V(z)$ whenever $z_h \to z$. The basic approach will be as follows. First we establish tightness of the continuous time (rescaled) processes defined in the previous section and characterize their subsequential limits. Then we define a time transformation for the limit processes to revert back to the original scale. We will show that the time transformed versions of the limit processes have the same laws as those of the various processes in the diffusion control problem. Using this characterization result we will show that, given a sequence of admissible controls $\{U^h, h > 0\}$, the $\limsup$ of the corresponding cost functions is bounded above by the cost for an admissible control for the diffusion control problem. This will establish that $\limsup_{h\to 0} V^h(z_h) \leq V(z)$ whenever $z_h \to z$. Finally we prove convergence of the value functions by proving the reverse inequality. The main idea of this proof is to select a near optimal control for the limit diffusion control problem and to construct from this an admissible control for the controlled Markov chain which is asymptotically near optimal.

We begin by introducing the following "relaxed control" formulation which arises naturally in the weak convergence arguments for convergence of the cost functions.

Relaxed control formulation. Let $\tilde{\mathcal{M}}$ denote the space of all Borel measures ϑ on $[0,p] \times [0,\infty)$ such that if $\vartheta(d\alpha, dt) = \vartheta_t(d\alpha)\nu(dt)$, then: (i) ϑ_t is a probability measure on $[0,p]$ for ν-almost every t, and (ii) $\nu(a,b] \leq b-a$ for all $0 \leq a \leq b < \infty$. Let $\mathcal{M}$ be the subset of $\tilde{\mathcal{M}}$ consisting of ϑ that satisfy for all $t \geq 0$, $\vartheta([0,p] \times [0,t]) = t$. Given a probability system Φ and initial condition $z \in \mathbb{S}_\ell$, let $\bar{\mathcal{A}}_p(\Phi, z)$ be the set of all processes $\bar{U} \equiv (m, M, N)$ where M and N are as introduced below (2.1), $m \in \mathcal{M}$ a.s. and $m(A \times [0,t])$ is $\mathcal{F}_t$-adapted for all $t \in [0,\infty)$, $A \in \mathcal{B}[0,p]$. Set $C(t) \doteq \int_{[0,p]} \alpha m_t(d\alpha)$ where m_t, a probability measure on $[0,p]$, is defined by the relation $m(d\alpha, dt) = m_t(d\alpha)dt$.

Let Z be defined via (2.7) with (C, M, N) as above and τ be given by (2.2). Define for $\bar{U} \in \bar{\mathcal{A}}_p(\Phi, z)$

$$\bar{J}(z, \bar{U}) \doteq I\!\!E \int_{[0,p]\times[0,\tau)} e^{-\beta t} f(\alpha) m(d\alpha, dt),$$

and let

$$\bar{V}(z) \doteq \sup_{\Phi} \sup_{\bar{U} \in \bar{\mathcal{A}}(\Phi, z)} \bar{J}(z, \bar{U}).$$

The following lemma establishes the equivalence between the relaxed control formulation and the precise control formulation.

Lemma 2.4.1 *For all $z \in \mathbb{S}$, $\bar{V}(z) = V(z)$.*

Proof. The inequality $V(z) \leq \bar{V}(z)$ is immediate since every exact control can be expressed as a relaxed control. Consider now the reverse inequality. Let Φ be a probability system and $\bar{U} = (\bar{m}, M, N) \in \bar{\mathcal{A}}(\Phi, z)$ be such that $\bar{V}(z) \leq \bar{J}(z, \bar{U}) + \epsilon$. From the boundedness of the cost function it follows that, without loss of generality, we can assume that there is a $T \in (0, \infty)$ such that $M(t) = M(t \wedge T)$ and $N(t) = N(t \wedge T)$ for all $t \in (0,\infty)$, and $\bar{m}_t(d\alpha) = \delta_p$ for all $t \geq T$. Also, T can be chosen large enough so that $f_*(p)e^{-\beta T}/\beta < \epsilon$.

Let Z be defined via (2.7) with $C(t) \doteq \int_{[0,p]} \alpha \bar{m}_t(d\alpha)$, and τ as before. Then

$$\bar{J}(z, \bar{U}) \leq I\!E \int_{[0,p]\times[0,T\wedge\tau)} e^{-\beta t} f(\alpha) m(d\alpha, dt) + \epsilon. \tag{2.43}$$

Also, by modifying m, M, and N if needed, we can assume that

$$M(t) = M(T \wedge \tau), \quad N(t) = N(T \wedge \tau) \text{ and } \bar{m}_t(d\alpha) = \delta_p \text{ for all } t \geq T \wedge \tau. \tag{2.44}$$

Following the proof of Theorem 1.2.1 in [6] one can show that there exists a sequence of exact controls $C_n \in \mathcal{A}_p(\Phi, z)$ which satisfy

$$\sup_{0\leq t\leq T_1} \left| \int_{[0,p]\times[0,t]} e^{-\beta t} f(\alpha) m(d\alpha, dt) - \int_{[0,t]} e^{-\beta t} f(C_n(t)) dt \right| \to 0 \text{ a.s., and} \tag{2.45}$$

$$\sup_{0\leq t\leq T_1} \left| \int_{[0,t]} (C_n(s) - C(s)) ds \right| \to 0 \text{ a.s.} \tag{2.46}$$

as $n \to \infty$ for all $T_1 \in (0, \infty)$. In fact the cited theorem shows that, for each n, C_n can be chosen such that it takes values in a finite set and there is a sequence $0 < t_1^n < t_2^n \cdots$ such that C_n is constant over $[t_k^n, t_{k+1}^n)$ for all $k \in I\!N_0$.

Let Z_n be defined via (2.7) with C replaced by C_n and M and N as introduced above. A straightforward application of Gronwall's inequality and (2.46) shows that for each $T_1 \in (0, \infty)$, there is a $c \equiv c(T_1) \in (0, \infty)$ such that

$$\sup_{0\leq t\leq T_1} I\!E|Z_n(t) - Z(t)| \leq c \sup_{0\leq t\leq T_1} I\!E\left| \int_{[0,t]} (C_n(s) - C(s)) ds \right|. \tag{2.47}$$

Hence $Z_n \to Z$, in probability, uniformly on $[0, T]$.

If $\tau = \infty$, $Z(t) \in \mathbb{S}^o$ for all $t \geq 0$. Thus, (2.47) implies that there exists N_0 such that if $n > N_0$ then $Z_n(t) \in \mathbb{S}^o$ for all $t \geq 0$, and therefore $\tau_n = \infty$ for all $n > N_0$. Then clearly $\tau_n \to \tau$ a.s. as $n \to \infty$ on the set $\{\tau = \infty\}$. Next note that, almost surely on the set $\{\tau < \infty\}$ and for every $\delta > 0$, there exist $t \in [\tau, \tau+\delta)$ and $\epsilon > 0$ such that $\mathrm{dist}(Z(t), \mathbb{S}) > \epsilon$. This is because, in view of (2.44), $Z(\tau+t)$, $t \geq 0$ is described via (2.7) with initial condition $Z(\tau)$ and $M \equiv N \equiv 0$. If $Z(\tau) = 0$ the property is satisfied trivially since $C(t) = p$ for all $t \geq \tau$. Otherwise, the property follows from a standard argument based on the law of the iterated logarithm for Brownian motion (cf. pages 260-261, [32]). This, along with the convergence of Z_n to Z shows that $\tau_n \doteq \inf\{t : Z_n(t) \notin \mathbb{S}^o\}$ converges to τ a.s. as $n \to \infty$. In proving this statement we also use the observation that if $Z(t-) \notin \mathbb{S}$ then $Z(t) \notin \mathbb{S}$. Thus $\tau_n \wedge T \to \tau \wedge T$. Combining these observations with (2.45) we obtain

$$\mathbb{E} \int_{[0,\tau_n \wedge T]} e^{-\beta t} f(C_n(t)) dt \to \mathbb{E} \int_{[0,p]\times[0,\tau\wedge T]} e^{-\beta t} f(\alpha) m(d\alpha, dt). \tag{2.48}$$

The result now follows on using this observation in (2.43). ■

Next note that the space $\tilde{\mathcal{M}}$ can be metrized using the Prohorov metric in the usual way (see pages 263-264 of [32]). Furthermore, with this metric $\tilde{\mathcal{M}}$ is a compact space and a sequence $\vartheta_n \in \tilde{\mathcal{M}}$ converges to ϑ if and only if for all continuous functions ψ on $[0,p] \times [0,\infty)$ with compact support,

$$\int_{[0,p]\times[0,\infty)} \psi(\alpha, t) m_n(d\alpha, dt) \to \int_{[0,p]\times[0,\infty)} \psi(\alpha, t) m(d\alpha, dt). \tag{2.49}$$

We now define $\tilde{\mathcal{M}}$-valued random variables $\hat{m}^h$ by the relation

$$\hat{m}^h(A \times [0,t]) \doteq \int_{[0,t]} 1_A(\hat{C}^h(s)) d\hat{T}^h(s), \quad A \in \mathcal{B}([0,p]), t \in [0,\infty).$$

Noting that the right side above is equal to $\int_{[0,t]} (\int_A \delta_{\hat{C}^h(s)}(d\alpha)) d\hat{T}^h(s)$, where δ_x is the probability measure concentrated at x, we can write $\hat{m}^h(d\alpha, dt)$ as $\hat{m}^h_t(d\alpha)\hat{\nu}^h(dt)$ where $\hat{m}^h_t$ and $\hat{\nu}^h$ are given by, for $A \in \mathcal{B}([0,p])$ and $0 \leq a \leq b < \infty$, $\hat{m}^h_t(A) = \delta_{\hat{C}^h(t)}(A)$, $\hat{\nu}^h(a,b] = \hat{T}^h(b) - \hat{T}^h(a)$.

Convergence of the Time Rescaled Processes. We begin by showing that the processes $\hat{E}_1^h$ and $\hat{E}_2^h$ converge weakly to the 0 process as $h \to 0$.

Lemma 2.4.2 *Let $\hat{E}_i^h, i = 1, 2$ be as defined above* (2.27)*. Then $\hat{E}_i^h$ converges in probability to 0 in $D([0,\infty) : \mathbb{R}^2)$.*

Proof. The local consistency condition (2.18) and property (2.22) imply that $E_{i,n}^h$ is an $\{\mathcal{F}_n^h\}$-martingale. As $\hat{n}^h(t)$ is a bounded stopping time (cf. (2.31)) and the increments of $E_{i,n}^h$ are bounded, it follows from the optional sampling theorem that the continuous time process $\hat{E}_i^h(t)$ is an $\{\hat{\mathcal{F}}^h(t)\}$-martingale, the trace of the quadratic variation of which is given by $\mathrm{Tr}\langle \hat{E}_i^h \rangle(t) = \sum_{k=0}^{\hat{n}^h(t)\wedge\eta_h - 1} \mathbb{E}[|\delta z_k^h - hv_i|^2 \, 1_{\{I_k^h = i\}} | \mathcal{F}_{k-1}^h]$. Finally, applying Doob's inequality, (2.19) and the observation that the maximum number of steps of either singular control in the first $\hat{n}^h(t)$ steps is t/h, we have for $i = 1, 2$ $\mathbb{E}[\sup_{s \le t} |\hat{E}_i^h(s)|]^2 \le 4\mathbb{E}\mathrm{Tr}\langle \hat{E}_1^h \rangle(t) \ \le O(h^2)(t/h) = O(h)$. The result now follows. ■

Define the process $\hat{A}^h$ by $\hat{A}^h(t) \doteq \int_{[0,t)} \hat{C}^h(s) d\hat{T}^h(s)$. Let $\bar{\mathbb{R}}$ denote the one point compactification of $\mathbb{R}$. The following proposition gives the tightness of the various time rescaled processes. The proof is similar to that of Theorem 5.3 of [33] and is therefore omitted.

Proposition 2.4.3 *Let $\hat{H}^h \doteq (\hat{Z}^h, \hat{T}^h, \hat{A}^h, \hat{M}^h, \hat{N}^h, \hat{R}^h, \hat{B}^h, \hat{S}^h)$. Then $\{(\hat{H}^h, \hat{\tau}_1^h, \hat{m}^h),$ $h > 0\}$ is a tight family in $D([0,\infty) : E) \times \bar{\mathbb{R}} \times \mathcal{M}$ where $E = \mathbb{S}_\ell^{h+} \times \mathbb{R}_+^6 \times \mathbb{R}^4$.*

We now turn our attention to characterizing subsequential limit points of $\{(\hat{H}^h, \hat{\tau}_1^h, \hat{m}^h),$ $h > 0\}$. Suppose that the initial condition sequence $\{z_h\}$ converges to some $z \in \mathbb{S}_\ell$. Slightly abusing notation, let h index a weakly convergent subsequence of $(\hat{H}^h, \hat{\tau}_1^h, \hat{m}^h)$ with weak limit, $(\hat{H}, \hat{\tau}_1, \hat{m})$, where $\hat{H} \doteq (\hat{Z}, \hat{T}, \hat{A}, \hat{M}, \hat{N}, \hat{R}, \hat{B}, \hat{S})$, given on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Let $\hat{\mathcal{F}}^*(t) \doteq \sigma(\hat{H}(s), \hat{m}(A \times [0,s)) | A \in \mathcal{B}([0,p]), 0 \le s \le t)$ and let $\hat{\mathcal{F}}(t) \doteq \hat{\mathcal{F}}^*(t+) \vee \mathcal{N}$ where $\mathcal{N}$ denotes the collection of all $\mathbb{P}$-null sets.

Theorem 2.4.4 *The limit point $(\hat{H}, \hat{\tau}_1, \hat{m})$ has the following properties.*

1. *$\hat{T}$ is nondecreasing and Lipschitz continuous with Lipschitz coefficient 1.*

2. *There exists an $\{\hat{\mathcal{F}}(t)\}$-progressively measurable process $\hat{C}$ with $\hat{C}(t) \in [0,p]$ for all $t \geq 0$, such that*
$$\hat{B}(t) = \int_0^{t\wedge\hat{\tau}_1} \begin{pmatrix} r\hat{X}(s) - \hat{C}(s) \\ b\hat{Y}(s) \end{pmatrix} d\hat{T}(s). \tag{2.50}$$

3. *$\hat{S}_1(t) = 0$ for all $t \geq 0$ and $\hat{S}_2$ is a continuous $\{\hat{\mathcal{F}}(t)\}$-martingale with quadratic variation $\langle \hat{S}_2 \rangle_t = \int_0^t |\sigma\hat{Y}(s)|^2 d\hat{T}(s), t \geq 0$.*

4. *$\hat{M}$ and $\hat{N}$ are nondecreasing and continuous.*

5. *$\hat{R}$ is a vector of nondecreasing continuous processes which satisfy*
$$\int_0^\infty 1_{\{\hat{X}(t)<\ell\}} d\hat{R}_1(t) = 0, \quad \int_0^\infty 1_{\{\hat{Y}(t)<\ell\}} d\hat{R}_2(t) = 0. \tag{2.51}$$

6. *$\hat{Z}$ is a continuous process satisfying $I\!P[\hat{Z}(t) \in S_\ell] = 1$ for all $t \geq 0$ and*
$$\hat{Z}(t) = z + \hat{B}(t) + \hat{S}(t) + v_1\hat{M}(t) + v_2\hat{N}(t) - \hat{R}(t). \tag{2.52}$$

7. *Writing $\hat{m}(d\alpha, dt)$ as $\hat{m}_t(d\alpha)\hat{\nu}(dt)$ we have $\hat{\nu}(a,b] = \hat{T}(b) - \hat{T}(a)$, $0 \leq a \leq b < \infty$.*

8. *$\hat{C}(t) = \int_{[0,p]} \alpha\hat{m}_t(d\alpha)$ for $\hat{\nu}$-almost every $t \in [0,\infty)$.*

Proof. By appealing to Skorohod representation theorem and by relabeling the convergent subsequence we can assume without loss of generality that $\hat{H}^h \to \hat{H}$ a.s. The fact that the process $\hat{T}$ is nondecreasing and Lipschitz continuous with Lipschitz coefficient 1 follows easily from similar properties for $\hat{T}^h$. Since $|\hat{A}^h(t) - \hat{A}^h(s)| \leq p|\hat{T}^h(t) - \hat{T}^h(s)|$ it follows that $\hat{A}$ is absolutely continuous with respect to $\hat{T}$. Therefore there exists a $[0,p]$-valued process $\hat{C}$, progressively measurable with respect to $\{\hat{\mathcal{F}}^*(t)\}$ such that $\hat{A}(t) = \int_0^t \hat{C}(s)d\hat{T}(s)$. This fact, together with $(\hat{Z}^h, \hat{T}^h) \to (\hat{Z}, \hat{T})$ a.s. and an application of the dominated convergence theorem yield 2. We next show that $\hat{S}$ has continuous paths. First note that by local consistency ((2.12), (2.13)) there exists $\zeta \in (0, \infty)$ such that for all $u \geq 0, h \geq 0$ $j(\hat{S}^h, u) \doteq \sup_{t \leq u} |\hat{S}^h(t) - \hat{S}^h(t-)| \leq 2\zeta h$. Thus for h small enough $j(\hat{S}^h) \doteq \int_0^\infty e^{-u}(j(\hat{S}^h, u) \wedge 1) du \leq 2\zeta h$. Therefore, by Theorem 3.10.2 in [17] the limiting process $\hat{S}$ has continuous paths. One can check that the quadratic variation of $\hat{S}^h$, which is an $\{\hat{\mathcal{F}}^h(t)\}$-martingale, is given by

$$\langle \hat{S}^h \rangle(t) = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \int_0^{t \wedge \hat{\tau}_1^h} |\sigma \hat{Y}^h(s)|^2 d\hat{T}^h(s) + \hat{\delta}_2^h(t), \tag{2.53}$$

where due to (2.14) and using (2.31) the process $\hat{\delta}_2^h$ satisfies for all $m \geq 1$,

$$I\!\!E \sup_{s \leq t} |\hat{\delta}_2^h(s)|^m \to 0 \text{ as } h \to 0.$$

From (2.53) it follows that $\hat{S}_1(t) = 0$ for all $t \geq 0$. From (2.53) and the Burkholder-Gundy inequalities we also have

$$I\!\!E |\hat{S}_2^h(t)|^4 \leq \alpha[T^2 + I\!\!E \sup_{0 \leq u \leq t} |\hat{\delta}_2^h(u)|^2].$$

Thus the family $\{(\hat{S}_2^h(t))^2, h > 0\}$ is uniformly integrable. A standard argument (cf. pages 1457-1458 [33]) shows that $\hat{S}_2$ is an $\hat{\mathcal{F}}_t$-martingale with quadratic variation as given in 3. Part 4 is immediate on noting that $\hat{M}^h, \hat{N}^h$ are nondecreasing and since the maximum number of purchase or sales steps over $(\hat{n}^h(t), \hat{n}^h(t+s))$ is $s/h + 1$,

$$|\hat{M}^h(t+s) - \hat{M}^h(t)| \leq s + h, \;\; |\hat{N}^h(t+s) - \hat{N}^h(t)| \leq s + h.$$

From Definition 2.3.1 (3) it follows that (2.51) holds with $(\hat{X}, \hat{Y}, \hat{R})$ replaced by $(\hat{X}^h$, $\hat{Y}^h$, $\hat{R}^h)$. Also clearly $\hat{Z}^h \in (-\infty, \ell] \times (-\infty, \ell]$. Parts 5 and 6 are now immediate consequences of (2.32) and continuity properties of the Skorohod map (see (2.6)). Next, write $\hat{m}(d\alpha, dt)$ as $\hat{m}_t(d\alpha)\hat{\nu}(dt)$. Since $\hat{m}^h([0,p],(a,b]) = \hat{T}^h(b) - \hat{T}^h(a)$ for $0 \le a \le b < \infty$, taking limits yields $\hat{\nu}(a,b] = \hat{T}(b) - \hat{T}(a)$. This proves 7. Part 8 is immediate from the representation $\int_{(a,b]} \hat{C}^h(s)d\hat{T}^h(s) = \int_{(a,b]\times[0,p]} \alpha\hat{m}^h(d\alpha, ds)$, $0 \le a \le b < \infty$ and the fact that (see proof of 2) $\int_{(a,b]} \hat{C}^h(s)d\hat{T}^h(s)$ converges to $\int_{(a,b]} \hat{C}(s)d\hat{T}(s)$. ■

Time Inversion. We now define an inverse time transformation that will revert the limit processes back to the original time scale. We will see that the time inverted processes lead to an admissible control pair for the diffusion control problem in (2.7)-(2.9). The key step in returning to the original time scale is the following result analogous to Lemma 2.3.6.

Lemma 2.4.5 *$\hat{T}(t) \to \infty$ with probability 1 as $t \to \infty$.*

Proof. We will argue via contradiction. Suppose $I\!\!P[\sup_{t\ge 0}\hat{T}(t) < \infty] > 0$. Then there exist $\epsilon > 0$ and $T_0 > 0$ such that

$$I\!\!P[\sup_{t\ge 0}\hat{T}(t) < T_0 - 1] > \epsilon. \tag{2.54}$$

As in the proof of Lemma 2.3.6 we can find a $K \in (0,\infty)$ such that $\liminf_{h\to 0} I\!\!P[\hat{T}^h(T_0 + 2K) < T_0 - 1] \le \epsilon/2$. The weak convergence $\hat{T}^h \Rightarrow \hat{T}$ now implies $I\!\!P[\hat{T}(T_0 + 2K) < T_0 - 1] \le \epsilon/2$. This contradicts (2.54) and hence the result follows. ■

Let T be the inverse of $\hat{T}$, defined as $T(t) \doteq \inf\{s : \hat{T}(s) > t\}$. From Lemma 2.4.5 it follows that $T(t) < \infty$ almost surely for all $t \ge 0$. Since $\hat{T}(t)$ is nondecreasing and continuous it follows that $T(t)$ is nondecreasing and right-continuous. Also note the following properties for all $t \ge 0$:

$$\begin{gathered} T(t) \ge t, \quad \hat{T}(T(t)) = t, \quad T(\hat{T}(t)) \ge t, \\ T(t) \uparrow \infty \text{ as } t \uparrow \infty, \quad T(t) < \infty \text{ a.s.}, \quad \hat{T}(s) \in [0,t] \Leftrightarrow s \in [0, T(t)]. \end{gathered} \tag{2.55}$$

Let $\hat{H}$ be as in Theorem 2.4.4. Define $H(t) \doteq \hat{H}(T(t))$. We will use similar notation for the various components of H; for example, $Z(t) \doteq \hat{Z}(T(t))$, etc. Let $\tau_1 \doteq \hat{T}(\hat{\tau}_1)$. Then by (2.52), for $t \geq 0$

$$Z(t) = z + B(t) + S(t) + v_1 M(t) + v_2 N(t) - R(t). \tag{2.56}$$

Before characterizing the various terms in (2.56) we note that for $t \geq 0$, $\{T(s) < t\} = \{\hat{T}(t) > s\} \in \hat{\mathcal{F}}(t)$ since $\hat{T}(t)$ is $\hat{\mathcal{F}}(t)$-measurable. Therefore, since $\hat{\mathcal{F}}(t)$ is right-continuous $T(s)$ is an $\hat{\mathcal{F}}(t)$-stopping time for each $s \geq 0$. Let $\mathcal{F}_0(t) \doteq \hat{\mathcal{F}}(T(t))$ and note that $H(t) \doteq \hat{H}(T(t))$ and $m(A \times [0,t]) \doteq \hat{m}(A \times [0, T(t)])$ are $\mathcal{F}_0(t)$-measurable. Define $\mathcal{F}(t) \doteq \sigma(H(s), m(A \times [0,s]) : A \in \mathcal{B}([0,p]), 0 \leq s \leq t)$. Then $\mathcal{F}(t) \subseteq \mathcal{F}_0(t)$.

Theorem 2.4.6 *The processes in (2.56) have the following properties.*

1. $B(t) = \int_0^{t\wedge\tau_1} \begin{pmatrix} rX(s) - C(s) \\ bY(s) \end{pmatrix} ds.$

2. $S_1 \equiv 0$ *and* S_2 *is a continuous* $\{\mathcal{F}_0(t)\}$*-martingale with quadratic variation*

$$\langle S_2\rangle(t) = \int_0^{t\wedge\tau_1} |\sigma Y(s)|^2 ds. \tag{2.57}$$

   There exists an enlargement of the probability space $(\Omega, \mathcal{F}, I\!P)$ *and the filtration* $\{\mathcal{F}_0(t)\}$ *that supports a Wiener process* W *that is a martingale with respect to the enlarged filtration and such that*

$$S_2(t) = \int_0^{t\wedge\tau_1} \sigma Y(s) dW(s). \tag{2.58}$$

3. *The process* C *is* $\{\mathcal{F}_0(t)\}$*-progressively measurable with* $C(t) \in [0,p]$ *a.s. for all* $t \geq 0$*. In addition,* $M(0), N(0) \geq 0$ *and the processes* M *and* N *are right-continuous, nondecreasing and* $\{\mathcal{F}_0(t)\}$*-adapted. Hence defining* $\Phi \doteq (\Omega, \mathcal{F}, I\!P, \{\mathcal{F}_0(t)\}, W)$ *we*

have $U \equiv (C, M, N) \in \mathcal{A}_p(\Phi, z)$.

4. *For all* $t \geq 0$, $Z(t) \in (-\infty, \ell] \times (-\infty, \ell]$ *a.s.,* R *is a vector of nondecreasing right continuous processes and the representation* (2.7) *holds with* (X_ℓ, Y_ℓ, Z_ℓ) *there replaced by* (X, Y, Z).

Remark 2.4.7 *Note that Theorem 2.4.6 does not prove that* (Z, R) *is a solution to the Skorohod problem introduced in Definition 2.2.3, since in general* (2.8) *may fail to hold for the process* R. *However, note that if* (Z^*, R^*) *is the solution of* (2.7) *and* (2.8) *with* $U = (C, M, N)$ *as in (3) of Theorem 2.4.6 then by classical comparison results for (reflected) stochastic differential equations one can show that* $Z^*(t) \geq Z(t)$ *for all* t, *a.s. and so* $\tau^* \geq \tau$, *where* τ *is as in* (2.2) *and* τ^* *is defined by the right side of* (2.2) *with* Z *replaced by* Z^*. *This in particular shows that*

$$\int_{[0,p]\times[0,\tau]} e^{-\beta t} f(\alpha) dm_t(d\alpha) dt \leq \int_{[0,p]\times[0,\tau^*]} e^{-\beta t} f(\alpha) dm_t(d\alpha) dt. \tag{2.59}$$

Proof of Theorem 2.4.6. Part 1 is an immediate consequence of Theorem 2.4.4 (2) and Lemma 2.3.4 on noting that

$$\int 1_{[T(\tau_1-), T(\tau_1)]} d\hat{T}(s) = 0 \text{ a.s.} \tag{2.60}$$

Clearly, $S_1(t) \doteq \hat{S}_1(T(t)) \equiv 0$ a.s. From Theorem 2.4.4 we have that $\{\hat{S}_2(t), \hat{\mathcal{F}}(t)\}$ is a continuous martingale. Thus for all $n \geq 1$, $I\!E[\hat{S}_2(T(t) \wedge n) | \hat{\mathcal{F}}(T(s))] = \hat{S}_2(T(s) \wedge n)$ a.s. Also as $\hat{S}_2$ has continuous paths and $T(t) < \infty$ a.s. we have as $n \to \infty$ for all $t \geq 0$, $\hat{S}_2(T(t) \wedge n) \to \hat{S}_2(T(t)) = S_2(t)$ a.s. Furthermore, from Theorem 2.4.4 (3), there exists $\alpha \in (0, \infty)$ such that $I\!E|\hat{S}_2(T(t) \wedge n)|^2 \leq \alpha t$ for all $t \geq 0, n \in I\!N$. Hence, for each fixed t, the family $\{\hat{S}_2(T(t) \wedge n), n \geq 1\}$ is uniformly integrable and therefore $\hat{S}_2(T(t) \wedge n) \to \hat{S}_2(T(t))$ in L^1. Taking limits as $n \to \infty$, it follows that $I\!E[\hat{S}_2(T(t)) | \hat{\mathcal{F}}(T(s))] = \hat{S}_2(T(s))$; that is $I\!E[S_2(t) | \mathcal{F}_0(s)] = S_2(s)$. This proves that $\{S_2(t), \mathcal{F}_0(t)\}$ is a martingale. Although

T in general may fail to be continuous, $S_2(\cdot) \doteq \hat{S}_2(T(\cdot))$ has continuous paths a.s. This is a consequence of the fact that $\{S_2^h(\cdot)\}_{h>0}$ is tight in $D([0,\infty))$, and an argument similar to one for $\{\hat{S}_2^h\}$ in the proof of Theorem 2.4.4 shows that any weak limit point, $\tilde{S}_2$, of $\{S_2^h\}$ must have continuous paths a.s. Also, since $\hat{S}_2^h(\cdot) = S_2^h(\hat{T}^h(\cdot))$, we must have that if $(\tilde{S}_2, \hat{S}_2, \hat{T})$ is a limit point of the tight sequence $(S_2^h, \hat{S}_2^h, \hat{T}^h)$ then $\hat{S}_2(t) = \tilde{S}_2(\hat{T}(t))$ and thus from (2.55) $S(t) \doteq \hat{S}_2(T(t)) = \tilde{S}_2(t)$. Thus we have shown that S_2 is a continuous $\mathcal{F}_0(t)$-martingale. We next consider its quadratic variation. By the Burkholder-Davis-Gundy inequalities (c.f. Theorem 3.3.28 in [29]) there exists a constant α independent of n such that

$$I\!E[|\hat{S}_2^2(T(t) \wedge n)|^2] \leq I\!E[(\sup_{0\leq s\leq T(t)} |\hat{S}_2(s \wedge n)|)^4] \leq \alpha(\alpha_1^2 t^2).$$

Thus the families $\{\hat{S}_2^2(T(t) \wedge n), n \geq 1\}$ and $\{\langle \hat{S}_2 \rangle (T(t) \wedge n), n \geq 1\}$ are uniformly integrable for each fixed $t \geq 0$. Since $\hat{S}_2^2$ (respectively $\langle \hat{S}_2 \rangle$) has continuous paths and $T(t) < \infty$ almost surely, $\hat{S}^2(T(t) \wedge n) \to \hat{S}_2^2(T(t))$ (respectively $\langle \hat{S}_2 \rangle (T(t) \wedge n) \to \langle \hat{S}_2 \rangle (T(t))$) almost surely as $n \to \infty$. By the uniform integrability, this convergence also holds in the L^1 sense. Thus

$$I\!E[\hat{S}_2^2(T(t) \wedge n) - \langle \hat{S}_2 \rangle (T(t) \wedge n) | \hat{\mathcal{F}}(T(s))] \to I\!E[\hat{S}_2^2(T(t)) - \langle \hat{S}_2 \rangle (T(t)) | \hat{\mathcal{F}}(T(s))]. \quad (2.61)$$

The above relation and the fact that $\hat{S}_2^2 - \langle \hat{S}_2 \rangle$ is an $\hat{\mathcal{F}}_t$-martingale now show that $I\!E[S_2^2(t) - \langle \hat{S} \rangle (T(t)) | \mathcal{F}_0(s)] = S_2^2(s) - \langle \hat{S} \rangle (T(s))$. Thus the quadratic variation of S_2 is given by $\langle S_2 \rangle (t) = \langle \hat{S}_2 \rangle (T(t))$. The representation (2.57) now follows on using Theorem 2.4.4, Lemma 2.3.4 and (2.60). By the martingale representation theorem (e.g. Theorem 3.4.2 in [29]) it now follows that there exists a one-dimensional Brownian motion W, possibly defined on an enlarged probability space, that is a martingale with respect to an enlargement of the filtration $\{\mathcal{F}_0(t)\}$ and is such that (2.58) holds.

