To the Graduate Council:

I am submitting herewith a dissertation written by Kirill Yevgenyevich Yakovlev entitled "Analysis of Solvability and Applications of Stochastic Optimal Control Problems Through Systems of Forward-Backward Stochastic Differential Equations." I have examined the final electronic copy of this dissertation for form and content and recommend that it be accepted in partial fulfillment of the requirements for the degree of Doctor of Philosophy, with a major in Mathematics.

Dr. Jie Xiong
Major Professor

We have read this dissertation
and recommend its acceptance:

Dr. Chuck Collins

Dr. Suzanne Lenhart

Dr. Phillip Daves

Accepted for the Council:

Carolyn R. Hodges
Vice Provost and Dean of Graduate School

(Original signatures are on file with official student records.)

Analysis of Solvability and Applications of Stochastic Optimal Control Problems Through Systems of Forward-Backward Stochastic Differential Equations.

A Dissertation
Presented for the
Doctor of Philosophy
Degree
The University of Tennessee, Knoxville

Kirill Yevgenyevich Yakovlev
May 2012

Dedication

I dedicate this work to my wife Jennifer and my sons Nikolai and Filip.

Acknowledgments

I would like to thank my advisor, Dr. Jie Xiong, for his participation and guidance that he provided during my years at the University of Tennessee. His deep knowledge of the subject and ability to conduct extensive research has played an important role in my degree. I greatly appreciate the time he has devoted helping me obtain the results and complete the components of my dissertation. I greatly appreciate all the advice he has given me to navigate through the mathematical and administrative labyrinths to arrive to the completion of this project. I appreciate his collaboration on the paper that became the basis for the first chapter of this dissertation.

It is impossible to overvalue the role that my wife, Jennifer, has played in helping me to obtain my goals. Her infinite patience, physical and mental support allowed me to finish my dissertation while going through the busiest time of our lives. Her advice on how to handle different aspects of multitasking life and work has been indispensable. Her participation in tracking deadlines and meeting all the requirements turned out to be extremely important.

My parents should share a great part of this success. Their firm emphasis on education and immediate participation in educational process allowed me to develop the necessary set of characteristics required for achieving educational goals. Their encouragement to obtain the most out of my abilities allowed me to obtain the highest degree level possible. It is necessary to mention that even being very far away from them I could sense the connection and strong support in the ventures that I have entered during my tenure in the United States.

I would like to express a sincere appreciation to the University of Tennessee Department of Mathematics faculty, administration, and staff. Many of them have invested their time and effort to provide me with the necessary knowledge for obtaining my degree. I also would like to express my appreciation to the University of Tennessee Department of Mathematics for providing the financial support and opportunities for obtaining valuable teaching experience. I would also like to thank them for putting trust in me to teach mathematics at the higher education level. In particular I would like to thank Pam Armentrout for her mentoring from the first day I set foot on the University of Tennessee campus. Her advice helped me to resolve numerous issues that appeared during my years at Tennessee.

My great appreciation goes to the members of my committee Chuck Collins, Suzanne Lenhart and Phillip Daves, that took the time to evaluate my dissertation. Separately, I would also like to thank my collaborators on the paper that became the basis for the first chapter of this dissertation: Chuck Collins, Suzanne Lenhart, Seema Nanda, and Jiongmin Yong.

Special thanks to Mike Saum. He provided his deep knowledge in programming and mathematics to help implement numerical algorithms. It is hard to overestimate his contribution to this work.

Abstract

A stochastic metapopulation model is investigated. The model is motivated by a deterministic model previously presented to model the black bear population of the Great Smoky Mountains in east Tennessee. The new model involves randomness and the associated methods and results differ greatly from the deterministic analogue.

A stochastic differential equation is studied and the associated results are stated and proved. Connections between a parabolic partial differential equation and a system of forward-backward stochastic differential equations is analyzed. A "four-step" numerical scheme and a Markovian iterative type numerical scheme are implemented. Algorithms and programs in the programming languages C and R are provided. Convergence speed and accuracy is compared for two numerical methods. Moreover, simulation results are presented and discussed.

Contents

1 Introduction **1**

1.1 Preliminaries . . . 1

1.2 Background Information . . . 2

1.2.1 Probabilistic Measure Theory . . . 2

1.2.2 Stochastic Analysis . . . 2

1.2.3 Regarding Stochastic Differential Equations . . . 4

1.2.4 The Four-Step Scheme . . . 8

1.3 Outline of topics . . . 9

2 Optimal Control of Harvesting in a Stochastic Metapopulation Model. **11**

2.1 Abstract . . . 11

2.2 Introduction . . . 11

2.3 Diffusion Approximation . . . 13

2.4 Existence of the Optimal Control . . . 19

2.5 Numerical Solution . . . 25

2.6 Numerical Results . . . 26

2.7 Conclusions . . . 28

3 Forward-Backward SDE in Stochastic Optimal Control Theory. Existence and Uniqueness of the Solution. Numerical Scheme. **34**

3.1 Introduction . . . 34

3.2 Formulation and preliminaries . . . 35

3.3 Solvability of FBSDE . . . 36

3.3.1 One-Dimensional Case . . . 37

3.3.2 Multidimensional Case . . . 49

3.4 Application . . . 55

3.4.1 Optimal Control Problem . . . 55

3.4.2 Numerical Solution of the FBSDE 57
3.4.3 Numerical Results 61
3.5 Conclusions 62

4 Summary **71**

Bibliography **73**

Appendices **77**

A Codes **78**

B Graphs **95**

Vita **102**

List of Figures

2.1 Population results of optimal control problem in the deterministic case, i.e. with no randomness. 29

2.2 Population results of the optimal control problem using one sample path and variance of the population growth $c = 0.1$. 30

2.3 Population results of the optimal control problem using one sample path and the variance of the population growth $c = 0.5$. 31

2.4 Population results of the optimal control problem with the variance of the mast availability c gradually increased from 0 to 0.9 by 0.1. 32

2.5 Ten sample paths of the population and the harvest rates when variance of the mast availability is $c = .25$. 33

3.1 Convergence for $p(t)$ and $X(t)$. 62

3.2 Population results of optimal control problem in the deterministic case, i.e. with no randomness. 64

3.3 Population results of optimal control with $c = 0.1$. 65

3.4 Population results of optimal control with $c = 0.5$. 66

3.5 Ten sample path simulation of population and harvesting rate when mast availability variance $c = 0.1$. 67

3.6 Ten sample path simulation of population and harvesting rate when mast availability variance $c = 0.2$. 68

3.7 Ten sample path simulation of population and harvesting rate when mast availability variance $c = 0.5$. 69

3.8 Population results of the optimal control problem with the variance of the mast availability c gradually increased from $c = 0$ to $c = 0.5$ by 0.1. 70

B.1 Population results of the optimal control problem using one sample path and the variance of the population growth $c = 0.2$. 96

B.2 Population results of the optimal control problem using one sample path and the variance of the population growth $c = 0.3$. 97
B.3 Population results of the optimal control problem using one sample path and the variance of the population growth $c = 0.4$. 98
B.4 Population results of the optimal control problem using one sample path and the variance of the population growth $c = 0.6$. 99
B.5 Population results of the optimal control problem using one sample path and the variance of the population growth $c = 0.7$. 100
B.6 Population results of the optimal control problem using one sample path and the variance of the population growth $c = 0.8$. 101

Chapter 1

Introduction

Backward stochastic differential equations were considered in stochastic optimal control problems in the 1970s by Bismut [1973] in the linear case and have been studied extensively by researchers since the early 1990s. They have proven to be invaluable in the areas of optimal control and financial theory. In this chapter, we will state necessary mathematical notions and basic assumptions for probability and stochastic optimal control theory. The underlying stochastic analytic methods will be reviewed and discussed. The formulation of backward and forward backward differential equations will be discussed and methods of obtaining solutions will be presented.

1.1 Preliminaries

All results and computations are assumed to occur on an appropriate probability space $(\Omega, \mathcal{F}, \mathcal{P})$ and a collection of increasing sub-σ-fields, $\mathcal{F}_t$, such that $\mathcal{F}_0$ contains all null sets (with respect to $\mathcal{P}$) and $\mathcal{F}_t$ is right continuous unless otherwise stated. A stochastic process with an index, such as time, can be written in the form X_t or $X(t)$. These notations will be used interchangeably, depending on the situation. The notation X_t is convenient to write succinctly, particularly in cases where multiple variables are present. And, the $X(t)$ notation is helpful when an emphasis on the indexing variable is desired. All stochastic processes $X(t)$ will be assumed to be $\mathcal{F}_t$-measurable.

1.2 Background Information

1.2.1 Probabilistic Measure Theory

In our work there is a need for precise usage of mathematical probability. Feller wrote a classical text on probability, which is divided into 2 volumes, Feller [1957a] and Feller [1957b]. This text was originally produced many years ago and is a commonly used basic text on probability theory. Kallenberg [2002] wrote a more modern text on probability that has been highly referenced recent times. Kallenberg has accomplished a great feat by creating a reasonably succinct text on all current branches of probability. The reader is assumed to have a basic knowledge of probability theory and is recommended to refer to Feller or Kallenberg as needed.

Definition 1.1 (Tightness)**.** Billingsley [1999] A family of probability measures, $\{\Pi_t\}_{t\in T}$ on a measurable space on some indexing set T is *tight* if for all $\epsilon > 0$, there exists a compact subset K of the measurable space such that

$$\Pi_t(K) > 1 - \epsilon \quad \text{for all } t \in T.$$

Definition 1.2 (Relative Compactness)**.** Billingsley [1999] A family of probability measures, $\{\Pi_t\}_{t\in T}$ on a measurable space on some indexing set T is *relatively compact* if any subsequence of $\{\Pi_t\}_{t\in T}$ contains a weakly convergent subsequence.

The celebrated Prohorov's theorem provides a useful relationship between tightness and relative compactness. One can use the books by Billingsley [1999] or Ethier and Kurtz [1985], which are standard texts on probabilistic measure theory text with an in depth view on Prohorov's theorem, as a reference for this theorem which is stated below.

Theorem 1.3 (Prohorov's Theorem)**.** *Billingsley [1999] A family of probability measures is tight if and only if it is relatively compact.*

1.2.2 Stochastic Analysis

A stochastic process is a random variable that allows for the consideration of the passage of time. Stochastic processes are used to provide stochastic models for various occurrences in areas such as biology and economics where there is a need for an aspect of randomness. Often, the methods needed for the stochastic models are quite different from the deterministic analogues. An important distinction between stochastic and deterministic models is the "deterministic models predict an outcome with absolute certainty, whereas stochastic models provide only the probability of an outcome" according to Allen [2003].

Definition 1.4 (Stochastic Process)**.** Yong and Zhou [1999] A *stochastic process* is a family of random variables, $X_t, t \in \mathcal{I}$ from $(\Omega, \mathcal{F}, P)$ to $\mathbb{R}^m$ where $\mathcal{I}$ is a nonempty index set and $(\Omega, \mathcal{F}, P)$ is a probability space.

Martingales represent a subclass of a stochastic processes. The most prolific example of a martingale is a Brownian motion. Martingales and semimartingales are defined here and will be used throughout this body of work. Kallenberg [2002] stresses the "importance of martingale methods and ideas can hardly be exaggerated" and play an essential part in the study of advanced probabilistic topics.

Definition 1.5 (Martingale)**.** Yong and Zhou [1999] A real-valued process X_t is called an $\{\mathcal{F}_t\}_{t\geq 0}$-martingale if it is $\{\mathcal{F}_t\}_{t\geq 0}$-adapted and for any $t \geq 0$, $E(X_t)$ exists with the property that

$$E(X_t|\mathcal{F}_s) = X_s \quad a.s. \quad \text{for all } 0 \leq s \leq t.$$

Definition 1.6 (Continuous Semimartingale)**.** Xiong [2008] A d-dimensional process X_t is a continuous martingale if

$$X_t = X_0 + M_t + A_t$$

where the $M^1, \ldots, M^d$ are continuous local martingales and the $A^1, \ldots, A^d$ are continuous finite variation processes.

The following theorem shows that a martingale which is square-integrable can be represented as a stochastic integral with respect to a Brownian motion. This is Theorem 3.17 in Xiong [2008].

Theorem 1.7 (Martingale Representation Theorem)**.** *Xiong [2008] Let* $M^i \in \mathcal{M}^{2,c}$ *and* $\Psi_{ij} : \mathbb{R}_+ \times \Omega \to \mathbb{R}$ *for* $i, j = 1, 2, 3, \ldots, d$ *be predictable processes such that*

$$< M^i, M^j >_t = \int_0^t \sum_{k=1}^d \Psi_{ik}(s)\Psi_{jk}(s)\, ds.$$

If

$$det(\Psi(s)) \neq 0 \quad a.s. \quad \forall s,$$

then there exists a d-dimensional Brownian motion B_t (on the original stochastic basis) such that

$$M_t^i = \sum_{k=1}^{d} \int_0^t \Psi_{ik}(s) dB_s^k.$$

Below we state Itô's formula. It is widely used in stochastic analysis, and it is the analog of the chain rule in calculus (Xiong [2008]). It is an essential result when considering applications as it demonstrates how "semimartingales are transformed under smooth mappings (Kallenberg [2002])."

Theorem 1.8 (Itô's Formula)**.** *Xiong [2008] If X_t is a d-dimensional continuous semimartingale and $F \in C_b^2(\mathbb{R}^d)$ where $C_b^2(\mathbb{R}^d)$ is the set of all bounded differentiable functions with bounded partial derivatives up to order 2, then,*

$$F(X_t) = F(X_0) + \sum_{i=1}^{d} \int_0^t \partial_i F(X_s)\, dM_s^i + \sum_{i=1}^{d} \int_0^t \partial_i F(X_s)\, dA_s^i$$

$$+\frac{1}{2} \sum_{i,j=1}^{d} \partial_{ij}^2 F(X_x)\, d < M^i, M^j >_s .$$

where the M^i's and A^j's are the martingales and continuous processes, respectfully, of the Doob-Meyer decomposition of X_t and $< M^i, M^j >$ represents the quadratic variation of M^i and M^j.

1.2.3 Regarding Stochastic Differential Equations

A stochastic differential equation (often abbreviated SDE) is a generalization of an ordinary differential equation involving an aspect of randomness. However, as is stated by Yong and Zhou [1999], SDEs can be quite complicated due to the involvement of the Itô integral. This implys SDEs involve more complex and specialized methods than the traditional ordinary differential equation counterparts. When possible, the ordinary differential equation comparison will be made along with comments as to why traditional methods of solving the equations fail.

Definition 1.9 (Stochastic Differential Equation)**.** Xiong [2008] A *stochastic differential equation* is an equation of the form,

$$\begin{cases} d\,X(t) = b(X(t))\,dt + \sigma(X(t))\,dB_t \\ \\ X(0) = \xi \end{cases}$$

where b is a continuous function mapping $\mathbb{R}^d$ to $\mathbb{R}^d$, written $b : \mathbb{R}^d \to \mathbb{R}^d$, and σ is a continuous map from $\mathbb{R}^d$ to $\mathbb{R}^{d\times m}$ ($\sigma : \mathbb{R}^d \to \mathbb{R}^{d\times m}$), B_t is a Brownian motion and $\xi \in \mathbb{R}^d$.

It is possible to define the above SDE in more generality, as is accomplished in Yong [2002]. However, a consideration on $\mathbb{R}^d$ will suit the purposes of this work. It is necessary to note that W_t representing a Wiener process is sometimes used in lieu of the Brownian motion in the literature.

The above SDE is derived from the assumption that many natural occurrences are affected by white noise. For example, a radio transmission may be altered due to white noise. Without this white noise, the transmission could be modeled by the ordinary differential equation

$$\frac{dX(t)}{dt} = b(X(t))$$

where b is a real function defined as above. But, with the white noise, we have,

$$\frac{dX(t)}{dt} = b(X(t)) + \sigma(X(t))\eta_t$$

where η is an m-dimensional white noise. The above equation is known as a stochastic differential equation (as opposed to an ordinary differential equation). Consider taking an "integral" on both sides of the above equation and you can see that the process has the form

$$X(t) = X(0) + \int_0^t b(X(s))\,ds + \int_0^t \sigma(X(s))\,dB_s$$

since the stochastic integral of the white noise is a Brownian motion. One can also write this in differential form as

$$dX_t = b(X_t)\,dt + \sigma(X_t)\,dB_t$$

Thus, we can view the process as being decomposed into a martingale part and a variation part. The solution, $X(t)$ can be viewed as a "strong solution" if the above equation holds under certain conditions. We can have a "weak solution" if equality in law is obtained. See Xiong [2008] for a more precise explanation of stochastic differential equations.

"Backward stochastic differential equations are terminal value problems of stochastic differential equations involving the Itô stochastic integral." Yong [2002] A Backward stochastic differential equation (BSDE) is an SDE that reveals 2 previously unknown processes.

Definition 1.10 (Backward Stochastic Differential Equation)**.** Yong and Zhou [1999] Let $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0}, P)$ be a complete filtered probability space on which an m-dimensional Brownian motion with natural filtration, $\{\mathcal{F}_t\}_{t\geq 0}$ is defined. The system

$$\begin{cases} dY(t) = h(Y(t), Z(t))\, dt + Z(t)\, dB_t, \\ \\ Y(T) = \xi \end{cases}$$

is a *backward stochastic differential equation.*

Note that $Y(t)$ and $Z(t)$ are the unknown processes required to be $\{\mathcal{F}_t\}$-adapted and $t \in [0, T)$. The supplementary process $Z(t)$ is necessary to preserve the non anticipative solution.

Here, forward-backward stochastic differential equations (FBSDEs) are defined.

Definition 1.11 (Forward-Backward Stochastic Differential Equation)**.** Yong [2002] Let $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0}, P)$ be a complete filtered probability space on which an m-dimensional Brownian motion with natural filtration, $\{\mathcal{F}_t\}_{t\geq 0}$ is defined. The system

$$\begin{cases} dX(t) = b(X(t), Y(t), Z(t))\, dt + \sigma(X(t), Y(t), Z(t))\, dB_t, \\ \\ dY(t) = h(X(t), Y(t), Z(t))\, dt + Z(t)\, dB_t, \\ \\ X(0) = x, \qquad Y(T) = g(X(T)) \end{cases}$$

is a *forward-backward stochastic differential equation.*

Note that $X(t)$, $Y(t)$, and $Z(t)$ are the unknown processes required to be $\{\mathcal{F}_t\}$-adapted and the functions b, σ, h, and g are nonrandom functions. The goal is to find a triple of processes, $X(t)$ and $Y(t)$ and $Z(t)$ such that $X(t)$ satisfies a (forward) SDE and $Y(t)$ satisfies a BSDE. The process $Z(t)$ is necessary to find adapted solutions.

The Pontryagin Maximum Principle is a result from control theory stating conditions for a control and a trajectory to be optimal for ordinary differential equations. The Pontryagin-type Stochastic Maximum Principle is a stochastic analog of the Pontryagin Maximum Principle. It says that if there exists an optimal pair of processes $(\bar{x}, \bar{u})$ and a pair of processes

$(p(t), q(t))$ satisfying an adjoint backward system of SDE, then $\mathcal{H}(t, \bar{x}, \bar{u}) = \max_{u \in \mathcal{U}} \mathcal{H}(t, \bar{x}, u)$, which is known as the Maximum Condition.

Consider a stochastic controlled system

$$\begin{cases} dx(t) = b(t, x(t), u(t))dt + \sigma(t, x(t), u(t)dW(t), \quad t \in [0, T], \\ x(0) = x_0 \end{cases} \tag{1.1}$$

and cost functional given by

$$J(u(\cdot)) = E\{\int_0^T f(t, x(t), u(t))dt + h(x(T))\}$$

where $W(t)$ is an m-dimensional Brownian motion starting at zero.

Following Yong and Zhou, we review the conditions for a Stochastic Maximum Principle, which include assuming that $\{\mathcal{F}_t\}_{t \geq 0}$ is the natural filtration generated by $W(t)$ augmented by all the P-null sets in $\mathcal{F}$. Thus, the "system noise is the only source of uncertainty" and "past information about the noise is available (Yong and Zhou [1999])." We have a separable metric space (U, d), real number $T > 0$, and measurable maps b, σ, f, and h. Suppose there exists a constant $L > 0$ and a modulus of continuity $\bar{\omega} \ : [0, \infty) \to [0, \infty)$ such that for $\varphi(t, x, u) = b(t, x, u),\ \sigma(t, x, u),\ f(t, x, u),\ h(x)$, we have

1. $|\varphi(t, x, u) - \varphi(t, \hat{x}, \hat{u})| \leq L|x - \hat{x}| + \bar{\omega}(d(u, \hat{u})), \quad \forall t \in [0, T],\ x, \hat{x} \in \mathbb{R}^n,\ u, \hat{u} \in U$
2. $|\varphi(t, 0, u)| \leq L, \quad \forall (t, u) \in [0, T] \times U.$

Also, assume the maps b, σ, f, and h are C^2 in x and there is a $L > 0$ and a modulus of continuity $\bar{\omega} \ : [0, \infty) \to [0, \infty)$ such that for $\varphi = b, \sigma, f, h$, it follows that

1. $|\varphi_x(t, x, u) - \varphi_x(t, \hat{x}, \hat{u})| \leq L|x - \hat{x}| + \bar{\omega}(d(u, \hat{u}))$,
2. $|\varphi_{xx}(t, x, u) - \varphi_{xx}(t, \hat{x}, \hat{u})| \leq +\bar{\omega}(|x - \hat{x}| + d(u, \hat{u}))$,

$\forall t \in [0, T],\ x, \hat{x} \in \mathbb{R}^n,\ u, \hat{u} \in U$ (Yong and Zhou [1999]).

Theorem 1.12 (Stochastic Maximum Principle)**.** *Yong and Zhou [1999] Let* $(\bar{x}(\cdot), \bar{u}(\cdot))$ *be an optimal pair. Then, there are pairs of processes*

$$\begin{cases} (p(\cdot), q(\cdot)) \in L^2_{\mathcal{F}}(0, T; \mathbb{R}^n \times (L^2_{\mathcal{F}}(0, T; \mathbb{R}^n)^m) \\ (P(\cdot), Q(\cdot)) \in L^2_{\mathcal{F}}(0, T; \mathbb{S}^n \times (L^2_{\mathcal{F}}(0, T; \mathbb{S}^n)^m) \end{cases} \tag{1.2}$$

where

$$\begin{cases} q(\cdot) = (q_1(\cdot), \ldots, q_m(\cdot), \quad Q(\cdot) = (Q_1(\cdot), \ldots, Q_m(\cdot) \\ q_j(\cdot) \in L^2_{\mathcal{F}}(0,T;\mathbb{R}^n, \quad Q_j(\cdot_\in L^2_{\mathcal{F}}(0,T;\mathbb{S}^n, \quad 1 \leq j \leq m, \end{cases} \tag{1.3}$$

satisfying the first-order and second-order adjoint equations (3.8) *and* (3.9) *from pages 115 and 116 of Yong and Zhou [1999] respectively such that*

$$\mathcal{H}(t, \bar{x}(t), \bar{u}(t)) = \max_{u \in \mathcal{U}} \mathcal{H}(t, \bar{x}(t), u), \quad a.e.t \in [0,T], \quad P - a.s.$$

1.2.4 The Four-Step Scheme

In their original article Ma et al. [1994] studied the solvability of Forward-Backward Stochastic Differential equations explicitly by solving a corresponding parabolic PDE. They proposed the Four-Step Scheme, which is a reverse of the Feynman-Kac-type Formulae.

1. Find $z(t,x,y,u) = u\sigma(t,x,y,z(t,x,y,u))$, s.t. $\forall (t,x,y,u) \in [0,T] \times \mathbb{R}^n \times \mathbb{R}^k \times \mathbb{R}^{k \times n}$

2. Use function z obtained above to solve the following parabolic system of PDE for $\theta(t,x)$:

$$\begin{cases} \theta^l_t + \frac{1}{2} tr[\theta^l_{xx}(\sigma\sigma^T)(t,x,\theta,z(t,x,\theta,\theta_x))] + < b(t,x,\theta,z(t,x,\theta,\theta_x)), \theta^l_x > \\ \quad -h^l(t,x,\theta,z(t,x,\theta,\theta_x)) = 0, \ (t,x) \in (0,T) \times \mathbb{R}^n, \ 1 \leq l \leq k, \\ \theta(T,x) = g(x), \ x \in \mathbb{R}^n \end{cases}$$

3. Use θ and z obtained in Steps 1-2 to solve the following forward SDE:

$$X(t) = x + \int_0^t \tilde{b}(s, X(s))ds + \int_0^t \sigma(s, X(s))dW(s)$$

   where

$$\begin{cases} \tilde{b}(t,x) = b(t,x,\theta,z(t,x,\theta,\theta_x)), \\ \tilde{\sigma}(t,x) = \sigma(t,x,\theta,z(t,x,\theta,\theta_x)) \end{cases}$$

4. Set

$$\begin{cases} p(t) = \theta(t, X(t)), \\ q(t) = z(t, X, \theta(t,X), \theta_x(t,X)). \end{cases}$$

This gives a solutions to forward backward stochastic differential equations by solutions to partial differential equations, which are most importantly adapted to the filtration.