The $\{\mathcal{F}_0(t)\}$-progressive measurability (respectively, adaptedness) of C (respectively, M and N) follows from the $\{\hat{\mathcal{F}}(t)\}$-progressive measurability of $\hat{C}$ (respectively, adaptedness of $\hat{M}$ and $\hat{N}$). Also, since $\hat{C}$ takes values in $[0,p]$, the same is true of C. Right continuity of M and N is a consequence of the fact that $\hat{M}$ and $\hat{N}$ are continuous and T is right continuous. This proves 3. Part 4 is once more an immediate consequence of Theorem 2.4.4 (part (5)) and Lemma 2.3.4. ■

Until now the parameters ℓ, p have been fixed and thus excluded from the notation. It is convenient to include these parameters in the notation for the remainder of this section.

Convergence of the Value Functions. Let $z \in \mathbb{S}_\ell$ and let $\{z_h, h > 0\}$ be a sequence with $z_h \in \mathbb{S}_\ell^h$ such that $z_h \to z$ as $h \to 0$. Our main goal in this section is to show that $V^h(z_h) \to V_{\ell,p}(z)$ as $h \to 0$. We begin with the following proposition.

Proposition 2.4.8 *Let $\{z_h\}, z$ be as above. Then* $\limsup_{h\to 0} V^h(z_h) \le V_{\ell,p}(z)$.

Proof. Fix for each $h > 0$, $U^h \equiv \{U_n^h, n \ge 1\} \in \mathcal{A}^h(z_h)$. In order to prove the proposition it suffices to show that

$$\limsup_{h\to 0} J^h(z_h, U^h) \le V_{\ell,p}(z). \tag{2.62}$$

Using Lemma 2.3.7 and boundedness of f, we can find for each $\epsilon \in (0,\infty)$, a $c \equiv c(\epsilon) \in (0,\infty)$ such that

$$J^h(z_h, U^h) \le I\!E \int_{[0,\hat{\tau}_1^h \wedge c]\times[0,p]} e^{-\beta \hat{T}^h(t)} f(\alpha) \hat{m}^h(d\alpha, dt) + \frac{\epsilon}{2}. \tag{2.63}$$

Let $(\hat{H}^h, \hat{\tau}_1^h, \hat{m}^h)$ be as in Proposition 2.4.3 and $(\hat{H}, \hat{\tau}_1, \hat{m})$ be one of its weak limit points. Once more, as in the proof of Theorem 2.4.4 we can assume, by relabeling and appealing to the Skorohod representation theorem, that $(\hat{H}^h, \hat{\tau}_1^h, \hat{m}^h) \to (\hat{H}, \hat{\tau}_1, \hat{m})$ a.s. Taking limits as $h \to 0$ in (2.63), we have

$$\limsup_{h\to 0} J^h(z_h, U^h) \le I\!\!E \int_{[0,p]\times[0,\hat{\tau}_1\wedge c]} e^{-\beta \hat{T}(t)} f(\alpha)\hat{m}(d\alpha, dt) + \frac{\epsilon}{2}.$$

As $\epsilon > 0$ and $c = c(\epsilon)$ are arbitrary,

$$\limsup_{h\to 0} J^h(z_h, U^h) \le I\!\!E \int_{[0,p]\times[0,\hat{\tau}_1]} e^{-\beta \hat{T}(t)} f(\alpha)\hat{m}(d\alpha, dt). \tag{2.64}$$

Let, as before, $\hat{\tau} \doteq \inf\{t \ge 0 : \hat{Z}(t) \notin \mathbb{S}^o\}$. Recall that $\hat{\tau}_1 \ge \hat{\tau}$. Then clearly

$$I\!\!E\Big[1_{\{\hat{\tau}=\infty\}} \int_{[0,p]\times[0,\hat{\tau}_1]} e^{-\beta \hat{T}(t)} f(\alpha)\hat{m}(d\alpha, dt)\Big] = I\!\!E\Big[1_{\{\hat{\tau}=\infty\}} \int_{[0,p]\times[0,\hat{\tau}]} e^{-\beta \hat{T}(t)} f(\alpha)\hat{m}(d\alpha, dt)\Big]. \tag{2.65}$$

Now suppose that $\hat{\tau} < \infty$. Let τ^* denote the first point of increase of $\hat{T}$ in $[\hat{\tau}, \hat{\tau}_1]$. More precisely, let $\tau^* \doteq \inf\{t \in [\hat{\tau}, \infty) : \hat{T}(t+\delta) > \hat{T}(t) \text{ for all } \delta > 0\} \wedge \hat{\tau}_1$. Note that

$$\begin{aligned} & I\!\!E\Big[1_{\{\hat{\tau}<\infty\}} \int_{(\hat{\tau},\hat{\tau}_1]} e^{-\beta \hat{T}(t)} \Big(\int_{[0,p]} f(\alpha) d\hat{m}_t(d\alpha)\Big) d\hat{T}(t)\Big] \\ = \; & I\!\!E\Big[1_{\{\hat{\tau}<\infty\}} \int_{[\tau^*,\hat{\tau}_1]} e^{-\beta \hat{T}(t)} \Big(\int_{[0,p]} f(\alpha) d\hat{m}_t(d\alpha)\Big) d\hat{T}(t)\Big]. \end{aligned} \tag{2.66}$$

We now show that the above quantity is equal to 0. Consider the evolution of the process $\hat{Z}$ over the interval $[\hat{\tau}, \tau^*]$. By definition $\hat{Z}(\hat{\tau}) \notin \mathbb{S}^o$. Since $\hat{T}$ is constant over this time interval, we see from Theorem 2.4.4 (2), (3) that $\hat{B}_1$ and $\hat{S}$ are both constant over this interval and since neither v_1 nor v_2 can push the process into the interior of $\mathbb{S}$, we see that $\hat{Z}(\tau^*) \notin \mathbb{S}^o$. Define $s^* \doteq \hat{T}(\tau^*)$. Since τ^* is a point of increase of $\hat{T}$ we have $T(s^*) = T(\hat{T}(\tau^*)) = \tau^*$. Thus $Z(s^*) = \hat{Z}(\tau^*) \notin \mathbb{S}^o$, where Z is defined by (2.56). Consider first the case $Z(s^*) \neq 0$; then from (2.57), $\langle S_2 \rangle$ is strictly increasing at s^*. From this it follows that for all $\delta > 0$ there exists $s_\delta \in [s^*, s^* + \delta]$ such that $\mathrm{dist}(Z(s_\delta), \mathbb{S}) > 0$, i.e. $\mathrm{dist}(\hat{Z}(T(s_\delta)), \mathbb{S}) > 0$. Now since $\mathbb{S}^h \to \mathbb{S}$ and $\hat{Z}^h \to \hat{Z}$ as $h \to 0$ we have $\mathrm{dist}(\hat{Z}^h(T(s_\delta)), \mathbb{S}^h) > 0$ for all h small enough. Therefore, by definition of $\hat{\tau}_1^h$

we must have $\hat{\tau}_1^h \leq T(s_\delta)$ for all h small enough. This implies $\hat{\tau}_1 \leq T(s_\delta)$. Now, taking $\delta \to 0$ and using the right continuity of T at s^* it follows that $\hat{\tau}_1 \leq T(s^*) = \tau^*$. Hence the quantity in (2.66) is equal to 0.

Finally consider the case when $Z(s^*) = 0$ (and $\hat{\tau} < \infty$). Let $s^{**} \doteq \inf\{s > s^* | Z(s^*) \neq 0\}$. From the dynamics of Z (see (2.56)) it follows that for every $\delta > 0$, there exists $s_\delta \in [s^{**}, s^{**} + \delta]$ such that $\text{dist}(Z(s_\delta), \mathbb{S}) > 0$. Arguing as before, we have $\hat{\tau}_1 \leq T(s^{**})$. Define $m_t(d\alpha) \doteq \hat{m}_{T(t)}(d\alpha)$ for $t \geq 0$. Since $C(t) = 0$ for $t \in [s^*, s^{**}]$ we get that $m_t = \delta_0$ for t in this interval. Thus since $f(0) = 0$, we have $\int_{[0,p]} f(\alpha) dm_t(d\alpha) = 0$ for all $t \in [s^*, s^{**}]$. Combining this with the fact that $[\tau^*, \hat{\tau}_1] \subset [T(s^*), T(s^{**})]$ we now see that the expression in (2.66) is 0. Thus

$$\mathbb{E}\Big[1_{\{\hat{\tau}<\infty\}} \int_{[0,p]\times[0,\hat{\tau}_1]} e^{-\beta \hat{T}(t)} f(\alpha) \hat{m}(d\alpha, dt)\Big] = \mathbb{E}\Big[1_{\{\hat{\tau}<\infty\}} \int_{[0,p]\times[0,\hat{\tau}]} e^{-\beta \hat{T}(t)} f(\alpha) \hat{m}(d\alpha, dt)\Big]. \tag{2.67}$$

Combining (2.64), (2.65), and (2.67) we now get

$$\limsup_{h\to 0} J^h(z_h, U^h) \leq \mathbb{E} \int_{[0,\hat{\tau}]} e^{-\beta \hat{T}(t)} (\int_{[0,p]} f(\alpha) d\hat{m}_t(d\alpha)) d\hat{T}(t).$$

We next consider the time inversion. Recall that $\tau \doteq \inf\{t : Z(t) \notin \mathbb{S}^o\}$. Note that $\tau \geq \hat{T}(\hat{\tau})$. Using this inequality and Lemma 2.3.4 we have

$$\mathbb{E} \int_{[0,\hat{\tau}]} e^{-\beta \hat{T}(t)} (\int_{[0,p]} f(\alpha) d\hat{m}_t(d\alpha)) d\hat{T}(t) \leq \mathbb{E} \int_{[0,\tau]} e^{-\beta t} (\int_{[0,p]} f(\alpha) dm_t(d\alpha)) dt.$$

The inequality (2.62) now follows from the above inequality and Remark 2.4.7. ■

We now proceed to the proof of the the reverse inequality

$$\liminf_{h\to 0} V^h(z_h) \geq V_{\ell,p}(z). \tag{2.68}$$

We begin with the following lemma which allows to restrict attention to controls that have several simplifying features.

Lemma 2.4.9 *Let Φ be a probability system and $U \in \mathcal{A}_p(\Phi, z)$ be a control with corresponding cost function $J_\ell(z, U)$. Then for any $\delta > 0$ there exists $U_\delta \in \mathcal{A}_p(\Phi, z)$ such that $|J_\ell(z, U) - J_\ell(z, U_\delta)| < \delta$ and U_δ satisfies the following.*

1. *There exists $T \in (0, \infty)$ such that $M_\delta(t) = M_\delta(T), N_\delta(t) = N_\delta(T)$, and $C_\delta(t) = 0$ for all $t \geq T$.*

2. *There exists $L \in (0, \infty)$ such that*

$$\sup_{0 \leq t < \infty} \sup_{\omega} (M_\delta(t \wedge \tau, \omega) + N_\delta(t \wedge \tau, \omega)) \leq L.$$

3. *There exist $\eta, \theta \in (0, \infty)$ and $K \in I\!N$ such that $C(t), M(t), N(t)$ take values in the finite set $\{k\eta : k = 0, 1, 2, \ldots K\}$. Furthermore, C, M, and N are piecewise constant with possible time points of change being $\{0, \theta, 2\theta, 3\theta, \ldots\}$.*

4. *There exists a $\gamma \in (0, \infty)$ such that θ is an integer multiple of γ and the chosen control $U = (C, M, N)$ satisfies the following equality for $m \geq 1$:*

$$\begin{aligned} & I\!P[(C(m\theta), \delta M(m\theta), \delta N(m\theta)) = k\eta | U(s), s < m\theta; W(s), s \leq m\theta] \\ = \; & I\!P[(C(m\theta), \delta M(m\theta), \delta N(m\theta)) = k\eta | U(n\theta), n < m; W(l\gamma), l\gamma \leq m\theta], \end{aligned} \tag{2.69}$$

   where $k = (k_1, k_2, k_3)$ and k_1, k_2, k_3 are integers, at most one of which is nonzero.

5. *Denoting for $m \geq 1$, $\Psi(m) \doteq \{C(n\theta), \delta M(n\theta), \delta N(n\theta), n < m\}$, $k \doteq (k_1, k_2, k_3)$, and $\mathcal{W}(m) \doteq \{W(l\gamma), l\gamma \leq m\theta\}$ rewrite the above probability as*

$$\begin{aligned} & I\!P[C(m\theta) = k_1\eta, \delta M(m\theta) = k_2\eta, \delta N(m\theta) = k_3\eta | \Psi(m), \mathcal{W}(m)] \\ \doteq \; & q_{m,k}(\Psi(m), z, \mathcal{W}(m)). \end{aligned} \tag{2.70}$$

Denote $\mathbb{P}[U(0) = (k_1\eta, k_2\eta, k_3\eta)]$ by $q_{0,k}(z)$. For each $m \geq 0$, the function $q_{m,k}$ can be chosen so that the function $(z, w) \to q_{m,k}(\psi, z, w)$ is continuous for every ψ.

Proof. Let $\epsilon > 0$ be arbitrary and let $T \in (0, \infty)$ be such that $f_*(p)e^{-\beta T}/\beta < \epsilon$. Consider $\tilde{U} = (\tilde{C}, \tilde{M}, \tilde{N})$ given by: $\tilde{C}(t) = C(t)1_{t<T}$; $\tilde{M}(t) = M(t \wedge T)$ and $\tilde{N}(t) = N(t \wedge T)$, $t \geq 0$. Clearly $\tilde{U} \in \mathcal{A}_p(z)$ and it is easy to check that $|J_\ell(z, U) - J_\ell(z, \tilde{U})| < \epsilon$. This proves 1. Henceforth we will assume, without loss of generality, that 1 holds for U in the statement of the lemma. Using (2.7) and the bounds on the state and control space it is easy to show $\sup_{0\leq t\leq T}[M(t \wedge \tau, \omega) + N(t \wedge \tau, \omega)] \leq c_1 + c_2 \sup_{0\leq t\leq T} |W(t)|$, where c_1 and c_2 are nonnegative constants that may depend on T. Let $L \in (0, \infty)$ be large enough so that $c_2 \mathbb{E} \sup_{0\leq t\leq T} |W(t)|/(L - c_1) < \epsilon$. Define $\tilde{U}$ by $\tilde{C} \equiv C$, $\tilde{M}(t) \doteq M(t) \wedge L$, $\tilde{N}(t) \doteq N(t) \wedge L$. Let $\tilde{Z}$ be the corresponding controlled process and $\tilde{\tau}$ the corresponding hitting time. Let T be as in part 1 and define $A \doteq \{\sup_{0\leq t\leq T}[M(t \wedge \tau) + N(t \wedge \tau)] < L\}$. Then

$$J_\ell(z, \tilde{U}) = \mathbb{E}[1_A \int_{[0,\tilde{\tau}]} e^{-\beta t} f(\tilde{C}(t))dt] + \mathbb{E}[1_{A^c} \int_{[0,\tilde{\tau}]} e^{-\beta t} f(\tilde{C}(t))dt]. \tag{2.71}$$

Using the bound on f, choice of L and Markov's inequality, the second term on the right side of the above inequality is bounded by ϵ. Also, since on the set A, $M(t) < L$ and $N(t) < L$ for all $t \leq \tau \wedge T$, we have that the evolution of $\tilde{Z}$ is the same as that of Z. Therefore $\tilde{\tau}$ in the first expression on the right side of (2.71) can be replaced by τ. This shows that $|J_\ell(z, \tilde{U}) - J_\ell(z, U)| \leq 2\epsilon$ and hence 2 follows.

We now consider 3. Let $U \equiv (C, M, N)$ be an admissible control satisfying properties 1 and 2 above and let Z be the solution to (2.7) under (C, M, N) defined on some probability system. Following Theorem 1.2.1 of [6] (see comments below (2.46)) we can assume without loss of generality that C takes values in a finite set, is RCLL and piecewise constant with finitely many points of change over $[0, T]$. We also assume without loss of generality (by modifying controls if needed) that $M(t) = M(t \wedge \tau)$, $N(t) = N(t \wedge \tau)$ and

$C(t) = C(t)1_{t<\tau} + p1_{t\geq\tau}$. Fix $\eta, \theta \in (0, \infty)$ and define the piecewise constant processes $C_{\eta,\theta}$, $M_{\eta,\theta}$, and $N_{\eta,\theta}$ as follows. For $m = 0$, define $M_{\eta,\theta}(m\theta) = M_{\eta,\theta}(0) = k\eta$ if $M(0) \in [k\eta, k\eta + \eta)$. For $m \geq 1$, set $\delta M_{\eta,\theta}(m\theta) = k\eta$ if $M(m\theta) - M(m\theta - \theta) \in [k\eta, k\eta + \eta)$. By property 2, we only need to consider the finite set $\{k\eta : k = 0, 1, 2, \ldots, K\}$ where K is some positive integer. Then let $M_{\eta,\theta}(t) = M_{\eta,\theta}(m\theta)$ for $t \in [m\theta, m\theta + \theta)$. Define $N_{\eta,\theta}$ analogously based on N. Define $C_{\eta,\theta}(m\theta) = k\eta$ if $C(m\theta) \in [k\eta, k\eta + \eta)$ and $C_{\eta,\theta}(t) = C_{\eta,\theta}(m\theta)$ for $t \in [m\theta, m\theta + \theta)$. The constructed process $U_{\eta,\theta} \equiv (C_{\eta,\theta}, M_{\eta,\theta}, N_{\eta,\theta})$ is an admissible control. Let $Z_{\eta,\theta}$ denote the solution to (2.7) under this control, defined on some probability system, and let $\tau_{\eta,\theta}$ denote the first time this process exits $\mathbb{S}^o_\ell$. Choose a sequence (η_k, θ_k) such that $\eta_k, \theta_k \to 0$ as $k \to \infty$. Denote Z_{η_k,θ_k} by Z_k. Similar abbreviations are used for $U_{\eta_k,\theta_k}, \tau_{\eta_k,\theta_k}$. One can easily check that $(Z_k, U_k) \to (Z, U)$ in $D([0,T], I\!R^2 \times [0,p] \times [0,L] \times [0,L])$ (in probability) as $k \to \infty$. If $\tau = \infty$, the uniform convergence $Z_k \to Z$ implies that there exists K_0 such that for all $k > K_0$ we have $Z_k(t) \in \mathbb{S}^o$ for all $t \geq 0$, and thus $\tau_k = \infty$ for all $k > K_0$. Therefore $\tau_k \to \tau$ a.s. as $k \to \infty$ on the set $\{\tau = \infty\}$. Next note that, almost surely on the set $\{\tau < \infty\}$, for every $\delta > 0$ there exists $t \in [\tau, \tau + \delta)$ and $\epsilon > 0$ such that $\mathrm{dist}(Z(t), \mathbb{S}) > \epsilon$. This is because, on this set, by our choice of U, $Z(\tau + t)$, $t \geq 0$ is described via (2.7) with $M \equiv N \equiv 0$ and initial condition $Z(\tau)$. In the case $Z(\tau) = 0$ the property is satisfied trivially since $C(t) = p$ for all $t \geq \tau$. Otherwise, the property follows from an argument analogous to proof of Theorem 9.4.3 of [32](see pages 260-261). Next, recalling that $Z_k \to Z$, $Z(t)$ is continuous for all $t \geq \tau$ and the observation that $Z(\tau-) \in (\mathbb{S}^o_\ell)^c \Rightarrow Z(\tau) \in (\mathbb{S}^o_\ell)^c$, we conclude $\tau_k \wedge T \to \tau \wedge T$ in probability. The convergence of $J(z, U_k)$ to $J(z, U)$ now follows. This proves 3.

The proofs of 4 and 5 are quite standard and we only provide a sketch; the reader is referred to the proof of Theorem 10.3.1 of [32] (pages 285-287) for details. Assume that U satisfies properties 1 through 3 and let $\gamma > 0$. Part 4 is essentially a consequence of the martingale convergence theorem on noting that the σ-fields $\mathcal{G}_\gamma \doteq \sigma\{U(n\theta), n <$

$m; W(l\gamma), l\gamma \leq, \theta\}$ increase to the σ-field $\mathcal{G} \doteq \sigma\{U(n\theta), n < m; W(s), s \leq, \theta\}$ as $\gamma \downarrow 0$. The main idea is to define controls U_γ and controlled processes Z_γ recursively over intervals $[m\theta, (m+1)\theta)$ by using the right side of (2.69) in defining the law of U_γ over $[m\theta, (m+1)\theta)$. Proving the weak convergence of (Z_γ, C_γ) to (Z, C) is straightforward. The convergence of hitting times is argued as in the proof of part 3. Finally, part 5 is proved by convolving $q_{m,k}$, defined in (2.70), in the (z, w) variables by a parametrized family of mollifiers and arguing weak convergence of the resulting controlled pair to (Z, C) as the mollifying parameter approaches a suitable limit. Convergence of stopping times is argued, once more, as in 3. ■

Construction of asymptotically near optimal admissible controls for MDP. Fix a probability system Φ, $z \in \mathbb{S}_\ell$, and a sequence $\{z_h\}$ such that $z_h \in \mathbb{S}^h$ and $z_h \to z$ as $h \to 0$. Let $\epsilon > 0$ be arbitrary. Let $U \in \mathcal{A}_p(\Phi, z)$ be such that U satisfies properties 1 through 5 of Lemma 2.4.9 and $V_{\ell,p}(z) - \epsilon \leq J_\ell(z, U)$. For each $h > 0$, we construct from U an admissible control sequence $\{U^h_n, n \geq 0\}$ for the MDP in Definition 2.3.1 with initial condition z_h such that the cost for U^h asymptotically agrees with the cost of U. We outline below the main steps in the construction of such a control sequence. Let $\mathcal{K} \doteq \{(k_1, k_2, k_3) : k_i = 0, 1, \ldots, K; i = 1, 2, 3$ such that at most one of k_1, k_2, k_3 is positive$\}$.

Step 1. We begin by taking a random draw, denoted by $\kappa = (\kappa_1, \kappa_2, \kappa_3)$ from the probability distribution $\{q_{0,k}(z_h), k \in \mathcal{K}\}$. Set $\tilde{U}^h_0 = \kappa\eta$, $Z^h_0 = z_h$, and $\Psi^h(1) \doteq \tilde{U}^h_0$. Also set $\tilde{n}_0 = 0$. Note that at most one of κ_2 and κ_3 will be nonzero. If both κ_2 and κ_3 are 0, set $n_1 = 0$, skip step 2 below and go to step 3. Otherwise proceed to Step 2.

Step 2 (A). Recall the kernel p_h defined in (2.21). If $\kappa_2 > 0$ let $U^h_0 = (1, 0)$ and take a random draw, denoted by Z^h_1 from $p_h(Z^h_0, U^h_0, d\tilde{z})$. We express this as "the chain takes a purchase control step". If $Z^h_1 \in \partial^h_{\mathrm{R}}$, we set $U^h_1 = (3, 0)$ and draw Z^h_2 from $p_h(Z^h_1, U^h_1, d\tilde{z})$, i.e. the chain takes a reflection step. Otherwise set $U^h_1 = (1, 0)$ and draw Z^h_2 from $p_h(Z^h_1, U^h_1, d\tilde{z})$. Define (U^h_n, Z^h_n), $n = 1, 2, \ldots$ recursively by either taking a purchase control step or, if needed, a reflection step, until a total of $[\kappa_2\eta/h]$ purchase control steps have been taken. Denote by n_1 the index of the state after the last purchase control has been exercised.

Step 2 (B). If $\kappa_3 > 0$ let $U_0^h = (2,0)$ (that is, the chain takes a sales control step) and proceed as in Step 2(A) above, alternating sales control steps and reflection steps (when needed) until $[\kappa_3\eta/h]$ sales control steps have been taken. Again, let n_1 denote the index of the state after the last sales control has been exercised.

Step 3. If $Z_{n_1}^h \in \partial_{\mathrm{R}}^h$ set $U_{n_1}^h = (3,0)$ and the chain takes a reflection step. Otherwise set $U_{n_1}^h = (0,\kappa_1\eta)$ and draw $Z_{n_1+1}^h$ from $p_h(Z_{n_1}^h, U_{n_1}^h, d\tilde{z})$, i.e. the chain takes a diffusion step with $c = \kappa_1\eta$. Let t_n^h be as defined below (2.23). Define $(U_n^h, Z_{n+1}^h, n = n_1, n_1 + 1, n_1 + 2, \ldots, \tilde{n}_1 - 1$ recursively, where $\tilde{n}_1 \doteq \inf\{n : t_n^h \geq \theta\}$, as follows. If $Z_n^h \in \partial_{\mathrm{R}}^h$, set $U_n^h = (3,0)$; otherwise set it to be $(0,\kappa_1\eta)$. Draw Z_{n+1}^h from $p_h(Z_n^h, U_n^h, d\tilde{z})$.

Step 4. Next we define the "pre-Wiener process" that is needed to obtain the control at the next step. Let $\{\nu_n, n \geq 1\}$ be an i.i.d. sequence of $N(0,1)$ random variables, independent of $(U_n^h, Z_{n+1}^h)_{n=0}^{\tilde{n}_1-1}$. Define S_n^h for $n \leq \tilde{n}_1 - 1$ as in (2.26); here we only consider the second component $S_{n,2}^h$. Set $\mathcal{S}_0^h \equiv 0$ and for $\tilde{n}_0 < n \leq \tilde{n}_1 - 1$,

$$\mathcal{S}_n^h \doteq \frac{S_{n+1,2}^h - S_{n,2}^h}{\sigma Y_n^h} 1_{\{|Y_n^h| \neq 0\}} + \nu_n \sqrt{\Delta_n^h}\, 1_{\{|Y_n^h| = 0\}}.$$

Next define $W_{\tilde{n}_0}^h = 0$ and $W_n^h \doteq W_{\tilde{n}_0}^h + \sum_{i=0}^{n-1} \mathcal{S}_i^h$. Now define for $0 \leq t \leq \theta$,

$$W^h(t) \doteq W_{n^h(t)}^h - W_{\tilde{n}_0}^h. \tag{2.72}$$

Finally define $\mathcal{W}^h(1) \doteq \{W^h(l\gamma), l \in \mathbb{N}_0, l\gamma \leq \theta\}$.

Step 5. Suppose we have, for $j = 1, \cdots, m$, defined $\tilde{n}_j \doteq \inf\{n : t_n^h \geq j\theta\}$; (Z_{n+1}^h, U_n^h), $n = 0, 1, \cdots, \tilde{n}_j - 1$; $\Psi^h(j)$ and $\mathcal{W}^h(j)$. Consider now the case $j = m+1$. Take a random draw, denoted once more by $\kappa = (\kappa_1, \kappa_2, \kappa_3)$, from $\{q_{m,k}(\Psi^h(m), z_h, \mathcal{W}^h(m)),\ k \in \mathcal{K}\}$. Set

$\tilde{U}^h_m = \kappa\eta$ and $\Psi^h(m+1) = (\tilde{U}^h_0, \cdots \tilde{U}^h_m)$. Follow step 2 with $\tilde{U}^h_0$ replaced by $\tilde{U}^h_m$ and the starting index of U^h replaced with $\tilde{n}_m$. Denote by n_{m+1} the index of the state obtained after the last singular control step in step 2. Follow step 3 with n_1 replaced by n_{m+1}. Let $\tilde{n}_{m+1} = \inf\{n : t^h_n \geq (m+1)\theta\}$. This defines $(Z^h_{n+1}, U^h_n), i = 0, 1, \ldots, \tilde{n}_{m+1} - 1$. Define $W^h(t) - W^h(m\theta)$, for $t \in [m\theta, (m+1)\theta)$, by the right side of (2.72) as in Step 4 with $\tilde{n}_0, \tilde{n}_1$ replaced by $\tilde{n}_m, \tilde{n}_{m+1}$ respectively. Now set $\mathcal{W}^h(m+1) \doteq \{W^h(l\gamma), l \in I\!N, l\gamma \leq (m+1)\theta\}$.

Noting that $\tilde{n}_m$ is strictly increasing in m, we obtain the the controlled chain $\{(Z^h_n, U^h_n), n = 0, 1, 2, \ldots\}$ via the recursion:

$$(\{(Z^h_{n+1}, U^h_n)\}_{n=0}^{\tilde{n}_m - 1}, \Psi^h(m), \mathcal{W}^h(m)) \to (\{(Z^h_{n+1}, U^h_n)\}_{n=0}^{\tilde{n}_{m+1} - 1}, \Psi^h(m+1), \mathcal{W}^h(m+1)).$$

The main step in the proof of (2.68) is showing that if interpolated processes (Z^h, U^h) using the above control sequence are defined as below (2.25) and W^h is defined by (2.72) then as $h \to 0$,

$$(Z^h, U^h, W^h) \text{ converges weakly in } D([0,\infty) : I\!R^5) \text{ to } (Z, U, W), \tag{2.73}$$

where W is a standard Brownian motion and Z is defined by (2.7) with the initial condition $Z(0-) = z$. This convergence is established by proving convergence over the period $[j\theta, (j+1)\theta)$ for each j in a recursive manner. Note that given the initial condition $Z(j\theta-) = z$ and the control value $U(j\theta) = k\eta$, the dynamics of Z for $t \in [j\theta, (j+1)\theta)$ are particularly simple and are given as

$$\begin{aligned} X(t) &= x + (1-\mu)\kappa_3\eta - (1+\lambda)\kappa_2\eta + \int_{j\theta}^t (rX(s) - \kappa_1\eta)ds, \\ Y(t) &= y + \kappa_2\eta - \kappa_3\eta + \int_{j\theta}^t bY(s)ds + \int_{j\theta}^t \sigma Y(s)dW(s). \end{aligned} \tag{2.74}$$

The following lemma provides the convergence of (Z^h, U^h, W^h) over one fixed period

$[j\theta, (j+1)\theta)$ given the initial data at $j\theta$. The proof follows via straightforward weak convergence arguments and thus is omitted.

Lemma 2.4.10 *Fix $z \in \mathbb{S}_\ell$ and let $k = (k_1, k_2, k_3) \in \mathcal{K}$. Let $(Z(t), W(t))$ given on some probability system Φ be defined for $t \in [0, \theta]$ by (2.7) with $Z(0-) = z$ and $(C(t), M(t), N(t)) = k\eta$ for $t \in [0, \theta]$. Consider a sequence $\{z_h\}$ such that $z_h \in \mathbb{S}^h$ and $z_h \to z$ as $h \to 0$. Define $\tilde{n}_1$ and the sequence $\{U_n^h, Z_n^h\}_{n=0}^{\tilde{n}_1}$ via steps 2 and 3 and $\{S_n^h, W_n^h\}_{n=0}^{\tilde{n}_1}$ by step 4. Let $\{\delta_h\}$ be a sequence of nonnegative reals such that $\delta_h \to 0$ as $h \to 0$. Define, for $t \in [0, \theta]$, the interpolated process (Z^h, W^h, U^h, E^h) as before (see (2.72) and below (2.26)) with the change that $\Delta^h(Z_0^h, U_0^h)$ is replaced by $\delta_h + \Delta^h(Z_0^h, U_0^h)$. Denote the laws of (Z^h, W^h, U^h) and (Z, W, U) on $D([0, \theta] : I\!R^5)$ by Π_h^{k,δ_h} and Π^k, respectively. Then $\Pi_h^{k,\delta_h} \to \Pi^k$ as $h \to 0$.*

In the following proposition we show that the cost, $J^h(z_h, U^h)$, corresponding to the above constructed control sequence, converges to $J_\ell(z, U)$ as $h \to 0$. The desired inequality in (2.68) then follows since $V^h(z_h) \geq J^h(z_h, U^h)$, $J_\ell(z, U) \geq V_{\ell,p}(z) - \epsilon$ and $\epsilon > 0$ is arbitrary.

Proposition 2.4.11 *Let $\epsilon > 0$ be arbitrary and fix $z \in \mathbb{S}_\ell$. Let Φ be a probability system and $U \equiv U(\epsilon) \in \mathcal{A}_p(\Phi, z)$ be such that U satisfies properties 1 through 5 of Lemma 2.4.9 and $V_{\ell,p}(z) - \epsilon \leq J_\ell(z, U)$. Let for each $h > 0$, $\{U_n^h\}$ be an admissible control sequence as constructed via steps 1 through 5 above. Then $J^h(z_h, U^h) \to J_\ell(z, U)$ as $h \to 0$ and consequently (2.68) holds.*

Proof. For $t \geq 0$, let $n^h(t), Z^h(t), M^h(t), N^h(t), C^h(t)$ be as defined below (2.25). Define $U^h \equiv (M^h, N^h, C^h)$ and let W^h be as in (2.72). We begin by establishing (2.73). Define for $j \in I\!N_0$ and $t \in [0, \theta)$

$$U_j^h(t) \doteq U^h(t + j\theta),\ Z_j^h(t) \doteq Z^h(t + j\theta),\ W_j^h(t) \doteq W^h(t + j\theta) - W^h(j\theta)$$

and set $(U_j^h(\theta), Z_j^h(\theta), W_j^h(\theta)) = (U_j^h(\theta-), Z_j^h(\theta-), W_j^h(\theta-))$. Define processes U_j, Z_j, W_j, $j \in I\!N_0$ in a similar manner. Recall the sequence $\{U_j^h\}$ constructed above Lemma

2.4.10 and let $\zeta_j^h \doteq (\tilde{U}_j^h, U_j^h, W_j^h, Z_j^h)$, $\zeta_j \doteq (\tilde{U}_j, U_j, W_j, Z_j)$, where $\tilde{U}_j \doteq (C(j\theta), \delta M(j\theta), \delta N(j\theta))$, $j \in I\!N_0$. Due to the piecewise constant feature of the control U, in order to prove (2.73), it suffices to show that:

$$\text{For all } n \in I\!N_0,\ \Upsilon_h^n \doteq \{\zeta_j^h\}_{j=0}^n \text{ converges weakly to } \Upsilon \doteq \{\zeta_j\}_{j=0}^n \text{ as } h \to 0. \quad (2.75)$$

We will prove (2.75) via induction (on n). The case $n = 0$ is immediate from Lemma 2.4.10 and continuity of the kernel $q_{0,k}$ on noting that for $k \in \mathcal{K}$ and $E \in \mathcal{B}(D([0,\theta], I\!R^5)$

$$\begin{aligned} I\!P(\tilde{U}_0^h = k\eta, (U_0^h, W_0^h, Z_0^h) \in E) &= q_{0,k}(z_h)\Pi_h^{k,0}(E), \\ I\!P(\tilde{U}_0 = k\eta, (U_0, W_0, Z_0) \in E) &= q_{0,k}(z)\Pi^k(E). \end{aligned}$$

Suppose now that (2.75) holds for $n = 0, \cdots, m$ and consider the case $n = m+1$. Denote the law of Υ_n^h, Υ_n by ϖ_n^h and ϖ respectively. By induction hypothesis $\varpi_m^h \to \varpi_m$ as $h \to 0$. Furthermore, ϖ_{m+1}^h can be expanded in terms of ϖ_m^h as follows.

$$d\varpi_{m+1}^h(\upsilon_{m+1}) = \sum_{k\in\mathcal{K}} q_{m+1,k}(z_h, \tilde{u}^{*,m}, w^{*,m})\Pi_h^{k,\delta_h(\upsilon_m)}(\varsigma_{m+1})d\varpi_m^h(\upsilon_m), \quad (2.76)$$

where $\upsilon_m = \{\varsigma_j\}_{j=0}^m$; $\varsigma_j = (\tilde{u}_j, u_j, w_j, z_j)$; $\tilde{u}_j = k\eta, k \in \mathcal{K}$; $(u_j, w_j, z_j) \in D([0,\theta] : I\!R^5)$; $\tilde{u}^{*,m} = \{\tilde{u}_j\}_{j=0}^m$; $w^{*,m} = \{w_j(l\gamma), l \in I\!N, l\gamma \le \theta\}_{j=0}^m$ and δ_h is a measurable map from the state space of Υ_m to $[0,1]$ satisfying $0 \le \delta_h(\upsilon_m) \le \tilde{\Delta}_*^h$.

From the continuity properties of the kernel $\{q_{m+1,k}\}$ and the weak convergence of ϖ_m to ϖ, we have for all continuous and bounded functions F_1, F_2 defined on suitable spaces, as $h \to 0$,

$$\begin{aligned} &\int F_1(\upsilon_m) \sum_{k\in\mathcal{K}} q_{m+1,k}(z, \tilde{u}^{*,m}, w^{*,m})(\int F_2(\zeta_{m+1})d\Pi^k(\zeta_{m+1}))d\varpi_m^h(\upsilon_m) \\ \to\ &\int F_1(\upsilon_m) \sum_{k\in\mathcal{K}} q_{m+1,k}(z, \tilde{u}^{*,m}, w^{*,m})(\int F_2(\zeta_{m+1})d\Pi^k(\zeta_{m+1}))d\varpi_m(\upsilon_m). \quad (2.77) \end{aligned}$$

Next, from Lemma 2.4.10, for all sequences $\{\delta_h\}$ converging to 0 and compact sets E (of Euclidean space of appropriate dimension), as $h \to 0$,

$$\sup_{k\in\mathcal{K},\tilde{u}^{*,m}\in\eta\mathcal{K}^{m+1},w^{*,m}\in E}\left|\int F_2(\zeta_{m+1})q_{m+1,k}(z_h,\tilde{u}^{*,m},w^{*,m})d\Pi_h^{k,\delta_h}(\zeta_{m+1})\right.$$
$$\left. - \int F_2(\zeta_{m+1})q_{m+1,k}(z_h,\tilde{u}^{*,m},w^{*,m})d\Pi^k(\zeta_{m+1})\right| \to 0. \tag{2.78}$$

The weak convergence of ϖ_{m+1}^h to ϖ_{m+1} now follows on combining (2.76), (2.77) and (2.78). This proves (2.73).

We now address convergence of the cost functions. First let T be as in Lemma 2.4.9. Recall $\eta^h = \inf\{n \geq 0 : Z_n^h \in \partial^h\}$ and $\tau^h = t_{\eta^h}^h$. Note that $\tau^h = \inf\{t \geq 0 : Z^h(t) \in \partial^h\}$ due to the piecewise constant nature of $Z^h(t)$.

Let $\tau = \inf\{t \geq 0 : Z(t) \notin \mathbb{S}^o\}$. It can be shown in a manner similar to that used in the proof of Lemma 2.4.1 that $\tau^h \to \tau$ as $h \to 0$ on the set $\{\tau = \infty\}$. Also, on the set $\{\tau < \infty\}$ for every $\delta > 0$ there exists $t \in [\tau, \tau + \delta)$ and $\epsilon > 0$ such that $\text{dist}(Z(t), \mathbb{S}) > \epsilon$. Furthermore, $|Z^h(t) - Z(t)|$ uniformly on $[0, T]$ and $\partial^h \to \partial\mathbb{S}$ as $h \to 0$. Together these three facts imply $\tau^h \wedge T \to \tau \wedge T$ as $h \to 0$. Therefore, since $(z_h, Z^h, U^h, W^h, \tau^h) \to (z, Z, U, W, \tau)$, by the dominated convergence theorem

$$J^h(z_h, U^h) = I\!E \int_{[0,\tau^h\wedge T)} e^{-\beta t} f(C^h(t))dt \to I\!E \int_{[0,\tau\wedge T)} e^{-\beta t} f(C(t))dt = J_\ell(z, U). \quad \blacksquare$$

Combining Corollary 2.2.5 and Propositions 2.4.8, 2.4.11 we have the following theorem.

Theorem 2.4.12 *Let $z \in \mathbb{S}$ and let $\{z_h, h > 0\}$ be a sequence with $z_h \in \mathbb{S}_\ell^h$ such that $z_h \to z$ as $h \to 0$. Then* $\lim_{p\to\infty} \lim_{\ell\to\infty} \lim_{h\to 0} V^h(z^h) = \lim_{p\to\infty} \lim_{\ell\to\infty} V_{\ell,p}(z) = V(z)$.

2.5 Computational Methods for the MDP

The convergence results in the previous section ensure that for small values of h, the MDP defined in Section 3 provides a good approximation to the diffusion control problem defined in Section 2. In this section we outline the numerical methods for solving the MDP. Specifically, we provide the algorithm through which we compute the value function (2.25) and the associated optimal control for each initial state $z_h \in \mathbb{S}^h_\ell$. In practice, we fix a value of h and use the associated MDP to provide approximations to the diffusion control problem. Thus, for the remainder of the section, we will take h as a fixed value and suppress it from the notation.

Specifying the controlled Markov chain. In Section 3, we specified a choice of transition probabilities and interpolation intervals which satisfy the local consistency criteria; see (2.15) - (2.17) and (2.20). Many variations of this choice are possible; when specifying the particular controlled Markov chain, consideration must be given to the numerical implementation. For example, note that the neither the denominators of the probabilities nor the interpolation intervals in (2.15) depend on the value of c. This was accomplished by allowing the self transition (x, y) to (x, y). Also, we have separated the pure diffusion effects from the effects of the consumption control. That is, as consumption always decreases wealth, we associate it with the transition from (x, y) to $(x - h, y)$ only. Recall that $\Delta(z, u) = 0$ for all z if $u = (3, c)$; that is, the interpolation interval is 0 if reflection occurs. Hence, using (2.24), a reflection step incurs no cost and thus $V(\ell + h, y) = V(\ell, y), V(x, \ell + h) = V(x, \ell)$. It is a consequence of Definition 2.3.1 that $I\!P[I_n = 3 | Z_n \in \partial_{\mathrm{R}} \setminus \partial^h] = 1$ for all n; that is, reflection is the only admissible action for states in the reflecting boundary. Therefore, by adjusting the transition probabilities associated with the diffusion and singular controls it is possible to eliminate states in the reflecting boundary without affecting the cost function. This modification helps in speeding up the convergence of the numerical scheme.

In what follows, we will assume that the reflecting boundary states have been eliminated and the appropriate adjustments to the transition probabilities made. Thus, the state space of the controlled Markov chain used in the numerical schemes is given by $\mathbb{S}_{\text{MDP}} \doteq \mathbb{S}_{\ell}^{h+} \setminus \partial_{\text{R}}$ and the control space is $\mathcal{U}_{\text{MDP}} \doteq \{0, 1, 2\} \times [0, p]$.

Dynamic programming equation. Let $U \equiv \{U_n, n = 0, 1, 2, \ldots\}$ be an admissible control sequence (see Definition 2.3.1) for the MDP with state space $\mathbb{S}_{\text{MDP}}$, control space $\mathcal{U}_{\text{MDP}}$, and initial state z. For the numerical methods it is convenient to work with the cost function

$$J(z, U) = I\!\!E \sum_{n=0}^{\eta-1} e^{-\beta t_n} f(C_n) \tilde{\Delta}(Z_n, C_n)\, 1_{\{I_n = 0\}}, \tag{2.79}$$

which is asymptotically equivalent to (2.24). Recall that the value function is given as $V(z) = \sup_{U \in \mathcal{A}(z)} J(z, U)$.

We now present the dynamic programming equation that characterizes the value function. We begin by introducing the class of feedback controls. A *feedback control* is a measurable function $\mathbf{u} : \mathbb{S}_{\text{MDP}} \to \mathcal{U}_{\text{MDP}}$. We write $\mathbf{u} = (\mathbf{i}, \mathbf{c})$ where $\mathbf{i}$ and $\mathbf{c}$ are the two coordinates of the function $\mathbf{u}$. Using such a function one can construct an admissible control pair (Z_n, U_n) recursively by setting $Z_0 = z_0$, $U_n = \mathbf{u}(Z_n), n \geq 0$, and

$$I\!\!P[Z_{n+1} \in \cdot | Z_0, \ldots, Z_n, U_0, \ldots, U_n] = p(Z_n, U_n, \cdot).$$

With an abuse of terminology we will refer to this sequence $\{U_n\}$ as a feedback control as well. Note that $U_n \equiv (I_n, C_n) = (\mathbf{i}(Z_n), \mathbf{c}(Z_n))$.

If $U = \{U_n, n = 0, 1, 2, \ldots\}$ is a feedback control then one can easily check that the pair (Z_n, U_n) is a Markov chain from which it follows that for all $z \in \mathbb{S}_{\text{MDP}}$,

$$J(z, U) = \sum_{\tilde{z} \in \mathbb{S}_{\text{MDP}}} r(z, \mathbf{u}(z), \tilde{z}) J(\tilde{z}, U) + f(\mathbf{c}(z)) \Delta^h(z, \mathbf{u}(z)) \tag{2.80}$$

where $r(z, \mathbf{u}(z), \tilde{z}) = e^{-\beta \Delta(z, \mathbf{u}(z))} p(z, \mathbf{u}(z), \tilde{z})$. Observing that $J(z, U) = 0$ for all $z \in \partial^h$, the summation above can be taken over $\tilde{z} \in \mathbb{S}^* \doteq \mathbb{S}_{\text{MDP}} \setminus \partial^h$.

We can write the above equality in a matrix form as follows. Let $|\mathbb{S}^*| = s$ and fix an ordering of all the states in $\mathbb{S}^*$, i.e. $\mathbb{S}^* = \{z_1, \ldots, z_s\}$. Let $F(\mathbf{u})$ be an $s \times 1$ vector whose i^{th} entry is $f(\mathbf{c}(z_i))\Delta(z_i, \mathbf{u}(z_i)), i = 1, \ldots, s$. Let $R(\mathbf{u})$ be the $s \times s$ matrix with the $(i,j)^{\text{th}}$ entry as $r(z_i, \mathbf{u}(z_i), z_j)$. Finally let $J(\mathbf{u})$ be the $s \times 1$ vector with the i^{th} entry being $J(z_i, U)$. Then using these matrices, (2.80) can be written as:

$$J(\mathbf{u}) = R(\mathbf{u})J(\mathbf{u}) + F(\mathbf{u}). \tag{2.81}$$

Next for $u \in \mathcal{U}_{\text{MDP}}$ let $R(u)$ be the $s \times s$ matrix with $(i,j)^{\text{th}}$ entry $r(z_i, u, z_j)$. From standard arguments (cf. Section 5.8 in [32]) it follows that the value function V satisfies the following dynamic programming equation:

$$V = \sup_{u \in \mathcal{U}_{\text{MDP}}} R(u)V + F(u), \tag{2.82}$$

where in the above equation V is interpreted as an $s \times 1$ vector whose the i^{th} entry is $V(z_i)$, and the supremum on the righthand side above is taken row by row.

The following contraction property is central in the characterization of the value function via the dynamic programming equation in (2.82). The proof of the following lemma relies on the fact that the cost is of the discounted form with a strictly positive discount factor at all diffusion control steps and although the discount is zero for singular control steps, such steps tend to push the process towards the boundary of the domain and thus cannot occur "too often".

Lemma 2.5.1 *For all feedback controls* $\mathbf{u}$, $R^n(\mathbf{u}) \to 0$ *as* $n \to \infty$.

Proof. Let $\mathbf{u}$ be a feedback control and denote $R^n(\mathbf{u})$ by simply R^n with entries r^n_{ij}, $i, j = 1, \ldots, s$. It suffices to show $\sum_{j=1}^s r^n_{ij} \to 0$ as $n \to \infty$, for each $i = 1, \ldots, s$.

Let (Z_n, U_n) be the controlled Markov chain associated with feedback control $\mathbf{u}$ and transition kernel as defined in Section 3 with the modifications discussed in this section, and let $\Delta_k \equiv \Delta(Z_k, U_k)$ be the associated interpolation intervals. Let $\eta \doteq \inf\{n : Z_n \in \partial^h\}$. A simple calculation yields for all $i = 1, \dots, s$:

$$\sum_{j=1}^{s} r_{ij}^n = I\!\!E_i[e^{-\beta \sum_{k=0}^{n-1} \Delta_k} \sum_{j=1}^{s} 1_{\{Z_n = z_j\}}],$$

where $I\!\!E_i$ denotes the expectation given that $Z_0 = z_i$.

Since the states in ∂^h are not included in $\mathbb{S}^*$ and $p(z, \mathbf{u}(z), z) = 1$ for $z \in \partial^h$, we have $1_{\{Z_n = z_j\}} = 0$ for $j = 1, \dots, s$ when $n \geq \eta$. Thus we have

$$\sum_{j=1}^{s} r_{ij}^n = I\!\!E_i[\, 1_{\{n<\eta\}} \sum_{j=1}^{s} e^{-\beta \sum_{k=0}^{n-1} \Delta_k} 1_{\{Z_n = z_j\}}]. \tag{2.83}$$

Fix $a \in \mathbb{Z}_+$; conditions on a will be specified later. Define:

$$\tilde{d} \doteq \#\{\theta \in \{1, 2, \dots, [n/a]\} : U_m = (0, \cdot) \text{ for some } m \in [(\theta-1)a, \theta a)\}.$$

Set $\tilde{d}' = [n/a] - \tilde{d}$. The integer a is used to group the steps of the chain from 1 to n together into intervals. The quantity $\tilde{d}$, ($\tilde{d}'$) counts the number of such intervals with at least one diffusion step (respectively, no diffusion steps). By (2.14) there is a $\delta > 0$ such that $\Delta_n \geq \delta$ for all diffusion steps (i.e. all n such that $U_n = (0, \cdot)$). Also, recall that $\Delta_n = 0$ if step n is not a diffusion step. Combining these observations we have:

$$I\!\!E_i[\, 1_{\{n<\eta\}} e^{-\beta \sum_{k=0}^{n-1} \Delta_k} \sum_{j=1}^{s} 1_{\{Z_n = z_j\}}] \leq e^{-\beta \frac{\delta}{4} [\frac{n}{a}]} + I\!\!E_i[\, 1_{\{n<\eta\}} \, 1_{\{\tilde{d}' > \frac{3}{4}[\frac{n}{a}]\}}]. \tag{2.84}$$

We will utilize the behavior of the singular controls to bound the second term on the righthand side of the line above. Let $E_1, E_2, \dots, E_{\tilde{d}'}$ denote the intervals containing no diffusion steps, each of size a. Let K_d denote the number of purchase control steps in E_d; then $a - K_d$ is the number of sales control steps in E_d. Due to the finiteness of the state space, the maximum number of successive transitions to the left is bounded; in particular, it is bounded by $B \doteq 2(\ell/h+1)$. Similarly, B is a bound on the maximum number of downward transitions in a row. From (2.17) we see that each sales control always pushes the chain downward. Thus, the application of too many sales controls in a row will cause the chain to hit the boundary. However, by (2.16) a purchase control potentially pushes the chain upward. Similarly, a purchase control always pushes the chain to the left, while a sales control has the potential to push the chain to the right. Thus, in order to avoid hitting the boundary, the number of sales controls must be properly balanced by the number of purchase controls. More precisely, if $n < \eta$ we must have $|K_d - (a - K_d)| < B$; that is, $(a-B)/2 < K_d < (a+b)/2,\ d = 1, \ldots, \tilde{d}'$. For $m \in E_d$ define $\tilde{L}_m \doteq 1_{\{Z_{m+1}-Z_m=(-h,0)',U_m=(1,0)\}}$. The random variable $\tilde{L}_m$ indicates if the chain moves strictly to the left at step m given that a purchase control is applied. Since on E_d there are no diffusion steps and movement to the left is only possible at purchase control steps, the number of increments δZ_k equal to $(-h,0)'$ on E_d is given by $\sum_{m \in E_d} \tilde{L}_m \doteq L_d$.

Let ϵ be chosen to satisfy $0 < \epsilon < p/2$. Recall that at a purchase control step the chain moves to the left with probability $q \doteq \lambda/(1+\lambda)$. Thus by Cramer's theorem (see Theorem 2.1.24 [13]) there exists a $\kappa \equiv \kappa(\epsilon)$ such that

$$I\!P[L_d < K_d(q-\epsilon)] \leq I\!P[|\frac{L_d}{K_d} - q| > \epsilon] \leq e^{-K_d\kappa} \leq e^{-\kappa(a-B)/2}, \tag{2.85}$$

where the last inequality follows from the bound on K_d. We claim that for each $d = 1, \cdots, \tilde{d}'$, $\{\eta < n\} \cap \{L_d > qK_d/2\} = \emptyset$. To see the claim suppose that $L_d > qK_d/2$. Then the number of upward steps ($\delta Z_k = (\cdot, h)'$) in E_d, given by $K_d - L_d$, is at most $(1-q/2)K_d$. The number of downward steps ($\delta Z_k = (\cdot, -h)'$) in E_d equals the number of sales control steps, $a - K_d$. Thus using the bounds on K_d we have

$$\#\{\text{down steps in } E_d\} - \#\{\text{up steps in } E_d\} \geq a - K_d - (1-q/2)K_d \ \geq\ aq/4 - (1+q/4)B,$$

which is greater than B for $a > B(4+q)/q$. Henceforth fix such an a. On the other hand, on the set $\{\eta < n\}$ we must have $|\#\{\text{down steps in } E_d\} - \#\{\text{up steps in } E_d\}| < B$; otherwise, the chain would hit the boundary. This leads to a contradiction and thus the claim holds. Combining this with (2.85) we have that

$$\begin{aligned}
I\!\!E_i[\,1_{\{n<\eta\}}\,1_{\{\tilde{d}'>\frac{3}{4}[\frac{n}{a}]\}}] &= I\!\!E_i[\,1_{\{n<\eta,\tilde{d}'>\frac{3}{4}[\frac{n}{a}]\}}\prod_{d=1}^{\tilde{d}'} 1_{\{L_d \le qK_d/2\}}] \\
&\le I\!\!E_i[\prod_{d=1}^{\frac{3}{4}[\frac{n}{a}]} 1_{\{L_d \le qK_d/2\}}] \\
&\le e^{-\kappa\frac{a-B}{2}\frac{3}{4}[\frac{n}{a}]}.
\end{aligned}$$

Finally, by (2.83), (2.84), and the above, we have $\sum_{j=1}^{s} r_{ij}^n \le e^{-\beta\frac{\delta}{4}[\frac{n}{a}]} + e^{-\kappa\frac{a-B}{2}\frac{3}{4}[\frac{n}{a}]}$. The result now follows on noting that the term on the right approaches 0 as $n \to \infty$. ■

An immediate consequence of the lemma (cf. Section 2.3 of [32]) is the following.

Theorem 2.5.2 *For any feedback control* $\mathbf{u}$*,* $J(\mathbf{u})$ *is the unique solution to the equation* $v = R(\mathbf{u})v + F(\mathbf{u})$. *Furthermore, the value function* $\{V(z), z \in \mathbb{S}^*\}$ *defined below* (2.79) *is the unique solution of* (2.82)*. Denoting the* arg max *for the* i^{th} *row maximization on the right side of* (2.82) *by* $\mathbf{u}(z_i)$ *and the control sequence corresponding to the feedback control* $\mathbf{u}$ *by* $U = \{U_n, n = 0, 1, 2, \ldots\}$*, we have that* U *is an optimal control, i.e.* $J(z, U) = V(z)$ *for all* $z \in \mathbb{S}^*$.

From the above theorem it follows that in order to compute the value function and the optimal control it suffices to solve the equation (2.82).

Numerical Methods. We will use classical iterative methods to find the optimal control by solving the dynamic programming equation (2.82). A sketch of the algorithm is provided here. Details can be found in Chapter 6 of [32].

The following theorem provides the basis for the numerical approximation of the optimal control. We refer the reader to Theorem 6.2.1 in [32] for a proof.

Theorem 2.5.3 *Let* $\mathbf{u}_0$ *be a feedback control. Define a sequence of feedback controls* $\{\mathbf{u}_n, n \geq 1\}$ *and costs* $\{J(\mathbf{u}_n), n \geq 1\}$ *recursively as follows. Given* $\mathbf{u}_n$*, define*

$$
\begin{aligned}
J(\mathbf{u}_n) &= R(\mathbf{u}_n)J(\mathbf{u}_n) + F(\mathbf{u}_n), && (2.86)\\
\mathbf{u}_{n+1} &\doteq \arg\max_{u \in \mathcal{U}_{MDP}} R(u)J(\mathbf{u}_n) + F(u), && (2.87)
\end{aligned}
$$

where the $\arg\max$ *on the righthand side is computed row by row. Then* $J(\mathbf{u}_n) \to V$ *as* $n \to \infty$.

Given some control, (2.87) provides a way of "updating" the control in the search for the optimal control. However, this requires solving (2.86) to obtain the cost associated with the given control. Finding an exact solution to this equation can be numerically intensive since it involves the inversion of an $s \times s$ matrix. Thus, we use instead an approximation to the cost function $J(\mathbf{u}_n)$ in (2.87). The following theorem provides a method for obtaining such an approximation. We refer the reader to Theorem 6.2.2 in [32] for the proof.

Theorem 2.5.4 *Let* $\mathbf{u}$ *be an admissible feedback control. Then for any initial* $s \times 1$ *vector* $\tilde{J}_0$ *the sequence defined recursively by:*

$$
\tilde{J}_{n+1} = R(\mathbf{u})\tilde{J}_n + F(\mathbf{u}), \tag{2.88}
$$

converges to $J(\mathbf{u})$.

The numerical method for finding the optimal control is obtained by combining Theorems 2.5.3 and 2.5.4 as follows.

Policy iteration: Having determined an approximation to $J(\mathbf{u}_n)$, denoted as $\tilde{J}(\mathbf{u}_n)$, one obtains $\mathbf{u}_{n+1}$ by solving the minimization problem in (2.87) by replacing $J(\mathbf{u}_n)$ there by $\tilde{J}(\mathbf{u}_n)$.

Value iteration: Given $\mathbf{u}_n$, iterate (2.88) a large number of times (say m) with $R(\mathbf{u})$ there replaced by $R(\mathbf{u}_n)$ and initial value $\tilde{J}_0$ replaced by $\tilde{J}(\mathbf{u}_{n-1})$. Set $\tilde{J}(\mathbf{u}_n) \doteq \tilde{J}_m$.

The numerical algorithm alternates between policy iterations and value iterations until some suitable stopping criterion is met. Several modifications of (2.88) are often used to improve numerical efficiency; see Section 6.2.4 of [32] for details.