1.3 Outline of topics

Stochastic models are well suited to be applied to biological populations, for which randomness is a common characteristic. It is often the case that deterministic models are sufficient for models involving very large populations. Conversely, when the populations in question are small, there is a possibility of population extinction. When a possibility of extinction is present, a stochastic model is preferable as is mentioned in Allen [2003]. Moreover, a random component can also be used to consider environmental aspects or other variations that may include randomness.

The black bear population of the Great Smoky Mountains is an excellent example of a situation where a stochastic population model is beneficial. There are interesting environmental factors, such as shortage of food and poaching practices. Also, the bear population is a relatively small population when compared with other biological populations and systems which makes the bear population a perfect candidate for a stochastic model. Moreover, there are sections of the environment where hunting is allowable and other sections where hunting is prohibited. Thus, the black bear population of East Tennessee is a very rich topic of study. But, it is important to mention the models and techniques in this work are not restricted to the black bear population and could be utilized in similar situations involving other populations.

In Chapter 2, a stochastic metapopulation model is investigated. This model could be implemented to study the behaviors of the black bear population in the Smoky Mountains or the model could be used in other similar situations. It is assumed there exist two adjacent regions through which the bears migrate. One region allows for controlled hunting practices and the other region is a protected area where harvesting is illegal. An optimal control problem for a model with SDEs is investigated. Stochasticity is injected into a previously used deterministic model through random growth rate and the related existence results are stated and proven.

At the end of chapter 2, the numerical implementation is discussed. Simulations are presented for various components of the source data. The sample paths are generated for each simulation with differing random seeds. Also, comparisons with the previously studied deterministic models are made.

The work in Chapter 3 addresses the numerical scheme regarding existence and uniqueness of the solution of a particular FBSDE in stochastic optimal control theory. A numerical method for a class of forward-backward stochastic differential equations (FBSDEs) is formulated and analyzed. There is a need to provide solutions in a situation where the coefficients of state and adjoint equations may not be Lipschitz continuous. In fact, in general most FBSDE systems derived in stochastic optimal control theory will fail to poses Lipschitz condition on the backward drift coefficient. The one dimensional and multidimensional cases are considered separately. Existence and uniqueness is considered and established. The results are used to revisit the bear metapopulation model. Numerical methods are described and discussed.

Lastly, the precise numerical implementation in the programming languages C and R are described for the reader in the Appendices. This code could be adapted for various situations or could be used as a reference to those interested in this work.

Chapter 2

Optimal Control of Harvesting in a Stochastic Metapopulation Model.

2.1 Abstract

We consider a metapopulation model for a single species inhabiting two bounded contiguous regions where movement of the population across the shared boundary is allowed. The population in one of the bounded regions can be harvested. We introduce stochastic growth rates for the two populations in a system of ordinary differential equations that model the population dynamics in these two regions. We derive the resulting stochastic control problem with harvesting in one region as the control. The existence of an optimal control is established by solving an associated quasi-linear-quadratic optimal control problem. We present numerical simulations to illustrate several scenarios.

2.2 Introduction

Consider two bounded regions with a shared boundary, allowing cross boundary movement of a population. The population in one of the bounded regions is subject to harvesting. This metapopulation model is motivated by an interest to study the spatial and temporal behavior of species, such as the black bear population in the Smoky Mountains, which are generally confined to a protected area such as a national park where poaching is prohibited. However in an adjoining region, such as a forest, limited hunting licenses would be permitted. Movement outside of the two bounded regions could happen due to forays for food during lean times when there is a shortfall in the protected areas. The model is general enough to be used for the study of single species populations of other habitat in regions that can be defined similarly.

We first present a deterministic model with harvesting control. Our state variables $P(t)$, $F(t)$, and $O(t)$ represent population densities in the Park, Forest and outside of Park/Forest area respectively. The resulting system of ordinary differential equations is given by:

$$\frac{dP(t)}{dt} = rP(t)\left(1 - \frac{P(t)}{K}\right) - e_P P(t) + e_F\left(1 - \frac{P(t)}{K}\right)F(t) \tag{2.2.1}$$

$$\frac{dF(t)}{dt} = rF(t)\left(1 - \frac{F(t)}{K}\right) - e_F F(t) + e_P\left(1 - \frac{F(t)}{K}\right)P(t) - u(t)F(t) \tag{2.2.2}$$

$$\frac{dO(t)}{dt} = (e_F + e_P)P(t)\frac{F(t)}{K} \tag{2.2.3}$$

The initial populations $P(0)$, $F(0)$, and $O(0)$ are known. The function $u(t)$ is a harvesting rate in the forest patch. Setting $u(t) \equiv 0$ represents the case of no harvesting. All of the parameters in the above equations are assumed to be positive and constant.

Since the food sources of the park and forest are similar, we take the population growth rate r (the net of birth and death rates) to be the same for both regions. The populations in the park and in the forest are assumed to follow logistic growth. The carrying capacity is assumed to be the same in both regions, though this would clearly depend upon the size of each region.

Emigration from one area to another depends only on the population in that area, hence giving a linear second term in the P and F differential equations. The emigration rate from park to forest, e_P, could be different from the rate from forest to park, e_F, as the proportion of the park boundary connected to the forest may be different from the proportion of the forest boundary connected to the park.

Immigration into a region is determined by the carrying capacity of that region and population density of the emigrating region. Hence the third terms in the P and F differential equations are quadratic. Implicitly, the immigration takes place uniformly across the boundary. The growth rate r is affected by the mast, i.e., acorns and nuts, that are a major source of nutrition for black bears. Twenty years of data show no discernable pattern in the mast production from one year to another [Beeman and Pelton 1970; Brody and Pelton 1998]. Therefore we assume that the distribution of the mast over the years is random. This results in a system of stochastic differential equations (SDEs).

For a background on stochastic models applied to biological populations, see books by Allen [2003], Ricciardi [1977], and Ludwig [1974]. For more specific examples of optimal harvesting in stochastic environments, see articles: Hanson and Ryan [1988], Ludwig and Varah [1979a], and Ludwig and Varah [1979b].

Salinas et al. [2005] have proposed a similar metapopulation model with harvesting in the forest and in the park (only at a low level). They also consider the case of a periodic carrying capacity. They investigated optimal harvesting control strategies to minimize the population in the outside region and the costs associated with the controls. Minimizing the population in the outside region arose from the desire to keep bears in the park and in the forest. We generalize this work to consider an optimal control problem for a model with SDEs, with a goal of minimizing the outside population O and the cost of harvesting. Our deterministic model is very similar to that used by Salinas et al. [2005].

The random nature of the food availability is better represented by a stochastic differential equation and so this generalization makes sense. Limited methodology is available for solving stochastic optimal control problems even though stochastic optimal control problems are of great interest. Currently they are solved through associated systems of forward-backward stochastic differential equations treated with the "Four-Step Scheme" [Yong 2002; Zhang 2004]. This rigorous approach is due to irreversibility of time for the terminal value SDE problem.

In a deterministic setting for an optimal control problem, one could apply Pontryagin's Maximum Principle, which we are not able to do in a system of SDEs. Application of the Stochastic Maximum Principle to the resulting system of SDEs is not valid here due to the nonlinear structure of the SDEs with respect to the state variables. We instead consider a quasi-linear-quadratic (QLQ) stochastic optimal control problem studied in Yong [2002], which is used to establish existence of an optimal control.

The rest of the paper is organized in the following way. In Section 2, we discretize and introduce stochasticity to the model. We derive an underlying system of SDEs for the model. In Section 3, we formulate a stochastic optimal control problem and prove existence of a stochastic optimal control via solving an associated QLQ problem. Section 4 provides details on numerical implementation. Section 5 presents the results of some numerical simulations and conclusions follow in Section 6.

2.3 Diffusion Approximation

This section can be rather challenging to an unexposed reader. The underlying system of stochastic differential equations for our model is derived in this section. Asymptotic analysis of the discrete time process presented below is pedagogical in nature and the interested reader may look at Xiong [2008] and Ethier and Kurtz [1985]. On initial reading, we suggestion that those with a weak background in asymptotic analysis skip to Proposition 2.3 where the resulting system of stochastic differential equations is presented.

To model the random growth rate resulting from randomness in mast we consider the discrete version of the population model and rescale the process.

We scale time by a factor n and the population size by a factor $\sqrt{n}$. We assume that the carrying capacity of the environment is $K = \sqrt{n}$. Thus, our discretized version of park population becomes:

$$P_m^n = P_0^n + \sum_{i=0}^{m-1} \left(rP_i^n \left(1 - \frac{P_i^n}{\sqrt{n}}\right) - e_P P_i^n + e_F \left(1 - \frac{P_i^n}{\sqrt{n}}\right) F_i^n \right) \frac{1}{n} \tag{2.3.1}$$

where P_m^n is a scaled version of the population density in the park and $m = 1, 2, ..., n$. We define a new variable that describes the park population density as

$$X_1^n(t) = \frac{1}{\sqrt{n}} P_{[nt]}^n.$$

We also suppose that the growth rate r takes the form

$$r(\eta_i^n) = a + c\sqrt{n}\eta_i^n$$

where η_i^n, $i = 1, 2, \cdots$, are i.i.d. random variables with mean 0 and variance 1. Thus, $r(\eta_i^n)$ is a random variable with a fixed mean a, and a large variance $c\sqrt{n}$. We rewrite equation (2.3.1) in terms of $X_1^n(t)$ to obtain

$$\begin{aligned} X_1^n(t) &= X_1^n(0) + \sum_{i=0}^{[nt]-1} \frac{c}{\sqrt{n}} \eta_i^n \frac{1}{\sqrt{n}} P_i^n \left(1 - \frac{P_i^n}{\sqrt{n}}\right) \\ &\quad + \sum_{i=0}^{[nt]-1} \frac{a}{n} \frac{1}{\sqrt{n}} P_i^n \left(1 - \frac{P_i^n}{\sqrt{n}}\right) \\ &\quad - \sum_{i=0}^{[nt]-1} \frac{e_F}{n} \frac{1}{\sqrt{n}} P_i^n + \sum_{i=0}^{[nt]-1} \frac{e_P}{n} \left(1 - \frac{P_i^n}{\sqrt{n}}\right) \frac{1}{\sqrt{n}} F_i^n \end{aligned}$$

We rename the summands and write:

$$X_1^n(t) \equiv X_1^n(0) + M_t^{1,n} + A_t^{11,n} - A_t^{12,n} + A_t^{13,n} \tag{2.3.2}$$

Here $M^{1,n}$ will be shown to be a martingale term.

We keep the last three term separate until it is beneficial for simplicity of notation to combine them as

$$A_t^{1,n} = A_t^{11,n} + A_t^{12,n} + A_t^{13,n}.$$

Similarly, we derive equations for the population density in the forest $X_2^n(t)$ and the population outside region, $X_3^n(t)$. Thus

$$\begin{aligned} X_2^n(t) &= X_2^n(0) + \sum_{i=0}^{[nt]-1} \frac{c}{\sqrt{n}}\eta_i^n \frac{1}{\sqrt{n}} F_i^n \left(1 - \frac{F_i^n}{\sqrt{n}}\right) \\ &\quad + \sum_{i=0}^{[nt]-1} \frac{a}{n}\frac{1}{\sqrt{n}} F_i^n \left(1 - \frac{F_i^n}{\sqrt{n}}\right) \\ &\quad - \sum_{i=0}^{[nt]-1} \left(\frac{e_F + u_t}{n}\right) \frac{1}{\sqrt{n}} F_i^n + \sum_{i=0}^{[nt]-1} \frac{e_P}{n} \left(1 - \frac{F_i^n}{\sqrt{n}}\right) \frac{1}{\sqrt{n}} F_i^n \\ &\equiv X_2^n(0) + M_t^{2,n} + A_t^{2,n}. \end{aligned}$$

Notice that $X_2^n(t)$ has an extra term that contains the control u_t and $M_t^{2,n}$ is a martingale term. We don't separate $A_t^{2,n}$ into its subparts and keep it as one process here. Since $O(t)$ does not depend directly on r, which is stochastic, the martingale term is not required for equation

$$\begin{aligned} X_3^n(t) &= X_3^n(0) + \sum_{i=0}^{[nt]-1} \left(\frac{e_F + e_P}{n}\right) F_i^n P_i^n \\ &\equiv X_3^n(0) + A_t^{3,n}. \end{aligned}$$

We desire to obtain convergence of this system to a solution of some system of SDEs. We are going to use the notion of tightness to do so. We only provide details on how to obtain the tightness of $X_1^n(t)$. The other two are done similarly.

Proposition 2.1. *Suppose that* $\sup_n \mathbb{E}|\eta_i^n|^4 < \infty$. *Then* $\{X_1^n(t) : \ n \geq 1\}$ *is tight in* $\mathbf{D}([0,T],[0,1])$*, the space of Cadlag functions.*

Before proving Proposition 2.1, we need the following lemma.

Lemma 2.2. *Let* $\{A^n\}$ *be a family of continuous stochastic processes with* $A_0^n = 0$. *For any*

$\omega \in \Omega$, and $s, t \in [0, T]$ if there exist positive constants K_1, K_2 such that

$$|A_t^n - A_s^n| \leq K_1|t-s| + \frac{K_2}{n}, \ \ \textit{for all} \ \ n \geq 1$$

then $\{A^n\}$ is tight.

Proof. Let $\{t_i\}_{i \geq 1}$ be a dense subset of $[0, T]$. For each i, we have

$$|A_{t_i}^n| \leq K_1 T + K_2$$

and hence, $\{A_{t_i}^n\}$ is tight in $\mathbb{R}$. Therefore, $\{A_{t_i}^n, i \geq 1\}$ is tight in $\mathbb{R}^\infty$, which is a metric space with metric

$$\rho(a, b) = \sum_i 2^{-i}(|a_i - b_i| \wedge 1).$$

Then by Prohorov's Theorem and Skorohod's Theorem [Ethier and Kurtz 1985] we can assume that $\{A_{t_i}^n, i \geq 1\}$ is convergent to $\{A_{t_i}, i \geq 1\}$ almost surely. Note that,

$$|A_{t_i} - A_{t_j}| = \lim_{n \to \infty} |A_{t_i}^n - A_{t_j}^n| \leq K_1|t_i - t_j|.$$

Thus, A_t can be extended to a process defined for all $t \in [0, T]$. Then for almost all $\omega \in \Omega$, for all $t \in [0, T]$, and $|t_i - t| < \epsilon$

$$\begin{aligned}
|A_t^n - A_t| &\leq |A_t^n - A_{t_i}^n| + |A_{t_i}^n - A_{t_i}| + |A_{t_i} - A_t| \\
&\leq 2K_1\epsilon + \frac{K_2}{n} + |A_{t_i}^n - A_{t_i}| \\
\sup_{t \leq T} |A_t^n - A_t| &\leq 2K_1\epsilon + \frac{K_2}{n} + \max_{1 \leq i \leq k} |A_{t_i}^n - A_{t_i}| \\
\overline{\lim_{n \to \infty}} \sup_{t \leq T} |A_t^n - A_t| &\leq 2K_1\epsilon \to 0
\end{aligned}$$

This proves the tightness of A^n, which concludes the proof of Lemma 2.2 □

We now proceed to proving Proposition 2.1.

Proof of proposition 2.1. Note that

$$\begin{aligned}
|A_t^{12,n} - A_s^{12,n}| &= \left| \sum_{i=0}^{[nt]-1} \frac{e_P}{n} \frac{1}{\sqrt{n}} P_i^n - \sum_{j=0}^{[ns]-1} \frac{e_P}{n} \frac{1}{\sqrt{n}} P_i^n \right| \\
&= \frac{e_P}{n} \left| \sum_{[ns]-1}^{[nt]-1} \frac{1}{\sqrt{n}} P_i^n \right| \\
&\leq \frac{e_P}{n} ([nt] - 1 - [ns] + 1) \max_{i \in ([ns]-1, [nt]-1)} \frac{P_i^n}{\sqrt{n}} \\
&\leq \frac{e_P}{n} ([nt] - [ns]), \text{ since } \max_{i \in ([ns]-1, [nt]-1)} \frac{P_i^n}{\sqrt{n}} \leq 1 \\
&\leq \frac{e_P}{n} (nt - ns + 1) \\
&\leq e_P |t - s| + \frac{e_P}{n}. \qquad (2.3.3)
\end{aligned}$$

We apply Lemma 2.2 to conclude that $A^{12,n}$ is tight. $\square$

Since $A^{11,n}$ and $A^{13,n}$ are defined similarly to $A^{12,n}$, their tightness follows similarly.

We still need to establish tightness of $M_t^{1,n}$. Notice that there exists a martingale $M_t^{1,n}$ with a quadratic variation process

$$[M^{1,n}]_t = \sum_{i=0}^{[nt]-1} \frac{c^2}{n} (\eta_i^n)^2 \left(\frac{1}{\sqrt{n}} P_i^n \right)^2 \left(1 - \frac{P_i^n}{\sqrt{n}} \right)^2.$$

If we can show that this martingale $M_t^{1,n}$ is tight then we can conclude that X_1^n is tight. Notice that η_i^n and P_i^n are two independent random variables. Then the Meyer process defined in Xiong [2008] is:

$$\begin{aligned}
\left\langle M^{1,n} \right\rangle_t &= \sum_{i=0}^{[nt]-1} \mathbb{E} \left\{ \frac{c^2}{n} (\eta_i^n)^2 \right\} \mathbb{E} \left\{ \left(\frac{1}{\sqrt{n}} P_i^n \right)^2 \left(1 - \frac{P_i^n}{\sqrt{n}} \right)^2 |\mathcal{F}_{i-1} \right\} \\
&= \sum_{i=0}^{[nt]-1} \frac{c^2}{n} \mathbb{E} \left\{ \left(\frac{1}{\sqrt{n}} P_i^n \right)^2 \left(1 - \frac{P_i^n}{\sqrt{n}} \right)^2 |\mathcal{F}_{i-1} \right\}.
\end{aligned}$$

As in the proof of Proposition 2.1, we get the tightness of $\langle M^{1,n} \rangle_t$. Then we use the result from Jacod and Shiryaev [1987, Theorem VI.4.13] to conclude that the martingale $M_t^{1,n}$ is tight. We have shown that $A^{11,n}$, $A^{12,n}$, $A^{13,n}$, $M_t^{1,n}$ are tight, and thus $X_1^n(t)$ is tight.

Similarly, we can prove the tightness of $X_2^n(t)$ and $X_3^n(t)$. From now on, we denote

$$X^n(t) = (X_1^n(t), X_2^n(t), X_3^n(t)), \quad X(t) = (X_1(t), X_2(t)_t, X_3(t)). \tag{2.3.4}$$

Proposition 2.3. *X^n converges weakly to X in $C([0,T],(0,1)^2 \times (0,\infty))$ which is the unique solution to the following system of SDEs:*

$$\begin{cases} X_1(t) = X_1(0) + \int_0^t cX_1(s)(1-X_1(s))dW_s + \int_0^t aX_1(s)(1-X_1(s))ds \\ \qquad - \int_0^t e_P X_1(s)ds + e_F \int_0^t (1-X_1(s))X_2(s)ds \\ X_2(t) = X_2(0) + \int_0^t aX_2(s)(1-X_2(s))ds + \int_0^t cX_2(s)(1-X_2(s))dW_s \\ \qquad - \int_0^t e_F X_2(s)ds + e_P \int_0^t (1-X_2(s))X_1(s)ds - \int_0^t u_s X_2(s)ds \\ X_3(t) = X_3(0) + \int_0^t (e_F + e_P)X_1(s)X_2(s)ds. \end{cases} \tag{2.3.5}$$

Proof. Let X be a limit point of $\{X^n\}$. On a subsequence, we assume that X^n converges weakly to X. By the Skorohod Representation Theorem [Ethier and Kurtz 1985] we may assume that X^n converges to X in $C([0,T],(0,1)^2 \times (0,\infty))$ a.s. Using this assumption it is easy to show convergence a.s. of the following

$$\begin{array}{ll} A_t^{1,n} \to & \int_0^t aX_1(s)(1-X_1(s))ds - e_P \int_0^t X_1(s)ds + e_F \int_0^t (1-X_1(s))X_2(s)ds \\ & \equiv A_t^1 \\ A_t^{2,n} \to & \int_0^t aX_2(s)(1-X_2(s))ds - e_F \int_0^t X_2(s)ds + e_P \int_0^t (1-X_2(s))X_1(s)ds \\ & - \int_0^t u_s X_2(s)ds \equiv A_t^2 \\ A_t^{3,n} \to & \int_0^t (e_F + e_P)X_1(s)X_2(s)ds \equiv A_t^3 \end{array}$$

and the Meyer process

$$\left\langle M^{i,n}, M^{j,n} \right\rangle_t \to \int_0^t c^2 X_i(s)(1-X_i(s))X_j(s)(1-X_j(s))ds, \ i,j = 1,2$$

Thus,

$$M_t^i \equiv X_i(t) - A_t^i, \ i = 1,2$$

are martingales with

$$\left\langle M^{i,n}, M^{j,n} \right\rangle_t \to \int_0^t c^2 X_i(s)(1-X_i(s))X_j(s)(1-X_j(s))ds, \ i,j = 1,2. \tag{2.3.6}$$

By the martingale representation theorem [Xiong 2008, Theorem 3.17], there is a one-dimensional Brownian motion W_t such that

$$M_t^i = \int_0^t c^2 X_i(s)(1 - X_i(s))dW_s,\ i = 1, 2$$

This proves that $X(t)$ is a solution to the system of SDEs (2.3.5). $\square$

Due to the structure properties it is easy to prove the uniqueness for the solution to (2.3.5). Also, due to the structure properties the boundary

$$\{X(t) :\ i = 1, 2,\ X_i(t) = 0 \text{ or } 1;\ \text{or } X_3(t) = 0\}$$

is unattainable.

2.4 Existence of the Optimal Control

Motivated by the desire to keep the outside population low, we choose the following objective functional, $J(u) = \mathbb{E}\left(X_3(T) + \frac{1}{2}\int_0^T mu^2(t)dt\right)$, over the control space,

$$\mathcal{U}[0, T] = \{u : [0, T] \to [0, 1] \mid u(\cdot) \text{ is } \mathcal{F}_t\text{-progressively measurable}\}. \tag{2.4.1}$$

Here $u(t)$ is the control and m is a positive constant, balancing the relative importance of the cost of harvesting against the minimization of the outside population at the final time.

We write the state system given by(2.3.5) in a more compact differential form as:

$$\begin{cases} dX(t) &= b(X(t), u(t))dt + \sigma(X(t))dW(t) \\ X(0) &= x_0 \end{cases} \tag{2.4.2}$$

where the drift coefficient $b(X(t), u(t))$ is given by:

$$b(X(t), u(t)) = \begin{pmatrix} b_1(X(t), u(t)) \\ b_2(X(t), u(t)) \\ b_3(X(t), u(t)) \end{pmatrix}$$

with

$$\begin{cases} b_1(X(t),u(t)) = aX_1(t)(1-X_1(t)) - e_P X_1(t) \\ \qquad\qquad\qquad\quad + e_F(1-X_1(t))X_2(t), \\ b_2(X(t),u(t)) = aX_2(t)(1-X_2(t)) - e_F X_2(t) \\ \qquad\qquad\qquad\quad + e_P(1-X_2(t))X_1(t) - u(t)X_2(t), \\ b_3(X(t),u(t)) = (e_F+e_P)X_1(t)X_2(t) \end{cases}$$

and the diffusion coefficient $\sigma(X(t))$ is given by:

$$\sigma(X(t)) = \begin{pmatrix} cX_1(t)(1-X_1(t)) \\ cX_2(t)(1-X_2(t)) \\ 0 \end{pmatrix}.$$

In the deterministic case, we would show the existence of the optimal control and then proceed with Pontryagin's Maximum Principle to obtain a characterization of an optimal control. Similarly, we must first establish the existence for a stochastic optimal control. This usually results in an application of Pontryagin-type Stochastic Maximum Principle. Then we can solve an associated system of Forward-Backward SDEs (FBSDEs) to obtain the optimal control and optimal state trajectory. Methodology for solving such forward-backward systems of SDEs is relatively sparse. Few types of such systems of SDEs are known to have solutions at this time: linear forward-backward SDE in Yong [1999] or Ma and Yong [1999], partially decoupled systems of SDE such as in Zhang [2004], and systems of SDE that have Lipschitz (in all variables except time) drift and diffusion coefficients studied in Ma and Yong [1995]. In the last case, when the coefficients of the system of FBSDEs are also deterministic, we can obtain its solution through solving an associated multidimensional system of parabolic PDE's. Our model would result in a forward-backward system of SDEs with deterministic coefficients which could possibly be solved through an associated system of PDE's. The Stochastic Maximum Principle cannot be applied here due to the nonlinear structure of our model in the state variables.

To overcome this difficulty, we use a different approach to show existence of the optimal control. It is not only descriptive, but also a very practical method for obtaining the optimal control and the optimal state sample path. We reformulate our control problem as a special case of the QLQ stochastic optimal control problem. Such a problem was studied in Yong [2002]. We proceed with an adaptation of the general QLQ control problem to our model.

We reduce our problem by eliminating the $X_3(t)$ term from the system of SDEs by using:

$$X_3(T) = \int_0^T (e_F + e_P)X_1(t)X_2(t)dt, \tag{2.4.3}$$

Thus we can conduct the following reduction of the original optimal control problem. We define a new functional

$$J(u) = \mathbb{E}\left(\int_0^T (e_F + e_P)X_1(t)X_2(t) + \frac{1}{2}mu^2(t)dt\right) \tag{2.4.4}$$

and a new state equation

$$\begin{cases} dX(t) &= (\hat{b}(t,X(t)) + b_0(X(t))u(t))dt + \hat{\sigma}(t,X(t))dW(t) \\ X(0) &= X_0 \end{cases} \tag{2.4.5}$$

where $\hat{b}$ is

$$\begin{Bmatrix} aX_1(t)(1-X_1(t)) - e_P X_1(t) + e_F(1-X_1(t))X_2(t) \\ aX_2(t)(1-X_2(t)) - e_F X_2(t) + e_P(1-X_2(t))X_1(t) \end{Bmatrix}$$

with $b_0 = (0, -X_2(t))$ and $\hat{\sigma} = (cX_1(t)(1-X_1(t)), cX_2(t)(1-X_2(t)))$.

In the previous section we saw that $0 < X_1(t) < 1$ and $0 < X_2(t) < 1$. Thus it is clear that $\hat{b}, b_0, \hat{\sigma}$ are measurable and there exists a constant $L > 0$ such that for all $t \in [0,T]$ and for all $x, \hat{x} \in (0,1)^2$,

$$\begin{cases} |\hat{b}(t,x) - \hat{b}(t,\hat{x})| \le L|x - \hat{x}| \\ |\hat{\sigma}(t,x) - \hat{\sigma}(t,\hat{x})| \le L|x - \hat{x}| \\ |b_0(x) - b_0(\hat{x})| \le L|x - \hat{x}| \\ \{\hat{b}(t,0), \hat{\sigma}(t,0), b_0(0)\} \le L. \end{cases} \tag{2.4.6}$$

Under the above observation, for any $u(\cdot) \in \mathcal{U}[0,T]$, the state equation (2.4.5) admits a unique strong solution $X(\cdot)$ [Yong and Zhou 1999] and the cost functional is well-defined. Now we can formulate a special case of the QLQ control problem.

Problem 2.4 (QLQ)**.** Find $\bar{u} \in \mathcal{U}[0,T]$ such that

$$J(\bar{u}) = \inf_{u \in \mathcal{U}[0,T]} J(u).$$

Next, we show that this optimal control problem has a solution. More importantly, in doing so we will be able to construct an optimal control explicitly via the "Four Step Scheme" [Ma et al. 1994].

Proposition 2.5. *A QLQ problem associated with the objective functional (2.4.4) and a state equation (2.4.5) that satisfies (2.4.6) admits an optimal control.*

Proof. Let's start by introducing a particular terminal backward SDE,

$$\begin{cases} dY(t) & = -\left[(e_f+e_p)X_1(t)X_2(t)+\frac{1}{2}mu^2(t)\right]dt+z(t)dW(t) \quad t\in[0,T] \\ Y(T) & = 0 \end{cases} \tag{2.4.7}$$

Note that for a given pair $(X_0, u(t)) \in \mathbb{R}^2 \times \mathcal{U}[0,T]$ this equation becomes a linear backward stochastic differential equation (BSDE). System (2.4.5) admits a unique strong solution, $X(\cdot) \equiv X(\cdot, x_0, u(\cdot))$. So, given $(X(\cdot), u(\cdot))$, equation (2.4.7) admits a unique adapted solution [Yong and Zhou 1999, Theorem 2.2 p.349], $(Y(\cdot), Z(\cdot)) \equiv (Y(\cdot; X_0, u(\cdot))), Z(\cdot; X_0, u(\cdot))),$ which depends on $(X_0, u(t))$ through $(X(\cdot), u(\cdot))$. Notice that

$$0 = Y(T) = Y(0) - \int_0^T \{(e_F+e_P)X_1(t)X_2(t)+\frac{1}{2}mu^2(t)\}dt + \int_0^T Z(t)dW(t). \tag{2.4.8}$$

Here, $Y(\cdot)$ is $\{\mathcal{F}_t\}_{t\geq 0}$-adapted, which means that each $Y(t)$ is $\mathcal{F}_t$-measurable. In particular, $Y(0)$ is $\mathcal{F}_0$-measurable. Since $\mathcal{F}_0$ is the trivial σ-field $\{\emptyset, \mathbb{R}^2\}$ augmented by all the P-null sets, $Y(0)$ is almost surely a constant (or equivalently, one has $\mathbb{E}[Y(0)\,|\,\mathcal{F}_0] = Y(0)$. Now, taking conditional expectation $\mathbb{E}[\cdot\,|\,\mathcal{F}_0]$ on the above, we have

$$\begin{aligned} 0 &= \mathbb{E}\{Y(0) - \textstyle\int_0^T\{(e_F+e_P)X_1(t)X_2(t)+\frac{1}{2}mu^2(t)\}dt + \int_0^T Z(t)dW(t) \;\big|\; \mathcal{F}_0\} \\ &= Y(0) - \mathbb{E}\textstyle\int_0^T\{(e_F+e_P)X_1(t)X_2(t)+\frac{1}{2}mu^2(t)\}dt \\ &= Y(0) - J(u). \end{aligned}$$

Thus,

$$Y(0) = J(u) = \mathbb{E}\int_0^T \{(e_F+e_P)X_1(t)X_2(t)+\frac{1}{2}mu^2(t)\}dt. \tag{2.4.9}$$

Now we can rewrite the cost functional as:

$$J(u(\cdot)) = Y(0; X_0, u(\cdot))). \tag{2.4.10}$$

Rewriting it in this way allows us to obtain an optimal control problem with cost functional

(2.4.10) and an FBSDE:

$$\begin{cases} dX(t) &= (\hat{b}(t,X(t)) + b_0(X(t))u(t))dt + \hat{\sigma}(t,X(t))dW(t) \\ dY(t) &= -\left[(e_f+e_p)X_1(t)X_2(t) + \frac{1}{2}mu^2(t)\right]dt + Z(t)dW(t) \\ Y(T) &= 0, \quad X(0) = X_0. \end{cases} \tag{2.4.11}$$

We conjecture that an optimal control can be found as a state feedback. If this is true, the system (2.4.11) will be a closed system that only involves $(X(\cdot), Y(\cdot), Z(\cdot))$. Using Ma et al. [1994], there exists an appropriate function $\theta(\cdot,\cdot)$ such that the following relationship holds:

$$Y(t) = \theta(t, X(t)), \quad t \in [0,T]. \tag{2.4.12}$$

Here $\theta(\cdot,\cdot)$ depends on the control $u(\cdot) \in \mathcal{U}[0,T]$ through the state $X(\cdot; X_0, u(\cdot))$ and we need to choose such a $u(\cdot)$ to minimize $Y(0)$. We apply Itô's formula to $\theta(t, X(t; Y_0, u(\cdot)))$ and with use of (2.4.11) we obtain:

$$dY(t) = -\left[(e_f+e_p)X_1X_2 + \frac{1}{2}mu^2(t)\right]dt + Z(t)dW(t)$$

which is equal to

$$\begin{aligned} d\theta(t,X(t)) = \{&\theta_t(t,X(t)) + < \theta_x(t,X(t)), b(t,X(t)) + b_0(t,X(t))u(t) > \\ &+ \frac{1}{2} < \theta_{xx}(t,X(t))\sigma(t,X(t)), \sigma(t,X(t)) > \}dt \\ &+ < \theta_x(t,X(t)), \sigma(t,X(t)) > dW(t). \end{aligned}$$

By equating corresponding drift terms and by suppressing the arguments, we obtain the following parabolic PDE:

$$\begin{aligned} 0 &= \theta_t + < \theta_x, b + b_0 u > + \frac{1}{2} < \theta_{xx}\sigma, \sigma > + (e_f+e_p)x_1x_2 + \frac{m}{2}u^2 \\ &= \theta_t + < \theta_x, b > + \frac{1}{2} < \theta_{xx}\sigma, \sigma > + b_0^T\theta_x u + (e_f+e_p)x_1x_2 + \frac{m}{2}u^2. \end{aligned}$$

Notice that for simplicity of notation we use lower case letters x_1, x_2 in the PDE since they are spatial variables here. By rearranging terms we can rewrite the above equation as a backward parabolic PDE with a terminal condition that $\theta(\cdot,\cdot)$ should satisfy:

$$\begin{cases} \theta_t + < \theta_x, b > + \frac{1}{2} < \theta_{xx}\sigma, \sigma > + (e_f + e_p)x_1x_2 - \frac{1}{2m}|b_0^T\theta_x|^2 \\ \quad + \frac{m}{2}[u + \frac{b_0^T\theta_x}{m}]^2 = 0 \\ \theta(T,x) = 0 \ \ x \in \mathbb{R}^2. \end{cases} \tag{2.4.13}$$

By a standard maximum principle for parabolic PDE's, the smallest $\theta(\cdot,\cdot)$ solution should be the solution of the following problem:

$$\begin{cases} \theta_t + < \theta_x, b > + \frac{1}{2} < \theta_{xx}\sigma, \sigma > + (e_f + e_p)x_1x_2 - \frac{1}{2m}|b_0^T\theta_x|^2 = 0 \\ \theta(T,x) = 0 \ \ x \in \mathbb{R}^2. \end{cases} \tag{2.4.14}$$

The classical solvability of a similar type of quasi-linear parabolic PDE with a quadratic gradient term is discussed in During and Jungel [2005], and thus this equation has a unique classical solution.

In this case we need to find $\bar{u}$ such that $\inf_{u\in\mathcal{U}} J(u) = J(\bar{u})$. Then using the θ PDE and the format of $\bar{u}$, we obtain

$$\begin{aligned} J(u) &= \mathbb{E}\left(\int_0^T (e_F + e_P)X_1(t)X_2(t) + \frac{1}{2}mu^2(t)dt\right) \\ &= \theta(0, x(0)) - \mathbb{E}\{\theta(T, x(T))\} + \\ &+ \mathbb{E}\int_0^T \theta_t + < \theta_x, b > + \frac{1}{2} < \theta_{xx}\sigma, \sigma > + (e_f + e_p)x_1x_2 \\ &- \frac{1}{2m}|b_0^T\theta_x|^2 + \frac{m}{2}\left[u + \frac{b_0^T\theta_x}{m}\right]^2 dt \\ &= \theta(0, x(0)) + \frac{m}{2}\mathbb{E}\int_0^T \left[u + \frac{b_0^T\theta_x}{m}\right]^2 dt \\ &\geq \theta(0, x(0)) = J(\bar{u}) \end{aligned}$$

and we conclude that $\bar{u} = -b_0^T\theta_x$ is an optimal control. □

Having explicitly obtained an optimal control we proceed to solve a closed-loop system of forward SDEs to obtain the corresponding optimal state trajectories

$$\begin{cases} dX(t) &= (\hat{b}(t, x(t)) - b_0(x(t))b_0^T(x(t))\theta_x(t, x(t)))dt + \hat{\sigma}(t, X(t))dW(t) \\ X(0) &= x_0 \end{cases}. \tag{2.4.15}$$

2.5 Numerical Solution

In order to obtain the optimal control and the optimal paths for population state variables, we need to solve the system of forward-backward system of equations (2.4.11). In Section 3 we have established that the "Four-Step Scheme" from Yong [2002] is applicable. Namely, we set up a PDE (2.4.14) that links forward variables $X(t)$ with backward variable $y(t)$ through some function $y(t) = \theta(t, X(t))$. And $\theta(t, X(t))$ satisfies equation (2.4.14) which is a quasi-linear parabolic PDE with a quadratic gradient term. From the proof in the last section, we know that the optimal control depends on θ through its spatial derivative via $\bar{u}(t, X) = b_0^T \theta_x$. Thus when we know a complete profile of θ we know the value of the optimal control at any value of $X(t)$. Then, we can proceed with solving the closed-loop forward equation (2.4.15).

To solve the PDE (2.4.14) we employ a finite difference method in time and space. We rewrite the equation (2.4.14) as $\theta_t = -F(x, \theta)$, where

$$F(x, \theta) = b^T \theta_x + \frac{1}{2}\sigma^t \theta_{xx} \sigma + (e_f + e_p) x_1 x_2 - \frac{1}{2m} |b_0^t \theta_x|^2. \tag{2.5.1}$$

Since we only need the profile of $\theta(t, x)$ for $0 \leq t \leq T$ and $x \in [0, 1]^2$, we can solve the PDE in a bounded spatial domain with appropriate zero-flux boundary conditions.

In this restricted domain, we choose a spacing Δx and discretize in the x_1 and x_2 with $(x_1(i), x_2(j))$ representing the (i, j) point in the discretization. We also chose a time step ΔT, with corresponding discrete time values t_k. Let $\theta_{i,j}^k$ represent an approximation to $\theta(t_k, x_1(i), x_2(j))$. Then we can use a centered difference to approximate $F((x_1(i), x_2(j)), \theta_{i,j}^k)$ by $F_{i,j}^k$ via

$$\begin{aligned}
F_{i,j}^k &= b_1 \frac{\theta_{i+1,j}^k - \theta_{i-1,j}^k}{2\Delta x} + b_2 \frac{\theta_{i,j+1}^k - \theta_{i,j+1}^k}{2\Delta x} \\
&+ \frac{1}{2}\sigma_1^2 \frac{\theta_{i+1,j}^k - 2\theta_{i,j}^k + \theta_{i-1,j}^k}{\Delta x^2} \\
&+ \sigma_1 \sigma_2 \frac{\theta_{i+1,j+1}^k - \theta_{i+1,j-1}^k - \theta_{i-1,j+1}^k + \theta_{i-1,j-1}^k}{4\Delta x^2} \\
&+ \frac{1}{2}\sigma_2^2 \frac{\theta_{i,j+1}^k - 2\theta_{i,j}^k + \theta_{i,j-1}^k}{\Delta x^2} \\
&+ (e_p + e_f) x_1(i) x_2(j) - \frac{1}{2m}(x_2(j)^2 \left[\frac{\theta_{i,j+1}^k - \theta_{i,j-1}^k}{2\Delta x}\right]^2 .
\end{aligned}$$

We continue the discretization, in time, using the Euler scheme. Since the PDE (2.4.14) is backward in time, we solve for $\theta_{i,j}^{k-1}$ obtaining

$$\theta_{i,j}^{k-1} = \theta_{i,j}^{k} + F_{i,j}^{k}\Delta t, \tag{2.5.2}$$

with $0 \leq i, j \leq l$ and $0 \leq k \leq n$. The terminal condition is set to $\theta_{i,j}^{n} = 0$ and for the boundary conditions, we set the boundary values of $\theta_{i,j}^{k},\ i, j = 0, l$ to the values of θ at the closest internal node.

Since we are using an explicit method (in time), we need to maintain the stability of the method by satisfying a $Courant - Friedrichs - Lewy$ (CFL) condition, which in our case, we set $\Delta t < \frac{1}{2}\Delta x^2$. For more complete information on how to set CFL conditions, we refer the reader to Richtmyer and Morton [1967]. After carrying out this numerical approximation we obtain an approximate profile of θ in the domain $[0, T] \times (0, 1)^2$. Now, we can reference correct values of θ whenever we need them.

The next step is to solve the forward SDE. We use the stochastic Euler scheme. Complete details on the stochastic Euler scheme forward in time along with information on the CFL condition can be found in Kloeden and Platen [1992]. Note that for the stochastic Euler scheme we need not only time increments Δt but also Wiener updates ΔW. Time stepping is the same as in the deterministic Euler scheme used above. Using a different random seed to generate Brownian motion produces a different sample path for each experiment. This allows us to observe different sample paths of population densities X_1, X_2 and u. After discretization in time the forward system of equations (with $X \in (0, 1)^2$) we get:

$$X^k = X^{k-1} + (b(X^{k-1}) - b_0(X^{k-1})b_0(X^{k-1})^T\theta_x^{k-1})\Delta t + \sigma(X^{k-1})\Delta W. \tag{2.5.3}$$

Each step of Euler approximation requires $\theta_x(t_k, X_1(t_k), X_2(t_k))$ to be updated. Because of the discretization we only have the values of θ_x at the specified nodes and they do not necessarily coincide with the values of $X_1(t_k), X_2(t_k)$ obtained with the approximation of the solution to (2.5.3). To overcome this difficulty we interpolate for $\theta_x(t_k, X_1(t_k), X_2(t_k))$ using known values of θ closest to it.

2.6 Numerical Results

A major contribution in this paper is the introduction of randomness in the population model and in the resulting stochastic control problem. We are interested in comparing results from the stochastic problem with those from the corresponding deterministic problem.

In the numerical simulations presented we choose the initial population densities to be $X_1(0) = 0.6$ (park population) and $X_2(0) = 0.5$ (forest population). Each sample path is generated with a different random seed.

To simulate the deterministic case we take our variance c to be 0. The results for this numerical experiment are given in Figure 2.1. This solution coincides to the deterministic case of Salinas et al. in their experiment when the same parameters are used.

In the next experiment we set the variance to a small value, $c = 0.1$. Figure 2.2 shows that the solutions for this value, keeping other parameters the same, did not change the dynamics of the population significantly compared to the deterministic case. We observed a similar decrease in the park population density and a slight increase of population density in the forest as expected. Clearly the the diffusion term does not contribute much to the change from the deterministic case. The harvesting rate in the deterministic and the stochastic case do not show much differences during the span of 5 years.

We wanted to see if setting the variance parameter to a higher value would cause more dramatic differences between deterministic and stochastic population densities. More importantly, we should see if the behavior of the optimal stochastic control is different compared to the deterministic optimal control.

For the next experiment we chose a larger value $c = 0.5$. In Figure 2.3 we can see the sample paths for this implementation. We observed that the shapes for sample paths of population in the park and forest are different from the previous two experiments. Larger parameter c indicates larger changes in the availability of the forage, which causes the populations densities change with larger fluctuations. We can also see more dramatic effects of a larger c in the control. There are spikes of harvesting activity during the years with high food availability and lower harvesting for the years with low food availability. Overall, the behavior of the control resembles the behavior of the control during the previous two implementations. It is also important that the terminal population densities for the park and forest are different from the deterministic case by a larger margin than in the previous experiment. The final population density in the forest is approximately 0.7 which is significantly higher than approximately 0.57 in the deterministic case. So, higher variance in the experiments can pull the population densities much higher or much lower than the experiments with lower variance. The deviations of stochastic optimal control from deterministic optimal control are smaller than deviations of the stochastic population densities from deterministic population densities.

For the next experiment we did 10 consecutive runs of the code continuously increasing the variance by a 0.1 increment and then graphed all of them on the same set of axis. It is displayed in Figure 2.4. This figure supports our previous observations that higher variance parameter pulls the population further away from the deterministic trajectory. The reader can clearly see the differences in sample paths and the role that variance plays in each case.

For the last experiment we used a fixed variance. It is displayed in Figure 2.5. The difference in sample paths is solely caused by the stochastic aspect of our model. We can see that the model is consistent and produce similar results for the different sample paths.

2.7 Conclusions

The contributions of this chapter were to demonstrate optimal control of a stochastic metapopulation model, to justify the procedure mathematically and to illustrate the results numerically. We formulated a stochastic control problem for a metapopulation model with a stochastic growth rate for the populations in the park and in the forest. The existence of an optimal control required a nonstandard approach due to the nonlinearity of the state variable terms in the model. Solving an appropriate stochastic QLQ problem enables us to obtain this existence result.

A rigorous treatment of such an optimal control problem for an ecological model is quite novel. The tools used here, especially the stochastic QLQ technique, can be applicable to many other ecological problems.

As the result of our simulations we have shown that the stochastic population trajectories can differ dramatically from its deterministic counterparts. Most significant differences are displayed when the variance in mast availability is high. A manager who is adjusting the harvest levels might need to raise the harvest level given in the deterministic optimal control case to be higher to give effectiveness due to the possible randomness in mast. We realize that the amount of stochastic variation in this problem may not realistically represent the changes of mast in the motivating example of bears in the Great Smoky Mountains National Park. The changes in the mast in this park may not occur as frequently under normal conditions as in this stochastic model. But our results do indicate that the randomness in the mast should be taken into account when managing the harvesting of the bears in the nearby forests, since the level of variance strongly affect the optimal harvesting control.

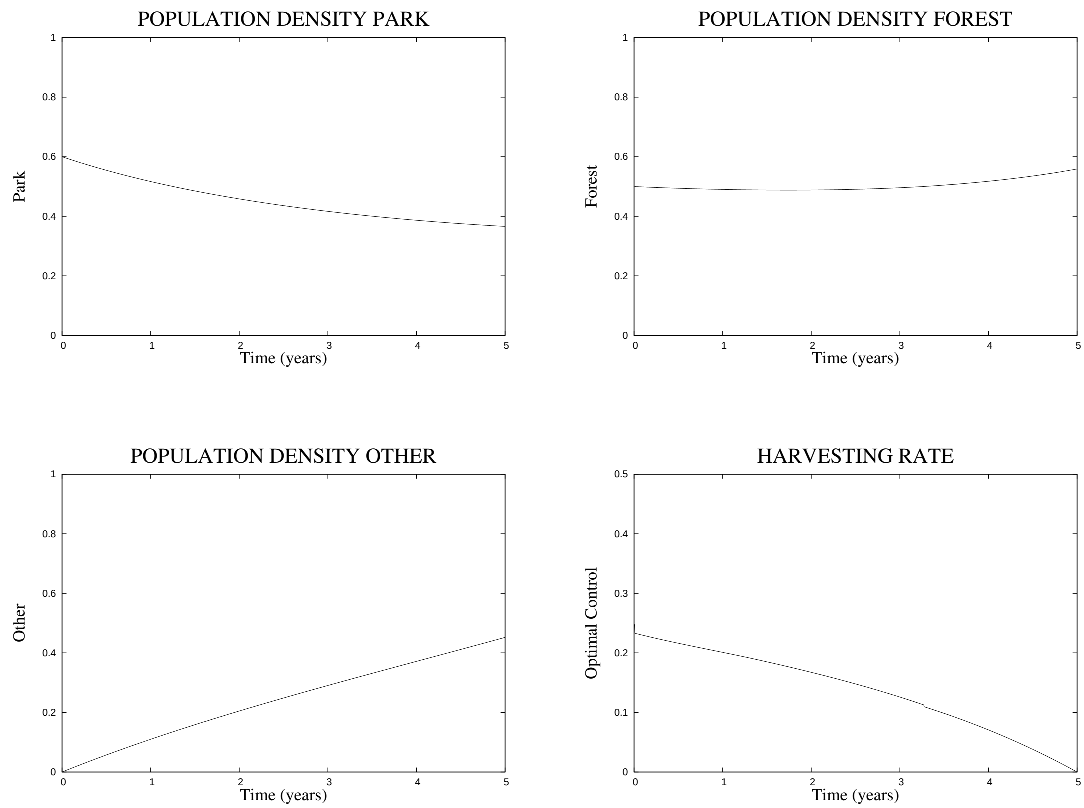

Figure 2.1: Population results of optimal control problem in the deterministic case, i.e. with no randomness.

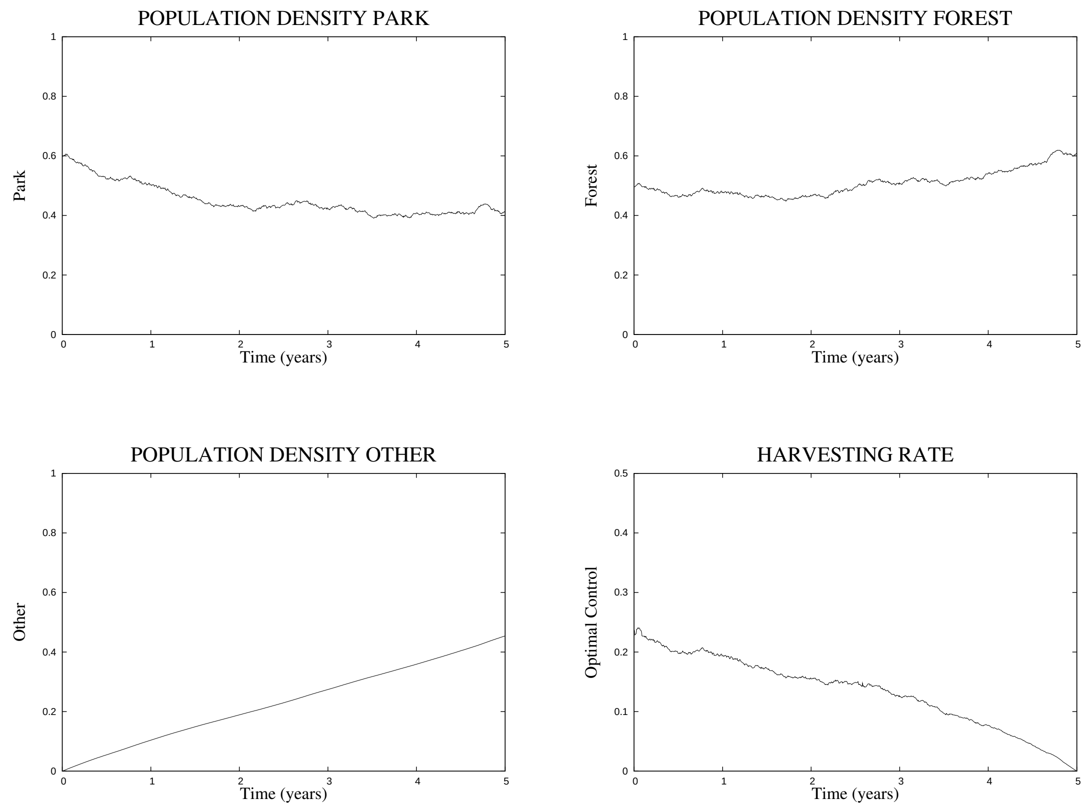

Figure 2.2: Population results of the optimal control problem using one sample path and variance of the population growth $c = 0.1$.