2.6 Numerical Study

We now present the results of a small pilot study using the method described in Section 2.5. We consider one of the examples in [50]. As in that reference, we set $r = 0.07$, $b = 0.12$, $\sigma = 0.40$, and $\beta = 0.10$. We consider the case $\lambda = \mu = 0.01$ and the utility function $f(c) = 2\sqrt{c}$. We take $\ell = 10$ as in [50] and $p = 10$. The discretization parameter is taken to be $h = 0.25$. (Note that [50] uses $h = 0.025$.)

To implement the numerical algorithm, we choose an initial feedback control matrix $\mathbf{u}_0$ given by, for $z \in \mathbb{S}^*$,

$$\mathbf{u}_0(z) = \begin{cases} (0, p), & x \geq 0, y \geq 0, \\ (1, 0), & x \geq 0, y < 0, \\ (2, 0), & \text{otherwise.} \end{cases}$$

Based on this control, the no-transaction region is the first quadrant of $\mathbb{R}^2$, and we always exercise the maximum amount of possible consumption. For $z \in \mathbb{S}^*$ we take $\tilde{J}_0(z)$ to be 75% of the value function computed in the absence of transaction costs; see equation (2.5) in [12].

We ran the algorithm described in the previous section. Figure 2.3 displays the first quadrant of the state space and illustrates the optimal control for this region. We see that the no-transaction region looks roughly like a cone. Consumption states are represented by the dots, purchase states by the plus signs, and sales states by the X's. The estimated

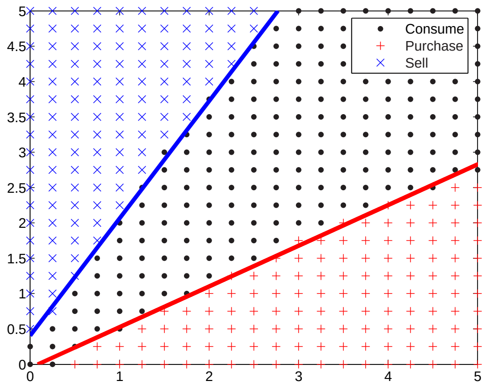

Figure 2.3: Numerically computed optimal control.

boundaries of the no-transaction region (the solid lines in the figure) are given by the lines $y = 0.575x - 0.050$ (the boundary of the "buy" region) and $y = 1.659x + 0.405$ (the boundary of the "sell" region.) The estimated sell boundary of the no-transaction region is similar to that obtained by Tourin and Zariphopoulou (see Figure 2.1 in [50]). However, the slope of our buy boundary appears to be lower than the slope illustrated in Figure 2.1 of [50]. A possible reason for this could be the difference in the discretization parameter. We used $h = 0.25$ to produce the test results provided here. Within the no-transaction region, consumption remains at a fairly constant percentage of wealth, 11.5%, which is very close to the constant percentage of consumption in the case of no transaction costs (see Theorem 2.1 in [12]). We also compare the value function computed by the algorithm versus the value function in the case of no transaction costs (again, see Theorem 2.1 in [12]). In general, the optimal value for an initial state computed in the presence of transaction costs is roughly 97% of the optimal value for the same state in the absence of transaction costs. However, when the initial wealth is small, this percentage tends to be lower (roughly 80% to 90%).

Chapter 3

Existence of Optimal Controls for Singular Control Problems with State Constraints

3.1 Introduction

This chapter is concerned with a general class of singular control problems with state constraints. In our setting, state constraints refers to the requirement that the controlled diffusion process take values in a closed convex cone at all times (see (3.3)). We consider an infinite horizon discounted cost of the form (3.4). The main objective of this chapter is to establish existence of an optimal control.

Classical compactness arguments which are used for establishing existence of optimal controls for problems with absolutely continuous control terms (cf. [6]) do not naturally extend to singular control problems. For one-dimensional models one can typically establish existence constructively, by characterizing the optimal controlled process as a reflected diffusion (cf. [2, 3, 23]). In higher dimensions, one approach is through studying regularity of solutions of the variational inequalities associated with singular control problems and the smoothness of the corresponding free boundary. Such smoothness results are the starting points in the characterization of the optimal controlled process as a constrained diffusion with reflection at the free boundary. Except for specific models (cf. [46, 47]), this approach encounters substantial difficulties, even for linear dynamics (cf. [51]). A key difficulty is that little is known about the regularity of the free boundary in higher dimensions. Alternative approaches for establishing existence of optimal controls based on compactness arguments are developed in [40, 25, 15]. The first of these papers considers linear dynamics while the last two consider models with nonlinear coefficients. In all three papers the state space is all of $\mathbb{R}^d$, i.e. there are no state constraints. It is important to note that, in our model, although the drift and diffusion coefficients are constant, the state constraint requirement introduces a (non-standard) nonlinearity in the dynamics. To the best of our knowledge, the current work is the first to address existence of an optimal control for a general multi-dimensional singular control problem with state constraints. While our method does not provide any characterization of the optimal control, it is quite general and should be applicable for other families of singular control problems (with or without state constraints).

State constraints are a natural feature in many practical applications of singular control. A primary motivation for the problems considered in this chapter arises from applications in controlled queueing systems such as the problem studied in Chapter 4. Under "heavy traffic conditions", formal diffusion approximations of a broad family of queuing networks with scheduling control, lead to the so-called Brownian control problems (BCPs) (cf. [20]). The BCP in turn can be transformed using techniques introduced in Harrison and Van-Mieghem [21] to a singular control problem with state constraints. We refer the reader to [1] for a concise description of connections between Brownian control problems and the class of singular control problems studied in [1] and the current chapter. In Section 4.3 we indicate how the results of the current chapter lead to existence of optimal controls for BCPs. State constraints arise in numerous other applications. For example, see Chapter 2 of the current work and Duffie, Fleming, Soner, and

Zariphopoulou [14] (and references therein) for control problems with state constraints in mathematical finance.

In Section 3.2 we introduce the singular control problem of interest. The main result of this chapter (Theorem 3.2.3) establishes existence of an optimal control. An important application of such a result lies in establishing connections between singular control problems and so-called obstacle/optimal stopping problems (see [45]). Indeed, in Chapter 4, using the existence result of the current chapter we will establish equivalence between a two-dimensional singular control problem that arises from the so-called criss-cross network, and an optimal stopping problem. Connections between singular control and optimal stopping, in addition to being of intrinsic mathematical interest, have important practical implications. Singular control problems rarely admit closed form solutions and therefore numerical approximation methods are necessary. Although numerical schemes for singular control problems are notoriously hard, optimal stopping problems have many well studied numerical schemes (cf. [32]). Exploiting connections between singular control and optimal stopping is expected to lead to simpler and more efficient numerical solution methods. In Chapter 4, using the above equivalence as a starting point we develop a numerical scheme for a scheduling control problem for a two dimensional queuing network model.

The basic idea in the proof of Theorem 3.2.3 is quite simple. For a given initial condition w, we choose a sequence of controls $\{U_n\}$ such that the corresponding cost sequence $\{J(w, U_n)\}$, converges to the value function $V(w)$. We then show that there is an admissible control U such that $\liminf J(w, U_n) \geq J(w, U)$, which completes the proof of the Theorem. Indeed, for problems with absolutely continuous controls this is the standard compactness argument (cf. [6]); one argues that that the sequence $\{U_n\}$ is tight in a suitable topology, picks a weak limit point U, and establishes the above inequality for this U using straightforward weak convergence arguments. The key difficulty in singular control problems is arguing compactness of the control sequence in a suitable topology;

the usual Skorohod topology on $D([0,\infty))$ is unsuitable as is suggested by the main result (Proposition 3.3.3) of Section 3.3. This result shows that for a typical discontinuous control U one can construct a sequence of continuous controls $\{U_n\}$ such that the costs for U_n converge to that for U; however, clearly U_n cannot converge to U in the usual Skorohod topology on $D([0,\infty))$. A powerful technique for bypassing this tightness issue, based on suitable stretching of time scale was introduced in [33]. Although such time transformation ideas go back to the work of of Meyer and Zheng [42] (see also Kurtz [31]), the papers [33, 37] were the first to use such ideas in stochastic control problems. A similar technique was also recently used in [7]. This "time stretching" technique is at the heart of our proof. Time transformation for the n-th control U_n is defined in such a way that, viewed in the new time scale, the process U_n is Lipschitz continuous with Lipschitz constant 1. Tightness in $D([0,\infty))$ (with the usual Skorohod topology) of the time transformed control sequence is then immediate. Finally, in order to obtain the candidate U for the above inequality, one must revert, in the limit, to the original time scale. This crucial step is achieved through Lemmas 3.4.2 and 3.4.3. The proof of the main result then follows via standard martingale characterization arguments and the optional sampling theorem.

Proof of Theorem 3.2.3 is facilitated by the result (Proposition 3.3.3) that the infimum of the cost over all admissible controls is the same as that over all admissible controls with continuous sample paths a.s. Although it may be possible to prove Theorem 3.2.3 without appealing to such a result, we believe that the result is of independent interest and it simplifies the proof of the main result considerably. The main difficulty in the proof of Proposition 3.3.3 is that if one approximates an arbitrary RCLL admissible control by a standard continuous approximation (cf. [39]), state constraints may be violated. Making sure that the continuous approximation is chosen in a manner that state constraints are satisfied is the key idea in the proof.

The chapter is organized as follows. We define the control problem and state the main result in Section 3.2. In Section 3.3 we characterize the value function as the infimum of the cost over all continuous controls. Section 3.4 is devoted to the proof of the main result. Finally, we briefly describe connections with Brownian control problems and stochastic networks in Section 4.3 .

Notation. We will use the following notation. If $\mathcal{X} \subset I\!\!R^d$ and A is an $n \times d$ matrix, then $A\mathcal{X} \doteq \{Ax : x \in \mathcal{X}\}$. A set $\mathcal{C} \subset I\!\!R^d$ is a cone of R^d if $c \in \mathcal{C}$ implies $ac \in \mathcal{C}$ for all $a \geq 0$. A function $f : [0, \infty) \to I\!\!R^d$ is said to have increments in $\mathcal{X}$ if $f(0) \in \mathcal{X}$ and $f(t) - f(s) \in \mathcal{X}$ for all $0 \leq s \leq t$. A stochastic process is said to have increments in $\mathcal{X}$ if, with probability one, its sample paths have increments in $\mathcal{X}$.

3.2 Setting and Main Result

The basic setup is the same as in [1]. Let $\mathcal{W}$ (respectively $\mathcal{U}$) be a closed convex cone of $I\!\!R^k$ ($I\!\!R^p$) with non-empty interior. We consider a control problem in which a p-dimension control process U, whose increments take values in $\mathcal{U}$, keeps a k-dimensional state process $W(t) \doteq w + B(t) + GU(t)$ in $\mathcal{W}$, where G is a fixed $k \times p$ matrix of rank k ($k \leq p$) and B is a k-dimensional Brownian motion with drift b and covariance matrix Σ given on some filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}, I\!\!P)$. We will refer to $\Phi \doteq (\Omega, \mathcal{F}, \{\mathcal{F}_t\}, I\!\!P, B)$ as a *system.* We assume that $G\mathcal{U} \cap \mathcal{W}^o \neq \emptyset$. Fix a unit vector $v_0 \in (G\mathcal{U})^o \cap \mathcal{W}^o$. Pick $u_0 \in \mathcal{U}$ for which $Gu_0 = v_0$. We also require that there exist $\hat{v}_1 \in I\!\!R^k$, $\hat{u}_1 \in I\!\!R^p$ and $a_0 \in (0, \infty)$ such that

$$v \cdot \hat{v}_1 \geq a_0|v|, v \in G\mathcal{U},\ w \cdot \hat{v}_1 \geq a_0|w|, w \in \mathcal{W},\ u \cdot \hat{u}_1 \geq a_0|u|, u \in \mathcal{U}. \tag{3.1}$$

The vectors $u_0, v_0, \hat{u}_1, \hat{v}_1$ will be fixed for the rest of the chapter.

Definition 3.2.1 *(Admissible control) An $\{\mathcal{F}_t\}$-adapted p-dimensional RCLL process U is an admissible control for the system Φ and initial data $w \in \mathcal{W}$ if the following two* *conditions hold $I\!\!P$-a.s.:*

$$U \text{ has increments in } \mathcal{U}, \tag{3.2}$$

$$W(t) \doteq w + B(t) + GU(t) \in \mathcal{W}, t \geq 0. \tag{3.3}$$

By convention, $U(0-) = 0$ and $W(0-) = w$. The process W is referred to as the controlled process associated with U and the pair (W, U) is referred to as an admissible pair for Φ and w. Let $\mathcal{A}(w, \Phi)$ denote the class of all such admissible controls.

The cost associated with given system Φ, initial data $w \in \mathcal{W}$ and admissible pair (W, U) is given by

$$J(w, U) \doteq I\!\!E \int_0^\infty e^{-\gamma t} \ell(W(t)) dt + I\!\!E \int_{[0,\infty)} e^{-\gamma t} h \cdot dU(t), \tag{3.4}$$

where $\gamma \in (0, \infty)$, $h \in I\!\!R^p$, and $\ell : \mathcal{W} \to [0, \infty)$ is a continuous function for which there exist constants $c_{\ell,1}, c_{\ell,2}, c_{\ell,3} \in (0, \infty)$ and $\alpha_\ell \in [0, \infty)$, depending only on ℓ, such that

$$c_{\ell,1}|w|^{\alpha_\ell} - c_{\ell,2} \leq \ell(w) \leq c_{\ell,3}(|w|^{\alpha_\ell} + 1), w \in \mathcal{W}. \tag{3.5}$$

We remark that the assumption on ℓ made above is weaker than that made in [1]. We also assume that $h \cdot \mathcal{U} \doteq \{h \cdot u : u \in \mathcal{U}\} \subset I\!\!R_+$.

The *value function* of the control problem for initial data $w \in \mathcal{W}$ is

$$V(w) = \inf_{\Phi} \inf_{U \in \mathcal{A}(w,\Phi)} J(w, U), \tag{3.6}$$

where the outside infimum is taken over all probability systems Φ. Lemma 4.4 of [1] shows that V is finite everywhere. The following assumption will be needed for the main result of the chapter.

Condition 3.2.2 *(i) Either $\alpha_\ell > 0$ or there exists $a_1 \in (0, \infty)$ such that $h \cdot u \geq a_1 |u|$*

for all $u \in \mathcal{U}$.

(ii) There exists $c_G \in (0,\infty)$ such that $|Gu| \geq c_G|u|$ for all $u \in \mathcal{U}$.

The following theorem, which guarantees the existence of an optimal control for the above control problem, is the main result of this chapter. The proof is postponed until Section 3.4.

Theorem 3.2.3 *Suppose that Assumption 3.2.2 holds. For all $w \in \mathcal{W}$ there exists a system Φ^* and a control $U^* \in \mathcal{A}(w, \Phi^*)$ such that $V(w) = J(w, U^*)$.*

3.3 Restriction to Continuous Controls

The main result of this section is Proposition 3.3.3, in which we show that in (3.6) it is enough to consider the infimum over the class of admissible controls with continuous paths. The use of continuous controls will play an important role in the time rescaling ideas used in the convergence proofs of Section 3.4.

The proof of Proposition 3.3.3 involves choosing an arbitrary control and constructing continuous approximations to it. We show that the cost functions associated with the approximating controls approach the cost function of the chosen control as the approximation parameter approaches its limit. The main difficulty of the proof lies in constructing *admissible* approximating controls, in particular, constructing the controls so that state constraints are satisfied. Such a construction is achieved via use of the Skorohod map, which is made precise in the following lemma. We refer the reader to Lemma 4.1 in [1] for a proof. We recall that for $T \geq 0$ and $\phi \in \mathcal{D}([0,\infty) : \mathbb{R}^k)$, $|\phi|^*_T$ denotes $\sup_{0\leq t\leq T} |\phi(t)|$.

Lemma 3.3.1 *There exist maps $\Gamma : \mathcal{D}([0,\infty) : \mathbb{R}^k) \to \mathcal{D}([0,\infty) : \mathbb{R}^k)$ and $\hat{\Gamma} : \mathcal{D}([0,\infty) : \mathbb{R}^k) \to \mathcal{D}([0,\infty) : \mathbb{R}_+)$ with the following properties. For any $\phi \in \mathcal{D}([0,\infty) : \mathbb{R}^k)$ with $\phi(0) \in \mathcal{W}$ define $\eta \doteq \hat{\Gamma}(\phi)$ and $\psi \doteq \Gamma(\phi)$. Then for all $t \geq 0$:*

1. $\eta(t) \in I\!R_+$, *and* η *is nondecreasing and RCLL.*

2. $\psi(t) \in \mathcal{W}$, *and* $\psi(t) = \phi(t) + v_0\eta(t)$.

3. *If* $\phi(t) \in \mathcal{W}$ *for all* $t \geq 0$ *then* $\Gamma(\phi) = \phi$ *and* $\hat{\Gamma}(\phi) = 0$.

Furthermore, the maps Γ *and* $\hat{\Gamma}$ *are Lipschitz continuous in the following sense. There exists* $\kappa \in (0, \infty)$ *such that for all* $\phi_1, \phi_2 \in \mathcal{D}([0,\infty) : I\!R^k)$ *with* $\phi_1(0), \phi_2(0) \in \mathcal{W}$ *and all* $T \geq 0$*:*

$$|\Gamma(\phi_1) - \Gamma(\phi_2)|_T^* + |\hat{\Gamma}(\phi_1) - \hat{\Gamma}(\phi_2)|_T^* \leq \kappa|\phi_1 - \phi_2|_T^*. \tag{3.7}$$

Before stating the main result of this section, we present the following lemma which states that we can further restrict attention to controls satisfying certain properties. The proof is contained in that of Lemma 4.7 of [1] and therefore is omitted.

Lemma 3.3.2 *For* $w \in \mathcal{W}$ *and a system* Φ, *let*

$$\begin{aligned} \mathcal{A}'(w,\Phi) \quad &= \quad \{U \in \mathcal{A}(w,\Phi) : \lim_{t\to\infty} e^{-\gamma t} I\!E|W(t)|^r = 0 \text{ and } I\!E\int_0^\infty e^{-\gamma t}|W(t)|^r dt < \infty \\ &\qquad \text{for all } r > 0, \text{ where } W \text{ is the controlled process associated with } U\}. \end{aligned}$$

Then $V(w) = \inf_\Phi \inf_{U\in\mathcal{A}'(w,\Phi)} J(w,U)$.

Proposition 3.3.3 *Let* Φ *be a system and* $w \in \mathcal{W}$. *Denote by* $\mathcal{A}^c(w,\Phi)$ *the class of all controls* $U \in \mathcal{A}(w,\Phi)$ *such that, for a.e.* ω, $t \mapsto U_t(\omega)$ *is a continuous map. Then*

$$V(w) = \inf_\Phi \inf_{U\in\mathcal{A}^c(w,\Phi)} J(w,U). \tag{3.8}$$

Proof. Fix $w \in \mathcal{W}$ and a system Φ. Let $U \in \mathcal{A}'(w,\Phi)$ be s.t. $J(w,U) < \infty$. Define $U^d(t) \doteq \sum_{0\leq s\leq t} \Delta U(s)$, where $\Delta U(s) = U(s) - U(s-)$, and $U^c(t) \doteq U(t) - U^d(t)$. That is, U^c is the continuous part and $U^d(t)$ is the pure jump part of the control U. Note that both processes are RCLL with increments in $\mathcal{U}$. We construct a sequence of continuous processes to approximate U^d as follows. For each integer $k \geq 1$ and $t \geq 0$ set

$$U_k^c(t) \doteq k \int_{(t-1/k)^+}^t U^d(s)ds + k(1/k - t)^+ U^d(0).$$

Note that for each k, U_k^c is continuous with increments in $\mathcal{U}$, and as $k \to \infty$,

$$U_k^c(t) \to U^d(t) \text{ a.e. } t \in [0, \infty), \quad \text{a.s.} \tag{3.9}$$

Also, from (3.1) it follows that for any function f with increments in $\mathcal{U}$, $t \mapsto f(t) \cdot \hat{u}_1$ and $t \mapsto Gf(t) \cdot \hat{v}_1$ are nondecreasing functions. This observation implies that, for all $T \geq 0$ and $0 \leq t \leq T$

$$a_0 |U_k^c(t)| \leq U_k^c(t) \cdot \hat{u}_1 \leq U^d(t) \cdot \hat{u}_1 \leq U^d(T) \cdot \hat{u}_1 \leq |U^d(T)|. \tag{3.10}$$

$$a_0 |GU_k^c(t)| \leq |GU^d(T)|, \quad a_0 |GU^d(t)| \leq |GU^d(T)|. \tag{3.11}$$

Thus by (3.9) and the dominated convergence theorem, for all $p > 0$,

$$\int_0^T |U_k^c(t) - U^d(t)|^p dt \to 0 \quad \text{a.s. as} \quad k \to \infty.$$

This suggests that a natural choice for the approximating control sequence is $\{U^c + U_k^c\}$. However this control may not be admissible since the corresponding state process $\tilde{W}_k$ defined as $\tilde{W}_k(t) \doteq w + B(t) + GU^c(t) + GU_k^c(t), t \geq 0$ may violate state constraints. We now use the Skorohod map introduced in Lemma 3.3.1 to construct an *admissible* continuous control. Define, for $t \geq 0$, $\eta_k(t) \doteq \hat{\Gamma}(\tilde{W}_k)(t)$ and

$$W_k(t) \doteq \Gamma(\tilde{W}_k)(t) = w + B(t) + GU^c(t) + GU_k^c(t) + Gu_0\eta_k(t). \tag{3.12}$$

Consider $U_k \doteq U^c + U_k^c + u_0\eta_k$. It is easily checked that $U_k(0) \geq 0$, U_k is continuous,

$\{\mathcal{F}_t\}$-adapted, and has increments in $\mathcal{U}$. Also, by Lemma 3.3.1, $W_k(t) \in \mathcal{W}$ for all $t \geq 0$. Thus U_k is an admissible control under Definition 3.2.1. We will now turn our attention to the corresponding cost functions. We begin by proving that $W_k(t) \to W(t)$ a.s as $k \to \infty$. The main idea is to appeal to the Lipschitz property (3.7); however, (3.9) establishes only pointwise convergence of $\tilde{W}_k$ to W and so a direct application of (3.7) is not useful. Define for each $k \geq 1$,

$$\bar{W}_k(t) \doteq k \int_{(t-1/k)^+}^{t} W(s)ds + k(1/k - t)^+ W(0).$$

Since $W(t) \in \mathcal{W}$ for all $t \geq 0$, it follows that $\bar{W}_k(t) \in \mathcal{W}$ for all $t \geq 0$, and thus $\bar{\eta}_k \doteq \hat{\Gamma}(\bar{W}_k) = 0$. Recalling the definition of U_k^c and using the Lipschitz property (3.7) we have for $T \geq 0$ and $0 \leq t \leq T$,

$$|W_k(t) - \bar{W}_k(t)| \leq \kappa \sup_{0 \leq t \leq T} \{|B(t) - k \int_{(t-1/k)^+}^{t} B(s)ds| + |G||U^c(t) - k \int_{(t-1/k)^+}^{t} U^c(s)ds|\}.$$

From the sample path continuity of B and U^c, the right side of the inequality approaches 0, almost surely, as $k \to \infty$. Next, since W has RCLL paths, $\bar{W}_k(t) \to W(t-)$, a.s. for every $t > 0$. Combining the above observations we have that:

$$W_k(t) \to W(t) \text{ for almost every } t \in [0, \infty), \text{ a.s.} \tag{3.13}$$

We now show that the costs associated with controls U_k converge to the cost corresponding to control U. We first consider the component of the cost arising from ℓ. Using (3.1), we have along the lines of equations (4.10)-(4.12) of [1], that there is $c_1 \in (0, \infty)$ such that for $0 \leq t \leq T < \infty$,

$$|GU^c(t)| + |GU^d(t)| + |W(t)| \leq c_1(|w| + |W(T)| + |B|_T^*). \tag{3.14}$$

Writing $W_k = W_k - W + W$ and using Lemma 3.3.1, we have for all $k \geq 1$ and $0 \leq t \leq T < \infty$:

$$|W_k(t)| \leq \kappa(|GU_k^c|_T^* + |GU^d|_T^*) + |W(t)| \leq c_2(|w| + |W(T)| + |B|_T^*), \tag{3.15}$$

where the second inequality follows on combining (3.14) and (3.11). Recalling (3.5), we get for some $c_3 \in (0, \infty)$,

$$\ell(W_k(t)) \leq c_3(|w|^{\alpha_\ell} + |W(t)|^{\alpha_\ell} + (|B|_t^*)^{\alpha_\ell} + 1).$$

Finally, since $U \in \mathcal{A}'(w, \Phi)$, we have from the above estimate, (3.13) and dominated convergence theorem that, as $k \to \infty$,

$$I\!\!E \int_0^\infty e^{-\gamma t}\ell(W_k(t))dt \to I\!\!E \int_0^\infty e^{-\gamma t}\ell(W(t))dt. \tag{3.16}$$

We now consider the component of the cost function associated with h. Note that since $I\!\!E \int_{[0,\infty)} e^{-\gamma t} h \cdot dU(t) \leq J(w, U) < \infty$, we have that

$$I\!\!E \int_{[0,\infty)} e^{-\gamma t} h \cdot dU(t) = \gamma \int_{[0,\infty)} e^{-\gamma t} I\!\!E(h \cdot U(t))dt < \infty. \tag{3.17}$$

Next, for $t \geq 0$:

$$\begin{aligned} |Gu_0\eta_k(t)| &\leq c_4(|w| + |W_k(t)| + |B(t)| + |GU^c(t)| + |GU^d(t)|) \\ &\leq c_5(|w| + |W_k(t)| + |W(t)| + |B|_t^*)) \\ &\leq c_6(|w| + |W(t)| + |B|_t^*), \end{aligned}$$

where the first inequality follows from (3.12) and (3.11), the second from (3.14), and the third from (3.15). Since η_k is nondecreasing, the above display implies $|\eta_k|_t^* \leq$

$c_7(|w| + |W(t)| + |B|_t^*)$. Thus since $U \in \mathcal{A}'(w, \Phi)$ we have that

$$\gamma \mathbb{E} \int_{[0,\infty)} e^{-\gamma t}(h \cdot u_0)\eta_k(t)dt = \mathbb{E} \int_{[0,\infty)} e^{-\gamma t} h \cdot u_0 d\eta_k(t) < \infty. \tag{3.18}$$

Next,

$$\begin{aligned}
& \mathbb{E} \int_{[0,\infty)} e^{-\gamma t} h \cdot dU_k(t) \\
= \; & \mathbb{E}\Big(\int_{[0,\infty)} e^{-\gamma t} h \cdot dU^c(t) + \int_{[0,\infty)} e^{-\gamma t} h \cdot dU_k^c(t) + \int_{[0,\infty)} e^{-\gamma t} h \cdot u_0 d\eta_k(t) \Big) \\
= \; & \gamma \mathbb{E}\Big(\int_{[0,\infty)} e^{-\gamma t} h \cdot U^c(t)dt + \int_{[0,\infty)} e^{-\gamma t} h \cdot U_k^c(t)dt + \int_{[0,\infty)} e^{-\gamma t} h \cdot u_0 \eta_k(t)dt \Big),
\end{aligned} \tag{3.19}$$

where the last line follows on using (3.18); noting that $\mathbb{E}(h \cdot (U_k^c(t) + U^c(t))) \leq \mathbb{E}(h \cdot U(t))$ and recalling that $J(w, U) < \infty$. From (3.17), (3.18) and (3.19), it now follows that $\mathbb{E} \int_{[0,\infty)} e^{-\gamma t} h \cdot dU_k(t)$ is finite and equals $\gamma \mathbb{E} \int_{[0,\infty)} e^{-\gamma t} h \cdot U_k(t)dt$.

From (3.9) and (3.13) we get that as $k \to \infty$,

$$(h \cdot u_0)\eta_k(t) \to 0 \text{ and } h \cdot U_k^c(t) \to h \cdot U^d(t), \text{ a.e. } t, \text{ a.s.} \tag{3.20}$$

Recalling that $|\eta_k|_t^* \leq c_7(|w| + |W(t)| + |B|_t^*)$ and that $U \in \mathcal{A}'(w, \Phi)$, equations (3.20) and (3.18) imply that, as $k \to \infty$,

$$\mathbb{E} \int_{[0,\infty)} e^{-\gamma t} h \cdot u_0 d\eta_k(t) \to 0. \tag{3.21}$$

Since $h \cdot U_k^c(t) \leq h \cdot U^d(t)$ and $\mathbb{E} \int_{[0,\infty)} e^{-\gamma t} h \cdot U^d(t)dt \leq J(w, U) < \infty$, we get that as $k \to \infty$,

$$\mathbb{E} \int_{[0,\infty)} e^{-\gamma t} h \cdot dU_k^c(t) \to \mathbb{E} \int_{[0,\infty)} e^{-\gamma t} h \cdot U^d(t)dt. \tag{3.22}$$

Finally, taking limits as $k \to \infty$ in (3.19) yields,

$$\mathbb{E} \int_{[0,\infty)} e^{-\gamma t} h \cdot dU_k(t) \to \mathbb{E} \int_{[0,\infty)} e^{-\gamma t} h \cdot dU(t). \tag{3.23}$$

Combining (3.16) and (3.23) we have $J(w, U_k) \to J(w, U)$ as $k \to \infty$. This proves the result. ■

3.4 Existence of an Optimal Control

In this section we prove our main result (Theorem 3.2.3) which guarantees existence of an optimal control for the control problem of Section 3.2. Fix $w \in \mathcal{W}$. From Proposition 3.3.3 we can find a sequence of systems $\{\Phi_n\}$, with $\Phi_n = (\Omega_n, \mathcal{F}_n, \{\mathcal{F}_n(t)\}, \mathbb{P}_n, B_n)$, and a sequence of controls $\{U_n\}$ with $U_n \in \mathcal{A}^c(w, \Phi_n), n \geq 1$ such that $J(w, U_n) < \infty$ for each n and

$$V(w) = \lim_{n \to \infty} J(w, U_n), \tag{3.24}$$

where

$$J(w, U_n) \doteq \mathbb{E}_n \int_0^\infty e^{-\gamma t} \ell(W_n(t)) dt + \mathbb{E}_n \int_{[0,\infty)} e^{-\gamma t} h \cdot dU_n(t), \tag{3.25}$$

and $\mathbb{E}_n$ denotes expectation with respect to $\mathbb{P}_n$. Let W_n be the state process corresponding to U_n, i.e.

$$W_n(t) \doteq w + B_n(t) + GU_n(t), \tag{3.26}$$

with $W_n(t) \in \mathcal{W}$ for all $t \geq 0$.