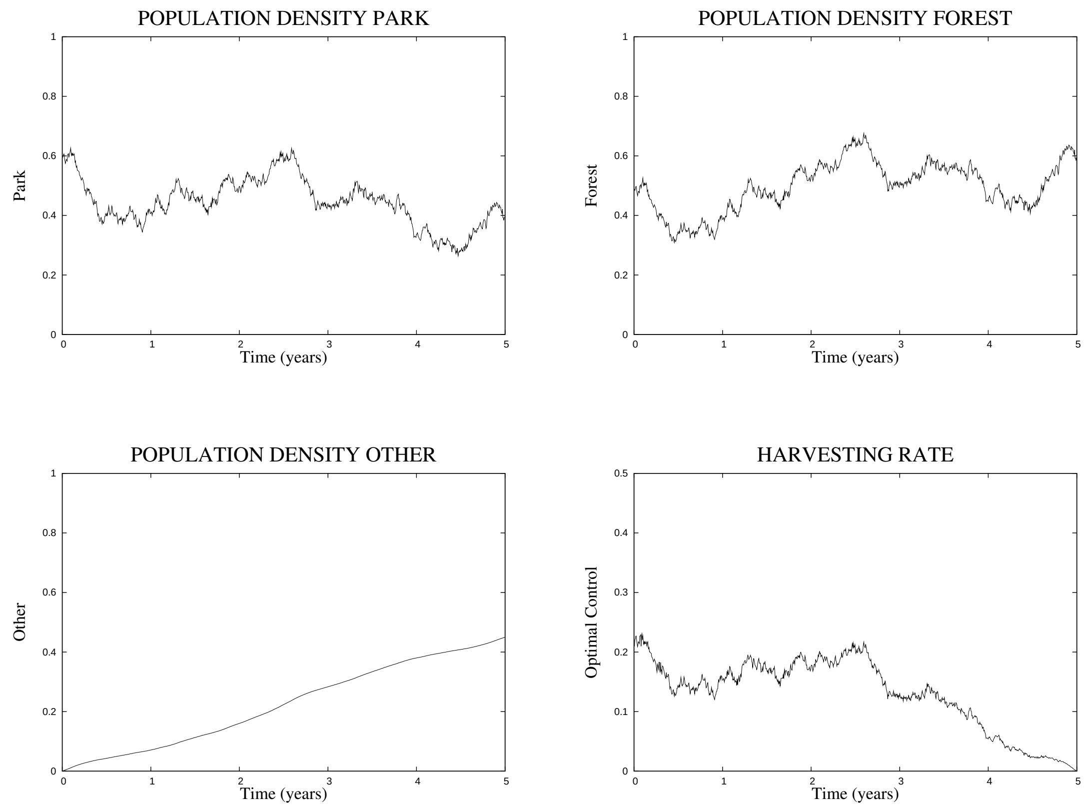

Figure 2.3: Population results of the optimal control problem using one sample path and the variance of the population growth $c = 0.5$.

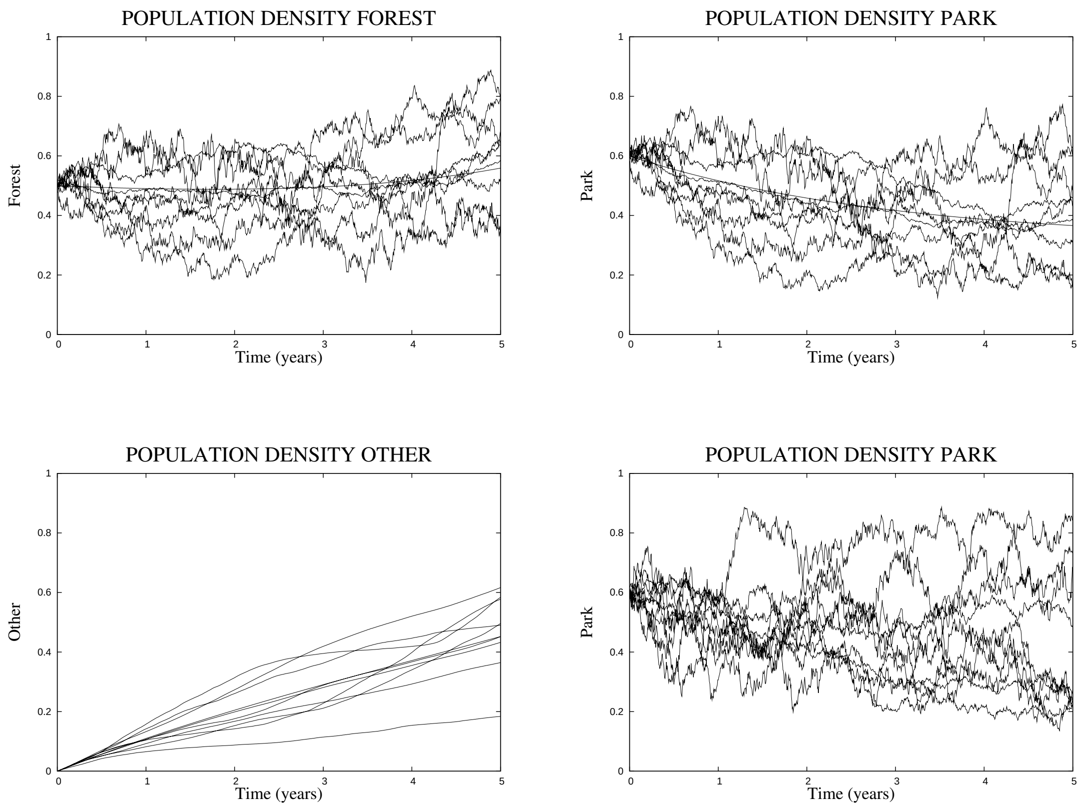

Figure 2.4: Population results of the optimal control problem with the variance of the mast availability c gradually increased from 0 to 0.9 by 0.1.

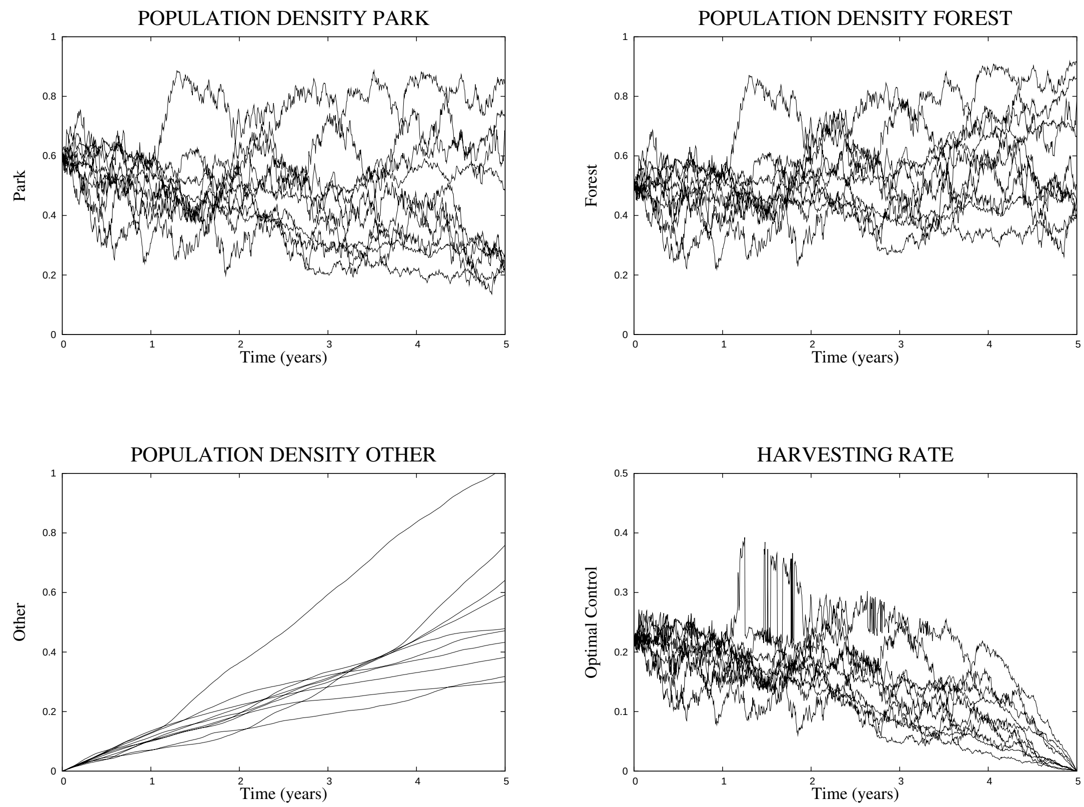

Figure 2.5: Ten sample paths of the population and the harvest rates when variance of the mast availability is $c = .25$.

Chapter 3

Forward-Backward SDE in Stochastic Optimal Control Theory. Existence and Uniqueness of the Solution. Numerical Scheme.

3.1 Introduction

In stochastic optimal control one often needs to solve a problem where coefficients of the state and adjoint equations are not necessarily Lipschitz. The approach developed by Ma et al. [1994] is limited in that sense. In this work we attempt to relax this restriction for a particular type of a forward-backward system of SDEs. Note that the type of FBSDEs that emerge from stochastic optimal control problems do not have necessary Lipschitz continuity conditions on the backward drift coefficient. In this work we make an attempt to relax this restriction for a particular type of the Forward-Backward system of SDEs.

Numerical methods for approximating solutions of this kind of systems of FBSDEs is limited. In their work Bender and Zhang [2008] introduce a Markovian type iteration method. This method appears to be more efficient for solving high dimensional systems of FBSDEs. In this work we provide adaptation of their method to the type of problems considered in stochastic optimal control theory.

The rest of this chapter is structured in the following way. In Section 3.2 we provide some preliminary information. Section 3.3 provides results for solvability of systems of FBSDEs where the drift coefficient of the backward equation is not Lipschitz. Section 3.4 formulates and solves a stochastic optimal control problem numerically.

3.2 Formulation and preliminaries

In the field of stochastic optimal control we often need to find an optimal control that is associated with the following optimal control problem. Consider an objective functional

$$J(u) = \mathbb{E}\left\{\int_0^T g(X(t), u(t))dt + h(X(T))\right\} \tag{3.2.1}$$

along with the state equations in the form of the system of forward SDEs

$$\begin{cases} dX(t) = k(X(t), u(t))dt + \sigma(X(t))dW_t \\ X(0) = x_0. \end{cases} \tag{3.2.2}$$

For the deterministic equivalent of this problem we would proceed by establishing existence and uniqueness of the optimal control explicitly. We would formulate the Hamiltonian, and use the Pontryagin Maximum Principle to determine the optimal control and solve for the optimal state trajectory.

Similarly if we can establish existence and uniqueness of the stochastic optimal control explicitly, which is usually done on case by case basis, we can proceed with formulation of the Hamiltonian and application of a stochastic version of the Pontryagin Maximum Principle. Such problems were studied in Yong and Zhou [1999]. For the above stochastic optimal control problem we can formulate the Hamiltonian in general form

$$\mathcal{H}(t, x, u, p, q) = < p, k(x, u) > +tr[q^*\sigma(x)] + g(x, u). \tag{3.2.3}$$

where $p(t)$ and $q(t)$ is a pair of adjoint variables. They satisfy the following BSDE

$$\begin{cases} dp(t) = -\{k_x(x, u)^*p(t) + \sum_{j=1}^m \sigma_x^j(x)^*q_j(t) - g_x(x, u)\}dt + q(t)dW_t \\ p(T) = -h(x(t)). \end{cases} \tag{3.2.4}$$

The Pontryagin-type Stochastic Maximum Principle says that if there exists an optimal pair $(\bar{x}, \bar{u})$ and a pair of processes $(p(t), q(t))$ satisfying system (3.2.4), then $\mathcal{H}(t, \bar{x}, \bar{u}) = \max_{u\in\mathcal{U}} \mathcal{H}(t, \bar{x}, u)$. Thus we obtain $\bar{u} = g(\bar{x}, p(t), q(t))$. Substituting $g(\bar{x}, p(t), q(t))$ into equations (3.2.2) and (3.2.4) we obtain the following system of forward-backward SDEs:

$$\begin{cases} dX(t) = b(t, X(t), p(t), q(t))dt + \sigma(t, X(t))dW_t \\ dp(t) = -\{b_x(t, X(t), p(t), q(t))^* p(t) + \sum_{j=1}^m \sigma_x^j(x)^* q_j(t) - f_x(X(t), p(t), q(t))\}dt + q(t)dW_t \\ X_0 = X(0) \quad p(T) = -h_x(X(T)) \end{cases} \tag{3.2.5}$$

where $(\Omega, \mathcal{F}, P)$ is a probability space carrying a standard m-dimensional Brownian motion W and $\mathcal{F}_t$ is the σ-algebra generated by W with X, p in $\mathbb{R}^n$ and q in $\mathbb{R}^{n\times m}$. We can assume that $b(t, X(t), p(t), q(t))$ and $\sigma(t, X(t))$ are Lipschitz continuous. But we notice that the drift coefficient

$$-\{b_x(t, X(t), p(t), q(t))^* p(t) + \sum_{j=1}^m \sigma_x^j(x)^* q_j(t) - f_x(X(t), p(t), q(t))\}$$

in the backward equation is not Lipschitz continuous under normal conditions. In fact, it is Lipschitz continuous only if b_x, σ_x are constants and $f_x(X, p, q)$ is at most linear in X, p and q (i.e. $f_x(X, p, q) = k_1 X + k_2 p + k_3 q$). Otherwise the classical result from Ma et al. [1994] does not apply.

In the next section we study the solvability of the forward-backward stochastic differential equation (3.2.5) when the drift coefficient of the backward stochastic differential equation is not Lipschitz.

3.3 Solvability of FBSDE

In their original article Ma et al. [1994] study the solvability of Forward-Backward Stochastic Differential equations explicitly via solving corresponding parabolic PDE. They proposed the so called "four-step" scheme."

1. Find $z(t, x, y, u) = u\sigma(t, x, y, z(t, x, y, u))$, s.t. $\forall (t, x, y, u) \in [0, T] \times \mathbb{R}^n \times \mathbb{R}^k \times \mathbb{R}^{k\times n}$
2. Use function z obtained above to solve the following parabolic system of PDE for $\theta(t, x)$:

$$\begin{cases} \theta_t^l + \frac{1}{2} tr[\theta_{xx}^l (\sigma\sigma^T)(t, x, \theta, z(t, x, \theta, \theta_x))] + < b(t, x, \theta, z(t, x, \theta, \theta_x)), \theta_x^l > \\ \quad -h^l(t, x, \theta, z(t, x, \theta, \theta_x)) = 0, \ (t, x) \in (0, T) \times \mathbb{R}^n, \ 1 \le l \le k, \\ \theta(T, x) = g(x), \ x \in \mathbb{R}^n. \end{cases}$$

3. Use θ and z obtained in Steps 1-2 to solve the following forward SDE:

$$X(t) = x + \int_0^t \tilde{b}(s, X(s))ds + \int_0^t \sigma(s, X(s))dW(s)$$

where

$$\begin{cases} \tilde{b}(t,x) = b(t,x,\theta,z(t,x,\theta,\theta_x)), \\ \tilde{\sigma}(t,x) = \sigma(t,x,\theta,z(t,x,\theta,\theta_x)). \end{cases}$$

4. Set

$$\begin{cases} p(t) = \theta(t, X(t)), \\ q(t) = z(t, X, \theta(t, X), \theta_x(t, X)). \end{cases}$$

A general drawback is that, to the best of our knowledge, this method is applicable to only a very small class of forward-backward SDEs. The authors were able to apply this result to only specific FBSDEs with drift and diffusion coefficients that are Lipschitz continuous in both the forward and backward equations. Thus it is not applicable to the forward-backward systems of SDEs that evolve from the class of stochastic optimal control problems discussed above. Our goal is to show that with minor modifications, a similar approach can be applied to the system of FBSDEs where the drift coefficient of the backward equation is not Lipschitz, in other words the stochastic optimal control problems of the type that we have formulated previously.

3.3.1 One-Dimensional Case

In this section we are going to use the following notation:

- $(\Omega \times \mathbb{T}) \equiv ([-\mathbb{R}, \mathbb{R}] \times [0, T])$ is a bounded domain.
- $\partial\Omega$ is a spatial boundary.

First we assume that $X(t), p(t)$ are in $\mathbb{R}$ and $q(t)$ is in $\mathbb{R}^m$. Then the forward-backward system of SDEs (3.2.5) becomes a system with 1-dimensional forward and backward equations. Using Feynman-Kac-type formulae that is discussed in Ma and Yong [1995] we know that if there exists a function $p(t) = \theta(t, X)$ such that $\theta(t, X)$ satisfies a particular parabolic PDE then the solution to the system (3.2.5) can be expressed through a classical solution to that parabolic PDE. So, let

$$p(t) = \theta(t, X(t)) \tag{3.3.6}$$

where $\theta(t, X)$ is some continuous and differentiable function in terms of time variable t and is twice differentiable in terms of space variable X. By applying Itô's formula to the equation (3.3.6) we obtain:

$$\begin{aligned} dp(t) &= d\theta(t, X(t)) \\ &= \{\theta_t + \tfrac{1}{2}\sigma^2(X(t))\theta_{xx} + b(t, X(t), p(t), q(t))\theta_x\}dt + \theta_x\sigma(X(t))dW. \end{aligned} \tag{3.3.7}$$

Combining this with the backward equation in system 3.2.5 we obtain a parabolic PDE:

$$\begin{cases} \theta_t + \frac{1}{2}\sigma^2(x)\theta_{xx} + b(x, \theta, \theta_x)\theta_x - b_x(x, \theta, \theta_x)\theta + \sigma_x(x)\sigma(x)\theta_x = f_x(x, \theta, \theta_x) \\ \theta(T, x) = -h_x(x(T)). \end{cases} \tag{3.3.8}$$

Note that we need to obtain the solution of the PDE (3.3.8) globally in space. Before we obtain that result for the unbounded domain we are going to solve the problem on the bounded domain. Let

$$a_0 = \tfrac{1}{2}\sigma^2(x)$$
$$a_1 = b(x, \theta, \theta_x)\theta_x - b_x(x, \theta, \theta_x)\theta + \sigma_x(x)\sigma(x)\theta_x - f_x(x, \theta, \theta_x).$$

Let us consider an approximation of (3.3.8). We state a terminal-boundary value problem in a bounded region $(x, t) \in (\Omega \times \mathbb{T})$.

$$\begin{cases} \theta_t + a_0(x)\theta_{xx} + a_1(x, \theta, \theta_x) = 0 \\ \theta(t, x) = -h_x(x) \ \text{ where } \ (x, t) \in (\partial\Omega \times T) \cup (\Omega \times T)\} \end{cases} \tag{3.3.9}$$

where T is the terminal time. For the equation (3.3.9) we can show that there exists a unique classical solution using results from Ladyženskaja et al. [1968]. Before we apply their result we put some restrictions on a_0 and a_1. These coefficients mustsatisfy the following assumptions:

Assumption 3.1. *For all $x \in \mathbb{R}$ $\nu < \frac{1}{2}\sigma^2(x) < \mu$ where ν, μ are positive constants.*

Assumption 3.2. *The functions b, σ, and f are Lipschitz continuous with the Lipschitz constant L, and $|b(x, \theta, p)| \leq k(1 + |p|)(1 + |\theta|)$.*

Assumption 3.3. *All second order partial derivatives of b, σ, and f are bounded by a constant M.*

Assumption 3.4. *We have* $h_x(x) \in C^{2+\alpha}$ *and* $|h_x(x)| \leq K$ $\forall x \in \mathbb{R}, \alpha \in (0,1)$, *where* K *does not explicitly depend on* x, *and* $\lim_{|x|\to\infty} h_x(x) = 0$.

The next Theorem follows directly from Ladyženskaja et al. [1968, Theorem 2.9 p.23]. We will state it here without proof.

Theorem 3.5. *Let* $\theta(t,x)$ *be a classical solution of equation (3.3.9) in* Q_T. *Suppose that the functions* a_0 *and* a_1 *take finite values for any finite* θ, θ_x, *and* $(x,t) \in \bar{Q}_T$ *and that for* $(x,t) \in Q_T$ *and arbitrary* θ

$$\begin{gathered} a_0(x) \geq 0 \\ \theta a_1(x,\theta,0) \geq -k_1\theta^2 - k_2 \end{gathered} \tag{3.3.10}$$

where k_1, k_2 *are nonnegative constants. Then*

$$\max_{Q_T} |\theta(t,x)| \leq \inf_{\lambda > k_1} e^{\lambda T} \left[\max_{\Gamma_T} |\theta| + \sqrt{\frac{k_2}{\lambda - k_1}} \right]. \tag{3.3.11}$$

.

We also need the following version of theorem from Ladyženskaja et al. [1968, Theorem 4.1 p. 558]. It will be utilized as a lemma in this work.

Lemma 3.6. *Suppose the following conditions hold:*

1. $(x,t) \in \bar{Q}_T$ *and arbitrary* θ *satisfying conditions (3.3.10) of Theorem 3.5.*

2. *For* $(x,t) \in \bar{Q}_T$, $|\theta| \leq M$ *(where* M *is from the condition (3.3.11), and arbitrary* p *the functions* $a_0(x)$ *and* $a_1(x,\theta,p)$ *are continuous and differential with respect to* x, θ, p *and the following inequalities hold with a sufficiently small* ϵ *determined by the numbers* M, ν, μ, μ_1.

$$\begin{gathered} \nu \leq a_0(x) \leq \mu \;\; \text{where} \;\; \nu, \mu > 0 \\ \left|\tfrac{\partial a_0}{\partial p}\right| (1+|p|)^3 + |a_1| + \left|\tfrac{\partial a_1}{\partial p}\right| (1+|p|) \leq \mu_1 (1+|p|)^2 \\ \left|\tfrac{\partial a_0}{\partial x}\right| (1+|p|)^2 + \left|\tfrac{\partial a_1}{\partial p}\right| \leq [\epsilon + P(|p|)](1+|p|)^3, \;\; \epsilon > 0 \\ -\tfrac{\partial a_1}{\partial \theta} \leq [\epsilon + P(|p|)](1+|p|)^2 \end{gathered}$$

   where $P(\rho)$ *is a nonnegative continuous function that tends to zero as* $\rho \to \infty$.

3. *For* $(x,t) \in \bar{Q}_T$, $|\theta| \leq M$, *and arbitrary* $|p| \leq M_1$, *the functions* $a_0(x)$ *and* $a_1(x,\theta,p)$ *are continuously differentiable with respect to all of their arguments.*

Then (3.3.9) admits a unique classical solution $\theta(t,x)$.

Now we are going to show that equation (3.3.9) under Assumptions 3.1–3.4 satisfies conditions of Lemma 3.6. Let functions a_0 and a_1 be finite when θ, θ_x are finite. Then the first condition of the Theorem 3.5 is satisfied under Assumption 3.1. Using Assumption 3.2, it follows that

$$\begin{aligned}\theta a_1(x,\theta,0) &= \theta\left\{b(x,\theta,0)\cdot 0 - b_x(x,\theta,0)\theta + \sigma_x\sigma\cdot 0 - f_x(x,\theta,0)\right\}\\ &= -b_x(x,\theta,0)\theta^2 - f_x(x,\theta,0)\theta\\ &\geq -L\theta^2 - L\theta\\ &\geq -L_1\theta^2 - L_1\end{aligned}$$

and thus satisfies the second condition of the Theorem 3.5. Due to Assumption 3.4, we conclude that for equation (3.3.9) we have $\max_{Q_t}|\theta(t,x)| \leq M_1$, where M_1 is a constant that depends on M, L_1, L_2, λ. So, the first hypothesis is satisfied, and we can assume that $|\theta|$ is a priori bounded. In order to verify the second hypothesis of the above Lemma we need to verify that all of the inequalities hold for equation (3.3.9). The first inequality is satisfied because of Assumption 3.1. The second inequality is also satisfied using Assumptions 3.2 and 3.3. Since

$$\begin{aligned}\left|\frac{\partial a_1}{\partial p}\right| &= |b + \sigma_x\sigma - f_{xp}|\\ &\leq |b| + |\sigma_x||\sigma| + |f_{xp}|\\ &\leq k(1+|p|) + L + M\\ &\leq K(1+|p|).\end{aligned}$$

Also

$$\begin{aligned}|a_1| &= |bp - b_x\theta + \sigma_x\sigma p - f_x|\\ &\leq |b||p| + |b_x||\theta| + |\sigma_x||\sigma||p| + |f_x|\\ &\leq k(1+|p|)|p| + KM + L|p| + M\\ &\leq K(1+|p|)^2. \qquad (3.3.12)\end{aligned}$$

Combining two parts above and choosing $P(|p|) = \frac{1}{1+|p|}$, we get the desired second inequality of Lemma 3.6. The third inequality is also satisfied since

$$\begin{aligned}
-\frac{\partial a_1}{\partial \theta} &= -b_\theta \theta_x - b_{x\theta}\theta - b_x - f_{x\theta} \\
&\leq |b_\theta||p| + |b_{x\theta}||\theta| + |b_x| + |f_{x\theta}| \\
&\leq K|p| + K \leq K(1+|p|)
\end{aligned}$$

as above. Choose $P(|p|) = \frac{1}{1+|p|}$ and thus

$$-\frac{\partial a_1}{\partial \theta} \leq (\epsilon + P(|p|))(1+|p|)^2.$$

Therefore Lemma 3.6 is satisfied and we have a unique classical solution of equation (3.3.9).

Note that from Assumption 3.4 and the proof of Theorem 4.1 in Ladyženskaja et al. [1968, p. 558] it follows that $\theta(t,x)$ and its partial derivatives $\theta_t(t,x)$, $\theta_{tt}(t,x)$, $\theta_x(t,x)$, $\theta_{xx}(t,x)$, and $\theta_{xxx}(t,x)$ are all uniformly bounded $\forall\ R > 0$. We will use this fact in the proof of the next proposition.

Remember that we want to obtain the solution of equation (3.3.8). The above Theorem 3.5 and Lemma 3.6 allow us to do so. Let us prove the following proposition.

Proposition 3.7. *Under Assumptions 3.1–3.4 equation (3.3.8) admits a unique classical solution $\theta(t,x)$. Moreover $\theta(t,x)$ and $\theta_x(t,x)$ are bounded.*

Proof. Construct a sequence of parabolic PDE's of type (3.3.9) on a bounded domain $\Omega \times \mathbb{T} = [-R_i, R_i] \times [0,T]$, where R_i is an increasing sequence. By Lemma 3.6 there exists a sequence of classical solutions $\{\theta^i(t,x),\ i = 1,2,...\}$ to the sequence of equations above. Thus $\theta^i(t,x)$ possess all the necessary derivatives. Recall that all partial derivatives are also uniformly bounded $\forall\ R_i > 0$. Let $\{\theta^i(t,x)\}_{i=1,2,...} \equiv \mathbb{F}_1$. Since $\theta^i(t,x)$ are uniformly bounded by a constant that does not depend on x, $\mathbb{F}_1$ is equicontinuous. By the Arzela-Ascoli Theorem there exist a subsequence $\theta^{i,k}(t,x),\ k = 1,2,3...$ that is uniformly convergent to $\theta(t,x)$. Similarly we can construct $\{\theta^i_t(t,x)\}_{i=1,2,...} \equiv \mathbb{F}_2$, $\{\theta^i_x(t,x)\}_{i=1,2,...} \equiv \mathbb{F}_3$, and $\{\theta^i_{xx}(t,x)\}_{i=1,2,...} \equiv \mathbb{F}_4$, which are all equicontinuous as well. By the Arzela-Ascoli Theorem,

$$\begin{aligned}
\theta^{i,k}_t(t,x) &\to \theta_t(t,x) \ \text{ uniformly as } i,k \to \infty \\
\theta^{i,k}_x(t,x) &\to \theta_x(t,x) \ \text{ uniformly as } i,k \to \infty \\
\theta^{i,k}_{xx}(t,x) &\to \theta_{xx}(t,x) \ \text{ uniformly as } i,k \to \infty\ .
\end{aligned}$$

Thus $\theta(t,x)$ is a classical solution of parabolic PDE (3.3.8). Moreover $\theta(t,x)$ and $\theta_x(t,x)$ are bounded.

Now that we have existence we need to show the uniqueness of the solution to equation (3.3.6). Notice that under Assumptions 3.1 and 3.2 the term a_1 of equation (3.3.6) that contains the low term derivatives is Lipschitz continuous. Notice

$$\begin{aligned}|a_1(x,\theta,\theta_x) - a_1(x,\tilde{\theta},\tilde{\theta}_x)| &\le |b(x,\theta,\theta_x)\theta_x - b(x,\tilde{\theta},\tilde{\theta}_x)\tilde{\theta}_x| \\ &+ |b_x(x,\theta,\theta_x)\theta - b_x(x,\tilde{\theta},\tilde{\theta}_x)\tilde{\theta}| \\ &+ |\sigma_x\sigma\theta_x - \sigma_x\sigma\tilde{\theta}_x| \\ &+ |f_x(x,\theta,\theta_x) - f_x(x,\tilde{\theta},\tilde{\theta}_x)| \\ &:= I_1 + I_2 + I_3 + I_4\end{aligned}$$

where

$$\begin{aligned}I_1 &= |b(x,\theta,\theta_x)\theta_x - b(x,\tilde{\theta},\tilde{\theta}_x)\tilde{\theta}_x| \\ &\le |b(x,\theta,\theta_x)\theta_x - b(x,\theta,\theta_x)\tilde{\theta}_x| + |b(x,\theta,\theta_x)\tilde{\theta}_x - b(x,\tilde{\theta},\tilde{\theta}_x)\tilde{\theta}_x| \\ &\le |b(x,\theta,\theta_x)||\theta - \tilde{\theta}| + |\tilde{\theta}_x||b(x,\theta,\theta_x) - b(x,\tilde{\theta},\tilde{\theta}_x)|.\end{aligned} \tag{3.3.13}$$

Since $|\theta|$ and $|\theta_x|$ are bounded, b is Lipschitz and has linear growth in θ and θ_x and have

$$I_1 \le K(|\theta - \tilde{\theta}| + |\theta_x - \tilde{\theta}_x|).$$

Similarly,

$$\begin{aligned}I_2 &\le |b_x(x,\theta,\theta_x)||\theta - \tilde{\theta}| + |\tilde{\theta}||b_x(x,\theta,\theta_x) - b_x(x,\tilde{\theta},\tilde{\theta}_x)| \\ &\le K(|\theta - \tilde{\theta}|).\end{aligned}$$

Because of Assumption 3.1,

$$I_3 \le K(|\theta_x - \tilde{\theta}_x|)$$

and due to $f(x,\theta,\theta_x)$ being Lipschitz we get

$$I_4 \le K$$

by letting the generic constant K absorb the constant term from estimation of I_4. So,

$$a_1(x, \theta, \theta_x) \leq K(|\theta - \tilde{\theta}| + |\theta_x - \tilde{\theta}_x|$$

and $a_1(x, \theta, \theta_x)$ is Lipschitz continuous. We rewrite equation (3.3.6) in divergence form, and for simplicity of notation we let $g(x) = -h_x(x)$:

$$\begin{cases} -\theta_t - (\frac{1}{2}\sigma_x^2)_x + \sigma_x\sigma\theta_x = \{b(x, \theta, \theta_x)\theta_x - b_x(x, \theta, \theta_x)\theta + \sigma_x\sigma\theta_x - f_x(x, \theta, \theta_x)\} \\ \theta(T, x) = g(x). \end{cases} \tag{3.3.14}$$

Notice the $\sigma_x\sigma\theta_x$ term cancels out and equations (3.3.14) becomes:

$$\begin{cases} -\theta_t - (\frac{1}{2}\sigma^2\theta_x)_x = \{b(x, \theta, \theta_x)\theta_x - b_x(x, \theta, \theta_x)\theta - f_x(x, \theta, \theta_x)\} \\ \theta(T, x) = g(x). \end{cases} \tag{3.3.15}$$

Consider two solutions θ and $\hat{\theta}$ to equation (3.3.15). Let $h = \theta e^{\lambda t}$ and $\hat{h} = \hat{\theta} e^{\lambda t}$ where $\lambda > 0$. Note that

$$\begin{aligned} h_t &= \theta_t e^{\lambda t} + \lambda\theta e^{\lambda t} = \theta_t e^{\lambda t} + \lambda h, \ \hat{h}_t = \hat{\theta}_t e^{\lambda t} + \lambda\hat{h} \\ h_x &= \theta_x e^{\lambda t}, \ \hat{h}_x = \hat{\theta}_x e^{\lambda t} \\ h_{xx} &= \theta_{xx} e^{\lambda t}, \ \hat{h}_{xx} = \hat{\theta}_x e^{\lambda t}. \end{aligned}$$

Multiply equation (3.3.15) by $e^{\lambda t}$ to obtain:

$$\begin{cases} -\theta_t e^{\lambda t} - (\frac{1}{2}\sigma^2\theta_x)_x e^{\lambda t} = e^{\lambda t}\{b(x, \theta, \theta_x)\theta_x - b_x(x, \theta, \theta_x)\theta - f_x(x, \theta, \theta_x)\} \\ \theta(T, x)e^{\lambda T} = g(x)e^{\lambda T}. \end{cases} \tag{3.3.16}$$

Substitute h_t, h_x, h_{xx} into equation (3.3.16) to transform it into:

$$\begin{cases} -h_t + \lambda h - (\frac{1}{2}\sigma^2 h_x)_x & = \{b(x, he^{-\lambda t}, h_x e^{-\lambda t})h_x - b_x(x, he^{-\lambda t}, h_x e^{-\lambda t})h \\ & \quad -f_x(x, he^{-\lambda t}, h_x e^{-\lambda t})\} \\ h(T, x) = g(x)e^{\lambda T}. \end{cases} \tag{3.3.17}$$

Equation for $\hat{h}$ is defined similarly. We consider the difference $h - \hat{h}$, which satisfies the following PDE:

$$\begin{cases} -(h-\hat{h})_t + \lambda(h-\hat{h}) - (\frac{1}{2}\sigma^2(h-\hat{h})_x)_x & = \{b(x, he^{-\lambda t}, h_x e^{-\lambda t})h_x - b(x, \hat{h}e^{-\lambda t}, \hat{h}_x e^{-\lambda t})\hat{h}_x \\ & -b_x(x, he^{-\lambda t}, h_x e^{-\lambda t})h + b_x(x, \hat{h}e^{-\lambda t}, \hat{h}_x e^{-\lambda t})\hat{h} \\ & -f_x(x, he^{-\lambda t}, h_x e^{-\lambda t}) + f_x(x, \hat{h}e^{-\lambda t}, \hat{h}_x e^{-\lambda t})\} \\ h(T,x) - \hat{h}(T,x) = 0. \end{cases} \tag{3.3.18}$$

Multiply both sides of equation (3.3.18) by $(h-\hat{h})$ and integrate with respect to time and space. The LHS of the equation (3.3.18) becomes:

$$\begin{aligned} LHS &= \int_t^T \int_{\mathbb{R}} \left[-(h-\hat{h})_s(h-\hat{h}) + \lambda(h-\hat{h})^2 - \left(\frac{1}{2}\sigma^2(h-\hat{h})_x\right)_x (h-\hat{h}) \right] dxds \\ &= \int_{\mathbb{R}} \left[\frac{1}{2}(h-\hat{h})^2\right] dx + \int_t^T \int_{\mathbb{R}} \left[\lambda(h-\hat{h})^2 - \left(\frac{1}{2}\sigma^2(h-\hat{h})_x\right)_x (h-\hat{h}) \right] dxds \\ &\geq \lambda \int_t^T \int_{\mathbb{R}} (h-\hat{h})^2 dxds - \int_t^T \int_{\mathbb{R}} \left(\frac{1}{2}\sigma^2(h-\hat{h})_x\right)_x (h-\hat{h}) dxds. \end{aligned}$$

Use integration by parts with respect to x-variable on the second integral to get

$$\begin{aligned} LHS \geq\ & \lambda \int_t^T \int_{\mathbb{R}} (h-\hat{h})^2 dxds - \int_t^T \lim_{x\to\infty} \frac{1}{2}\sigma^2(h-\hat{h})_x(h-\hat{h}) \\ & - \lim_{x\to-\infty} \frac{1}{2}\sigma^2(h-\hat{h})_x(h-\hat{h}) ds + \int_t^T \int_{\mathbb{R}} \frac{1}{2}\sigma^2 \left((h-\hat{h})_x\right)^2 dxds. \end{aligned}$$

Since $\lim_{x\to\pm\infty} \frac{1}{2}(h-\hat{h})_x(h-\hat{h}) = 0$ and $\frac{1}{2}\sigma^2 > \mu$, we get

$$LHS \geq \lambda \int_t^T \int_{\mathbb{R}} (h-\hat{h})^2 dxds + \mu \int_t^T \int_{\mathbb{R}} \left((h-\hat{h})_x\right)^2 dxds.$$

Now

$$RHS = \int_t^T \int_{\mathbb{R}} (h-\hat{h})[b(x, he^{-\lambda t}, h_x e^{-\lambda t})h_x - b(x, \hat{h}e^{-\lambda t}, \hat{h}_x e^{-\lambda t})\hat{h}_x \tag{3.3.19}$$

$$- b_x(x, he^{-\lambda t}, h_x e^{-\lambda t})h + b_x(x, \hat{h}e^{-\lambda t}, \hat{h}_x e^{-\lambda t})\hat{h} \tag{3.3.20}$$

$$- f_x(x, he^{-\lambda t}, h_x e^{-\lambda t}) + f_x(x, \hat{h}e^{-\lambda t}, \hat{h}_x e^{-\lambda t})]dxds.$$

Since the function in the brackets is Lipschitz continuous and $\min_{[0,T]}(t) = 0$ we get the following inequality

$$
\begin{aligned}
RHS &\leq K \int_t^T \int_{\mathbb{R}} |h - \hat{h}|(|h - \hat{h}| + |(h - \hat{h})_x|) dxds \\
&= K \int_t^T \int_{\mathbb{R}} |h - \hat{h}|^2 dxds + K \int_t^T \int_{\mathbb{R}} |h - \hat{h}||(h - \hat{h})_x| dxds.
\end{aligned}
$$

Applying Cauchy's inequality with $\epsilon > 0$ to the second integral and combining like terms we obtain

$$
\begin{aligned}
RHS &\leq K \int_t^T \int_{\mathbb{R}} |h - \hat{h}|^2 dxds + K \int_t^T \int_{\mathbb{R}} \frac{1}{4\epsilon} |h - \hat{h}|^2 + \epsilon |(h - \hat{h})_x|^2 dxds \\
&\leq \frac{K}{4\epsilon} \int_t^T \int_{\mathbb{R}} |h - \hat{h}|^2 dxds + K\epsilon \int_t^T \int_{\mathbb{R}} |(h - \hat{h})_x|^2 dxds.
\end{aligned}
$$

Now we combine the RHS and the LHS to obtain

$$
\left(\frac{4\lambda\epsilon - K}{4\epsilon} \right) \int_t^T \int_{\mathbb{R}} |h - \hat{h}|^2 dxds + (\mu - K\epsilon) \int_t^T \int_{\mathbb{R}} |(h - \hat{h})_x|^2 dxds \leq 0.
$$

Choose ϵ such that

$$
\mu - K\epsilon > 0
$$

and choose λ s.t

$$
4\lambda\epsilon - K > 0. \tag{3.3.21}
$$

We conclude that

$$
\begin{aligned}
h - \hat{h} &= e^{\lambda t}(\theta - \hat{\theta}) = 0 \\
h_x - \hat{h}_x &= e^{\lambda t}(\theta_x - \hat{\theta}_x) = 0.
\end{aligned}
$$

Thus $\theta = \hat{\theta}$ and it is the unique solution to the equation (3.3.6). $\square$

Now that we have solution of the PDE (3.3.8) we can focus on the solution of the system (3.2.5) when X and p are one-dimensional.

Proposition 3.8. *Under Assumptions 3.1–3.4, the forward-backward system of SDEs (3.2.5) admits a solution $(X(t), p(t), q(t))$ and p and q are bounded.*

Proof. We have shown there exists a function $\theta(t, x)$ which is a solution to equation (3.3.6). It provides us with the connection between the forward variable X and backward variables p and q. Set $p(t) = \theta(t, X(t))$ and $q(t) = \theta_x(t, X(t))\sigma(X(t))$. Substituting them into the forward SDE in the system of equations (3.2.5) we get the following forward equation:

$$dX(t) = b(X(t), \theta(t, X(t)), \theta_x(t, X(t))\sigma(X(t))dt + \theta_x(X(t))\sigma(X(t))dW.$$

Since $\theta(t, X(t))$, $\theta_x(t, X(t))$ are uniformly bounded, and σ satisfies Assumption 3.1, we rewrite the previous equation as

$$dX(t) = \hat{b}(X(t))dt + \hat{\sigma}(X(t))dW.$$

Since b and σ are uniformly Lipschitz we conclude the above equation has a strong solution due to a classical existence theory. We verify that p and q are solutions to the backward system of equation (3.2.5). By applying Itô's formula to $p = \theta(t, X)$ we obtain

$$dp = d\theta = \theta_t dt + \theta_x dX + \frac{1}{2}\theta_{xx}(dX)^2.$$

Remember $dX(t) = b(t, X(t), \theta, \theta_x)dt + \sigma(t, X(t), \theta, \theta_x)dW$ and we can substitute it into above expression to obtain

$$dp(t) = \{\theta_t + b(t, X(t), \theta, \theta_x)\theta_x + \frac{1}{2}\theta_{xx}\sigma^2(t, X(t))\}dt + \sigma(t, X(t))\theta_x dW.$$

From the parabolic PDE we can change the drift term to

$$dp(t) = -\{b_x(t, X(t), \theta, \theta_x)\theta + \sigma_x(t, X(t))\sigma(t, X(t))\theta_x - f_x(t, X(t), \theta, \theta_x)\}dt + \theta_x\sigma(t, X(t))dW.$$

Substitute $p(t) = \theta(t, X(t))$ and $q(t) = \theta_x(t, X(t))\sigma(X(t))$ into the above equation to obtain

$$dp(t) = -\{b_x(t, X(t), p(t), q(t))p(t) + \sigma_x(t, X(t))q(t) - f_x(t, X(t), p(t), q(t))\}dt + q(t)dW$$

which is exactly the backward equation of the system of SDEs (3.2.5). Thus (X, p, q) is a solution to the system of FBSDEs with p and q bounded. $\square$

Proposition 3.9. *Under Assumptions 3.1–3.4 a solution* $(X(t), p(t), q(t))$ *is unique in a class of solutions where* $p(t)$ *and* $q(t)$ *are bounded.*

Proof. The drift coefficient of the backward equation is

$$
\begin{aligned}
|b_x(t, X(t), p(t), q(t))p(t) &+ \sigma_x(t, X(t))q(t) - f_x(t, X(t), p(t), q(t)) \\
&- b_x(t, \bar{X}(t), \bar{p}(t), \bar{q}(t))\bar{p}(t) + \sigma_x(t, \bar{X}(t))\bar{q}(t) - f_x(t, \bar{X}(t), \bar{p}(t), \bar{q}(t))| \\
&\leq |b_x(t, X(t), p(t), q(t))p(t) - b_x(t, \bar{X}(t), \bar{p}(t), \bar{q}(t))\bar{p}(t)| \\
&+ |\sigma_x(t, X(t))q(t) - \sigma_x(t, \bar{X}(t))\bar{q}(t)| \\
&+ |f_x(t, X(t), p(t), q(t)) - f_x(t, \bar{X}(t), \bar{p}(t), \bar{q}(t))| \\
&\leq |b_x(t, X(t), p(t), q(t))||p(t) - \bar{p}(t)| \\
&+ |\bar{p}(t)||b_x(t, X(t), p(t), q(t)) - b_x(t, \bar{X}(t), \bar{p}(t), \bar{q}(t))| \\
&+ |\sigma_x(t, X(t))||q(t) - \bar{q}(t)| \\
&+ |\bar{q}(t)||\sigma_x(t, X(t)) - \sigma_x(t, \bar{X}(t))| \\
&+ |f_x(t, X(t), p(t), q(t)) - f_x(t, \bar{X}(t), \bar{p}(t), \bar{q}(t))| \\
&\leq K(|X(t) - \bar{X}(t)| + |p(t) - \bar{p}(t)| + |q(t) - \bar{q}(t)|)
\end{aligned}
$$

where the last inequality follows from Assumptions 3.1 and 3.2 and the uniform boundedness of $p(t)$ and $q(t)$. Thus we conclude that the drift term of the backward SDE is Lipschitz continuous.

We need to show that the solution of the FBSDEs is of the particular form defined previously. Let $(X(t), p(t), q(t))$ be any solution and let

$$
\begin{aligned}
\tilde{p}(t) &= \theta(t, X(t)) \\
\tilde{q}(t) &= \sigma(t, X(t))\theta_x(t, X(t)).
\end{aligned}
$$

Apply Itô's formula to $p(t)$ to obtain:

$$
\begin{aligned}
d\tilde{p}(t) &= d\theta(t, X(t)) = \{\theta_t + \tfrac{1}{2}\theta_{xx}\sigma^2 + \theta_x b(t, X(t), p(t), q(t))\}dt + \theta_x \sigma dW \\
&= \{-\tfrac{1}{2}\sigma^2\theta_{xx} - b(X(t), \tilde{p}(t), \tilde{q}(t))\theta_x + b_x(X(t), \tilde{p}(t), \tilde{q}(t))\theta - \sigma_x\sigma\theta_x \\
&\quad + f_x(X(t), \tilde{p}(t), \tilde{q}(t)) + \tfrac{1}{2}\sigma^2\theta_{xx} + b(t, X(t), p(t), q(t))\theta_x\}dt + \theta_x\sigma dW \\
&= \{(b(t, X(t), p(t), q(t)) - b(X(t), \tilde{p}(t), \tilde{q}(t)))\theta_x + b_x(X(t), \tilde{p}(t), \tilde{q}(t))\theta - \sigma_x\tilde{q} \\
&\quad + f_x(X(t), \tilde{p}(t), \tilde{q}(t))\}dt + \tilde{q}(t)dW
\end{aligned}
$$

and

$$dp(t) = -\{b_x(t, X(t), p(t), q(t))p(t) + \sigma_x q(t) - f_x(X(t), p(t), q(t))\}dt + q(t)dW. \quad (3.3.22)$$

Then, by using Itô's formula again, we obtain

$$\begin{aligned} d|\tilde{p}(t) - p(t)|^2 = \ & \{2(\tilde{p}(t) - p(t))((b(t, X(t), p(t), q(t)) - b(X(t), \tilde{p}(t), \tilde{q}(t)))\theta_x \\ & +(b_x(X(t), \tilde{p}(t), \tilde{q}(t)) - b_x(X(t), p(t), q(t)))\theta \\ & +\sigma_x(q(t) - \tilde{q}(t)) + (f_x(X(t), \tilde{p}(t), \tilde{q}(t)) - f_x(t)) \\ & +(\tilde{q}(t) - q(t))^2)\}dt + 2(\tilde{p}(t) - p(t))(\sigma\theta_x - q(t))dW. \end{aligned}$$

Integrate both sides with respect to time and take an expectation to obtain:

$$\begin{aligned} \mathbb{E}|\tilde{p}(t) - p(t)|^2 \ & = -\mathbb{E}\int_t^T \{2(\tilde{p}(t) - p(t))((b(X(t), p(t), q(t)) - b(X(t), \tilde{p}(t), \tilde{q}(t)))\theta_x \\ & +(b_x(X(t), \tilde{p}(t), \tilde{q}(t)) - b_x(X(t), p(t), q(t)))\theta + \sigma_x(q(t) - \tilde{q}(t)) \\ & +(f_x(X(t), \tilde{p}(t), \tilde{q}(t)) - f_x(X(t), p(t), q(t))) + (\tilde{q}(t) - q(t))^2)\}ds \end{aligned}.$$

Since b and f are Lipschitz continuous, σ_x is bounded, and θ_x is bounded we obtain:

$$\begin{aligned} \mathbb{E}|\tilde{p}(t) - p(t)|^2 + \int_t^T \mathbb{E}|\tilde{q}(t) - q(t))|^2 ds \ & \leq K\int_t^T \mathbb{E}\{|\tilde{p}(t) - p(t)|(|p(t) - \tilde{p}(t)| \\ & \quad +|\tilde{q}(t) - q(t)|)\}ds \\ & \leq K\int_t^T \mathbb{E}\{|\tilde{p}(t) - p(t)|^2 + |\tilde{q}(t) - q(t)|^2\}ds \end{aligned}$$

where K is some generic constant. Then by Gronwall's inequality we conclude that

$$p(t) = \tilde{p}(t) \quad q(t) = \tilde{q}(t) \text{ a.e. } t \in [0, T] \ \mathbf{P}\text{-a.s.}$$

So the solution has to posses the form constructed previously.