Time rescaling. For each $n \geq 1$ and $t \geq 0$ define

$$\tau_n(t) \doteq t + U_n(t) \cdot \hat{u}_1 \tag{3.27}$$

Since $\hat{u}_1 \cdot U_n$ is continuous and nondecreasing, τ_n is continuous and strictly increasing.

Also for $0 \leq s \leq t$,

$$\tau_n(t) - \tau_n(s) \geq t - s, \quad \tau_n(t) - \tau_n(s) \geq a_0 |U_n(t) - U_n(s)|. \tag{3.28}$$

The rescaled time process, $\hat{\tau}_n$, is given by $\hat{\tau}_n(t) \doteq \inf\{s \geq 0 : \tau_n(s) > t\}$. Note that $\hat{\tau}_n$ is continuous and strictly increasing. Also, $t = \hat{\tau}_n(\tau_n(t)) = \tau_n(\hat{\tau}_n(t))$, $\hat{\tau}_n(t) \leq t \leq \tau_n(t)$, and $\hat{\tau}_n(s) < t$ if and only if $\tau_n(t) > s$.

We define the time rescaled processes via $\hat{B}_n(t) \doteq B_n(\hat{\tau}_n(t))$, $\hat{U}_n(t) \doteq U_n(\hat{\tau}_n(t))$, and $\hat{W}_n(t) \doteq W_n(\hat{\tau}_n(t))$. From (3.26), for $t \geq 0$

$$\hat{W}_n(t) = W_n(\hat{\tau}_n(t)) = w + B_n(\hat{\tau}_n(t)) + GU_n(\hat{\tau}_n(t)) = w + \hat{B}_n(t) + G\hat{U}_n(t). \tag{3.29}$$

Also, from (3.28), for $0 \leq s \leq t$,

$$\hat{\tau}_n(t) - \hat{\tau}_n(s) \leq t - s, \quad a_0 |\hat{U}_n(t) - \hat{U}_n(s)| \leq t - s. \tag{3.30}$$

Let $\mathcal{E}$ denote the space of continuous functions from $[0, \infty)$ to $I\!R^k \times [0, \infty) \times I\!R^k \times \mathcal{U} \times \mathcal{W}$, endowed by the usual topology of uniform convergence on compacts. Note that for each $n \geq 1$ $(B_n, \hat{\tau}_n, \hat{B}_n, \hat{U}_n, \hat{W}_n)$. is a random variable with values in the Polish space $\mathcal{E}$. We next consider tightness of the family $\{(B_n, \hat{\tau}_n, \hat{B}_n, \hat{U}_n, \hat{W}_n), n \geq 1\}$.

Lemma 3.4.1 *The family $\{(B_n, \hat{\tau}_n, \hat{B}_n, \hat{U}_n, \hat{W}_n), n \geq 1\}$ is tight.*

Proof. Clearly, $\{B_n\}$ is tight. Tightness of $\{(\hat{\tau}_n, \hat{U}_n)\}$ follows from (3.30). Since $\hat{B}_n(t)$ is the composition of $B_n(\cdot)$ and $\hat{\tau}_n(\cdot)$, tightness of $\{\hat{B}_n\}$ follows from tightness of $\{(B_n, \hat{\tau}_n)\}$. Finally, tightness of $\{\hat{W}_n\}$ follows from (3.29) and tightness of $\{(\hat{\tau}_n, \hat{U}_n, \hat{B}_n)\}$. ■

Choose a convergent subsequence of $\{(B_n, \hat{\tau}_n, \hat{B}_n, \hat{U}_n, \hat{W}_n), n \geq 1\}$ (also indexed by n) with limit $(B', \hat{\tau}, \hat{B}, \hat{U}, \hat{W})$ defined on some probability space. Clearly, B' is a (b, Σ)-Brownian motion with respect to its own filtration. By the Skorohod representation theorem there exists a probability space $(\Omega^*, \mathcal{F}^*, I\!\!P^*)$ on which are defined a sequence of processes $\{(B'_n, \hat{\tau}'_n, \hat{B}'_n, \hat{U}'_n, \hat{W}'_n), n \geq 1\}$ and a process $(B'', \hat{\tau}', \hat{B}', \hat{U}', \hat{W}')$, such that $(B'_n, \hat{\tau}'_n, \hat{B}'_n, \hat{U}'_n, \hat{W}'_n) \stackrel{d}{=} (B_n, \hat{\tau}_n, \hat{B}_n, \hat{U}_n, \hat{W}_n)$, $(B'', \hat{\tau}', \hat{B}', \hat{U}', \hat{W}') \stackrel{d}{=} (B', \hat{\tau}, \hat{B}, \hat{U}, \hat{W})$, and $(B'_n, \hat{\tau}'_n, \hat{B}'_n, \hat{U}'_n, \hat{W}'_n) \to (B'', \hat{\tau}', \hat{B}', \hat{U}', \hat{W}')$ almost surely as $n \to \infty$. To simplify notation, we will assume (without loss of generality) that

$$(B_n, \hat{\tau}_n, \hat{B}_n, \hat{U}_n, \hat{W}_n) \to (B', \hat{\tau}, \hat{B}, \hat{U}, \hat{W}), \text{ almost surely } (I\!\!P^*) \text{ as } n \to \infty. \tag{3.31}$$

The following lemma is central in the time rescaling ideas used later in the section.

Lemma 3.4.2 *Suppose that Assumption 3.2.2 holds. Then there exists $\alpha^* \in (0, \infty)$ such that for all $t \geq 0$,*

$$\limsup_{n\to\infty} I\!\!E_n |U_n(t)|^{\alpha^*} < \infty. \tag{3.32}$$

Proof. From Assumption 3.2.2 we have that either $\alpha_\ell > 0$ or there exists $a_1 \in (0, \infty)$ such that $h \cdot u \geq a_1 |u|$ for all $u \in \mathcal{U}$. Suppose first that the latter condition holds. Then for all $t \geq 0$,

$$J(w, U_n) \geq \gamma I\!\!E_n \int_0^\infty e^{-\gamma t} h \cdot U_n(t) dt \geq \gamma e^{-\gamma(t+1)} I\!\!E_n (h \cdot U_n(t)) \geq \gamma a_1 e^{-\gamma(t+1)} I\!\!E_n |U_n(t)|.$$

Thus in this case (3.32) holds with $\alpha^* = 1$. Next suppose that $\alpha_\ell > 0$. From Assumption 3.2.2 and (3.26), we have:

$$c_G |U_n(t)| \leq |G U_n(t)| \leq |W_n(t)| + |B_n(t)| + |w|,$$

which implies that for some $c_1 \in (0, \infty)$

$$|U_n(t)|^{\alpha_\ell} \leq c_1 (|W_n(t)|^{\alpha_\ell} + |B_n(t)|^{\alpha_\ell} + |w|^{\alpha_\ell}).$$

Therefore, using moment properties of B_n we have for some $c_2 \in (0, \infty)$

$$\mathbb{E}_n |U_n(t)|^{\alpha_\ell} \le c_2(\mathbb{E}_n |W_n(t)|^{\alpha_\ell} + t^{\alpha_\ell} + 1). \tag{3.33}$$

Combining the above estimate with (3.5) we get

$$\limsup_{n\to\infty} \int_0^\infty e^{-\gamma s} \mathbb{E}_n |U_n(s)|^{\alpha_\ell} ds < \infty. \tag{3.34}$$

Finally,

$$\int_0^\infty e^{-\gamma t} \mathbb{E}_n |U_n(t)|^{\alpha_\ell} dt \ge \int_0^\infty e^{-\gamma t} \mathbb{E}_n (\hat{u}_1 \cdot U_n(t))^{\alpha_\ell} dt \ge e^{-\gamma(t+1)} a_0^{\alpha_\ell} \mathbb{E}_n |U_n(t)|^{\alpha_\ell}.$$

The inequality (3.32) now follows with $\alpha^* = \alpha_\ell$ on combining the above inequality with (3.34). ■

The following lemma, a consequence of Lemma 3.4.2, gives a critical property of $\hat{\tau}$.

Lemma 3.4.3 *Suppose that Assumption 3.2.2 holds.*

$$\hat{\tau}(t) \to \infty \text{ as } t \to \infty, \mathbb{P}^*\text{-a.s.} \tag{3.35}$$

Proof. Fix $M > 0$ and consider $t \in (M, \infty)$. Since $\hat{\tau}_n(t) < M$ if and only if $\tau_n(M) > t$ we have by (3.27),

$$\{\hat{\tau}_n(t) < M\} = \{M + U_n(M) \cdot \hat{u}_1 > t\} \subset \{|U_n(M)| > (t - M)\}.$$

Recall the constant α^* in Lemma 3.4.2. The above relation and an application of Markov's inequality yield for all $t > M$

$$\mathbb{P}^*[\hat{\tau}_n(t) < M] \le \mathbb{P}^*[|U_n(M)|^{\alpha^*} > (t - M)^{\alpha^*}] \le \frac{1}{(t - M)^{\alpha^*}} \mathbb{E}^* |U_n(M)|^{\alpha^*}.$$

Thus by the weak convergence $\hat{\tau}_n \Rightarrow \hat{\tau}$,

$$\begin{aligned} I\!P^*[\lim_{t\to\infty} \hat{\tau}(t) < M] &\leq \lim_{t\to\infty} \limsup_{n\to\infty} I\!P^*[\hat{\tau}_n(t) < M] \\ &\leq \lim_{t\to\infty} \frac{1}{(t-M)^{\alpha^*}} \limsup_{n\to\infty} I\!E^*|U_n(M)|^{\alpha^*}. \end{aligned}$$

The right side of the last inequality is 0 by Lemma 3.4.2. Since $M > 0$ is arbitrary, the result follows. ■

We now introduce an inverse time transformation which allows us to revert back to the original time scale. For $t \geq 0$ define $\tau(t) \doteq \inf\{s \geq 0 : \hat{\tau}(s) > t\}$. The following properties are easily checked.

- $\tau(t) < \infty$ a.s. for all $t \geq 0$. (This follows from Lemma 3.4.3.)
- τ is strictly increasing and right-continuous.
- $\tau(t) \geq t \geq \hat{\tau}(t)$. In particular, $\tau(t) \to \infty$ a.s. as $t \to \infty$.
- $0 \leq \hat{\tau}(s) \leq t \Leftrightarrow 0 \leq s \leq \tau(t)$, and $\hat{\tau}(\tau(t)) = t, \tau(\hat{\tau}(t)) \geq t$.

The time transformed processes are defined as $B^*(t) \doteq \hat{B}(\tau(t))$, $U^*(t) \doteq \hat{U}(\tau(t))$, $W^*(t) \doteq \hat{W}(\tau(t))$, $t \geq 0$. By (3.31) and (3.29) we have $\hat{W}(t) = w + \hat{B}(t) + G\hat{U}(t)$ for all $t \geq 0$, a.s., which implies

$$W^*(t) = \hat{W}(\tau(t)) = w + \hat{B}(\tau(t)) + G\hat{U}(\tau(t)) = w + B^*(t) + GU^*(t).$$

Note that U^* is RCLL with increments in $\mathcal{U}$ and $W^*(t) \in \mathcal{W}$ for all $t \geq 0$.

We next introduce a suitable filtration on $(\Omega^*, \mathcal{F}^*, I\!P^*)$. For $t \geq 0$ define the σ-fields $\hat{\mathcal{F}}'(t) \doteq \sigma\{(\hat{B}(s), \hat{U}(s), \hat{W}(s), \hat{\tau}(s)), 0 \leq s \leq t\}$ and $\hat{\mathcal{F}}_t \equiv \hat{\mathcal{F}}(t) \doteq \hat{\mathcal{F}}'(t+) \vee \mathcal{N}$, where $\mathcal{N}$ denotes the family $I\!P^*$-null sets. Then $\{\hat{\mathcal{F}}_t\}$ is a right-continuous, complete filtration. For any $s, t \geq 0$, $\{\tau(s) < t\} = \{\hat{\tau}(t) > s\} \in \hat{\mathcal{F}}(t)$. Therefore, since $\{\hat{\mathcal{F}}_t\}$ is right-continuous, $\tau(s)$ is an $\{\hat{\mathcal{F}}_t\}$-stopping time for any $s \geq 0$. For each $t \geq 0$, define the

σ-field $\mathcal{F}^*(t) \doteq \hat{\mathcal{F}}(\tau(t))$. Since τ is nondecreasing, $\{\mathcal{F}_t^*\}$ is a filtration. Clearly, $\hat{B}$ and $\hat{U}$ are $\{\hat{\mathcal{F}}_t\}$-adapted; therefore B^* and U^* are $\{\mathcal{F}_t^*\}$-adapted (cf. Proposition 1.2.18 of [29]). We show in Lemma 3.4.6 below that B^* is an $\{\mathcal{F}_t^*\}$-Brownian motion with drift b and covariance matrix Σ. Before stating this result, we present the following change of variables formula which we will use in the convergence analysis. We refer the reader to Theorem IV.4.5 of [43] for a proof.

Lemma 3.4.4 *Let a be a $\mathbb{R}_+$-valued, right-continuous function on $[0,\infty)$ such that $a(0) = 0$. Let c be its right-inverse, i.e. $c(t) \doteq \inf\{s \geq 0 : a(s) > t\}$, $t \geq 0$. Assume that $c(t) < \infty$ for all $t \geq 0$. Let f be a nonnegative Borel measurable function on $[0,\infty)$, and let F be a $\mathbb{R}_+$-valued, right-continuous, nondecreasing function on $[0,\infty)$. Then*

$$\int_{[0,\infty)} f(s)dF(a(s)) = \int_{[0,\infty)} f(c(s-))dF(s), \tag{3.36}$$

with the convention that the contribution to the integrals above at 0 is $f(0)F(0)$. In particular, taking $F(s) = s, s \geq 0$,

$$\int_{[0,\infty)} f(s)da(s) = \int_{[0,\infty)} f(c(s))ds. \tag{3.37}$$

Remark 3.4.5 *Recall that $\hat{B}_n(t) = B_n(\hat{\tau}_n(t))$. It follows from continuity and almost sure convergence of $(B_n, \hat{\tau}_n, \hat{B}_n) \to (B', \hat{\tau}, \hat{B})$ that $\hat{B}(t) = B'(\hat{\tau}(t))$ a.s. Thus, $B^*(t) \doteq \hat{B}(\tau(t)) = B'(\hat{\tau}(\tau(t))) = B'(t)$ a.s. In particular, B^* is a (b,Σ)-Brownian motion with respect to its own filtration. The following lemma shows that, in fact, B^* is a Brownian motion with respect to the larger filtration $\{\mathcal{F}_t^*\}$.*

Lemma 3.4.6 *B^* is an $\{\mathcal{F}_t^*\}$-Brownian motion with drift b and covariance matrix Σ.*

Proof. For any infinitely differentiable function $f : \mathbb{R}^k \to \mathbb{R}$ with compact support, define

$$Af(x) \doteq \sum_{i=1}^k b_i \frac{\partial}{\partial x_i} f(x) + \frac{1}{2}\sum_{i=1}^k \sum_{j=1}^k \sigma_{ij} \frac{\partial^2}{\partial x_i \partial x_j} f(x), \tag{3.38}$$

where the entries of b are denoted b_i and those of Σ are denoted σ_{ij}. First suppose

$$\begin{aligned} &I\!\!E^*[g(\hat{B}(s_m), \hat{U}(s_m), \hat{W}(s_m), \hat{\tau}(s_m), s_m \leq t, m = 1, \ldots, q) \\ &\quad \times \{f(\hat{B}(t+s)) - f(\hat{B}(t)) - \int_t^{t+s} Af(\hat{B}(u)) d\hat{\tau}(u)\}] = 0, \end{aligned} \tag{3.39}$$

for al $s, t \geq 0$, continuous bounded functions g (on a suitable domain), positive integers $q \geq 1$, and sequences $\{s_m\}$. Define for $t \geq 0$,

$$\hat{Y}_f(t) \doteq f(\hat{B}(t)) - \int_0^t Af(\hat{B}(u)) d\hat{\tau}(u).$$

Then by equation (3.39), $\hat{Y}_f$ is an $\{\hat{\mathcal{F}}'_t\}$-martingale, and therefore $\hat{Y}_f$ is an $\{\hat{\mathcal{F}}_t\}$-martingale as well. Recall that $\tau(s)$ is an $\{\hat{\mathcal{F}}_t\}$-stopping time such that $\tau(s) < \infty$ a.s. for all $s \geq 0$. Since f and Af are bounded (by some $c > 0$),

$$I\!\!E^*|\hat{Y}_f(\tau(t))| \leq I\!\!E^*|f(\hat{B}(\tau(t)))| + I\!\!E^* \int_0^{\tau(t)} |Af(\hat{B}(u))| d\hat{\tau}(u) \leq c + cI\!\!E^*|\hat{\tau}(\tau(t))| = c(1+t).$$

In addition, we have for any $T \in (0, \infty)$,

$$\begin{aligned} I\!\!E^*[|\hat{Y}_f(T)|\, 1_{\{\tau(t) > T\}}] &\leq I\!\!E^*[|\hat{Y}_f(T)|\, 1_{\{\hat{\tau}(T) \leq t\}}] \\ &\leq I\!\!E^*[\{|f(\hat{B}(T))| + \int_0^T |Af(\hat{B}(u))| d\hat{\tau}(u)\}\, 1_{\{\hat{\tau}(T) \leq t\}}] \\ &\leq cI\!\!E^*[(1 + \hat{\tau}(T))\, 1_{\{\hat{\tau}(T) \leq t\}}] \\ &\leq c(1+t) I\!\!P^*[\hat{\tau}(T) \leq t] \end{aligned}$$

The last term above approaches 0 as $T \to \infty$ by Lemma 3.4.3. Therefore by the optional sampling theorem (cf. Theorem 2.2.13 in [17]), we have for $s \leq t$,

$$I\!\!E^*[\hat{Y}_f(\tau(t)) | \mathcal{F}^*(s)] = I\!\!E^*[\hat{Y}_f(\tau(t)) | \hat{\mathcal{F}}(\tau(s))] = \hat{Y}_f(\tau(s)),$$

that is, $\hat{Y}_f(\tau(t))$ is an $\{\mathcal{F}_t^*\}$-martingale. Now

$$\begin{aligned}\hat{Y}_f(\tau(t)) &= f(\hat{B}(\tau(t))) - \int_0^\infty Af(\hat{B}(u))\, 1_{\{0\le u<\tau(t)\}} d\hat{\tau}(u) \\ &= f(B^*(t)) - \int_0^\infty Af(\hat{B}(\tau(u)))\, 1_{\{0\le \tau(u)<\tau(t)\}} du \\ &= f(B^*(t)) - \int_0^t Af(B^*(u))du,\end{aligned}$$

where we have used Lemma 3.4.4 and the fact that τ is strictly increasing. Thus,

$$\begin{aligned}&\mathbb{E}^*[f(B^*(t+s)) - f(B^*(t)) - \int_t^{t+s} Af(B^*(u))du|\mathcal{F}^*(t)] \\ =\ &\mathbb{E}^*[\hat{Y}_f(\tau(t+s)) - \hat{Y}_f(\tau(t))|\mathcal{F}^*(t)],\end{aligned}$$

which is 0 for any $s, t \ge 0$ since $\hat{Y}_f(\tau(t))$ is an $\{\mathcal{F}_t^*\}$-martingale. Therefore, B^* is an $\{\mathcal{F}_t^*\}$-Brownian motion with drift b and covariance Σ. Hence in order to prove the lemma it suffices to prove (3.39).

Recall that B_n is an $\{\mathcal{F}_n(t)\}$-Brownian motion with drift b and covariance Σ. Let f be as above and define $Y_{f,n}(t) \doteq f(B_n(t)) - \int_0^t Af(B_n(u))du$. Then $Y_{f,n}$ is an $\{\mathcal{F}_n(t)\}$-martingale for each $n \ge 1$.

Fix $t \ge 0$ and note that $\{\hat{\tau}_n(s) < t\} = \{\tau_n(t) > s\} = \{t + U_n(t) \cdot \hat{u}_1 > s\} \in \mathcal{F}_n(t)$ for all $s \ge 0$, $n \ge 1$. Thus for each $s \ge 0$, $\hat{\tau}_n(s)$ is an $\{\mathcal{F}_n(t)\}$-stopping time. Define $\hat{Y}_{f,n}(t) \doteq Y_{f,n}(\hat{\tau}_n(t))$ for $t \ge 0$. Since $\hat{\tau}_n(t)$ is an $\{\mathcal{F}_n(t)\}$-stopping time bounded by t, we have by the optional sampling theorem (see Problem 1.3.24 in [29]) that for any $s \ge 0$

$$\mathbb{E}_n[\hat{Y}_{f,n}(t+s)|\mathcal{F}_n(\hat{\tau}_n(t))] = \mathbb{E}_n[Y_{f,n}(\hat{\tau}_n(t+s))|\mathcal{F}_n(\hat{\tau}_n(t))] = Y_{f,n}(\hat{\tau}_n(t)) = \hat{Y}_{f,n}(t).$$

This implies that for any bounded $\mathcal{F}_n(\hat{\tau}_n(t))$-measurable function ξ_n,

$$\mathbb{E}_n[\xi_n\{\hat{Y}_{f,n}(t+s) - \hat{Y}_{f,n}(t)\}] = 0. \tag{3.40}$$

Now for any $s \leq t$, the random variables $B_n(\hat{\tau}_n(s)), U_n(\hat{\tau}_n(s)), W_n(\hat{\tau}_n(s))$ are $\mathcal{F}_n(\hat{\tau}_n(s))$-measurable (cf. Proposition 1.2.18 in [29]). Also, $\hat{\tau}_n(s)$ is $\mathcal{F}_n(\hat{\tau}_n(s))$-measurable (cf. Problem 1.2.13 in [29]). Since $\hat{\tau}_n(s) \leq \hat{\tau}_n(t)$ for $s \leq t$, we have $\mathcal{F}_n(\hat{\tau}_n(s)) \subset \mathcal{F}_n(\hat{\tau}_n(t))$. Thus,

$$g(B_n(\hat{\tau}_n(s_m)), U_n(\hat{\tau}_n(s_m)), W_n(\hat{\tau}_n(s_m)), \hat{\tau}_n(s_m), 0 \leq s_m \leq t, m = 1, \ldots, q),$$

is $\mathcal{F}_n(\hat{\tau}_n(s))$-measurable, for all bounded continuous functions g with appropriate domain, positive integers $q \geq 1$, and sequences $\{s_m\}$. Therefore, from (3.40)

$$\begin{aligned}&I\!\!E_n[g(\hat{B}_n(s_m), \hat{U}_n(s_m), \hat{W}_n(s_m), \hat{\tau}_n(s_m), 0 \leq s_m \leq t, m = 1, \ldots, q)\\&\quad \times\{\hat{Y}_{f,n}(t+s) - \hat{Y}_{f,n}(t)\}] = 0,\end{aligned}$$

Recalling our use of the Skorohod representation theorem above (3.31), we have

$$\begin{aligned}&I\!\!E^*[g(\hat{B}_n(s_m), \hat{U}_n(s_m), \hat{W}_n(s_m), \hat{\tau}_n(s_m), 0 \leq s_m \leq t, m = 1, \ldots, q)\\&\quad \times\{\hat{Y}_{f,n}(t+s) - \hat{Y}_{f,n}(t)\}] = 0. \qquad (3.41)\end{aligned}$$

Another application of Lemma 3.4.4 shows that:

$$\int_0^t Af(\hat{B}_n(u))d\hat{\tau}_n(u) = \int_0^{\hat{\tau}_n(t)} Af(B_n(u))du.$$

This implies

$$\begin{aligned}\hat{Y}_{f,n}(t) = Y_{f,n}(\hat{\tau}_n(t)) &= f(B_n(\hat{\tau}_n(t))) - \int_0^{\hat{\tau}_n(t)} Af(B_n(u))du\\&= f(\hat{B}_n(t)) - \int_0^t Af(\hat{B}_n(u))d\hat{\tau}_n(u). \qquad (3.42)\end{aligned}$$

Combining (3.41) and (3.42) we have

$$\begin{aligned}
&I\!\!E^*[g(\hat{B}_n(s_m), \hat{U}_n(s_m), \hat{W}_n(s_m), \hat{\tau}_n(s_m), 0 \le s_m \le t, m = 1, \ldots, q) \\
&\quad \times \{f(\hat{B}_n(t+s)) - f(\hat{B}_n(t)) - \int_t^{t+s} Af(\hat{B}_n(u)) d\hat{\tau}_n(u)\}] = 0. \qquad (3.43)
\end{aligned}$$

Finally, recall that $(B_n, \hat{\tau}_n, \hat{B}_n, \hat{U}_n, \hat{W}_n) \to (B', \hat{\tau}, \hat{B}, \hat{U}, \hat{W})$ $I\!\!P^*$-a.s. as $n \to \infty$. An application of the bounded convergence theorem yields (3.39) on taking $n \to \infty$ in (3.43) (cf. Lemma 2.4 of [11]). This completes the proof. ■

As an immediate consequence we have:

Corollary 3.4.7 *Let* $\Phi^* \doteq (\Omega^*, \mathcal{F}^*, I\!\!P^*, \{\mathcal{F}_t^*\}, B^*)$. *Then* $U^* \in \mathcal{A}(w, \Phi^*)$.

We now show that U^* is an optimal control by studying convergence of the cost functions $J(w, U_n)$, thus completing the proof of the main result.

Proof of Theorem 3.2.3. Let $\{U_n\}$ and U^* be as above. By Lemma 3.4.4 we have that the cost corresponding to the admissible pair (W_n, U_n) is given by

$$\begin{aligned}
J(w, U_n) &\doteq I\!\!E_n \int_0^\infty e^{-\gamma t} \ell(W_n(t)) dt + \gamma I\!\!E_n \int_0^\infty e^{-\gamma t} h \cdot U_n(t) dt \\
&= I\!\!E_n \int_0^\infty e^{-\gamma \hat{\tau}_n(t)} \ell(W_n(\hat{\tau}_n(t))) d\hat{\tau}_n(t) + \gamma I\!\!E_n \int_0^\infty e^{-\gamma \hat{\tau}_n(t)} h \cdot U_n(\hat{\tau}_n(t)) d\hat{\tau}_n(t) \\
&= I\!\!E^* \int_0^\infty e^{-\gamma \hat{\tau}_n(t)} \ell(\hat{W}_n(t)) d\hat{\tau}_n(t) + \gamma I\!\!E^* \int_0^\infty e^{-\gamma \hat{\tau}_n(t)} h \cdot \hat{U}_n(t) d\hat{\tau}_n(t). \qquad (3.44)
\end{aligned}$$

Since $(\hat{\tau}_n, \hat{U}_n, \hat{W}_n) \to (\hat{\tau}, \hat{U}, \hat{W})$ $I\!\!P^*$-a.s., we have (cf. Lemma 2.4 of [11]) for all $u \ge 0$ and $N \ge 1$,

$$\int_{[0,u)} [N \wedge e^{-\gamma \hat{\tau}_n(t)} \ell(\hat{W}_n(t))] d\hat{\tau}_n(t) \to \int_{[0,u)} [N \wedge e^{-\gamma \hat{\tau}(t)} \ell(\hat{W}(t))] d\hat{\tau}(t), I\!\!P^*\text{-a.s.},$$

as $n \to \infty$. Thus, we have, almost surely ($I\!\!P^*$),

$$\liminf_{n\to\infty} \int_0^\infty e^{-\gamma\hat{\tau}_n(t)} \ell(\hat{W}_n(t)) d\hat{\tau}_n(t) \geq \int_0^u [N \wedge e^{-\gamma\hat{\tau}(t)} \ell(\hat{W}(t))] d\hat{\tau}(t).$$

Taking limits as $N \to \infty$ and $u \to \infty$ on both sides yields

$$\liminf_{n\to\infty} \int_0^\infty e^{-\gamma\hat{\tau}_n(t)} \ell(\hat{W}_n(t)) d\hat{\tau}_n(t) \geq \int_0^\infty e^{-\gamma\hat{\tau}(t)} \ell(\hat{W}(t)) d\hat{\tau}(t), I\!\!P^*\text{-a.s.} \tag{3.45}$$

Similarly,

$$\liminf_{n\to\infty} \gamma \int_{[0,\infty)} e^{-\gamma\hat{\tau}_n(t)} h \cdot \hat{U}_n(t) d\hat{\tau}_n(t) \geq \gamma \int_{[0,\infty)} e^{-\gamma\hat{\tau}(t)} h \cdot \hat{U}(t) d\hat{\tau}(t), I\!\!P^*\text{-a.s.} \tag{3.46}$$

Therefore by (3.24), (3.44), Fatou's lemma, (3.45) and (3.46),

$$\begin{aligned}
V(w) &= \liminf_{n\to\infty} J(w, U_n) \\
&= \liminf_{n\to\infty} \{I\!\!E^* \int_0^\infty e^{-\gamma\hat{\tau}_n(t)} \ell(\hat{W}_n(t)) d\hat{\tau}_n(t) + \gamma I\!\!E^* \int_0^\infty e^{-\gamma\hat{\tau}_n(t)} h \cdot \hat{U}_n(t) d\hat{\tau}_n(t)\} \\
&\geq I\!\!E^* \liminf_{n\to\infty} \int_0^\infty e^{-\gamma\hat{\tau}_n(t)} \ell(\hat{W}_n(t)) d\hat{\tau}_n(t) + \gamma I\!\!E^* \liminf_{n\to\infty} \int_0^\infty e^{-\gamma\hat{\tau}_n(t)} h \cdot \hat{U}_n(t) d\hat{\tau}_n(t) \\
&\geq I\!\!E^* \int_0^\infty e^{-\gamma\hat{\tau}(t)} \ell(\hat{W}(t)) d\hat{\tau}(t) + \gamma I\!\!E^* \int_0^\infty e^{-\gamma\hat{\tau}(t)} h \cdot \hat{U}(t) d\hat{\tau}(t).
\end{aligned}$$

Applying Lemma 3.4.4 to the last line above and recalling $W^*(t) = \hat{W}(\tau(t))$ and $U^*(t) = \hat{U}(\tau(t))$ yields

$$V(w) \geq I\!\!E^* \int_0^\infty e^{-\gamma t} \ell(W^*(t)) dt + \gamma I\!\!E^* \int_0^\infty e^{-\gamma t} h \cdot U^*(t) dt.$$

The quantity on the right side above defines the cost function $J(w, U^*)$ for the admissible (by Corollary 3.4.7) pair (W^*, U^*). Thus we have $V(w) = J(w, U^*)$ and hence U^* is an optimal control. ■

3.5 Brownian Control Problems

In this section, as an application of Theorem 3.2.3, we prove existence of an optimal control for Brownian Control problems. Such control problems (cf. [20]) arise from formal diffusion approximations of multiclass queuing networks with scheduling control. Here we do not describe the underlying queuing problem but merely refer the reader to [7] where details on connections between a broad family of queuing network control problems and Brownian control problems can be found. Our presentation of BCPs is adapted from [21].

Let $\tilde{\Phi} \doteq (\Omega, \mathcal{F}, \{\mathcal{F}_t\}, I\!\!P, \tilde{B})$ be a system, where $\tilde{B}$ is an m-dimensional Brownian motion with drift $\tilde{b}$ and non-degenerate covariance matrix $\tilde{\Sigma}$. The problem data of the BCP consists of an $m \times n$ matrix R, a $p \times n$ matrix K (referred to, respectively as the input-output matrix and the capacity consumption matrix) and an initial condition $q \in I\!\!R^m_+$. The matrix K is assumed to have rank p ($p \le n$).

Definition 3.5.1 *(Admissible control for the BCP) An $\{\mathcal{F}_t\}$-adapted, n-dimensional RCLL process Y is an admissible control for the BCP associated with the system $\tilde{\Phi}$ and initial data $q \in I\!\!R^m_+$ if the following two conditions hold $I\!\!P$-a.s.:*

$$\begin{aligned} U(t) \doteq KY(t) \ &\textit{is nondecreasing with}\ U(0) \ge 0, \\ Q(t) \doteq q + \tilde{B}(t) + RY(t) \ge 0,\ &t \ge 0. \end{aligned} \tag{3.47}$$

Denote by $\tilde{\mathcal{A}}(q, \tilde{\Phi})$ the class of all admissible controls for the BCP associated with $\tilde{\Phi}$ and q. The goal of the BCP is to minimize the cost function

$$\tilde{J}(q, Y) \doteq I\!\!E \int_0^\infty e^{-\gamma t} \tilde{\ell}(Q(t))dt + I\!\!E \int_{[0,\infty)} e^{-\gamma t} h \cdot dU(t),$$

where $\gamma \in (0, \infty)$, $h \in I\!\!R^p_+$ and $\tilde{\ell} : I\!\!R^m_+ \to [0, \infty)$ is continuous. The value function for the BCP is $\tilde{V}(q) = \inf_{\tilde{\Phi}} \inf_{Y \in \tilde{\mathcal{A}}(q, \tilde{\Phi})} \tilde{J}(q, Y)$.

Under a continuous selection condition (see [21] or equation (3.3) of [1]), the BCP introduced above can be reduced to an equivalent control problem of the singular type (with state constraints). This reduction, referred to as the "Equivalent Workload Formulation" (EWF), is the main result of [21]. Under further conditions, this singular control problem with state constraints is of the form studied in the current chapter. Such sufficient conditions were presented in Section 3 of [1]; however we list them here for the reader's convenience. Let $\tilde{\ell}$ be linear, non-negative on $I\!R^m_+$, and assume that it vanishes only at zero. Define $\mathcal{B} \doteq \{x \in I\!R^n : Kx = 0\}$. Let $\mathcal{R} \doteq R\mathcal{B} \subset R^m$ and denote the dimension of $\mathcal{R}$ by r. The dimension of $\mathcal{M} \doteq \mathcal{R}^{\perp}$ is then $k \doteq m - r$. Let M be any $k \times m$ matrix whose rows span $\mathcal{M}$. By Proposition 2 of [21] there exists a $k \times p$ matrix G which satisfies $MR = GK$. The choice of G, in general, is not unique. We assume that the matrices M and G are of full rank and have nonnegative entries. We further assume that each column of G has at least one strictly positive entry. These assumptions are satisfied for a broad family of controlled queuing networks (see Section 3 of [1] and [7] for examples). Under these assumptions, Theorem 3.2.3 leads to the following result.

Theorem 3.5.2 *For every $q \in I\!R^m_+$, there exists a system $\tilde{\Phi}$ and $Y \in \tilde{\mathcal{A}}(q, \tilde{\Phi})$ such that $\tilde{J}(q, Y) = \tilde{V}(q)$.*

Remarks on the Proof. The proof is an immediate consequence of Theorem 3.2.3 and Proposition 3 of [21]. The latter proposition shows that for any admissible control for the EWF there exists a control for the BCP (and vice-versa) such that the costs coincide. Since an EWF under the above assumptions is a control problem of the form formulated in Section 3.2, existence of an optimal control for the EWF follows from Theorem 3.2.3. Using the equivalence result in Proposition 3 of [21], one then obtains an optimal control for the BCP.

Chapter 4

Numerical Scheme for a Brownian Control Problem through Optimal Stopping

4.1 Introduction

Stochastic networks have been an area of active research in recent years with applications in a variety of disciplines, including manufacturing, communications, and computing. In general, a stochastic network consists of a system in which customers (or jobs) arrive at random times and are placed in a series of buffers, where they await service by one or more servers. Servers may accept customers from multiple buffers and service completion times may depend on customer characteristics. Upon completion of service, a customer may exit the system or be redirected to another buffer, where he will await service by a different server.

A fundamental yet challenging problem with critical practical implications concerns control of multiclass stochastic networks. The main objective is to design a control policy for the network in order to optimize some criteria. Control can take a variety of forms, including control of rates of arrival and/or service, and routing or scheduling of jobs.

Optimization criteria can incorporate inventory holding costs, server idleness times, and other appropriate performance measures.

Excepting simple examples such control problems are quite intractable by classical queueing techniques, and thus suitable approximation methods are needed. One common approach for systems that are critically loaded (see [20]) uses *heavy traffic approximation* to replace, formally, a control problem for a stochastic network by one for a diffusion process. This leads to a challenging class of singular diffusion control problems with state constraints. The basic approach is to solve the diffusion control problem (say approximately by a suitable numerical procedure), and then use insights derived from the solution to obtain a near optimal network control policy.

In this chapter we study a two-dimensional controlled queueing network model, often referred to as the "criss-cross" network. The network is described in Section 4.2 and has been studied by several authors (see [24], [9], [34], [38], and [8]). The basic problem concerns the optimal scheduling of jobs in a two server, two customer queueing system. Even though the network is simple to describe, analysis of the control problem is quite subtle and depends heavily on the values of the model parameters. In the current work, we focus on the parameter regime IIb of [38], (Condition 4.6.1 of the current work), a regime which to date has not yielded an analytical solution.

Our goal in the present study is to develop a numerical scheme for the singular diffusion control problem corresponding to this network. One can adapt the numerical solution of this diffusion control problem in a straightforward way to develop a scheduling policy for the underlying network; however, this is not pursued in the current work.

We begin by presenting the Brownian control problem (BCP) that is obtained by taking a formal heavy traffic limit of the controlled queueing network. We next present an equivalent lower dimensional control problem that is of the singular type with state constraints. This reduced control problem, commonly referred to as the equivalent workload formulation (EWF), is the main subject of the study in this chapter.

Since no closed form analytical solutions are available, a standard approach, as carried out in Chapter 2, is through approximations by finite state controlled Markov chains. However, such schemes are numerically quite intensive and thus it is of interest to develop simpler alternative schemes that take into account specific problem features. In this chapter we exploit the linearity of the dynamics and convexity of the cost to reduce the singular control problem to an optimal stopping problem. This is the main result of this chapter, which is presented in Theorem 4.7.1. Although numerical schemes for singular control problems are notoriously hard, there are relatively simple schemes available for optimal stopping problems. In the last section of this chapter, we use such schemes to obtain near optimal control policies for the EWF corresponding to the criss-cross network. We show by examples that a numerical scheme based on a Markov chain approximation of the singular control problem can suffer from serious divergence problems, whereas the scheme based on optimal stopping is quite robust to the choice of initialization. We also illustrate how incorporating results and insights from a computational algorithm for the optimal stopping problem can substantially improve the numerical performance of the computational scheme based on the singular control problem.

4.2 The Criss-Cross Network

The network of interest is illustrated in Figure 4.1. We refer the reader to [8] for a precise mathematical formulation. Two types of customers (or jobs, packets, etc.), class 1 and class 2, arrive according to independent renewal processes to be served by server 1. Class 1 customers exit the system upon completion of service by server 1. Class 2 customers, after being served by server 1, are redesignated as class 3 and proceed to be served by server 2. They exit the system upon completion of service by server 2. Service times for each activity are i.i.d. with finite mean and variance. A controller needs to decide, at any given moment, should server 1 process class 1, class 2, or idle. Similarly, a corresponding

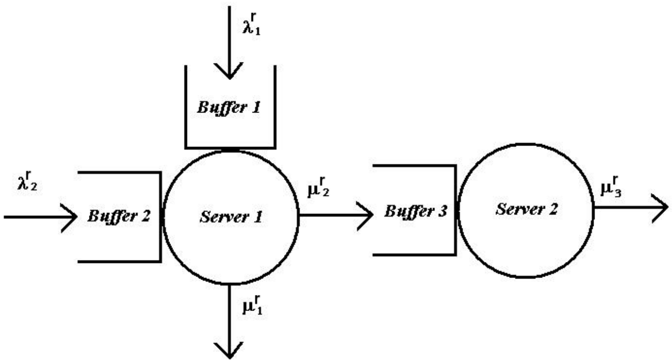

Figure 4.1: The criss-cross network

decision regarding the second server needs to be made. The goal of the controller is to implement a service scheduling policy that will minimize a discounted linear holding cost over an infinite time horizon.

4.3 The Brownian Control Problem

We now describe the diffusion control problem that arises on taking a formal heavy traffic limit of the above controlled network. Suppose that we have a sequence of such criss-cross networks, indexed by $r = 0, 1, 2, \ldots$. For the r^{th} network, assume that class i $(i = 1, 2)$ customers arrive at rate λ_i^r, and are served by server 1 at rate μ_i^r. Assume also that class 3 customers are served by server 2 at rate μ_3^r. Suppose there exist $\lambda_j \in (0, \infty)$, $j = 1, 2$, and μ_k, $k = 1, 2, 3$, such that

$$\lim_{r \to \infty} \lambda_j^r = \lambda_j, \quad \lim_{r \to \infty} \mu_k^r = \mu_k.$$

The following heavy traffic assumptions stipulate that the sequence of networks approach a critically loaded system as $r \to \infty$.

Condition 4.3.1 *Suppose there exist b_k, $k = 1, 2, 3$ such that $b_1 + b_2 = 1$ and*

$$\lim_{r\to\infty} r\frac{\lambda_1^r}{\mu_1^r} = b_1, \quad \lim_{r\to\infty} r\frac{\lambda_2^r}{\mu_2^r} = b_2, \quad \lim_{r\to\infty} r(1 - \frac{\lambda_2^r}{\mu_3^r}) = b_3. \tag{4.1}$$

Note that the above assumption implies in particular,

$$\frac{\lambda_1}{\mu_1} + \frac{\lambda_2}{\mu_2} = 1, \quad \frac{\lambda_2}{\mu_3} = 1. \tag{4.2}$$

Given the control problem for the r^{th} network, one arrives at the Brownian control problem described below by taking formal limits as $r \to \infty$; see [8] for details.

Let $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}, I\!P)$ be a filtered probability space equipped with a a three-dimensional $\{\mathcal{F}_t\}$-Brownian motion $\tilde{B}$ that starts from the origin with drift $\theta \doteq (\mu_1 b_1, \mu_2 b_2, \mu_3 b_3)'$ and covariance matrix

$$\Lambda \doteq \begin{pmatrix} 2\lambda_1 & 0 & 0 \\ 0 & 2\lambda_2 & -\lambda_2 \\ 0 & -\lambda_2 & 2\lambda_2 \end{pmatrix}.$$

We will refer to $\Phi \doteq (\Omega, \mathcal{F}, \{\mathcal{F}_t\}, I\!P, \tilde{B})$ as a *system.*

Definition 4.3.2 *An admissible control for the Brownian control problem (BCP) for initial condition $q \in I\!R^3_+$ and the system Φ is a three-dimensional $\{\mathcal{F}_t\}$-adapted, RCLL process $Y \equiv (Y_1, Y_2, Y_3)$ for which:*

$$\begin{aligned} &Q_k(t) \doteq q_k + \tilde{B}_k(t) + \mu_k Y_k(t) - \mu_2 Y_2(t)\, 1_{\{k=3\}} \geq 0 \text{ for all } t \geq 0, k = 1, 2, 3, \\ &Y_1 + Y_2 \text{ is nondecreasing with } Y_1(0) + Y_2(0) \geq 0, \\ &Y_3 \text{ is nondecreasing with } Y_3(0) \geq 0. \end{aligned} \tag{4.3}$$

We denote by $\tilde{\mathcal{A}}(q, \Phi)$ the set of all admissible controls for q and Φ. We refer to Q as the controlled process corresponding to Y.

Let $\gamma \in (0, \infty)$ and $h_k \in (0, \infty), k = 1, 2, 3$. The cost for exercising the control $Y \in \tilde{\mathcal{A}}(q, \Phi)$ is given by

$$\tilde{J}(q, Y) \doteq I\!\!E \int_0^\infty e^{-\gamma t} h \cdot Q(t) dt. \tag{4.4}$$

The value function for the BCP with initial condition q is

$$\tilde{V}(q) \doteq \inf_{\Phi} \inf_{Y \in \tilde{\mathcal{A}}(q,\Phi)} \tilde{J}(q, Y), \tag{4.5}$$

where the outside supremum is taken over all probability systems Φ.

Although we have not provided a precise description of the network control problem, we note that all the processes in Definition 4.3.2 have natural interpretations in terms of the underlying queueing system. In particular, the process Q_k is a surrogate for a suitably scaled process of queue length at buffer k. The requirement that Q be nonnegative stems from physical queueing considerations. Similarly, $Y_1(t) + Y_2(t)$ can be interpreted as the amount of idleness server 1 incurs over the period $[0, t]$, in an appropriate limiting sense. The process Y_3 corresponds to the idleness process for server 2. The weights h_k represent the holding costs per customer per unit time for buffer k. Finally, the discount factor γ weights costs incurred now relative to those incurred in the future.

4.4 Equivalent Workload Formulation of the BCP

We will consider the following regime, which corresponds to Case II of [38].

Condition 4.4.1 *Case II: $h_1\mu_1 - h_2\mu_2 + h_3\mu_2 > 0$.*

The control problem in the previous section does not fall under the classical framework of singular control since an admissible control is allowed to have paths with infinite total variation. However, using techniques from Harrison and van-Mieghem [21] it is possible to reformulate the above to a standard singular control problem with state constraints which, furthermore, lies in a lower dimensional space. This formulation is referred to in the literature as the Equivalent Workload Formulation (EWF). We present this workload problem below and discuss its connections with the BCP in the next section.

Let

$$M \doteq \begin{pmatrix} \frac{1}{\mu_1} & \frac{1}{\mu_2} & 0 \\ 0 & \frac{1}{\mu_3} & \frac{1}{\mu_3} \end{pmatrix}.$$

Define $B \doteq M\tilde{B}$, where $\tilde{B}$ is as in Section 4.3.

Definition 4.4.2 *An admissible control for the workload control problem for initial condition $w = (w_1, w_2) \in I\!R_+^2$ and the system Φ is a two-dimensional $\{\mathcal{F}_t\}$-adapted, RCLL process $I = (I_1, I_2)$ for which:*

$$W(t) \doteq w + B(t) + I(t) \geq 0, t \geq 0; \tag{4.6}$$

$$I \textit{ is nondecreasing with } I(0) \geq 0. \tag{4.7}$$

We refer to W as the controlled process corresponding to I. Let $\mathcal{A}(w, \Phi)$ denote the set of all admissible controls for w and Φ. The cost function corresponding to $I \in \mathcal{A}(w, \Phi)$ is given by

$$J(w, I) \doteq I\!E \int_0^\infty e^{-\gamma t} \hat{h}(W(t))dt, \tag{4.8}$$

where the continuous function $\hat{h} : I\!R_+^2 \to [0, \infty)$ is defined as

$$\hat{h}(z_1, z_2) = \begin{cases} (h_2\mu_2 - h_3\mu_2)z_1 + h_3\mu_3 z_2, & \mu_3 z_2 \geq \mu_2 z_1, \\ h_1\mu_1 z_1 + \frac{\mu_3}{\mu_2}(h_2\mu_2 - h_1\mu_1)z_2, & \mu_3 z_2 < \mu_2 z_1. \end{cases} \tag{4.9}$$

The value function of the workload control problem for initial condition w is

$$V(w) \doteq \inf_{\Phi} \inf_{I \in \mathcal{A}(w,\Phi)} J(w, I), \tag{4.10}$$

where the outside supremum is taken over all probability systems Φ. The goal in this reduced control problem is to find an optimal control, i.e. $I^* = (I_1^*, I_2^*) \in \mathcal{A}(w, \Phi)$ such that $V(w) = J(w, I^*)$. In the sections that follow, we develop a convergent numerical scheme which yields an approximation to the value function and a control whose cost attains this approximate value.

Before analyzing the workload control problem defined above, we discuss below connections between the workload control problem and the BCP.

4.5 Connections between EWF and BCP

For $q \in I\!R_+^3$ and a system Φ, let $Y \in \tilde{\mathcal{A}}(q, \Phi)$ be an admissible control for the BCP with corresponding controlled process Q. Define $I = (I_1, I_2)$ by $I_1 \doteq Y_1 + Y_2$ and $I_2 \doteq Y_3$, and let $w \doteq Mq$, $W(t) \doteq MQ(t), t \geq 0$. It is easy to check that $I \in \mathcal{A}(w, \Phi)$; that is, I is an admissible control for the workload control problem, with corresponding controlled process W. Furthermore, $J(w, I) \leq \tilde{J}(q, Y)$, and if $\tilde{J}(q, Y) = \tilde{V}(q)$ then we have $J(w, I) = V(w)$.

Conversely, given a control $I \in \mathcal{A}(w, \Phi)$ for some system Φ and $w \in I\!R_+^2$ one can obtain a control $Y \in \tilde{\mathcal{A}}(q, \Phi)$ for any $q \in I\!R_+^3$ satisfying $Mq = w$, such that $\tilde{J}(q, Y) = J(w, I)$, in the following manner (see [8] for details). Let W be the controlled processes corresponding to I. Define Y via:

$$\begin{aligned}
Y_1(t) &\doteq \begin{cases} -\frac{\tilde{B}_1(t)}{\mu_1}, & \mu_3 W_2(t) \geq \mu_2 W_1(t), \\ -\frac{\tilde{B}_3(t)}{\mu_2} + I_1(t) - \frac{\mu_3}{\mu_2} I_2(t), & \mu_3 W_2(t) < \mu_2 W_1(t); \end{cases} \\
Y_2(t) &\doteq \begin{cases} \frac{\tilde{B}_1(t)}{\mu_1} + I_1(t), & \mu_3 W_2(t) \geq \mu_2 W_1(t), \\ \frac{\tilde{B}_3(t)}{\mu_2} + \frac{\mu_3}{\mu_2} I_2(t), & \mu_3 W_2(t) < \mu_2 W_1(t); \end{cases} \\
Y_3(t) &\doteq I_1(t).
\end{aligned}$$

It can be easily checked that $Y \in \tilde{\mathcal{A}}(q, \Phi)$ and the costs agree, i.e. $\tilde{J}(q, Y) = J(w, I)$. This one-to-one correspondence shows that once an optimal (or near optimal) control policy for the WCP is determined, one can obtain an optimal (respectively, near optimal) control for the BCP.

4.6 A Further Reduction in Parameter Regime IIb

We now focus attention on the EWF control problem of Section 4.4. To simplify notation, for the remainder of this chapter we denote a system by $\Phi = (\Omega, \mathcal{F}, \{\mathcal{F}_t\}, I\!P, B)$, where B is a two-dimensional $\{\mathcal{F}_t\}$-Brownian motion with drift $M\theta$ and covariance matrix $M\Lambda M'$. In particular, for $I \in \mathcal{A}(w, \Phi)$ we have $W(t) = w + B(t) + I(t)$.

For the remainder of the chapter we assume that the network parameters satisfy the following condition.

Condition 4.6.1 *(Parameter regime IIb of [38]).*

$$h_1\mu_1 - h_2\mu_2 + h_3\mu_2 > 0, \quad h_2\mu_2 - h_3\mu_2 < 0, \quad h_2\mu_2 - h_1\mu_1 \geq 0. \tag{4.11}$$

The above parameter regime is one of the key issues left unresolved in [38]. In this regime an explicit solution of the BCP is not available and the determination of an asymptotically optimal control policy for the underlying network is an open problem. In the rest of this chapter we will develop numerical approaches for approximating the solution of the BCP by first establishing an equivalence between this control problem and a problem of optimal stopping.

It is easy to check that under Condition 4.6.1, $\hat{h}(z_1, z_2)$ is nondecreasing in z_2 for each fixed z_1; i.e. if $z_2 \leq \tilde{z}_2$ then $\hat{h}(z_1, z_2) \leq \hat{h}(z_1, \tilde{z}_2)$ for all z_1. We now use this monotonicity property to reduce the workload problem of Section 4.4 to a problem with a one-dimensional control.

Fix $w \in I\!R^2_+$ and a system Φ and define for $t \geq 0$,

$$I_2^*(t) \doteq -\min\{0, \inf_{0 \leq s \leq t}[w_2 + B_2(s)]\}. \tag{4.12}$$

Note that I_2^* is $\{\mathcal{F}_t\}$-adapted, RCLL (in fact, continuous), and nondecreasing with $I_2^*(0) = 0$. Furthermore $W_2^*(t) \doteq w_2 + B_2(t) + I_2^*(t) \geq 0$ for all $t \geq 0$. Let $I = (I_1, I_2) \in \mathcal{A}(w, \Phi)$ with corresponding controlled process $W = (W_1, W_2)$. Then defining $\tilde{I} \doteq (I_1, I_2^*)$ we have $\tilde{I} \in \mathcal{A}(w, \Phi)$ and $I_2(t) \geq I_2^*(t)$ for all $t \geq 0$, which implies $W_2(t) \geq W_2^*(t), t \geq 0$. Define $J_1(w, I_1) \doteq J(w, \tilde{I})$; that is,

$$J_1(w, I_1) = I\!E \int_0^\infty e^{-\gamma t} \hat{h}(W_1(t), W_2^*(t)) dt. \tag{4.13}$$

Since $\hat{h}(z_1, z_2)$ is nondecreasing in z_2 we have $J(w, I) \geq J_1(w, I_1)$ and therefore

$$V(w) = \inf_{\Phi} \inf_{I_1 \in \mathcal{A}_1(w_1, \Phi)} J_1(w, I_1), \tag{4.14}$$

where $\mathcal{A}_1(w_1, \Phi)$ is the set of all $\{\mathcal{F}_t\}$-adapted RCLL processes I_1 that are nondecreasing with $I_1(0) \geq 0$ and $W_1(t) \geq 0$ for all $t \geq 0$.

Thus under Condition 4.6.1 the main difficulty with the workload control problem lies in characterizing (or approximating) the first component of an optimal control. In preparation for a numerical scheme for obtaining a near optimal control policy, we present next an optimal stopping problem and our main equivalence result (Theorem 4.7.1) that connects this problem with the EWF.

4.7 Equivalence between Singular Control and Optimal Stopping Problems

Since $\hat{h}$ is convex its left-derivative in the x-direction exists at every point $z = (z_1, z_2) \in I\!R^2_+$ with $z_1 > 0$ and is given by

$$\hat{h}^-_x(z) \doteq \lim_{\delta \downarrow 0} \frac{\hat{h}(z_1, z_2) - \hat{h}(z_1 - \delta, z_2)}{\delta}.$$

It is easily checked that for $(z_1, z_2) \in (0, \infty) \times [0, \infty)$,

$$\hat{h}^-_x(z_1, z_2) = \begin{cases} h_2\mu_2 - h_3\mu_2, & \mu_3 z_2 \geq \mu_2 z_1, \\ h_1\mu_1, & \mu_3 z_2 < \mu_2 z_1. \end{cases} \tag{4.15}$$

Let $w = (w_1, w_2) \in I\!R^2_+$, $\tilde{\Phi} = (\tilde{\Omega}, \tilde{\mathcal{F}}, \{\tilde{\mathcal{F}}_t\}, \tilde{I\!P}, \tilde{B})$ be a system, and let $\tilde{I} \in \mathcal{A}(w, \tilde{\Phi})$. Denote the filtration $\sigma\{\tilde{B}(s), \tilde{I}(s), 0 \leq s \leq t\}$ by $\hat{\mathcal{F}}_t$. We refer to $\hat{\Phi} \doteq (\tilde{\Omega}, \tilde{\mathcal{F}}, \{\hat{\mathcal{F}}_t\}, \tilde{I\!P}, \tilde{B}, \tilde{I})$ as a *controlled system*. Let $\mathcal{S}(\hat{\Phi})$ be the set of all $\{\hat{\mathcal{F}}_t\}$-stopping times. Finally, let $\tilde{S} \doteq \inf\{t \geq 0 : \tilde{W}^0_1(t) \leq 0\}$, where $\tilde{W}^0_1(t) \doteq w_1 + \tilde{B}_1(t)$. The reward associated with stopping at time $\tilde{\sigma} \in \mathcal{S}(\hat{\Phi})$ is given by

$$\hat{J}(w, \tilde{\sigma}) \doteq \tilde{I\!E} \int_0^{\tilde{\sigma} \wedge \tilde{S}} e^{-\gamma t} \hat{h}^-_x(\tilde{W}^0_1(t), \tilde{W}^*_2(t)) dt, \tag{4.16}$$

where $\tilde{W}^*_2(t) \doteq w_2 + \tilde{B}_2(t) - \min\{0, \inf_{0 \leq s \leq t}[w_2 + \tilde{B}_2(s)]\}$. Consider the optimal stopping problem of choosing a stopping time $\tilde{\sigma}$ to maximize the reward in (4.16). Then the value function for the optimal stopping problem for initial condition w is defined as

$$u(w) \doteq \sup_{\hat{\Phi}} \sup_{\tilde{\sigma} \in \mathcal{S}(\hat{\Phi})} \hat{J}(w, \tilde{\sigma}), \tag{4.17}$$

where the outside supremum is taken over all controlled systems $\hat{\Phi}$.

The following theorem is our main result which establishes an equivalence between the workload (singular) control problem and the above optimal stopping problem. The proof is presented in Section 4.8. For a function f defined on a domain in $\mathbb{R}^2$, let f_x denote the directional derivative of f in the x-direction.

Theorem 4.7.1 *Let $w = (w_1, w_2) \in \mathbb{R}^2_+$ with $w_1 > 0$. Then $V_x(w)$ exists and equals $u(w)$.*

We now give a heuristic explanation of the above equivalence result. The optimal control in a singular control problem often takes the form of a "no action" region. No control is applied within this region, and when the state process hits the boundary of the no action region, enough control is applied to instantaneously push the process back inside the no action region. If the initial state falls outside the no action region, control is applied immediately to push the state to the boundary of the no action region.

For $w = (w_1, w_2) \in \mathbb{R}^2_+$ suppose $I_1 \in \mathcal{A}_1(w_1, \Phi)$ takes this no-action region form. Let I_2^* be as in (4.12) and let $W = (W_1, W_2^*)$ be the state process corresponding to w and $I = (I_1, I_2^*)$. Suppose there is a function g that defines the boundary between the action and no-action regions for the control I_1. That is, for $z = (z_1, z_2)$ no action is taken if $z_1 > g(z_2)$, otherwise control is applied. Then the (cumulative) amount of control applied by time t, $I_1(t)$, can be written as $I_1(t) = \sup_{0 \le s \le t} \xi(W(s))$ where

$$\xi(z_1, z_2) = \begin{cases} 0, & z_1 > g(z_2), \\ g(z_2) - z_1, & z_1 \le g(z_2). \end{cases}$$

Taking the left derivative of ξ in the x-direction we obtain

$$\xi_x^-(z_1, z_2) = \begin{cases} 0, & z_1 > g(z_2), \\ -1, & z_1 \le g(z_2). \end{cases}$$

Intuitively ξ_x^- specifies whether a control is applied ($\xi_x^- = -1$) or not ($\xi_x^- = 0$) and one can interpret $\{z : \xi_x^-(z) = 0\}$ as the no-action region. We can associate with I a stopping time $\sigma = \inf\{t \geq 0 : \xi_x^-(W(t)) = -1\}$; then σ is the first time at which control is applied. For the stopping rule σ, $\{z : \xi_x^-(z) = 0\}$ represents the continuation region while $\{z : \xi_x^-(z) = -1\}$ is the stopping region. Within the continuation region, the state process is uncontrolled; in particular, the evolution of the first coordinate $W_1(t)$ is given by $w_1 + B_1(t)$. This linearity of the dynamics suggests that the rate of change of the value function (in the x-direction) should be given in terms of the x-direction derivative of the cost function, and thus one expects the relation $u(w) = V_x(w)$.