Now to show the uniqueness of the solution we consider two solutions to the system of FBSDEs, $(X(t), p(t), q(t))$ and $(\hat{X}(t), \hat{p}(t), \hat{q}(t))$. Because $X(t)$ and $\hat{X}(t)$ satisfy the same forward SDE by a classical existence and uniqueness theory we conclude that $X(t) = \hat{X}(t)$ a.e. $t \in [0, T]$ **P**-a.s. If so, based on the previous calculation we can conclude that

$$p(t) = \hat{p}(t) \text{ and } q(t) = \hat{q}(t) \text{ a.e. } t \in [0, T] \ \mathbf{P}\text{-a.s.} \quad (3.3.23)$$

□

At this point we have obtained a solution to the one-dimensional system of FBSDEs that has both backward variables bounded, thus allowing us to conclude that this solution is unique in the class of solutions with bounded backward variables $p(t)$ and $q(t)$. In the next section we will extend this result to a multi-dimensional case.

3.3.2 Multidimensional Case

In this section we will extend our result to a multidimensional case. To avoid confusion we restate the problem for an n-dimensional case. In this section the notation is the following:

- $(\Omega \times \mathbb{T}) \equiv ([-\mathbb{R}, \mathbb{R}]^n \times [0, T])$ is a bounded domain.
- $\partial\Omega$ is a spatial boundary.

Multi-dimensional equivalent of equation (3.3.8) is written as following

$$\begin{cases} \theta^l_t + \sum_{i,j=1}^n \frac{1}{2}\sigma_i(x)\sigma_j(x)\theta^l_{x_j,x_i} + \sum_{i=1}^n b^i(x,\theta,z(x,\theta_x))\theta^l_{x_i} + \hat{b}^l(t,X,p(t),z(x,\theta_x)) = 0 \\ \theta(T,x) = -h(x(T)). \end{cases} \tag{3.3.24}$$

Let $a_{ij}(x) = \frac{1}{2}\sigma_i(x)\sigma_j(x)$ and $z(x,\theta_x) = \theta_x\sigma(x)$. To simplify notation we can rewrite the above system of equations as:

$$\begin{cases} \theta^l_t - \sum\limits_{i,j=1}^n a_{ij}(x)\theta^l_{x_ix_j} + \sum\limits_{i=1}^n b^i(x,\theta,\theta_x)\theta^l_{x_i} + \hat{b}^l(x,\theta,\theta_x) = 0 \\ \theta(T,x) = -h(x(T)). \end{cases} \tag{3.3.25}$$

Similar to the one dimensional case, in order to obtain solution of the above system we need to impose some restrictions on the coefficients of this system of FBSDEs. So the following assumptions are in effect for all of the subsequent results unless stated otherwise.

Assumption 3.10. *The functions b^i, σ_i and $\hat{b}^l$ are smooth functions and h is smooth and uniformly bounded.*

Assumption 3.11. *The following inequality holds.*

$$\nu\xi^2 \leq a_{ij}(x)\xi_i\xi_j \leq \mu\xi^2$$

where ν and μ are positive constants and $\xi_i, \xi_j \in \mathbb{R}$.

Assumption 3.12. *All second order partial derivatives of b, σ, and b^l are bounded by a constant M.*

Assumption 3.13. *For all $x \in \mathbb{R}^n$, $h_x(x) \in C^{2+\alpha}$ $||h_x(x)|| \leq K$ $\alpha \in (0,1)$, where K does not explicitly depend on x, and* $\lim_{|x|\to\infty} h_x(x) = 0$.

Before we can obtain a solution to the system of equations (3.3.25) we need to obtain a solution to a similar system of semi-linear PDE's in the parabolic cylinder $Q_t = \Omega \times [0,T]$ where the spatial domain $\Omega = B_R \in \mathbb{R}^n$ is a ball of radius R of the form

$$\begin{cases} \theta_t^l + \sum_{i,j=1}^n \frac{1}{2}\sigma_i(x)\sigma_j(x)\theta_{x_j,x_i}^l + \sum_{i=1}^n b^i(x,\theta,z(x,\theta_x))\theta_{x_i}^l + \hat{b}^l(t,X,p(t),z(x,\theta_x)) = 0 \\ \theta(t,x) = -h_x(x) \text{ where } (x,t) \in (\partial\Omega \times T) \cup (\Omega \times T)\}. \end{cases} \tag{3.3.26}$$

Solvability of this kind of system of quasi-linear PDE's is proved in Ladyženskaja et al. [1968, Theorem 7.1 p. 506]. In particular this Theorem states that there is an a priori bound in $\max_{Q_T} |u(t,x)| \leq M$ where M depends only on the value of $\theta(T,x)$, T, and some constants. In particular if:

1. $a_{ij}(x)\xi_i\xi_j \geq 0$ and
2. $\hat{b}^l(x,\theta,p)\theta^l \geq -c_1|\theta|^2 - c_2$, where c_1 and c_2 are positive constants

then,

$$\max_{Q_T} |\theta(t,x)| \leq \min_{c>c_1} e^{cT} \left\{ \max_{\Omega} |\theta(x,0)| + \sqrt{\frac{c_2}{c-c_1}} \right\} = M$$

this gives us an a priori estimate of $\max_{Q_T} |\theta(t,x)|$. Next we slightly adjust the Theorem from Ladyženskaja to our problem.

Lemma 3.14. *Suppose that the following conditions are satisfied:*

a) Conditions 1 and 2 are satisfied for $(t,x) \in \bar{Q}_T \backslash \Gamma_T$ and an arbitrary θ.

b) $a_{ij}(x)$ is uniformly elliptic, $|\theta| \leq M$ from above, and

$$\begin{aligned} &|a^i(x,\theta,p)| \leq \mu(|\theta|)(1+|p|) \\ &|\hat{b}(x,\theta,\theta_x)| \leq [\epsilon(|\theta|) + P(|p|,|\theta|)](1+|p|)^2 \\ &|\frac{\partial a_{ij}(x,\theta)}{\partial x_k}, \frac{\partial a_{ij}(x,\theta)}{\partial \theta^l}| \leq \mu(|\theta|) \end{aligned} \tag{3.3.27}$$

where $P(|p|,|\theta|) \to 0$ as $|p| \to 0$, $\epsilon(M)$ is sufficiently small number determined only by M, $\nu(M)$, and $\mu(M)$.

c) The map $-h(x(T))$ is bounded in $C^{2+\alpha}$.
Then, 3.3.26 admits a unique classical solution $\theta(t,x)$.

Now that we can solve the system of quasi-linear PDE (3.3.26) we can extend this result to the system (3.3.25) that is global in space.

Proposition 3.15. *Under Assumptions 3.10–3.13 the system (3.3.24) admits a unique solution $\theta(t,x)$, moreover $|\theta(t,x)|$ and $|\theta_x(t,x)|$ are bounded.*

Proof. Construct a sequence of parabolic PDE's of type (3.3.9) on a bounded domain $\Omega \times \mathbb{T} = [-R_i, R_i]^n \times [0,T]$, where R_i is an increasing sequence. By Lemma 3.6 there exists a sequence of classical solutions $\{\theta^i(t,x),\ i = 1,2,...\}$ to the sequence of equations above. Thus $\theta^i(t,x)$ posses all the necessary derivatives. Recall that $\theta^i(t,x)$, $\theta^i_t(t,x)$, $\theta^i_{tt}(t,x)$, $\theta^i_{x_j}(t,x)$, $\theta^i_{x_jx_l}(t,x)$ $j,l = 1,...,n$, and $\theta^i_{x_jx_lx_k}(t,x)$ $j,l,k = 1,...,n$ are also uniformly bounded $\forall\ R_i > 0$. Let $\{\theta^i(t,x)\}_{i=1,2,...} \equiv \mathbb{F}_1$, then since $\theta^i(t,x)$ are uniformly bounded by a constant that does not depend on x then $\mathbb{F}_1$ is equicontinuous. By the Arzela-Ascoli Theorem there exist a subsequence $\theta^{i,k}(t,x)$, $k = 1,2,3...$ that is uniformly convergent to $\theta^i(t,x)$. Similarly we construct $\{\theta^i_t(t,x)\}_{i=1,2,...} \equiv \mathbb{F}_2$, $\{\theta^i_{x_j}(t,x)\}_{i=1,2,...} \equiv \mathbb{F}_{3,j}$, and $\{\theta^i_{x_jx_l}(t,x)\}_{i=1,2,...} \equiv \mathbb{F}_{4,j,l}$ which are all equicontinuous as well. By the Arzela-Ascoli Theorem, we have

$$\begin{aligned}
\theta^{i,k}_t(t,x) \to \theta_t(t,x) &\ \text{ uniformly as } i,k \to \infty \\
\theta^{i,k}_{x,j}(t,x) \to \theta_x(t,x) &\ \text{ uniformly as } i,k \to \infty \\
\theta^{i,k}_{x_jx_l}(t,x) \to \theta_{x_jx_l}(t,x) &\ \text{ uniformly as } i,k \to \infty
\end{aligned}$$

Thus $\theta(t,x)$ is a classical solution of the parabolic PDEs (3.3.8). Moreover $\theta(t,x)$ and $\theta_x(t,x)$ are bounded. Uniqueness of the solution is the same as in one dimensional case. □

Now we have a bounded solution to the system of semi-linear PDE, i.e. we know θ and θ_x for all possible x and $t \in [0,T]$. What remains to solve for is a particular state $X(t)$. This can be done through solving a system of forward SDEs.

$$dX(t) = b(X(t), \theta(t, X(t), \theta_x(t, X(t))dt + \sigma(X(t))dW_t. \tag{3.3.28}$$

Because θ and θ_x are bounded, the coefficients $b(X(t), \theta(t, X(t), \theta_x(t, X(t))$ and $\sigma(X(t))$ are uniformly Lipschitz. Therefore equation (3.3.28) will have a unique solution. Also, $p(t) = \theta(t, X(t))$ and $q(t) = \theta_x(t, X(t))\sigma(X(t))$ are bounded.

Proposition 3.16. *Under Assumptions 3.10–3.13, the forward-backward system of SDEs 3.2.5 admits a solution (X(t),p(t),q(t)) where $p(t)$ and $q(t)$ are bounded.*

Proof. We have shown there exists a function $\theta(t,x)$ which is a solution to equation (3.3.24). It provides us with the connection between the forward variable $X(t)$ and backward variables $p(t), q(t)$. Set $p(t) = \theta(t, X(t))$ and $q(t) = \theta_x(t, X(t))\sigma(X(t))$. Substituting them into the forward SDE in the system of equations (3.2.5) we get the following forward:

$$dX(t) = b(X(t), \theta(t, X(t)), \theta_x(t, X(t))\sigma(X(t))dt + \theta_x(X(t))\sigma(X(t))dW.$$

since $\theta(t, X(t))$, $\theta_x(t, X(t))$ are uniformly bounded, and σ satisfies Assumption 3.10, we rewrite previous equation as

$$dX(t) = \hat{b}(X(t))dt + \hat{\sigma}(X(t))dW.$$

Since b and σ are uniformly Lipschitz we conclude that the above equation has a strong solution due to classical existence theory. We verify that $p(t), q(t)$ are solutions to the backward equation of system (3.2.5).

Apply Itô's formula to $p = \theta(t, X)$ to obtain

$$dp^l = d\theta^l(t, X(t)) = \theta^l_t dt + \sum_{i=1}^{n} \theta^l_{x_i} dX_i + \frac{1}{2} \sum_{i,j=1}^{n} \theta^l_{x_i x_j} \sigma^i \sigma^j < dX_i, dX_j > .$$

As a reminder $dX_i = b^i(t, X(t), \theta, \theta_x)dt + \sigma^i(t, X(t), \theta, \theta_x)dW$, substitute into above expression and obtain

$$dp^l = \left\{ \theta^l_t + \sum_{i=1}^{n} b^i(t, X(t), \theta, \theta_x)\theta^l_{x_i} + \frac{1}{2} \sum_{i,j=1}^{n} \theta_{x_i x_j} \sigma^i \sigma^j \right\} dt + \sum_{i=1}^{n} \sigma^i(t, X(t))\theta^l_{x_i} dW(t).$$

From the parabolic PDE we recognize that the drift term can be written in the following way

$$\begin{aligned} dp^l = & -\left\{ \sum_{i=1}^{n} b^l_{x_i}(t, X(t), \theta, \theta_x)\theta^i + \sum_{j=1}^{n} \sigma^i_{x_j}(t, X(t)) \sum_{i=1}^{n} \sigma^i(t, X(t))\theta^l_{x_i} - \sum_{i-1}^{n} f^l_{x_i}(t, X(t), \theta, \theta_x) \right\} dt \\ & + \sum_{i=1}^{n} \sigma^i(t, X(t))\theta^l_{x_i} dW \end{aligned}$$

substitute $p^l(t) = \theta^l(t, X(t))$ and $q^l(t) = \sum_{i=1}^{n} \theta^l_{x_i}(t, X(t))\sigma_i(X(t))$ into the above equation

$$
\begin{aligned}
dp^l = \quad & -\left\{ \sum_{i=1}^{n} b^l_{x_i}(t, X(t), p(t), q(t))p^i(t) + \sum_{j=1}^{n} \sigma^j_{x_j}(t, X(t))q^i(t) - \sum_{i=1}^{n} f^l_{x_i}(t, X(t), p(t), q(t)) \right\} dt \\
& +q^l(t)dW
\end{aligned}
$$

which is exactly the backward equation of the system of SDEs (3.2.5). Thus $(X(t), p(t), q(t))$ is a solution to the system of FBSDEs with $p(t), q(t)$ bounded. □

Proposition 3.17. *Under Assumptions 3.10–3.13, forward-backward system of SDEs 3.2.5 (X(t),p(t),q(t)) is unique in a class of solutions where $p(t)$ and $q(t)$ are bounded.*

Proof. The drift coefficient of the backward equation is

$$
\begin{aligned}
& |\sum_{i=1}^{n} b^l_{x_i}(t, X(t), p(t), q(t))p^i(t) + \sum_{j=1}^{n} \sigma^j_{x_j}(t, X(t))q^i(t) - \sum_{i=1}^{n} f^l_{x_i}(t, X(t), p(t), q(t)) \\
& \qquad - \sum_{i=1}^{n} b^l_{x_i}(t, \bar{X}(t), \bar{p}(t), \bar{q}(t))\bar{p}^i(t) + \sum_{j=1}^{n} \sigma^j_{x_j}(t, \bar{X}(t))\bar{q}^i(t) - \sum_{i=1}^{n} f^l_{x_i}(t, \bar{X}(t), \bar{p}(t), \bar{q}(t))| \\
& \quad \leq \sum_{i=1}^{n} |b^l_{x_i}(t, X(t), p(t), q(t))p^i(t) - b^l_{x_i}(t, \bar{X}, \bar{p}, \bar{q})\bar{p}^i| \\
& \qquad + \sum_{i=1}^{n} |\sigma^i_{x_i}(t, X(t))q^i(t) - \sigma^i_{x_i}(t, \bar{X}(t))\bar{q}^i(t)| \\
& \qquad + \sum_{i=1}^{n} |f^l_{x_i}(t, X(t), p(t), q(t)) - f^l_{x_i}(t, \bar{X}, \bar{p}, \bar{q})| \\
& \quad \leq \sum_{i=1}^{n} \{|b^l_{x_i}(t, X(t), p(t), q(t))||p^i(t) - \bar{p}^i(t)| \\
& \qquad +|\bar{p}^i(t)||b^l_{x_i}(t, X(t), p(t), q(t)) - b^l_{x_i}(t, \bar{X}(t), \bar{p}(t), \bar{q}(t))|\} \\
& \qquad + \sum_{i=1}^{n} |\sigma^i_{x_i}(t, X(t))||q^i(t) - \bar{q}^i(t)| + |\bar{q}^i(t)||\sigma^i_{x_i}(t, X(t)) - \sigma^i_{x_i}(t, \bar{X}(t))| \\
& \qquad + \sum_{i=1}^{n} |f^l_{x_i}(t, X(t), p(t), q(t)) - f^l_{x_i}(t, \bar{X}(t), \bar{p}(t), \bar{q}(t))| \\
& \quad \leq K \sum_{i=1}^{n} (|X^i(t) - \bar{X}^i(t)| + |p^i(t) - \bar{p}^i(t)| + |q^i(t) - \bar{q}^i(t)|)
\end{aligned}
$$

where the last inequality follows from Assumptions 3.10 and 3.11 and the uniform boundedness of $p(t)$ and $q(t)$. Thus we conclude that the drift term of the backward SDE is Lipschitz continuous.

We need to show that the solution of the FBSDEs is of the particular form defined previously. Let $(X(t), p(t), q(t))$ be any solution where

$$
\begin{aligned}
\tilde{p}^l(t) &= \theta^l(t, X(t)) \\
\tilde{q}^l(t) &= \sum_{i=1}^{n} \sigma^i(t, X(t))\theta^l_{x_i}.
\end{aligned}
$$

Apply Itô's formula to $\tilde{p}^l(t)$ to obtain:

$$\begin{aligned} d\tilde{p}^l &= \left\{ \theta^l_t + \sum_{i=1}^n b^i(t,X(t),p(t),q(t))\theta^l_{x_i} + \frac{1}{2}\sum_{i,j=1}^n \theta_{x_i x_j}\sigma^i\sigma^j \right\} dt + \sum_{i=1}^n \sigma^i(t,X(t))\theta^l_{x_i} dW(t) \\ &= \{-\frac{1}{2}\sum_{i,j=1}^n \theta_{x_i x_j}\sigma^i\sigma^j - \sum_{i=1}^n b^i(t,X(t),\tilde{p}(t),\tilde{q}(t))\theta^l_{x_i} + \sum_{i=1}^n b^l_{x_i}(t,X(t),\tilde{p}(t),\tilde{q}(t))\theta^i \\ &\quad - \sum_{j=1}^n \sigma^i_{x_j}(t,X(t)) \sum_{i=1}^n \sigma^i(t,X(t))\theta^l_{x_i} + \sum_{i-1}^n f^l_{x_i}(t,X(t),\tilde{p}(t),\tilde{q}(t)) + \frac{1}{2}\sum_{i,j=1}^n \theta_{x_i x_j}\sigma^i\sigma^j \\ &\quad + \sum_{i=1}^n b^i(t,X(t),p(t),q(t))\theta^l_{x_i}\} dt + \sum_{i=1}^n \sigma^i(t,X(t))\theta^l_{x_i} dW. \end{aligned}$$

Thus

$$\begin{aligned} d\tilde{p}^l &= \{\sum_{i=1}^n (b^i(t,X(t),p(t),q(t)) - \sum_{i=1}^n b^i(t,X(t),\tilde{p}(t),\tilde{q}(t)))\theta^l_{x_i} + \sum_{i=1}^n b^l_{x_i}(t,X(t),\tilde{p}(t),\tilde{q}(t))\theta^i \\ &\quad - \sum_{j=1}^n \sigma^i_{x_j}(t,X(t))\tilde{q}^j(t) + \sum_{i-1}^n f^l_{x_i}(t,X(t),\tilde{p}(t),\tilde{q}(t))\} dt + \tilde{q}^l(t) dW \end{aligned}$$

and

$$\begin{aligned} dp^l = &\ -\left\{ \sum_{i=1}^n b^l_{x_i}(t,X(t),p(t),q(t))p^i(t) + \sum_{j=1}^n \sigma^j_{x_j}(t,X(t))q^i(t) - \sum_{i=1}^n f^l_{x_i}(t,X(t),p(t),q(t)) \right\} dt \\ &+ q^l(t) dW. \end{aligned}$$

Then by using Itô's formula again we obtain

$$\begin{aligned} d|\tilde{p}^l(t) - p^l(t)|^2 = &\ \{2(\tilde{p}^l(t) - p^l(t))(\sum_{i=1}^n (b^i(t,X(t),p(t),q(t)) - \sum_{i=1}^n b^i(t,X(t),\tilde{p}(t),\tilde{q}(t)))\theta^l_{x_i} \\ &+ (\sum_{i=1}^n b^l_{x_i}(t,X(t),p(t),q(t)) - \sum_{i=1}^n b^l_{x_i}(t,X(t),\tilde{p}(t),\tilde{q}(t)))\theta^i \\ &+ \sigma_x(q - \tilde{q}) \\ &+ \sum_{j=1}^n \sigma^i_{x_j}(t,X(t))(q^j(t) - \tilde{q}^j(t)) \\ &+ (\tilde{q}^l(t) - q^l(t))^2)\} dt \\ &+ 2(\tilde{p}^l(t) - p^l(t))(\tilde{q}^l(t) - q^l(t)) dW. \end{aligned}$$

Integrate both sides with respect to time and take an expectation to obtain:

$$\begin{aligned} \mathbb{E}|\tilde{p}^l(t) - p^l(t)|^2 &= -\mathbb{E}\int_t^T \{2(\tilde{p}^l(t) - p^l(t))(\sum_{i=1}^n (b^i(t,X(t),p(t),q(t)) - \sum_{i=1}^n b^i(t,X(t),\tilde{p}(t),\tilde{q}(t)))\theta^l_{x_i} \\ &\quad + (\sum_{i=1}^n b^l_{x_i}(t,X(t),p(t),q(t)) - \sum_{i=1}^n b^l_{x_i}(t,X(t),\tilde{p}(t),\tilde{q}(t)))\theta^i \\ &\quad + \sigma_x(q - \tilde{q}) + \sum_{j=1}^n \sigma^i_{x_j}(t,X(t))(q^j(t) - \tilde{q}^j(t)) + (\tilde{q}^l(t) - q^l(t))^2)\} ds. \end{aligned}$$

Since b and f are Lipschitz continuous, σ_x is bounded, and θ_x is bounded we obtain:

$$\begin{aligned}\mathbb{E}|\tilde{p}^l(t) - p^l(t)|^2 + \int_t^T \mathbb{E}|\tilde{q}^l(t) - q^l(t))|^2 ds &\leq K \int_t^T \mathbb{E}\{|\tilde{p}^l(t) - p^l(t)|(|p^l(t) - \tilde{p}^l(t)| \\ &\quad +|\tilde{q}^l(t) - q^l(t)|)\}ds \\ &\leq K \int_t^T \mathbb{E}\{|\tilde{p}^l(t) - p^l(t)|^2 + |\tilde{q}^l(t) - q^l(t)|^2\}ds,\end{aligned}$$

where K is a constant. Then by Gronwall's inequality we conclude that

$$p^l(t) = \tilde{p}^l(t) \quad q^l(t) = \tilde{q}^l(t) \text{ a.e. } t \in [0,T] \ \mathbf{P}\text{-a.s.}$$

So the solution has to posses the previously established form.

To show the uniqueness of the solution we consider two solutions to the system of FB-SDEs, (X,p,q) and $(\hat{X},\hat{p},\hat{q})$. Because $X(t)$ and $\hat{X}(t)$ satisfy the same forward SDE by classical existence and uniqueness theory we conclude that $X(t) = \hat{X}(t)$ a.e. $t \in [0,T]$ $\mathbf{P}$-a.s. If so, we can conclude that $p(t) = \hat{p}(t)$ and $q(t) = \hat{q}(t)$ a.e. $t \in [0,T]$ $\mathbf{P}$-a.s based on the previous calculation. □

3.4 Application

3.4.1 Optimal Control Problem

In this section we consider the optimal control problem from the previous chapter. In particular, the state equation for the bear metapopulation model was

$$\begin{cases} dX(t) &= (\hat{b}(t,X(t)))dt + \hat{\sigma}(t,X(t))dW(t) \\ X(0) &= X_0 \end{cases} \tag{3.4.1}$$

where $\hat{b}$ is

$$\left\{ \begin{array}{c} aX_1(t)(1-X_1(t)) - e_P X_1(t) + e_F(1-X_1(t))X_2(t) \\ aX_2(t)(1-X_2(t)) - e_F X_2(t) + e_P(1-X_2(t))X_1(t) - X_2(t)u(t) \end{array} \right\}$$

with $\hat{\sigma} = (cX_1(t)(1-X_1(t)), cX_2(t)(1-X_2(t)))$. The functional is thus:

$$J(u) = \mathbb{E}\left(\int_0^T (e_F + e_P)X_1(t)X_2(t) + \frac{1}{2}mu^2(t)dt\right). \tag{3.4.2}$$

We use $f(X(t), u(t))$ to indicate integrand of objective functional $(e_F + e_P)X_1(t)X_2(t) + \frac{1}{2}mu^2(t)$.