Finally, observe that we infimize over admissible controls in defining $V(w)$, but we take the supremum over $\{\hat{\mathcal{F}}_t\}$-stopping times in defining $u(w)$. This can be explained by the reversal of sign in the action region from positive (for $\xi(z)$) to negative (for $\xi_x^-(z)$).

4.8 Proof of Theorem 4.7.1

This section is devoted to the proof of Theorem 4.7.1. The proof proceeds in three main steps, established in Lemmas 4.8.1, 4.8.3, and 4.8.5.

We begin with the following lemma which establishes convexity of the value function V.

Lemma 4.8.1 *$V : \mathbb{R}_+^2 \to [0, \infty)$ is convex.*

Proof. Fix $w^1, w^2 \in \mathbb{R}_+^2$ and let $\alpha \in (0,1)$. Set $\hat{w} = \alpha w^1 + (1-\alpha)w^2$. It suffices to show that $V(\hat{w}) \leq \alpha V(w^1) + (1-\alpha)V(w^2)$. Let $\epsilon > 0$ be arbitrary, and let Φ_1, Φ_2 be systems and $I^i \in \mathcal{A}(w^i, \Phi_i)$, $i = 1, 2$, be such that

$$J(w^i, I^i) \leq V(w^i) + \epsilon/2,\ i = 1, 2. \tag{4.18}$$

Denote by B^i the Brownian motion and by W^i the controlled process (corresponding to I^i) on Φ_i. Then $W^i = w^i + B^i + I^i$, $i = 1, 2$. Since I^1, I^2 are in general on different probability systems, one can not combine them directly to construct a control for the initial condition $\hat{w}$. The following argument allows us to assume without loss of generality that I^1 and I^2 are given on the same system.

Let $\mathcal{E}_0 = \mathcal{D}([0,\infty) : I\!R^2) \times \mathcal{C}([0,\infty) : I\!R^2)$ and $\{\mathcal{F}^0_t\}_{t\geq 0}$ be the canonical filtration on this system. Also let $\mathcal{F}^0 = \sigma\{\mathcal{F}^0_t : t \geq 0\}$. Denote the measure induced by I^i, B^i on $(\mathcal{E}_0, \mathcal{F}^0)$ by $I\!P^i_0$. Writing a canonical element in $\mathcal{E}_0$ as (u, b), we can decompose $I\!P^i_0$ as

$$I\!P^i_0(du\; db) = Q^i(b, du)Q(db),$$

where Q is the measure induced by a Brownian motion with drift $M\theta$ and covariance matrix $M\Lambda M'$.

Next let $\mathcal{E} = \mathcal{D}([0,\infty) : I\!R^2) \times \mathcal{D}([0,\infty) : I\!R^2) \times \mathcal{C}([0,\infty) : I\!R^2)$, $\{\mathcal{F}_t\}$ be the canonical filtration on this space, and $\mathcal{F} = \sigma\{\mathcal{F}_t : t \geq 0\}$. Denote a typical element in $\mathcal{E}$ as (u_1, u_2, b) and introduce the probability measure $I\!P$ on $(\mathcal{E}, \mathcal{F})$ by

$$I\!P(du_1\; du_2\; db) = Q^1(b, du_1)Q^2(b, du_2)Q(db).$$

Denoting the canonical coordinate processes on $\mathcal{E}$ by $\hat{I}_1, \hat{I}_2, \hat{B}$, we see that $\hat{\Phi} \doteq (\mathcal{E}, \mathcal{F}, \{\mathcal{F}_t\}, I\!P, \hat{B})$ is a system (cf. Lemma IV.1.2 of [27]), and $\hat{I}^i \in \mathcal{A}(w^i, \hat{\Phi})$ with corresponding controlled process $\hat{W}^i = w^i + \hat{B} + \hat{I}^i$. Furthermore, $J(w^i, \hat{I}^i) = J(w^i, I^i)$, $i = 1, 2$.

Now set $\hat{I} \doteq \alpha\hat{I}^1 + (1-\alpha)\hat{I}^2$. It is easy to check that $\hat{I} \in \mathcal{A}(\hat{w}, \hat{\Phi})$. From the convexity of $\hat{h}$ and recalling (4.18) we have,

$$\begin{aligned} V(\hat{w}) &\leq J(\hat{w}, \hat{I}) \\ &= I\!E\int_0^\infty e^{-\gamma t}\hat{h}(\alpha\hat{W}^1(t) + (1-\alpha)\hat{W}^2(t))dt \\ &\leq \alpha I\!E\int_0^\infty e^{-\gamma t}\hat{h}(\hat{W}^1(t))dt + (1-\alpha)I\!E\int_0^\infty e^{-\gamma t}\hat{h}(\hat{W}^2(t))dt \\ &\leq \alpha V(w^1) + (1-\alpha)V(w^2) + \epsilon. \end{aligned}$$

In the above display $I\!\!E$ denotes expectation with respect to $I\!\!P$. Since $\epsilon > 0$ is arbitrary, the result follows. ■

Let $w_1 > 0$, $w_2 \geq 0$, and $\delta > 0$ be such that $w_1 - \delta > 0$. Define

$$\Delta_- V(w) \doteq \lim_{\delta \downarrow 0} \frac{V(w_1 - \delta, w_2) - V(w_1, w_2)}{-\delta}. \tag{4.19}$$

Also, for $w_1 > 0$, $w_2 \geq 0$, and $\delta > 0$, define

$$\Delta^+ V(w) \doteq \lim_{\delta \downarrow 0} \frac{V(w_1 + \delta, w_2) - V(w_1, w_2)}{\delta}. \tag{4.20}$$

Existence of the above limits is a consequence of convexity of V (Theorem 24.1 [44]). The following lemma is also an immediate consequence of convexity of V. For a proof, see Theorem 24.1 of [44].

Lemma 4.8.2 *Let $w = (w_1, w_2) \in I\!\!R^2_+$ with $w_1 > 0$. Then $\Delta_- V(w) \leq \Delta^+ V(w)$.*

In the following lemma we establish the inequality $\Delta_- V(w) \geq u(w)$. To prove this result we first select a near-optimal control I_1 for the singular control problem which is admissible for initial condition w_1. From this we construct a control I_1^δ which is admissible for initial condition $w_1 - \delta$, by "bumping" I_1 by an amount δ at the first time the uncontrolled process $w_1 + B_1(t)$ falls below δ. The desired inequality follows by comparing the cost functions associated with I_1 and I_1^δ, using the near-optimality of I_1, and taking appropriate limits.

Lemma 4.8.3 *Let $w = (w_1, w_2) \in I\!\!R^2_+$ with $w_1 > 0$. Then $\Delta_- V(w) \geq u(w)$.*

Proof. Let $\epsilon > 0$ be arbitrary and let $\Phi = (\Omega, \mathcal{F}, \{\mathcal{F}_t\}, I\!\!P, B)$ and $I_1 \in \mathcal{A}_1(w_1, \Phi)$, with corresponding state process W_1, be such that

$$V(w) \leq J_1(w, I_1) \leq V(w) + \epsilon. \tag{4.21}$$

Next let $\tilde{\sigma}$ be a $\{\hat{\mathcal{F}}_t\}$-stopping time for some controlled system $\hat{\Phi} = (\tilde{\Omega}, \tilde{\mathcal{F}}, \{\tilde{\mathcal{F}}_t\}, \tilde{I\!P}, \tilde{B}, \tilde{I})$. Although $\tilde{\sigma}$ and I_1 are possibly defined on different probability spaces we can implement a construction analogous to that in Lemma 4.8.1 to define all processes on a common probability space. More precisely, let $\Omega^\dagger = \mathcal{D}([0,\infty) : I\!R^2) \times \mathcal{D}([0,\infty) : I\!R^2) \times \mathcal{C}([0,\infty) : I\!R^2)$, $\{\mathcal{F}_t^\dagger\}$ the canonical filtration on this space, and $\mathcal{F}^\dagger = \sigma\{\mathcal{F}_t^\dagger : t \geq 0\}$. Let $\tilde{I}^\dagger$, $I^\dagger$, $B^\dagger$, denote the canonical coordinate processes on this space and let $I\!P^\dagger$ be the probability measure on $(\Omega^\dagger, \mathcal{F}^\dagger)$ under which: $(\tilde{I}^\dagger, B^\dagger)$ has the same distribution as $(\tilde{I}, \tilde{B})$; $(I^\dagger, B^\dagger)$ has the same distribution as (I, B); and $B^\dagger$ is an $\{\mathcal{F}_t^\dagger\}$-Brownian motion. Then, letting $\Phi^\dagger = (\Omega^\dagger, \mathcal{F}^\dagger, \{\mathcal{F}_t^\dagger\}, I\!P^\dagger, B^\dagger)$, $I^\dagger \in \mathcal{A}(w, \Phi^\dagger)$ with corresponding controlled process $W^\dagger = w + B^\dagger + I^\dagger$. Let $W_2^{*,\dagger}(t) = -\min\{0, \inf_{0\leq s\leq t}[w_2 + B_2^\dagger(t)]\}$ and $J_1(w, I_1^\dagger) = I\!E^\dagger \int_0^\infty e^{-\gamma t}\hat{h}_x(W_1^\dagger(t), W_2^{*,\dagger}(t))dt$. Then clearly $J_1(w, I_1) = J_1(w, I_1^\dagger)$. Furthermore, there is an $\{\mathcal{F}_t^\dagger\}$-stopping time $\sigma^\dagger$ such that

$$(\sigma^\dagger, S^\dagger, B^\dagger) \stackrel{d}{=} (\tilde{\sigma}, \tilde{S}, \tilde{B}), \tag{4.22}$$

where $S^\dagger \doteq \inf\{t \geq 0 : W_1^{0,\dagger}(t) \leq 0\}$ and $W_1^{0,\dagger}(t) \doteq w_1 + B_1^\dagger(t)$. In particular, $\hat{J}(w, \tilde{\sigma}) = \hat{J}(w, \sigma^\dagger)$, where $\hat{J}(w, \sigma^\dagger) = I\!E^\dagger \int_0^{\sigma^\dagger \wedge S^\dagger} e^{-\gamma t}\hat{h}_x^-(\tilde{W}_1^{0,\dagger}(t), \tilde{W}_2^{*,\dagger}(t))dt$.

Let $w_1 > 0$, $w_2 \geq 0$, and $\delta > 0$ be such that $w_1 - \delta > 0$, and let $w_\delta = (w_1 - \delta, w_2)$. Define $S_\delta^\dagger \doteq \inf\{t \geq 0 : w_1 + B_1^\dagger(t) \leq \delta\}$. Since $B_1^\dagger$ has continuous paths a.s., $S_\delta^\dagger \uparrow S^\dagger$ a.s. $(I\!P^\dagger)$ as $\delta \downarrow 0$. Define for $t \geq 0$, $I_1^\delta(t) \doteq I_1^\dagger(t) + \delta\, 1_{\{t \geq \sigma^\dagger \wedge S_\delta^\dagger\}}$ and

$$\begin{aligned} W_1^\delta(t) &\doteq w_1 - \delta + B_1^\dagger(t) + I_1^\delta(t) \\ &= w_1 + B_1^\dagger(t) + I_1^\dagger(t) - \delta\, 1_{\{0 \leq t < \sigma^\dagger \wedge S_\delta^\dagger\}}. \end{aligned}$$

From definitions of $I_1^\dagger$ and $S_\delta^\dagger$ it follows that $I_1^\delta \in \mathcal{A}_1(w_1 - \delta, \Phi^\dagger)$ with corresponding controlled process W_1^δ. Thus we have

$$\begin{aligned}
V(w_\delta) &\leq J_1(w_\delta, I_1^\delta) \\
&= I\!E^\dagger \int_0^\infty e^{-\gamma t}\hat{h}(W_1^\delta(t), W_2^{*,\dagger}(t))dt \\
&= I\!E^\dagger \int_0^{\sigma^\dagger \wedge S_\delta^\dagger} e^{-\gamma t}\hat{h}(W_1^\dagger(t) - \delta, W_2^{*,\dagger}(t))dt \\
&+ I\!E^\dagger \int_{\sigma^\dagger \wedge S_\delta^\dagger}^\infty e^{-\gamma t}\hat{h}(W_1^\dagger(t), W_2^{*,\dagger}(t))dt \\
&= I\!E^\dagger \int_0^{\sigma^\dagger \wedge S_\delta^\dagger} e^{-\gamma t}[\hat{h}(W_1^\dagger(t) - \delta, W_2^{*,\dagger}(t)) - \hat{h}(W_1^\dagger(t), W_2^{*,\dagger}(t))]dt \\
&+ I\!E^\dagger \int_0^\infty e^{-\gamma t}\hat{h}(W_1^\dagger(t), W_2^{*,\dagger}(t))dt. \qquad (4.23)
\end{aligned}$$

Since $\hat{h}$ is convex we have (see Theorem 24.1 of [44]) that for $z_1 > 0, z_2 \geq 0$ and δ small enough

$$\hat{h}(z_1 - \delta, z_2) - \hat{h}(z_1, z_2) = -\delta \int_0^1 \hat{h}_x^-(z_1 - \theta\delta, z_2)d\theta. \qquad (4.24)$$

Then, recalling (4.21) it follows from the above equation and (4.23) that

$$V(w_\delta) \leq V(w) + \epsilon - \delta I\!E^\dagger \int_0^{\sigma^\dagger \wedge S_\delta^\dagger} e^{-\gamma t} \int_0^1 \hat{h}_x^-(W_1^\dagger(t) - \theta\delta, W_2^{*,\dagger}(t))d\theta dt. \qquad (4.25)$$

Let $\tilde{S}_\delta \doteq \inf\{w_1 + \tilde{B}_1(t) \leq \delta\}$ and $\tilde{W}_2^*(t) = -\min\{0, \inf_{0 \leq s \leq t}[w_2 + \tilde{B}_2(t)]\}$. By the above display and (4.22) we have

$$\begin{aligned}
\frac{V(w_\delta) - V(w)}{-\delta} &\geq \frac{-\epsilon}{\delta} + I\!E^\dagger \int_0^{\sigma^\dagger \wedge S_\delta^\dagger} e^{-\gamma t} \int_0^1 \hat{h}_x^-(w_1 - \theta\delta + B_1^\dagger(t) + I_1^\dagger(t), W_2^{*,\dagger}(t))d\theta dt \\
&= \frac{-\epsilon}{\delta} + \tilde{I\!E} \int_0^{\tilde{\sigma} \wedge \tilde{S}_\delta} e^{-\gamma t} \int_0^1 \hat{h}_x^-(w_1 - \theta\delta + \tilde{B}_1(t) + \tilde{I}_1(t), \tilde{W}_2^*(t))d\theta dt \\
&\geq \frac{-\epsilon}{\delta} + \tilde{I\!E} \int_0^{\tilde{\sigma} \wedge \tilde{S}_\delta} e^{-\gamma t} \int_0^1 \hat{h}_x^-(w_1 - \theta\delta + \tilde{B}_1(t), \tilde{W}_2^*(t))d\theta dt,
\end{aligned}$$

where the last line follows since $\hat{h}_x^-(z_1, z_2)$ is nondecreasing in z_1 (see Theorem 24.1 of

[44]) and $\tilde{I}_1 \geq 0$. Note that the second term on the right side of the last inequality above does not depend on ϵ. Thus letting $\epsilon \to 0$ yields

$$\frac{V(w_\delta) - V(w)}{-\delta} \geq \tilde{I\!E} \int_0^{\tilde{\sigma} \wedge \tilde{S}_\delta} e^{-\gamma t} \int_0^1 \hat{h}_x^-(w_1 - \theta\delta + \tilde{B}_1(t), \tilde{W}_2^*(t)) d\theta dt. \tag{4.26}$$

Next note that since $\hat{h}_x^-$ is nondecreasing in z_1 we have almost surely ($\tilde{I\!P}$) as $\delta \downarrow 0$,

$$\int_0^1 \hat{h}_x^-(w_1 - \theta\delta + \tilde{B}_1(t), \tilde{W}_2^*(t)) d\theta \uparrow \int_0^1 \hat{h}_x^-(w_1 + \tilde{B}_1(t), \tilde{W}_2^*(t)) d\theta = \hat{h}_x^-(w_1 + \tilde{B}_1(t), \tilde{W}_2^*(t)).$$

Since clearly $\tilde{S}_\delta \uparrow \tilde{S}$ a.s. ($\tilde{I\!P}$) as $\delta \downarrow 0$, it follows on taking limits in (4.26) that

$$\Delta_- V(w) = \lim_{\delta \downarrow 0} \frac{V(w_\delta) - V(w)}{-\delta} \geq \tilde{I\!E} \int_0^{\tilde{\sigma} \wedge \tilde{S}} e^{-\gamma t} \hat{h}_x^-(w_1 + \tilde{B}_1(t), \tilde{W}_2^*(t)) dt.$$

Since the system $\hat{\Phi}$ and $\tilde{\sigma} \in \mathcal{S}(\hat{\Phi})$ are arbitrary we have

$$\Delta_- V(w) \geq \sup_{\hat{\Phi}} \sup_{\tilde{\sigma} \in \mathcal{S}(\hat{\Phi})} \tilde{I\!E} \int_0^{\tilde{\sigma} \wedge S} e^{-\gamma t} \hat{h}_x^-(w_1 + \tilde{B}_1(t), \tilde{W}_2^*(t)) dt = u(w). \blacksquare$$

In Lemma 4.8.5 we establish the inequality $\Delta_+ V(w) \leq u(w)$. As in the proof of Lemma 4.8.3 we first select an admissible control for initial condition w_1. However, we now require that the chosen control be optimal (rather than just ϵ-optimal). Such a selection is made possible by the following lemma, which establishes the existence of an optimal control for the workload control problem of Section 4.4.

Lemma 4.8.4 *Let $w \in I\!R_+^2$. Then there exists a system Φ and $I^* \in \mathcal{A}(w, \Phi)$ such that $V(w) = J(w, I^*)$.*

Proof. In the notation of Chapter 3, let $\mathcal{W} = \mathcal{U} = I\!R_+^2$, G be the two-dimensional identity matrix, $\ell = \hat{h}$, and $h = 0$. Then ℓ satisfies equation (3.5) with $\alpha_\ell = 1$ and Condition 3.2.2 is satisfied with $c_G = 1$. Thus the result is an immediate consequence of Theorem 3.2.3. $\blacksquare$

Note that if $w = (w_1, w_2)$ and $I^* = (I_1^*, I_2^*) \in \mathcal{A}(w, \Phi)$ is an optimal control given on some system Φ, then in view of the discussion below (4.11), we can assume without loss of generality that I_2^* is given by the right side of (4.12). Thus, in particular, $I_1^* \in \mathcal{A}_1(w_1, \Phi)$ and $V(w) = J_1(w, I_1^*)$.

Starting from I_1^* we now construct a control I_1^δ which is admissible for initial condition $w_1 + \delta$. The constructed I_1^δ has the property that no control is applied until the first time at which $I_1^*(t) \geq \delta$, and from then on $I_1^\delta = I_1^* - \delta$. The desired inequality follows by comparing the associated cost functions, using the optimality of I_1^*, and taking appropriate limits. To obtain the desired limits, we rely heavily on the explicit form of $\hat{h}_x^-$ given by (4.15). Note the abuse of notation in our use of w_δ and I_1^δ in Lemma 4.8.5; these quantities differ from those used in Lemma 4.8.3.

Lemma 4.8.5 *Let $w = (w_1, w_2) \in I\!R_+^2$. Then $\Delta^+ V(w) \leq u(w)$.*

Proof. Let $I_1^* \in \mathcal{A}_1(w_1, \Phi)$, with corresponding state process $W_1^* = w_1 + B_1 + I_1^*$, be an optimal control given on some system $\Phi = (\Omega, \mathcal{F}, \{\mathcal{F}_t\}, I\!P, B)$:

$$V(w) = J_1(w, I_1^*) = I\!E \int_0^\infty e^{-\gamma t} \hat{h}(W_1^*(t), W_2^*(t))dt, \tag{4.27}$$

where W_2^* is the state process corresponding to I_2^* given by (4.12). Define $\sigma^* \doteq \inf\{t \geq 0 : I_1^*(t) > 0\}$. Defining $\hat{\mathcal{F}}_t \doteq \sigma\{B(s), I^*(s), 0 \leq s \leq t\}$ we see that $\hat{\Phi} = (\Omega, \mathcal{F}, \{\hat{\mathcal{F}}_t\}, I\!P, B, I^*)$ is a controlled system and $\sigma^* \in \mathcal{S}(\hat{\Phi})$. Let $\delta > 0$ and $w_\delta = (w_1 + \delta, w_2)$. Define $\tilde{W}_1(t) = w_1 + \delta + B_1(t)$ and $\sigma^\delta \doteq \inf\{t : W_1^*(t) \geq \tilde{W}_1(t)\}$. Note that

$$\sigma^\delta = \inf\{t : I_1^*(t) \geq \delta\} \geq \inf\{t : I_1^*(t) > 0\} = \sigma^*.$$

Also, since I_1^* is right-continuous, $\sigma^\delta \downarrow \sigma^*$ a.s. as $\delta \to 0$.

Note $\tilde{W}_1(\sigma^\delta) = \tilde{W}_1(\sigma^\delta-) \geq W_1^*(\sigma^\delta-)$, which implies $\tilde{W}_1(\sigma^\delta) \geq 0$. Define

$$I_1^\delta(t) \doteq (I_1^*(t) - \delta)\, 1_{\{t \geq \sigma^\delta\}}, t \geq 0. \tag{4.28}$$

It follows from the definition of I_1^* and σ^δ that I_1^δ is $\{\mathcal{F}_t\}$-adapted, RCLL, and nondecreasing with $I_1^\delta(0) \geq 0$. Next define for $t \geq 0$,

$$\begin{aligned} W_1^\delta(t) &\doteq w_1 + \delta + B_1(t) + I_1^\delta(t) \\ &= w_1 + B_1(t) + I_1^*(t) + (\delta - I_1^*(t))\, 1_{\{0 \leq t < \sigma^\delta\}}. \end{aligned}$$

Note that $W_1^\delta(t) \geq 0$ for all $t \geq 0$ and thus $I_1^\delta \in \mathcal{A}_1(w_1 + \delta, \Phi)$. Also note that

$$W_1^\delta(t) = \begin{cases} \tilde{W}_1(t), & 0 \leq t < \sigma^\delta, \\ W_1^*(t), & t \geq \sigma^\delta. \end{cases}$$

Since $I_1^\delta \in \mathcal{A}_1(w_1 + \delta)$ we have

$$\begin{aligned} V(w_\delta) &\leq J_1(w_\delta, I_1^\delta) \\ &= I\!\!E \int_0^\infty e^{-\gamma t} \hat{h}(W_1^\delta(t), W_2^*(t)) dt \\ &= I\!\!E \int_0^{\sigma^\delta} e^{-\gamma t} \hat{h}(\tilde{W}_1(t), W_2^*(t)) dt + I\!\!E \int_{\sigma^\delta}^\infty e^{-\gamma t} \hat{h}(W_1^*(t), W_2^*(t)) dt. \end{aligned} \tag{4.29}$$

Subtracting $V(w)$ from both sides and using (4.27) yields,

$$\begin{aligned} & V(w_\delta) - V(w) \\ \leq\ & I\!\!E \int_0^{\sigma^\delta} e^{-\gamma t} [\hat{h}(\tilde{W}_1(t), W_2^*(t)) - \hat{h}(W_1^*(t), W_2^*(t))] dt \\ =\ & I\!\!E \int_0^{\sigma^\delta} e^{-\gamma t} (\tilde{W}_1(t) - W_1^*(t)) \int_0^1 \hat{h}_x^-(W_1^*(t) + \theta(\tilde{W}_1(t) - W_1^*(t)), W_2^*(t)) d\theta dt, \end{aligned}$$

where the last line follows from the convexity of $\hat{h}$ (cf. Theorem 24.2 [44]). Since $\hat{h}_x^-$ is nondecreasing we have, on noting that $\tilde{W}_1(t) - W_1^*(t) \geq 0$ for $t < \sigma^\delta$,

$$
\begin{aligned}
V(w_\delta) - V(w) &\leq I\!\!E \int_0^{\sigma^\delta} e^{-\gamma t}(\tilde{W}_1(t) - W_1^*(t))\hat{h}_x^-(\tilde{W}_1(t), W_2^*(t))dt \\
&= I\!\!E \int_0^{\sigma^*} e^{-\gamma t}(\tilde{W}_1(t) - W_1^*(t))\hat{h}_x^-(\tilde{W}_1(t), W_2^*(t))dt \\
&+ I\!\!E \int_{\sigma^*}^{\sigma^\delta} e^{-\gamma t}(\tilde{W}_1(t) - W_1^*(t))\hat{h}_x^-(\tilde{W}_1(t), W_2^*(t))dt. \qquad (4.30)
\end{aligned}
$$

For $t < \sigma^*$, $I_1^*(t) = 0$ which implies $\tilde{W}_1(t) - W_1^*(t) = \delta$. Thus the first term on the right side of (4.30) is equal to $\delta I\!\!E \int_0^{\sigma^*} e^{-\gamma t}\hat{h}_x^-(\tilde{W}_1(t), W_2^*(t))dt$. On the other hand, for $\sigma^* < t < \sigma^\delta$, $0 < I_1^*(t) < \delta$ and so $0 \leq \tilde{W}_1(t) - W_1^*(t) < \delta$. Thus since $\hat{h}_x^-$ is bounded (by some $C > 0$) it follows that the second term on the right side of (4.30) is bounded from above by $\delta C I\!\!E[\sigma^\delta - \sigma^*]$. Next note that since $I_1^* \in \mathcal{A}_1(w_1, \Phi)$, $W_1^*(t) \geq 0$ for all $t \geq 0$ and thus $\sigma^* = \inf\{t \geq 0 : I_1^*(t) > 0\} \leq \inf\{t \geq 0 : w_1 + B_1(t) \leq 0\} = S$. Using these observations in (4.30) we have

$$
\begin{aligned}
\frac{V(w_\delta) - V(w)}{\delta} &\leq I\!\!E \int_0^{\sigma^*} e^{-\gamma t}\hat{h}_x^-(\tilde{W}_1(t), W_2^*(t))dt + C I\!\!E[\sigma^\delta - \sigma^*] \\
&= I\!\!E \int_0^{\sigma^* \wedge S} e^{-\gamma t}\hat{h}_x^-(\tilde{W}_1(t), W_2^*(t))dt + C I\!\!E[\sigma^\delta - \sigma^*] \\
&= I\!\!E \int_0^{\sigma^* \wedge S} e^{-\gamma t}\hat{h}_x^-(W_1^0(t), W_2^*(t))dt + F(\delta) + C I\!\!E[\sigma^\delta - \sigma^*], \qquad (4.31)
\end{aligned}
$$

where $W_1^0(t) = w_1 + B_1(t)$ and

$$
F(\delta) \doteq I\!\!E \int_0^{\sigma^* \wedge S} e^{-\gamma t}[\hat{h}_x^-(\tilde{W}_1(t), W_2^*(t)) - \hat{h}_x^-(W_1^0(t), W_2^*(t))]dt. \qquad (4.32)
$$

Since $\sigma^* \in \mathcal{S}(\hat{\Phi})$, the first term on the right side of (4.31) is bounded from above by $u(w)$. Also, since $\sigma^\delta \downarrow \sigma^* \geq 0$ a.s., $\sigma^\delta \leq \hat{S} \doteq \inf\{t \geq 0 : w_1 + B_1(t) \leq -1\}$, and $I\!\!E\hat{S} < \infty$, we have $I\!\!E[\sigma^\delta - \sigma^*] \to 0$ as $\delta \downarrow 0$. Thus to complete the proof of the lemma it suffices to show $\limsup_{\delta \downarrow 0} F(\delta) \le 0$.

Note $\tilde{W}_1(t) \ge W_1^0(t)$ and thus $\hat{h}_x^-(\tilde{W}_1(t), W_2^*(t)) \ge \hat{h}_x^-(W_1^0(t), W_2^*(t))$. From the expression for $\hat{h}_x^-$ in (4.15) we see that

$$\begin{aligned}
& \{\hat{h}_x^-(\tilde{W}_1(t), W_2^*(t)) - \hat{h}_x^-(W_1^0(t), W_2^*(t)) > 0\} \\
= & \{\tilde{W}_1(t) > (\mu_3/\mu_2)W_2^*(t) \ge W_1^0(t)\} \\
= & \{w_1 + \delta + B_1(t) > (\mu_3/\mu_2)(w_2 + B_2(t) + I_2^*(t)) \ge w_1 + B_1(t)\} \\
= & \{0 \le (\mu_3/\mu_2)(w_2 + B_2(t) + I_2^*(t)) - (w_1 + B_1(t)) < \delta\}.
\end{aligned}$$

For $t \ge 0$, define the events

$$\begin{aligned}
A^\delta(t) & \doteq \{0 \le (\mu_3/\mu_2)(w_2 + B_2(t) + I_2^*(t)) - (w_1 + B_1(t)) < \delta\} \\
A(t) & \doteq \{0 = (\mu_3/\mu_2)(w_2 + B_2(t) + I_2^*(t)) - (w_1 + B_1(t))\}.
\end{aligned}$$

Note that for each $t \ge 0$, $\lim_{\delta \downarrow 0} A^\delta(t) = A(t)$ and since Λ is non-degenerate, $P[A(t)] = 0$ for all $t \ge 0$. Therefore

$$\begin{aligned}
F(\delta) & = I\!\!E \int_0^\infty e^{-\gamma t}[\hat{h}_x^-(\tilde{W}_1(t), W_2^*(t)) - \hat{h}_x^-(W_1^0(t), W_2^*(t))] \, 1_{\{0 \le t \le \sigma^* \wedge S\}} dt \\
& \le I\!\!E \int_0^\infty e^{-\gamma t}[h_1\mu_1 - h_2\mu_2 + h_3\mu_2] \, 1_{A^\delta(t)} dt \\
& = \int_0^\infty e^{-\gamma t}[h_1\mu_1 - h_2\mu_2 + h_3\mu_2] I\!\!P[A^\delta(t)] dt,
\end{aligned}$$

where the last line follows from Fubini's theorem. Letting $\delta \to 0$ and noting that $I\!\!P[A^\delta(t)] \to I\!\!P[A(t)] = 0$ for all $t \ge 0$, we see that $\limsup_{\delta \to 0} F(\delta) \le 0$. This completes the proof of the lemma. ∎

Proof of Theorem 4.7.1. Combining the results of Lemmas 4.8.3, 4.8.2, and 4.8.5 we have

$$u(w) \le \Delta_- V(w) \le \Delta^+ V(w) \le u(w).$$

Thus for all $w = (w_1, w_2) \in I\!R^2_+$ with $w_1 > 0$, $V_x(w)$ exists and is equal to $u(w)$. ■

The following gives a characterization of an optimal stopping time in terms of an optimal control.