For this stochastic optimal control problem the system of adjoint equations is a system of backward SDEs and is set to be:

$$\begin{cases} dp(t) = -\left\{\hat{b}_x p(t) + \sigma_x(X(t))q(t) - f_x^T(X(t), u(t))\right\} dt + q(t)dW(t) \\ p(T) = 0. \end{cases} \tag{3.4.3}$$

Recall that for the above optimal control we can set the stochastic Hamiltonian (see Yong and Zhou [1999]) as:

$$H(t, X(t), u(t), p, q) = < p, \hat{b}(t, X(t), u(t)) > +tr[q(t)^T \sigma(X(t))] + f(t, X(t), u(t)) \tag{3.4.4}$$

Considering the derivative of the Hamiltonian with respect to control u

$$H_u(t, X(t), u(t), p, q) = -X_2(t)p_2(t) + m\, u(t) \tag{3.4.5}$$

and considering the convexity of the Hamiltonian we can solve for the optimal control

$$\bar{u}(t) = \frac{p_2(t)X_2(t)}{m}. \tag{3.4.6}$$

Thus, the system of forward-backward SDEs that claims the optimal solution to the problem takes the following general form:

$$\begin{cases} dX(t) = \tilde{b}(X(t), p(t))dt + \sigma(X(t))dW(t) \\ dp(t) = -\left\{\tilde{b}_x(X(t), p(t))p(t) + \sigma_x(X(t))q(t) - \tilde{f}_x^T(X(t), p(t))\right\} dt + q(t)dW(t) \\ X(0) = x_0 \quad p(T) = 0. \end{cases} \tag{3.4.7}$$

where $\sigma(X(t))$ is as above, $\tilde{b}(X(t), p(t))$ is

$$\left\{ \begin{array}{c} aX_1(t)(1 - X_1(t)) - e_P X_1(t) + e_F(1 - X_1(t))X_2(t) \\ aX_2(t)(1 - X_2(t)) - e_F X_2(t) + e_P(1 - X_2(t))X_1(t) - \dfrac{p_2(t)X_2^2(t)}{m}. \end{array} \right\}$$

$\tilde{f}_x(X(t), p(t))$ is equal to

$$\left\{ \begin{array}{c} (e_p + e_f)X_2(t) \\ (e_p + e_f)X_1(t) + \dfrac{p_2^2(t)X_2(t)}{m}, \end{array} \right\}$$

$\tilde{b}_x(X(t), p(t))$ is

$$\left\{ \begin{array}{cc} a(1 - 2X_1(t)) - e_P - e_F X_2(t) & e_F(1 - X_1(t)) \\ e_P(1 - X_2(t)) & a(1 - 2X_2(t)) - e_F - e_P X1(t) - \dfrac{2p_2(t)X_2(t)}{m}) \end{array} \right\}$$

and σ_x is

$$\left\{ \begin{array}{cc} c(1 - 2X_1(t)) & 0 \\ 0 & c(1 - 2X_2(t)) \end{array} \right\}.$$

From the previous section we know there exists a bounded solution of this system.

3.4.2 Numerical Solution of the FBSDE

Until recently there was a limited number of numerical methods that could be employed in solving such kind of forward-backward SDE systems. In the second chapter we have obtained a solution to the FBSDE system through the Four Step Scheme method presented in Yong and Zhou [1999]. The main idea there was that for systems where coefficients satisfy certain Lipschitz conditions there can be constructed an elliptic PDE such that its solution contains profiles for both forward and backward equations. This method proved to be very cumbersome. As the dimensions of the underlying forward and backward equations increase this method becomes very difficult to implement. Recently there was a sequence of papers that produced an alternative methodology involving a purely probabilistic approach. Main idea of alternate methodology was introduced by Gobet et al. [2005] and Bender and Denk [2007]. We are going to implement results presented in Bender and Zhang [2008]. First we show that our problem results in a FBSDE system that is consistent with the type of problems considered in the latter work.

From the structure of the Hamiltonian and the fact that $tr[q(t)^T\sigma(X(t))]$ does not depend on the control $u(t)$ we can see that the backward variable $q(t)$ does not enter into the optimal control, thus preventing $q(t)$ from entering into the forward equations. This results in the general form of the resulting FBSDE looking like:

$$\left\{\begin{array}{l} dX(t) = b(t, X(t), p(t))dt + \sigma(t, X(t), p(t))dW(t) \\ dp(t) = B(t, X(t), p(t), q(t))dt + q(t)dW(t) \\ X(0) = x_0, \quad p(T) = g(x(T)) \end{array}\right\}.$$

where

$$B(t, X(t), p(t), q(t)) = -\left\{b_x(X(t), p(t))p(t) + \sigma_x(X(t))q(t) - f_x^T(X(t), p(t))\right\}$$

and σ as above.

Extra requirement to the system of FBSDEs from Bender and Zhang [2008] are presented here in the form of the following two assumptions. The wording is preserved from the original paper.

Assumption 3.18. *The coefficients (b, σ, f) are uniformly Holder-$\frac{1}{2}$ continuous with respect to t.*

Assumption 3.19. *At least one of the following is satisfied:*

1. *The time duration T is small enough.*

2. *Weak coupling of $p(t)$ into the forward SDE, i.e. b_p and σ_p are small enough.*

The last assumption requires some clarification. The first part of Assumption 3.19 is more of an implementation issue. For large terminal T the above numerical scheme blows up due to error accumulation. An exact formula for permissible T to the best of our knowledge does not exist yet. It depends on the forward backward coefficients complexity and the choice of Δt. At this point T is determined numerically and the largest possible T is determined for which the method does not blow up. This can be a topic for future research to determine a particular formula for determining T and Δt. The second part of Assumption 3.19 can be easily satisfied for the type of FBSDE considered in this work. Recall that in the equation 3.4.7 the backward variable $p(t)$ enters into forward drift coefficient only. Thus, $\sigma_p = 0$. If $p(t)$ is bounded then for $b(t, X(t), p(t))$ that fluctuates slightly, $b_p(t, X(t), p(t))$ can be expected to be small. In the original work by Bender and Zhang [2008] no formula for the size of $b_p(t, X(t), p(t))$ is provided. It is suggested though that it should be reasonably small. More research is necessary to determine a particular formula.

From our findings about existence of the bounded solution to the type of FBSDE considered in earlier sections we can also satisfy condition that $b(t, X(t), p(t)), \sigma(t, X(t), p(t)), g(x(T))$

are deterministic and Lipschitz continuous functions. The function $B(t, X(t), p(t), q(t))$ satisfies findings of previous sections.

The FBSDE Markovian Iteration algorithm described below requires the following input: Λ – Number of sample paths; M – Maximum number of Markovian iterations; n – Number of time steps; T – Terminal time; σ – Stochastic volatility constant (different from $\sigma(X(T))$); $X0 := [X0_1, X0_2]$ – Initial state vector.

In addition, important variables used to store calculations are allocated as follows:

- dW – dimensioned $[\Lambda \times (n+1)]$: Random Walk in time for each path.
- X, p, q – dimensioned $[\Lambda \times (n+1) \times 2]$: X is the state vector for each path at each time step.
- U, V – dimensioned $[(n+1) \times 4]$: Least Squares regression coefficients for each component at each time step. Note that $U(x) = \sum_{i=0}^{3} \alpha_i \phi_i(x)$, where the basis functions $\{\phi_i(x)\}$ are defined as $\{1, x_1, x_2, x_1 x_2\}$, and α_i are the coefficients from the regression analysis. Similarly, $V(x) = \sum_{i=0}^{3} \beta_i \phi_i(x)$.

Note that p, q, U, V are intialized to 0 upon entering the procedure.

Algorithm 1 FBSDE Markovian Iteration

procedure MarkovIter$(\Lambda, M, n, T, \sigma, X0)$ ▷ Main Program
  $\Delta t \leftarrow T/n$
  $dW[:,:] \leftarrow \mathcal{N}(0,1) \cdot \sqrt{\Delta t}$ ▷ Different BM for each path
  $X_1 \leftarrow X0$ ▷ All paths start at the same state
  for $m \leftarrow 1, M$ **do** ▷ Markovian Iteration Loop
    for $i \leftarrow 2, n+1$ **do** ▷ Update X, Forward Stochastic Euler
      $X_i \leftarrow F_E(X_{i-1}, U_{i-1}(X_{i-1}), dW_i, \sigma(X_{i-1}))$
    end for
    for $i \leftarrow n, 2, -1$ **do** ▷ Backward Update
      $p_{i+1} \leftarrow U_{i+1}(X_{i+1})$ ▷ predict p_{i+1}
      $V_i(x) \leftarrow LSQReg\left(\frac{1}{h} p_{i+1} dW_{i+1}, X_i\right)$
      $q_i \leftarrow V_i(X_i)$ ▷ predict q_i
      $U_i(x) \leftarrow LSQReg\left(B(X_i, p_{i+1}, q_i), X_i\right)$
    end for
    $p_2 \leftarrow U_2(X_2)$ ▷ predict p_2
    $q_1 \leftarrow \mathbb{E}\left[\frac{1}{h} p_2 dW_2\right]$ ▷ estimate q_1 over Λ paths
    $p_1 \leftarrow \mathbb{E}\left[B(X_1, p_2, q_1)\right]$ ▷ estimate p_1 over Λ paths
  end for
end procedure

The Forward Stochastic Euler update function F_E above is a vector valued function. The backward update function B involves $B(t, X(t), p(t), q(t))$. Note also that unless explicitly stated otherwise, updates to variables are performed over all paths in parallel.

In the implementation of this scheme we solve the system (3.4.7) for the maximum time $T_{max} = 2.5$. We choose uniform time discretization with $n = 500$ time steps with uniform time step $\Delta t = 0.005$. Stochastic volatility coefficient σ varies for different experiments. The following is iterated over m-many Markovian iterations with maximum possible number of iterations $M = 100$. We use $\Lambda = 5000$. When error tolerance $||P_0^m - P_0^{m-1}|| < 10^{-5}$ is achieved then we terminate Markovian iterations.

In each Markovian iteration m the forward in time variable X is approximated by using the Stochastic Euler Scheme.

$$X_{j+1}^i = X_j^i + b(X_j^i, U_j(X_j^i))dt + \sigma(X_j, U_j(X_j^i))dW_j^i \tag{3.4.8}$$

where right hand side of the equation 3.4.8 is equivalent to F_E on line 7 of the FBSDE Markovian Iteration algorithm, dW_j^i is appropriate Brownian motion increments, $U_j(X_j^i)$ is an approximation of p_2 using LSQReg coefficients from $(m-1)^{th}$ Markovian iteration. In other words

$$U_j(X_j^i) = U_j^1 c + U_j^2 (X_1)_j^i + U_j^3 (X_2)_j^i + U_j^2 (X_1)_j^i (X_2)_j^i$$

Where $(X_1)_j^i$, $(X_2)_j^i$ and $i = 0, ..., \Lambda$ $j = 2, .., n$ are used to form monomial basis functions. This is equivalent to steps (6)-(8) in the algorithm (3.4.2).

Within the same Markovian iteration m we update the pair of backward variables $p(t), q(t)$. Terminal values for backward variables are set as $p_j^i = g(X_n^i)$ and $q_n^i = 0$. We move backwards in time $j = n-1, ..., 2$. Then we use $U_j(X_j^i)$ the same way as in forward component

$$p_{j+1}^i = U_{j+1}(X_{j+1}^i). \tag{3.4.9}$$

Notice that we use the same monomial functions as in the forward variable update. Using Monte-Carlo simulations we solve least squares minimization problem on line (8) of the algorithm (3.4.2). We use coefficients V_j^i to obtain approximation of backward variable q

$$q_{j+1}^i = V_{j+1}(X_{j+1}^i) = V_j^1 c + V_j^2 (X_1)_j^i + V_j^3 (X_2)_j^i + V_j^2 (X_1)_j^i (X_2)_j^i. \tag{3.4.10}$$

Similarly we solve least square problem on line (13) of the algorithm (3.4.2). This produces a set of coefficients U_j^i. These coefficients will be used in the next Markovian iteration step

$m + 1$. For the time $t = 1$ and $t = 0$ we set

$$\begin{aligned} p_1^i &= U_1(X_1^i) \\ q_0 &= \frac{1}{\Lambda} \sum_{i=1}^{\Lambda} \frac{1}{h} p_1^i dW_1^i \\ P_0^m &= \frac{1}{\Lambda} \sum_{i=1}^{\Lambda} p_1^i + B(X_0^i, p_1^i, q_0^i). \end{aligned} \tag{3.4.11}$$

This is a summary of the code provided in Appendix A. It is necessary to mention that we have tested several choices of basis functions and have found that their choice does not play a critical role. A simple monomial of degree 1 used in our implementation performed equally well as a monomial of degree 3. As a matter of fact, the use of more complicated function in least-squares played a negative role. The choice of monomial of higher order increased run time without significantly improving convergence rate. It is also our observation that the number of independent paths of Brownian motion were used in least-square regression had stronger effect on convergence than the choice of basis functions. With higher number of Brownian increments we were able to obtain faster convergence of Markovian iteration.

3.4.3 Numerical Results

Convergence of Markovian iteration is rather fast. In Figure 3.1, we show convergence obtained for m-many Markovian iterations for different variances c. For the deterministic case the convergence is almost instant. As the variance c gradually increased we can see that convergence slows down. The rate of convergence of this method is an advantage over other methods. In the previous chapter the time required to obtain solution of the parabolic PDE was higher.

We conducted several experiments to identify the solutions to the optimal control problem formulated in the previous sections. In the first experiment we tested the deterministic case where the variance $c = 0$. In Figure 3.2 one can see that we are able to obtain results that replicate populations for the deterministic case from Chapter 2. The difference is in the values for the optimal control. This can be explained by the differences in formulation of the optimal control problem in each implementation. Although the values of the optimal control problems in Chapter 2 are different from the values for optimal control in this chapter, the shape of the graph is similar. Note that at the terminal time both controls tend to zero.

Results in the Figures 3.3 and 3.4 are solutions to a stochastic system with variances set to $c = 0.1$ and $c = 0.5$ respectively. One can see the result of stochasticity here. Diffusion has a stronger effect on the paths of population for higher values of the variance.

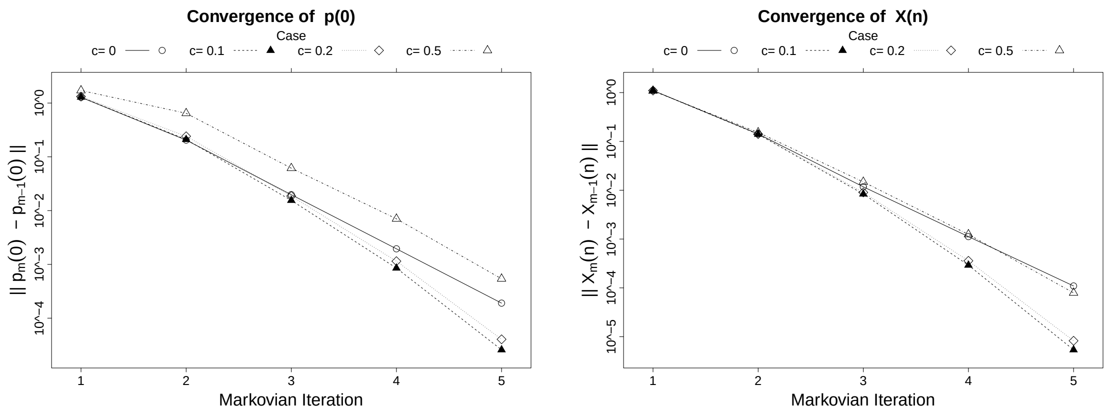

Figure 3.1: Convergence for $p(t)$ and $X(t)$.

In Figures 3.5–3.7 we show ten sample paths of population densities and harvesting rates when mast availability is $c = 0.1$, $c = 0.2$, and $c = 0.5$, respectfully. We can see that the harvesting rate is converging to 0 at the end of the period T for all three experiments. We can see that the population densities in the Park and Forest diverge by larger margins as the mast availability variance becomes larger.

In the Figure 3.8, we produced 5 sample paths for population densities and harvesting rates. We incrementally increase mast availability variance by 0.1 from a deterministic case with variance equal to 0.0 to a stochastic case with variance equal to 0.5. We can see how with all other parameters equal the sample paths deviate further from the deterministic graph with higher variance. Also, the paths for stochastic optimal controls become more volatile as well.

3.5 Conclusions

In this chapter we considered a solution to a general stochastic optimal control problem. We have shown that for this general type of problem the resulting forward-backward system of SDEs has a drift coefficient that rarely satisfies the Lipschitz continuity condition. In the standard theory on solvability of FBSDE systems, the Lipschitz condition is critical. In most cases both the drift and diffusion coefficients need to satisfy Lipschitz conditions in order to have solvability of such systems. Unfortunately, most FBSDE systems that emerge from stochastic optimal control problems do not have Lipschitz conditions satisfied for backwards drift. This significantly limits the number of problems that can be solved through an associated FBSDE. In this chapter we have successfully shown how to obtain

Lipschitz condition on a drift term of a system of backward SDEs. We have shown that if we can establish a bound on the backward variables through associated PDE then we can obtain unique solution to a system of FBSDEs. In order to achieve this objective we have further studied the connection between parabolic PDEs and systems of FBSDEs. In Ladyženskaja et al. [1968] a priori bounds are established for solutions of certain kind of parabolic PDEs on bounded domains. Using these results we were able to prove existence of solutions to systems of FBSDE on unbounded domains given some minimal assumptions presented in this chapter are met.

To support our findings we have solved the stochastic optimal control problem from Chapter 2 through a system of FBSDEs derived form Stochastic Maximum Principle. We set up an associated FBSDE system and estimated its solutions using a numerical method that uses a Markovian-type iteration. This technique is an adaptation from the original work by Bender and Zhang [2008]. This technique is purely probabilistic and requires estimation of conditional expectations at each iteration. This method proved to be superior in some aspects to the four-step scheme used previously. Rapid convergence was achieved with this method. The most significant advantage of this method is that it allows handling of multidimensional systems of FBSDEs. In our case the system had 2 dimensions in the forward variables and 2 dimensions in the backward variables. Stochastic optimal control problems that results in a high dimensional system of FBSDEs are hard to solve through a related parabolic PDE due to complexity of the implementation. The Markovian-type iteration is not limited in how many dimensions a system of FBSDEs have.

Significance of these findings is that a larger family of stochastic optimal control problems can be solved. Along with numerical schemes this methodology can be applicable to numerous kinds of problems that one may encounter in different fields like finance, engineering, and biology.

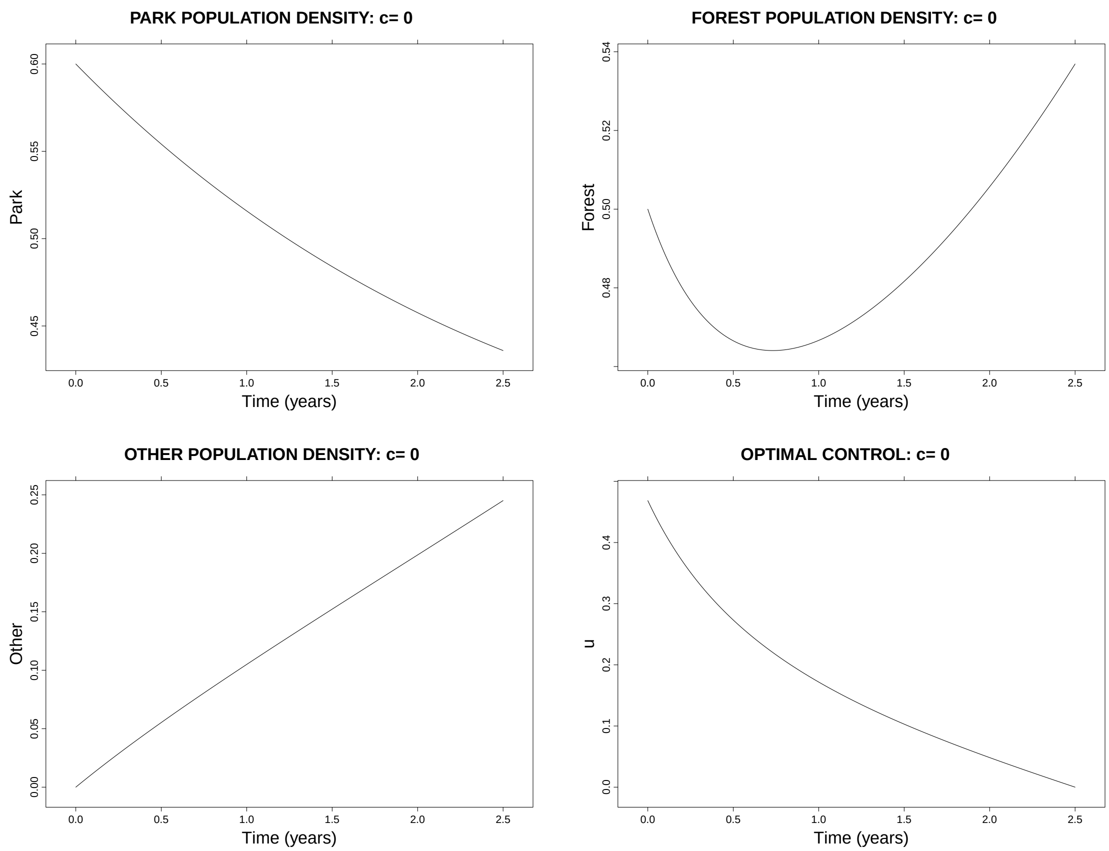

Figure 3.2: Population results of optimal control problem in the deterministic case, i.e. with no randomness.

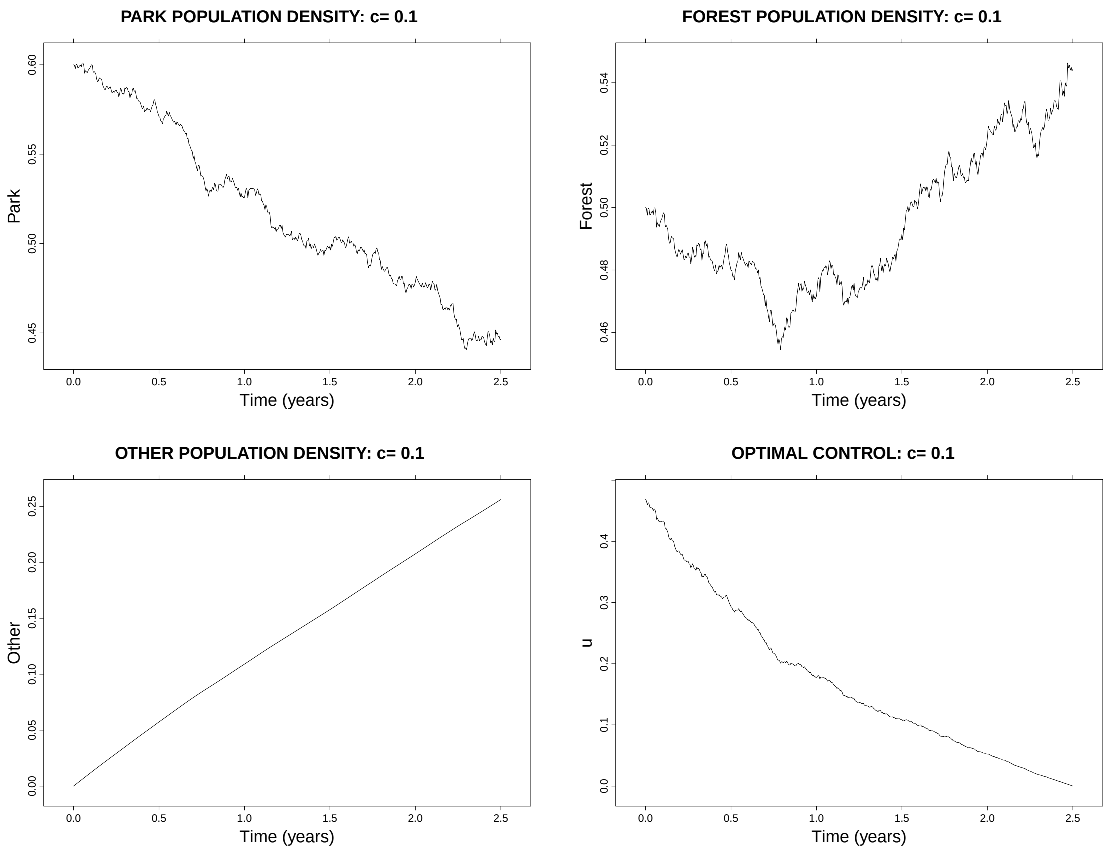

Figure 3.3: Population results of optimal control with $c = 0.1$.