Corollary 4.8.6 *Let $w \in I\!R^2_+$ and suppose Φ and $I^* = (I_1^*, I_2^*) \in \mathcal{A}(w, \Phi)$ are such that $V(w) = J(w, I^*)$. Let $\sigma^* \doteq \inf\{t \geq 0 : I_1^*(t) > 0\}$. Then $u(w) = \hat{J}(w, \sigma^*)$.*

Proof. Inequality (4.31) and the discussion that follows imply $\Delta^+ V(w) \leq \hat{J}(w, \sigma^*)$. The result is then a consequence of Theorem 4.7.1. ■

4.9 Numerical Study

In this section we present the results of a numerical study of the singular control problem of Section 4.4 and the optimal stopping problem of Section 4.7. We only present a sketch of the approximation schemes and computational algorithms. The development for the singular control problem is similar to that in Chapter 2. For an example of a Markov chain approximation for an optimal stopping problem we refer the reader to Section 5.8.1 in [32]. In the examples that follow, we fix the following values for the model parameters: $h_1 = h_2 = 1$, $h_3 = 3$, $\mu_1 = 0.5$, $\mu_2 = 2$, $\mu_3 = 1$, and $\gamma = 1$. We also set the drift of the Brownian motion B to be $(1, 1)'$ and its covariance matrix to be the two-dimensional identity matrix.

The development of the approximating Markov decision problem is similar for the singular control and the optimal stopping problem. As in Section 2.2 we first truncate the state space and replace the dynamical system by one which is reflected at the upper truncating boundary. We then discretize the truncated continuous state space with a grid. Thus, the state space for the approximating problems is given by $\mathbb{S} \doteq \{0, h, 2h, \ldots, \ell\} \times \{0, h, 2h, \ldots, \ell\}$, where $h > 0$ is the discretization parameter, and ℓ is a truncation parameter. (For the purposes of this section, we take h and ℓ as fixed values and omit them from all notation.) We then define a discrete time controlled Markov chain $\{(W_n, I_n), n = 0, 1, 2, \ldots\}$ on $\mathbb{S}$ with transition kernel $p(w, i, \tilde{w})$ that is locally consistent with the law of the continuous time process of interest. Although we denote the controlled Markov chains corresponding to both (singular control and optimal stopping) problems by the same symbols, we note that the controlled transition kernels for the two chains are quite different. We also define a sequence of interpolation intervals Δ_n; we take these to be 0 for control and reflection steps. Finally, we define an approximating Markov decision problem by defining an appropriate discrete version of the cost function. Note that we take $h = 0.02$ and $\ell = 1$ in the numerical schemes that follow.

We use classical iterative schemes to approximate numerically the optimal controls and value functions for the two MDPs. The computational algorithms are similar to the one in Section 2.5. Let $\mathbf{i} : \mathbb{S} \to \{0, 1\}$ be a feedback control and let $I_n = \mathbf{i}(W_n)$ be the associated control sequence. In the first control problem $I_n = 1$ indicates exercise of the singular control, and it corresponds to a stopping decision in the second control problem. Similarly, $I_n = 0$ represents a diffusion step in the first control problem and a continuation step in the second control problem. In particular, $p(w, 1, \tilde{w})$ equals $1_{\{\tilde{w}=w+he_1\}}$ for the first problem, and it equals $1_{\{\tilde{w}=w\}}$ for the second problem. This difference is key to the convergence properties of the two schemes. From the Markov property of $\{W_n\}$ under the feedback control $\mathbf{i}$, one has

$$J(w, I) = \sum_{\tilde{w} \in S} r(w, \mathbf{i}(w), \tilde{w}) J(\tilde{w}, I) + H(w)\Delta(w, \mathbf{i}(w)), w \in S, \tag{4.33}$$

where $r(w, \mathbf{i}(w), \tilde{w}) = e^{-\beta\Delta(w,\mathbf{i}(w))} p(w, \mathbf{i}(w), \tilde{w})$, and H is $\hat{h}$ in the singular control problem and $\hat{h}_x^-$ in the optimal stopping problem. As in Section 2.5, the above equation provides the basis for the computational algorithm. Starting from some initial control and cost, the algorithm alternates between sequences of value iterations and policy iterations until some convergence criteria is met.

Figure 4.2 displays the approximately optimal feedback controls which result from

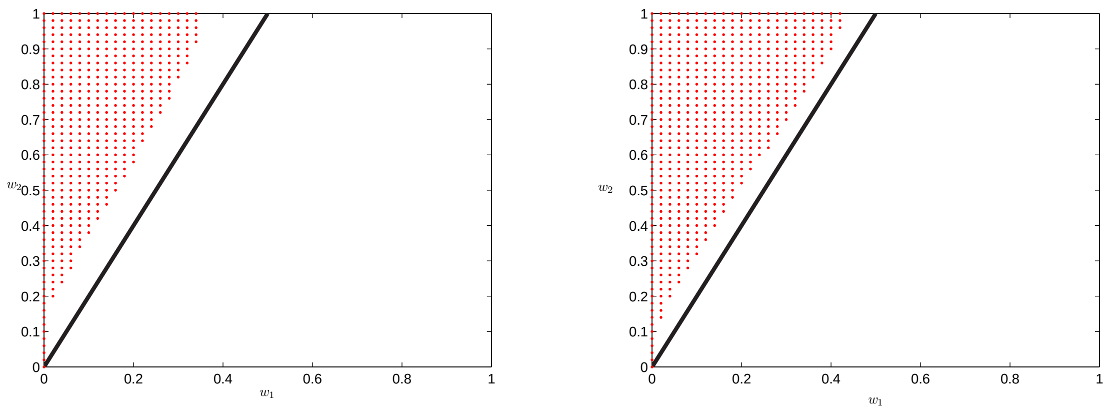

Figure 4.2: Comparison of approximately optimal controls. Left: optimal stopping MDP. Right: singular control MDP.

a numerical run for each of the MDPs. (In each of the feedback control plots in this section, dots represent the states w for which $\mathbf{i}(w) = 1$, i.e. the control states. The solid line is the line $\mu_3 w_2 = \mu_2 w_1$.) In the optimal stopping MDP, we see that the resulting control divides the state space into stopping and continuation regions. Furthermore, the boundary between the regions lies to the left of the line $\mu_3 w_2 = \mu_2 w_1$. Intuitively, since $\hat{h}_x^-(w) > 0$ for any state to the right of this line, it is always optimal to continue from that state. Once the state process is to the left of the line, the reward is negative. The optimal control allows negative reward to accumulate, but stops the process once the state is too far from the positive reward region. Similar considerations apply to the singular control MDP. In view of Corollary 4.8.6 we might expect the boundary between the stop/continuation regions in the optimal stopping MDP to be the same as the boundary between the action/no-action region in the singular control MDP. In Figure 4.2, we see that, while the boundaries do have similar shapes, they are not exactly equal. This difference can be attributed to discretization errors since the MDPs corresponding to the singular control problem and optimal stopping are equivalent only in an asymptotic sense as the discretization parameter h approaches 0.

We now discuss some numerical features of the algorithms used to produce the controls in Figure 4.2, particularly some advantages of incorporating computational results from the optimal stopping problem MDP in the numerical algorithm for the singular control MDP. The algorithm for the singular control problem is in general unstable and can encounter substantial convergence problems. The following example illustrates some of these difficulties. We take as an initial policy the feedback control $\mathbf{i}(w) = 1_{\{\mu_3 w_2 \geq \mu_2 w_1\}}$. There is no obvious choice for an initial guess (in the value iterations) for the cost function associated with this control and initial state $w \in \mathbb{S}$. Note that the cost function (4.13) is uniformly bounded (in control and state) from below by 0 and from above by the constant 3ℓ. We take the constant 0.75ℓ as our initial estimate for the cost for any initial state $w \in \mathbb{S}$. The feedback controls are updated through a sequence of policy iterations. Between two policy iterations, starting from a feedback control and an initial guess for the cost, we run 100 value iterations to get a better estimate of the associated cost.

Figure 4.3 displays the initial control, updated controls after the first and second policy iterations, and the control after 500 policy iterations. We see that even with a large number of value and policy iterations, and reasonable initial guesses for the control and cost function, the computational algorithm for the singular control approximation can produce results that drastically diverge from the optimal control illustrated in Figure 4.2. This divergence can be attributed to the truncation effects; we see in Figure 4.3 that applying singular control at the reflecting boundary $\partial_{\text{R}} \doteq \{w \in \mathbb{S} : w = (\ell, w_2)\}$ in the early stages of the computation contaminates the numerical results. To better understand the source of this divergence, consider an initial state on the reflecting boundary ∂_{R}. Suppose that $\mathbf{i}$ is a feedback control for which $\mathbf{i}(w) = 1$ for some $w \in \partial_{\text{R}}$ and that $\tilde{J}_0(w)$ is an estimate of the associated cost. Since exercising singular control at the reflecting boundary does not affect the coordinates of the state due to instantaneous reflection, we have $p(w, 1, w) = 1$ for $w \in \partial_{\text{R}}$. Also, since $\Delta(w, 1) = 0$ we see that after the k-th value iteration of an algorithm based on equation (4.33) we will have $\tilde{J}_k(w) = \tilde{J}_0(w)$, $k \geq 1$. Therefore, as long as $\tilde{J}_0(w)$ is not the true cost, the value iterations will not converge

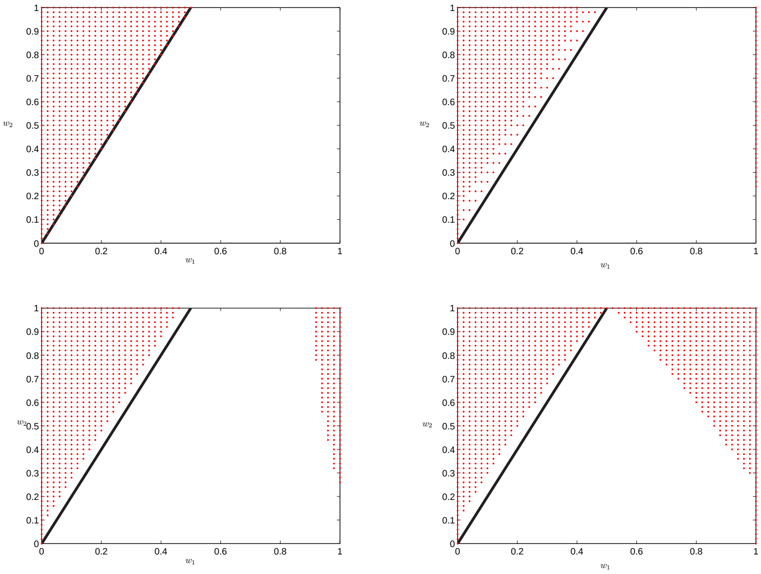

Figure 4.3: Divergence of numerical algorithm for singular control MDP. Top left: initial control. Top right: control after one policy iteration. Bottom left: control after two policy iterations. Bottom right: control after 500 policy iterations.

to the correct value. This is especially a problem when the initial guess for the cost associated with the initial control is not chosen well. In this case, it is possible that the first policy iteration determines it is cheaper to apply control at some of the states on the reflecting boundary ∂_{R}; we can see this happening in Figure 4.3. For these states, we can expect the estimated cost for the updated control to be poor even after a large number of value iterations. These errors then carry over to the next policy iteration and are compounded as the algorithm progresses.

The previous example shows that the computational algorithm for the singular control MDP may not converge if the initial guesses for the control and its associated cost are not chosen well. In contrast, a similar problem is not encountered by the numerical scheme for the optimal stopping MDP. In this algorithm, we can take advantage of the fact that if the control is to stop immediately in some initial state, then the reward associated with that state and control is 0. Since $p(w, 1, w) = 1$ and $\Delta(w, 1) = 0$, with an initial guess of $\tilde{J}_0(w) = 0$, the value iterations converge trivially to 0 which is the true cost for

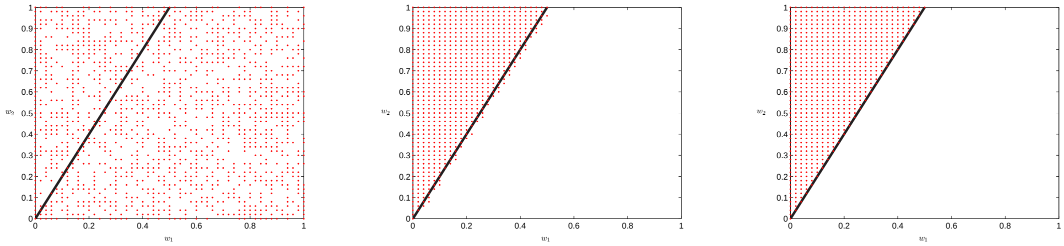

Figure 4.4: Convergence of algorithm for optimal stopping MDP. Left: initial control. Middle: control after one policy iteration. Right: control after two policy iterations.

that state and control. In addition, a positive reward is earned for states $w \in \partial_{\mathrm{R}}$ when the action is to continue. Thus, policy iterations will always choose to continue in such states, avoiding the boundary effects which arise in the algorithm for the singular control problem MDP. The example illustrates that the truncation boundary has a much more serious impact on the singular control problem than on the optimal stopping problem.

The robustness of the optimal stopping MDP algorithm to initial conditions is illustrated by the following numerical run which uses a randomly generated initial control and a cost of 0 for every state. After each policy iteration we perform a single value iteration to estimate the cost corresponding to the updated control. Figure 4.4 displays the initial control and updated controls after the first and second policy iterations. We see that even with dubious initial values and a single value iteration, the computational algorithm quickly returns to an appropriate guess for the control. After only the second policy iteration, the control is $1_{\{\mu_3 w_2 \geq \mu_2 w_1\}}$, a very reasonable initial guess for the optimal stopping problem. Furthermore, applying control ($\mathbf{i}(w) = 1$) in states on the reflecting boundary ∂_{R} does not contaminate the results as it did in the algorithm for the singular control MDP.

It is possible to ameliorate the numerical difficulties encountered by the scheme for the singular control MDP by imposing a penalty to make control costly when applied at states near the boundary ∂_{R}. However, this is an ad hoc solution and we instead take advantage of the equivalence with the optimal stopping problem to obtain good guesses for the initial control and associated cost for use in the singular control algorithm. In the study described in Figure 4.5 we take for the singular control algorithm the initial control as the near optimal control obtained from the optimal stopping algorithm. In addition, we use the near optimal cost from the optimal stopping algorithm to obtain an initial guess for the cost in the singular control scheme. Namely, guided by Theorem 4.7.1, we use the estimates of $u(w), w \in \mathbb{S}$ and numerical integration (e.g. trapezoid rule) to obtain an approximation for $V(w), w \in \mathbb{S}$. We take $V(\ell, w_2) = 1.5\ell$, $0 \leq w_2 \leq \ell$, as the initial condition for the numerical integration. Note that, in general, one expects that an approximation for V, obtained by first numerically approximating u and then integrating, will be poorer than one obtained by approximating the singular control problem by an MDP. However, the estimate obtained from the analysis of the optimal stopping problem can provide valuable initialization data for the MDP corresponding to the singular control problem. In this study, after each policy iteration we perform a single value iteration to estimate the cost corresponding to the updated control. Figure 4.5 displays the initial control, updated controls after the first two policy iterations, and the approximately optimal control. We see that by incorporating insights obtained from the optimal stopping problem we obtain good initial guesses and avoid the numerical difficulties described above. Figure 4.6 displays the initial guess for the value function obtained through numerical integration and the approximate optimal value function for the singular control MDP.

The numerical run which produced Figure 4.5 employed a single value iteration between policy iterations. Figure 4.7 displays the results of two numerical algorithms which perform 100 value iterations between policy iterations. The figure on the left displays the control after 500 policy iterations of the singular control algorithm used to produce Figure 4.3. In contrast, the figure on the right in Figure 4.7 displays the control after only 20 policy iterations of the singular control algorithm which incorporates insights from optimal stopping. We see that after 20 iterations, the control produced by the

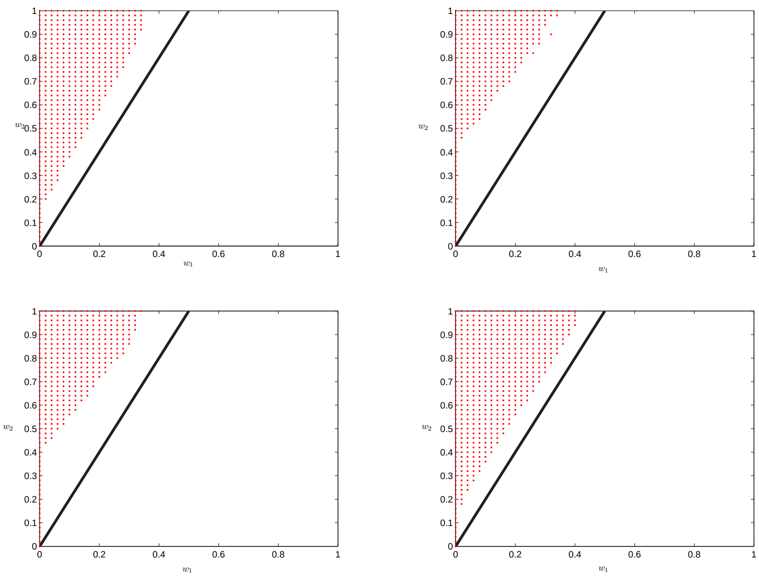

Figure 4.5: Algorithm for singular control MDP based on optimal stopping. Top left: initial control. Top right: control after one policy iteration. Bottom left: control after two policy iterations. Bottom right: approximately optimal control.

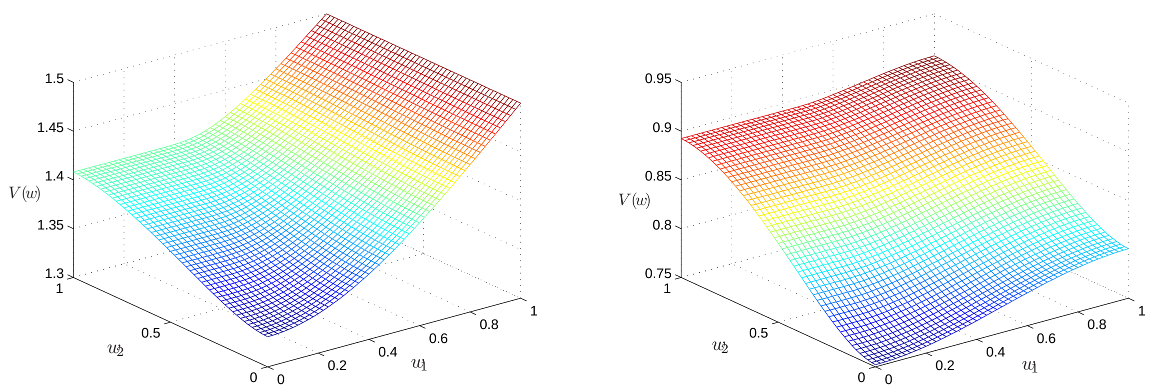

Figure 4.6: Approximate value function for singular control MDP. Left: initial guess based on optimal stopping and numerical integration. Right: approximate value function.

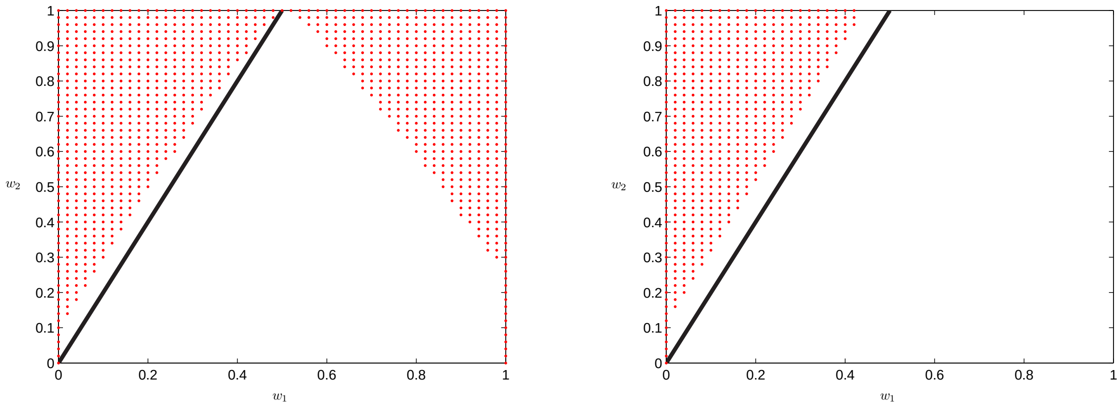

Figure 4.7: Comparison of singular control MDP algorithms. Left: control after 500 policy iterations of singular control algorithm. Right: control after 20 policy iterations of singular control algorithm which incorporates optimal stopping.

second algorithm is already close to the near optimal control, while after 500 iterations the singular control algorithm exhibits the divergence discussed earlier in this section.

Bibliography

[1] R. Atar and A. Budhiraja. Singular control with state constraints on unbounded domain. *Annals of Probability.* To appear.

[2] J. Bather and H. Chernoff. Sequential decisions in the control of a spaceship. volume 18, pages 181–207, Berkeley, CA. Fifth Berkeley Symposium on Mathematical Statistics and Probability, Vol. III: Physical Sciences, University of California Press.

[3] V. E. Benes, L. A. Shepp, and H. S. Witsenhausen. Some solvable stochastic control problems. *Stochastics*, 4(1):39–83, 1980.

[4] P. Billingsley. *Convergence of Probability Measures.* John Wiley & Sons, Inc., New York, second edition, 1999.

[5] F. Boetius. *Singular stochastic control and its relations to Dynkin game and entry-exit problems.* PhD thesis, University of Konstanz, 2001.

[6] V. S. Borkar. *Optimal Control of Diffusion Processes.* John Wiley & Sons, Inc., New York, 1989.

[7] A. Budhiraja and A. Ghosh. Diffusion approximations for controlled stochastic networks: An asymptotic bound for the value function. *Submitted for publication.*

[8] A. Budhiraja and A. Ghosh. A large deviations approach to asymptotically optimal control of crisscross network in heavy traffic. *Annals of Applied Probability*, 15(3):1887–1935, 2005.

[9] H. Chen, P. Yang, and D. Yao. Control and scheduling in a two-station queueing network. *Queueing Systems*, 18:301–332, 1994.

[10] M. G. Crandall, H. Ishii, and P.-L. Lions. User's guide to viscosity solutions of second order partial differential equations. *Bull. Amer. Math. Soc.*, 27:1–67, 1992.

[11] J. G. Dai and R. J. Williams. Existence and uniqueness of semimartingale reflecting Brownian motions in convex polyhedrons. *Theory of Probability and its Applications*, 40:1–40, 1995.

[12] M. H. A. Davis and A. R. Norman. Portfolio selection with transaction costs. *Mathematics of Operations Research*, 15:676–713, 1990.

[13] A. Dembo and O. Zeitouni. *Large Deviations Techniques and Applications.* Springer-Verlag, New York, second edition, 1998.

[14] D. Duffie, W. Fleming, H. M. Soner, and T. Zariphopoulou. Hedging in incomplete markets with HARA utility. *J. Economics Dynamics and Control*, 21(4–5).

[15] F. Dufour and B. Miller. Singular stochastic control problems. *SIAM Journal on Control and Optimization*, 43:708–730, 2004.

[16] P. Dupuis and H. Ishii. On Lipschitz continuity of the solution mapping to the Skorohod problem, with applications. *Stochastics*, 35:31–62, 1991.

[17] S. N. Ethier and T. G. Kurtz. *Markov Processes: Characterization and Convergence.* Wiley, New York, 1986.

[18] W. H. Fleming and H. M. Soner. *Controlled Markov Processes and Viscosity Solutions.* Springer-Verlag, New York, 1992.

[19] W.H. Fleming and R. Rishel. *Deterministic and Stochastic Optimal Control.* Springer-Verlag, Berlin, 1975.

[20] J. M. Harrison. *Brownian models of queueing networks with heterogeneous customer population*, pages 147–186. Springer-Verlag, New York, 1988.

[21] J. M. Harrison and J. A. Van Mieghem. Dynamic control of Brownian networks: State space collapse and equivalent workload formulation. *Annals of Applied Probability*, 7:747–771, 1997.

[22] J. M. Harrison and M. I. Reiman. Reflected Brownian motion on an orthant. *Annals of Applied Probability*, 9:302–308, 1981.

[23] J. M. Harrison and M. I. Taksar. Instantaneous control of Brownian motion. *Mathematics of Operations Research*, 8:439–453, 1983.

[24] J. M. Harrison and L. Wein. Scheduling network of queues: heavy traffic analysis of a simple open network. *Queueing Systems*, 5:265–280, 1900.

[25] U. G. Haussmann and W. Suo. Singular optimal controls I: Existence. *SIAM Journal on Control and Optimization*, 33:916–936, 1995.

[26] A. Hindy, C. Huang, and S. Zhu. Numerical analysis of a free-boundary singular control problem in financial economics. *J. Economics Dynamics and Control*, 21:297–327, 1997.

[27] N. Ikeda and S. Watanabe. *Stochastic Differential Equations and Diffusion Processes.* North-Holland Publishing Company, New York, 1981.

[28] I. Karatzas, J. Lehoczky, S. Sethi, and S. E. Shreve. Explicit solution of a general investment-consumption problem. *Mathematics of Operations Research*, 11:261–294, 1986.

[29] I. Karatzas and S. E. Shreve. *Brownian Motion and Stochastic Calculus.* Springer-Verlag, New York, 1988.

[30] N. V. Krylov. *Controlled Diffusion Processes.* Springer-Verlag, New York, 1979.

[31] T. G. Kurtz. Random time changes and convergence in distribution under the Meyer-Zheng conditions. *Annals of Applied Probability*, 19(3):1010–1034, 1991.

[32] H. J. Kushner and P. Dupuis. *Numerical Methods for Stochastic Control Problems in Continuous Time*. Springer-Verlag, New York, second edition, 2001.

[33] H. J. Kushner and L. F. Martins. Numerical methods for stochastic singular control problems. *SIAM Journal on Control and Optimization*, 29:1443–1475, 1991.

[34] H. J. Kushner and L. F. Martins. Heavy traffic analysis of a controlled multi-class queueing network via weak convergence methods. *SIAM Journal on Control and Optimization*, 34:1781–1797, 1996.

[35] H. Liu. Optimal consumption and investment with transaction costs and multiple risky assets. *J. of Finance*, 59:289–338, 2004.

[36] M. J. P. Magill and G. M. Constantinides. Portfolio selection with transaction costs. *J. of Economic Theory*, 13:245–263, 1976.

[37] L. F. Martins and H. J. Kushner. Routing and singular control for queueing networks in heavy traffic. *SIAM Journal on Control and Optimization*, 28:1209–1233, 1990.

[38] L. F. Martins, S. E. Shreve, and H. M. Soner. Heavy traffic convergence of a controlled multi-class queuing system. *SIAM Journal on Control and Optimization*, 34:2133–2171, 1996.

[39] J. L. Menaldi and M. Robin. On some cheap control problems for diffusion processes. *Trans. Amer. Math. Soc.*, (2):771–802, 1983.

[40] J. L. Menaldi and M. I. Taksar. Optimal correction problem of a multidimensional stochastic system. *Automatica*, 25(3):223–237, 1989.

[41] R. C. Merton. Optimum consumption and portfolio rules in a continuous-time model. *J. of Economic Theory*, 3:373–413, 1971.

[42] P. A. Meyer and W. A. Zheng. Tightness criteria for laws of semimartingales. *Ann. Inst. H. Poincare' Probab. Statist.*, 20:353–372, 1984.

[43] P. Protter. *Stochastic Integration and Differential Equations*. Springer-Verlag, New York, second edition, 2004.

[44] R. T. Rockafellar. *Convex Analysis*. Princeton University Press, Princeton, NJ, 1970.

[45] A. N. Shiryayev. *Optimal Stopping Rules*. Springer-Verlag, New York, 1978.

[46] S. E. Shreve and H. M. Soner. Regularity of the value function for a two-dimensional singular stochastic control problem. *SIAM Journal on Control and Optimization*, 27:876–917, 1989.

[47] S. E. Shreve and H. M. Soner. A free boundary problem related to singular stochastic control: the parabolic case. *Communications in Partial Differential Equations*, 16:373–424, 1991.

[48] S. E. Shreve and H. M. Soner. Optimal investment and consumption with transaction costs. *Annals of Applied Probability*, 4:609–692, 1991.

[49] P. E. Souganidis. Approximation schemes for viscosity solutions of Hamilton-Jacobi equations. *J. Diff. Eq.*, 56:345–390, 1985.

[50] A. Tourin and T. Zariphopoulou. *Viscosity solutions and numerical schemes for investment/consumption models with transaction costs*, pages 245–269. Cambridge University Press, Cambridge, 1997.

[51] S. A. Williams, Pao-Liu Chow, and José-Luis Menaldi. Regularity of the free boundary in singular stochastic control. *Journal of Differential Equations*, 111:175–201, 1994.

[52] J. Yong and X. Zhou. *Stochastic Controls: Hamiltonian Systems and HJB Equations.* Springer-Verlag, New York, 1999.