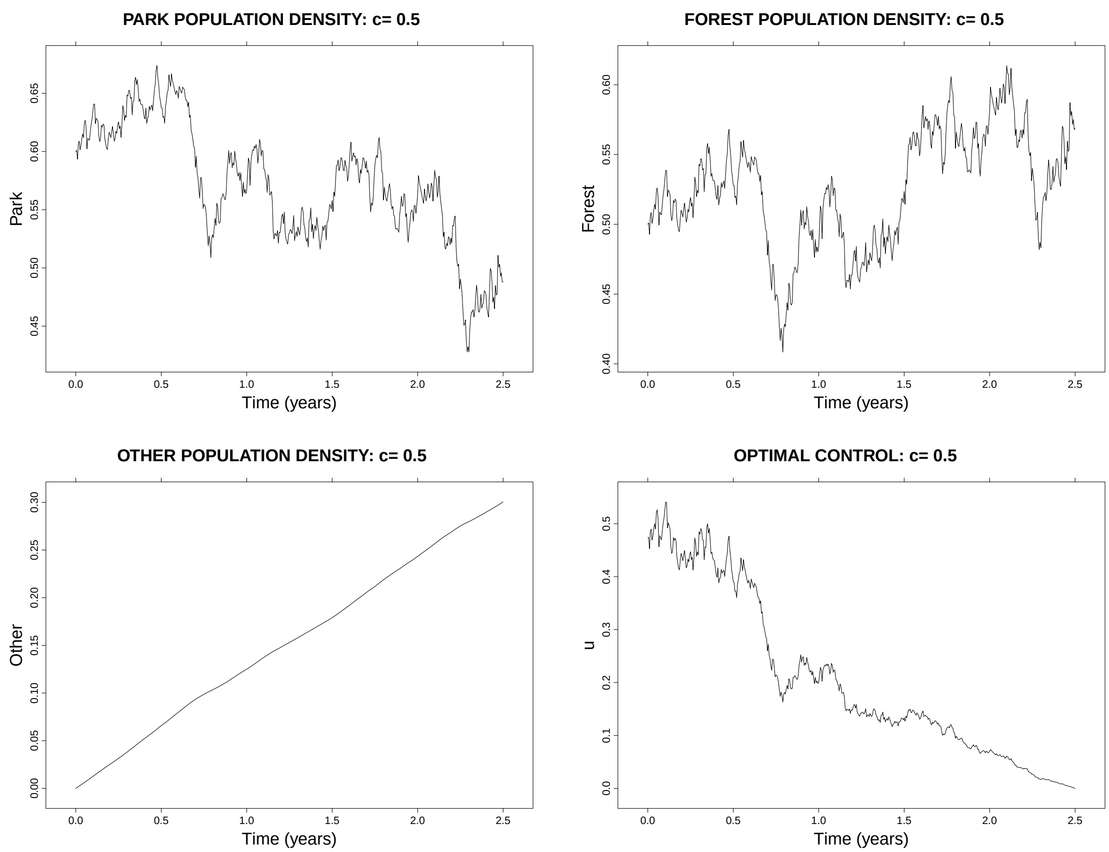

Figure 3.4: Population results of optimal control with $c = 0.5$.

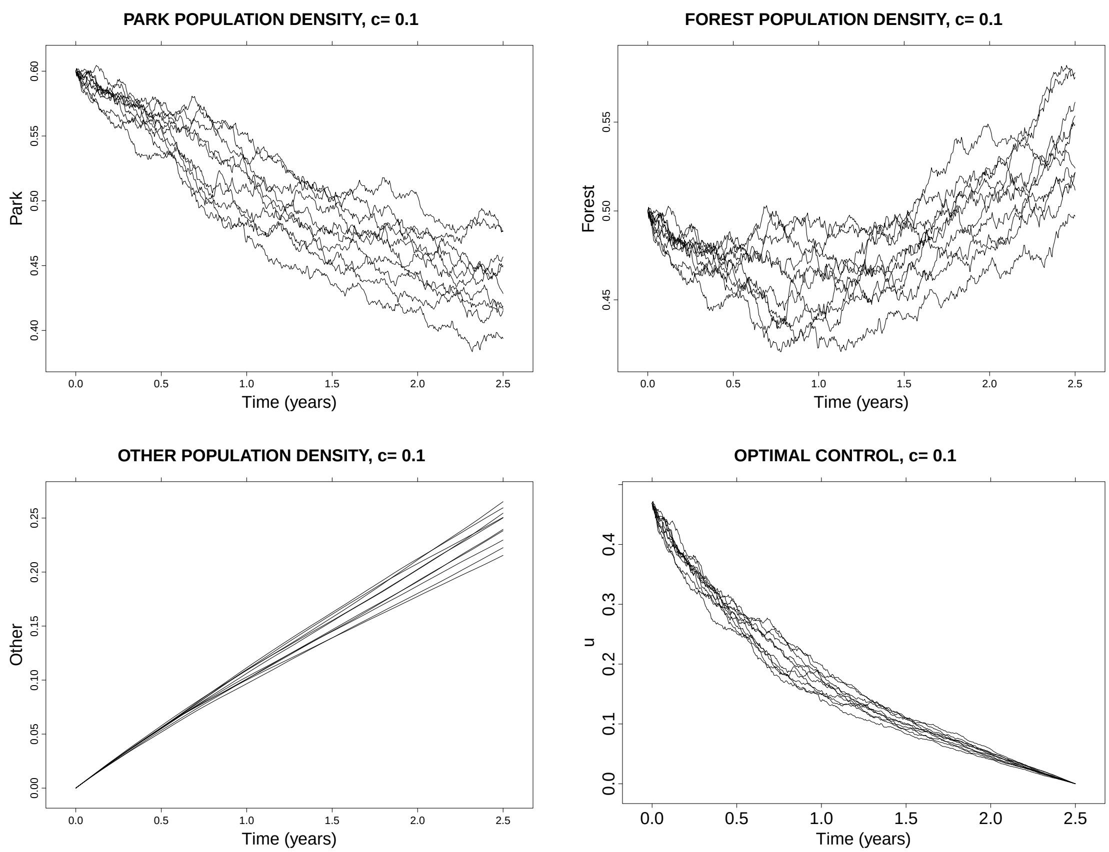

Figure 3.5: Ten sample path simulation of population and harvesting rate when mast availability variance $c = 0.1$.

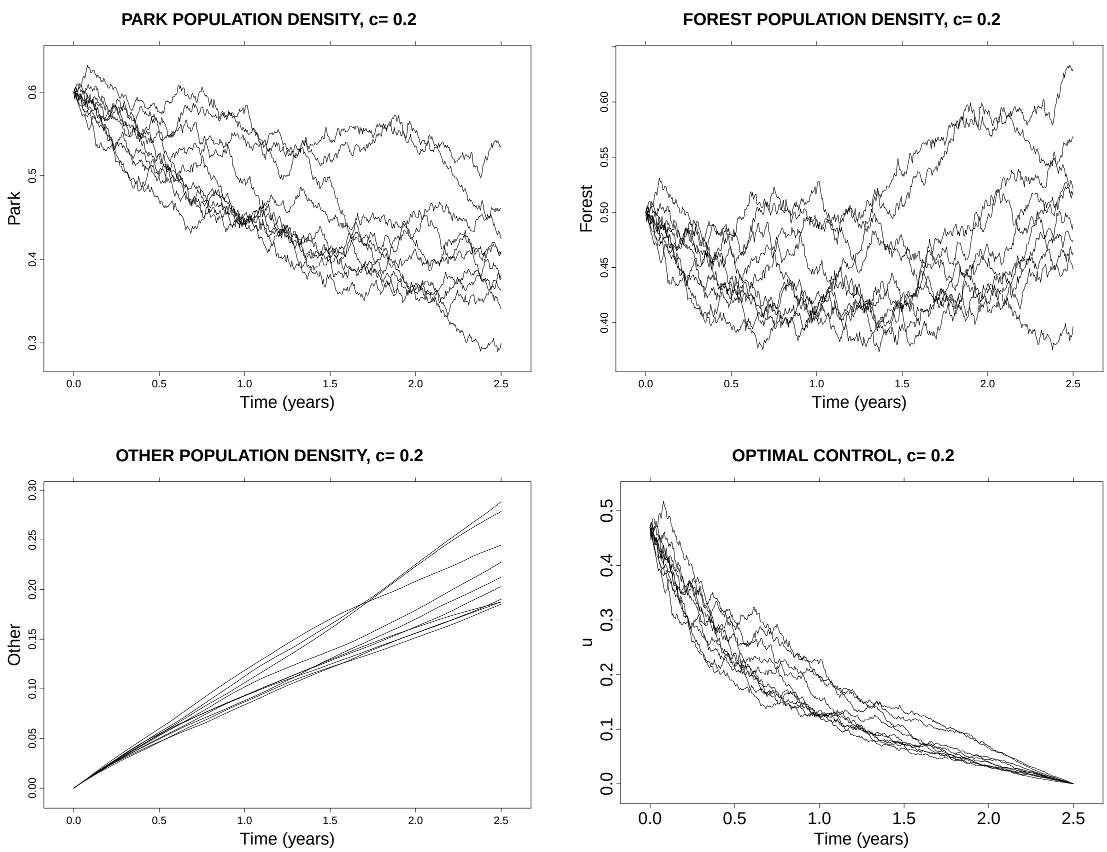

Figure 3.6: Ten sample path simulation of population and harvesting rate when mast availability variance $c = 0.2$.

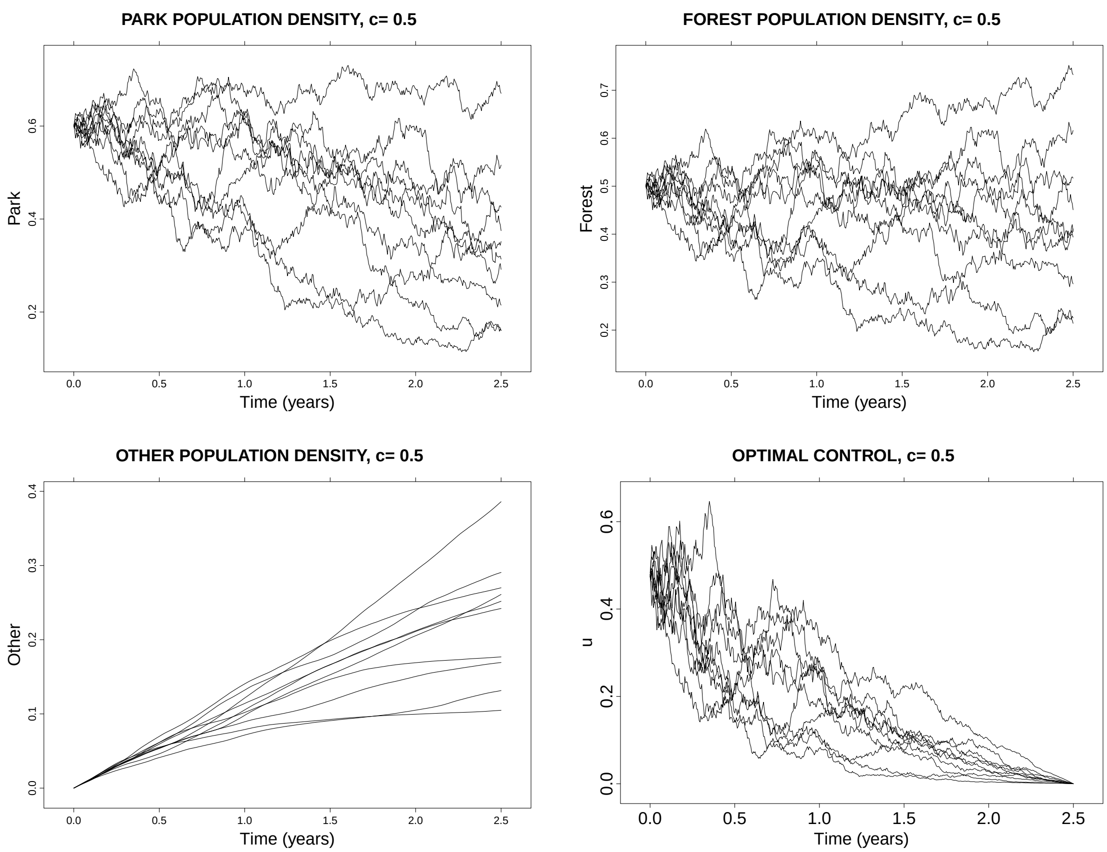

Figure 3.7: Ten sample path simulation of population and harvesting rate when mast availability variance $c = 0.5$.

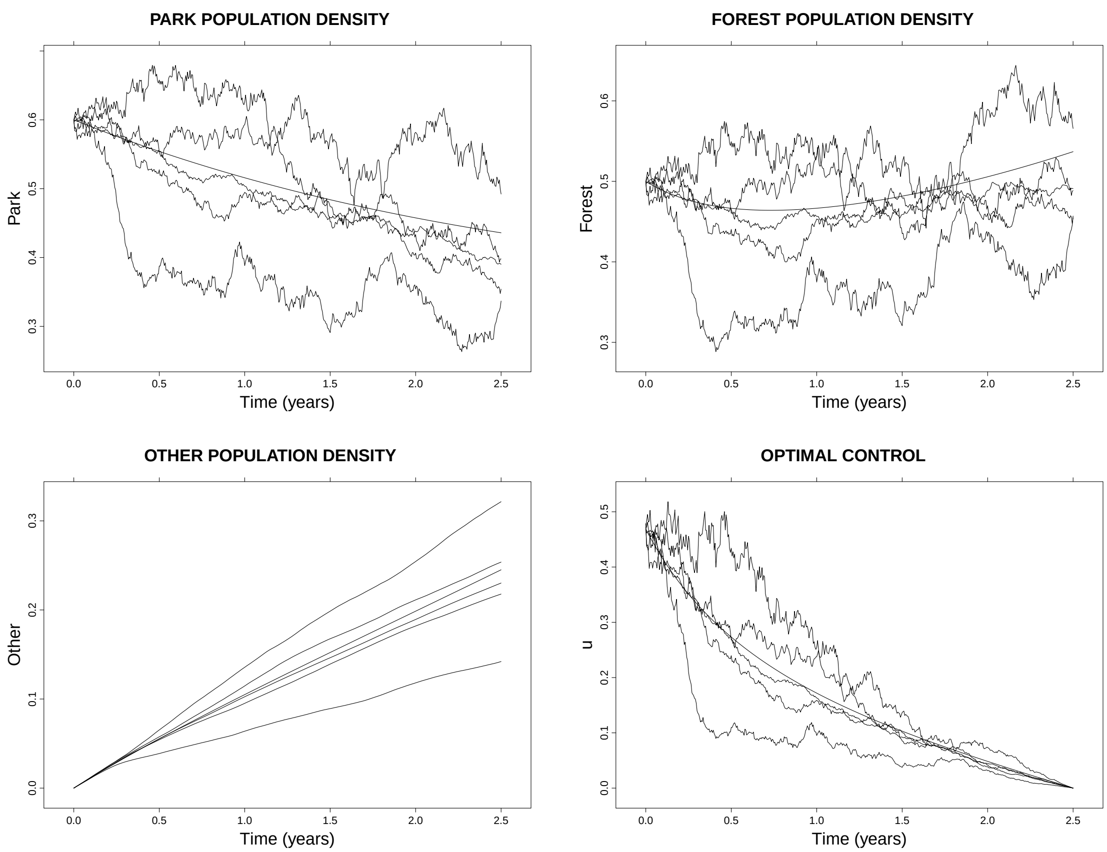

Figure 3.8: Population results of the optimal control problem with the variance of the mast availability c gradually increased from $c = 0$ to $c = 0.5$ by 0.1.

Chapter 4

Summary

This work studies of a stochastic metapopulation models. Similar to solving a deterministic optimal control problem one might be interested in solving a stochastic optimization control problem. Now, instead of having the underlying state equation be a deterministic differential equation, it becomes a stochastic one.

The stochastic optimal control problem resulting from a stochastic metapopulation model cannot be treated with standard deterministic techniques. Solvability of the stochastic optimal control problem can be hard to obtain. Unlike deterministic optimal control problems, there are a limited number of general techniques to be applied to stochastic optimal control problems. In this work we study the application of a Stochastic Pontryagin-type Maximum Principle. When the Stochastic Maximum Principle is applied we obtain an underlying system of forward-backward stochastic differential equations. Solutions to such system of FBSDEs solve the optimal control problem. Unfortunately, solvability of only small class of such systems has been established to date.

In Chapter 2 we considered a black bear population in the Great Smoky Mountains National Park. A random parameter was introduced through a growth rate coefficient. A QLQ stochastic optimal control problem with objective functional is set up such that its solution is an optimal harvesting rate along with optimal state trajectories. A four-step scheme studied in Yong and Zhou [1999] was an essential method in showing existence and uniqueness of solutions. A numerical implementation of the four-step scheme is conducted. Effects of using different initial conditions and variance of mast availability are examined.

In Chapter 3 we consider a larger family of stochastic optimal control problems that result in systems of FBSDEs not necessarily satisfying Lipschitz conditions on the drift term of the backward equation. Connections between the system of such FBSDEs and parabolic PDEs are studied. We have shown that under certain assumptions the solution to this kind of parabolic PDEs posses a priori bounds along with its derivatives being bounded. This transforms into bounds for both backward variables of the system of FBSDEs. Having both backwards terms bounded allows one to establish necessary conditions for Lipschitz continuity. In this chapter we revisit the black bear metapopulation model from Chapter 2. An alternative method of solving optimal control problem is derived. Using Pontryagin-type Maximum Principle a weakly coupled system of FBSDEs is generated. To solve this system of FBSDEs numerically we successfully applied an adaptation of Markovian iteration that was introduced in Bender and Zhang [2008]. We again explore different parameters for stochastic optimal control problem and compare results to results generated in Chapter 2.

Bibliography

Allen, L. J. (2003). *An Introduction to Stochastic Processes with Applications to Biology.* Prentice Hall.

Beeman, L. and Pelton, M. (1970). Seasonal foods and feeling ecology of the black bear in the smoky mountains. *Int. Conf. on Bear Research and Management 2, University of Calgary*, pages 75–107.

Bender, C. and Denk, R. (2007). A Forward Scheme for Backward SDEs. *Stochastic Processes and their Applications*, 117(1):1793–1812.

Bender, C. and Zhang, J. (2008). Time Discretization And Markovian Iteration For Coupled FBSDEs. *The Annals of Applied Probability*, 18(1):143–177.

Billingsley, P. (1999). *Convergence of Probability Measures.* Wiley, second edition.

Bismut, J. (1973). *Analyse convexe et probabilites.* Faculte des Sciences de Paris, second edition.

Brody, A. and Pelton, M. (1998). Seasonal changes in digestion in black bears. *Canad. J. Zool.*, 66:1482–1484.

During, B. and Jungel, A. (2005). Existence and uniqueness of solutions to a quasi-linear parabolic equation with quadratic gradients in the financial markets. *Nonlinear Analysis*, 62:519–544.

Ethier, S. N. and Kurtz, T. G. (1985). *Markov Processes: Characterization and Convergence.* Wiley Series in Probability and Statistics. John Wiley & Sons.

Feller, W. (1957a). *Introduction to Probability Theory and its Applications I.* Wiley, second edition.

Feller, W. (1957b). *Introduction to Probability Theory and its Applications II.* Wiley, second edition.

Gobet, E., Lemor, J.-P., and Warin, X. (2005). A Regression Based Monte Carlo Method to Solve Backaward Stochastic Differential Equations. *The Annals of Applied Probability*, 15(3):2172–2202.

Hanson, F. B. and Ryan, D. (1988). Optimal harvesting with density dependent random effects. *Natural Resource Modeling*, 2:439–455.

Jacod, J. and Shiryaev, A. N. (1987). *Limit Theorems for Stochastic Processes.* Comprehensive Studies in Mathematics. Springer-Verlag, New York.

Kallenberg, O. (2002). *Foundations of Modern Probability.* Springer-Verlag, New York, second edition.

Kloeden, P. E. and Platen, E. (1992). *Numerical Solutions of Stochastic Differential Equations.* Application of Mathematics. Springer-Verlag, New York.

Ladyženskaja, O. A., Solonnikov, V. A., and Ural'ceva, N. (1968). *Linear and Quasi-Linear Equations of Parabolic Type.* Nauka, Moscow.

Ludwig, D. (1974). *Stochastic Population Theories.* Lecture Notes in Biomathematics 3. Springer-Verlag, Berlin.

Ludwig, D. and Varah, J. M. (1979a). Optimal harvesting of a randomly fluctuating resource I: Application of perturbation methods. *SIAM J. Appl. Math.*, 37:166–184.

Ludwig, D. and Varah, J. M. (1979b). Optimal harvesting of a randomly fluctuating resource II: Numerical methods and results. *SIAM J. Appl. Math.*, 37:185–205.

Ma, J., Protter, P., and Yong, J. (1994). Solving Forward-Backward Stochastic Differential Equations Explicitly - A Four Step Scheme. *Probability Theory and Related Fields*, 98(3):339–359.

Ma, J. and Yong, J. (1995). Solvability of Forward-Backward SDEs and the Nodal Set of Hamilton-Jacobi-Bellman equations. *Chinese Ann. Math.*, B 16(3):279–298.

Ma, J. and Yong, J. (1999). *Forward-Backward Stochastic Differential Equations and Their Applications.* Heidelberg. Springer-Verlag, Berlin.

Ricciardi, L. M. (1977). *Diffusion Processes and Related Topics in Biology.* Lecture Notes in Biomathematics 14. Springer-Verlag.

Richtmyer, R. D. and Morton, K. W. (1967). *Difference Methods for Initial Value Problems.* Application of Mathematics. John Wiley and Son, New York.

Salinas, R. A., Lenhart, S., and Gross, L. J. (2005). Control of a metapopulation harvesting model for black bears. *Natural Resources Modeling*, 18(3):307–321.

Xiong, J. (2008). *An Introduction to Stochastic Filtering Theory.* Oxford University Press, Oxford.

Yong, J. (1999). Linear forward-backward stochastic differential equations. *Applied Mathematics and Optimization*, 39:93–119.

Yong, J. (2002). Stochastic optimal control and forward-backward stochastic differential equations. *Computational and Applied Mathematics*, 21:369–403.

Yong, J. and Zhou, X. Y. (1999). *Stochastic Controls Hamiltonian Systems and HJB equations.* Application of Mathematics. Springer-Verlag, New York.

Zhang, J. (2004). A numerical scheme for BSDEs. *The Annals of Applied Probability*, 14:459–488.

Appendices

Appendix A

Codes

Solution to Forward-Backward SDE using a Four-Step Scheme (`C` code)

```
/* version2.c                          */
/* 10/09/2010                          */

/*Declare Libraries used in the program*/
#include <math.h>
#include <stdlib.h>
#include <stdio.h>
#include <gsl/gsl_rng.h>
#include <gsl/gsl_randist.h>
#include <sys/time.h>

#undef max
#define max(a,b) ((a)>(b)?(a):(b))

/*Define random seed and gsl library parameters*/
struct timezone tz_dummy;
struct timeval t_snap;
unsigned long seed;
double sigma1 = 1.0;
const gsl_rng_type * T;
gsl_rng *r;

/*Routine and Function declarations*/
void weiner();
void rand_euler();
void init(void);
void beesig(int i, int j);
void thetaeval(void);
void theta_interp(double y1, double y2, int K);

/*Global variables/arrays declaration*/
double *K, *W, *dW; // random numbers, weiner process, weiner increments
double deltaT;      // time Delta
int nt;             //number of Weiner upgrades
double *X1,*X2,*Theta,*THETA,*Y1,*Y2; //forward variable grid, theta grid, backward variable grid
double *NTheta; //Duplicats
double b[3];
double sigma[3];
double dTHETA;
double delX1, delX2, delT;//deltas
double tmax=5.;//max time interval T
double Ef,Ep,a,c; //coefficients migration forest, migration park, mean, variance
int i,j,k;//indecies of spacial and time
int N,M;//number of partitions N for space and M for time step
int I=0;
```

```
/********************Input/Output files declaration******************************/
FILE *output1=NULL;
FILE *output2=NULL;
FILE *output3=NULL;
FILE *output4=NULL;
FILE *output5=NULL;
FILE *input=NULL;

/********************Main program body******************************/
int main()
{
  int l,p,z;

  /*Open Input/Output files*/
  output1=fopen("output1.dat","w");
  output2=fopen("output2.dat","w");
  output3=fopen("output3.dat","w");
  output4=fopen("output4.dat","w");
  output5=fopen("output5.dat","w");
  input=fopen("input.dat","r");


  N=20;
  M=1000;
  nt=M;

  /*Allocate Arrays*/
  X1=(double *) calloc(N+1, sizeof(double));
  X2=(double *) calloc(N+1, sizeof(double));
  Y1=(double *) calloc((nt+1), sizeof(double));
  Y2=(double *) calloc((nt+1), sizeof(double));
  Theta=(double *) calloc((N+1)*(N+1), sizeof(double));
  NTheta=(double *) calloc((N+1)*(N+1), sizeof(double));
  K=(double *) calloc((nt+1), sizeof(double));
  dW=(double *) calloc((nt+1), sizeof(double));
  W=(double *) calloc((nt+1), sizeof(double));
  THETA=(double *) calloc((N+1)*(N+1)*(nt+1), sizeof(double));

  /*More Inputs*/
  T = gsl_rng_minstd;
  r=gsl_rng_alloc (T);
  input=fopen("input.dat","r");
  fscanf(input,"%lf %lf %lf %lf %lf %lf",&Ef,&Ep,&a,&c,&Y1[0],&Y2[0]);

  int k;
  for(k=1; k<=10; k++) {
    init();
    for(z=nt-2; z>=0; z--) {
      thetaeval();
      for(l=0; l<=N; l++)
        for(p=0; p<=N; p++) {
          I=p+l*(N+1);
          THETA[I+z*(N+1)*(N+1)]=Theta[I];
          if(z==0) {
            if(p==N)
              fprintf(output5,"%lf %lf %lf\n\n",X1[l],X2[p],THETA[I+z*(N+1)*(N+1)]);
            else
              fprintf(output5,"%lf %lf %lf\n",X1[l],X2[p],THETA[I+z*(N+1)*(N+1)]);
          }
        }
    }

    weiner();
    rand_euler();
    fprintf(output1," \n");
    fprintf(output2," \n");
    fprintf(output3," \n");
    fprintf(output4," \n");
    c=c+0.1;
  }
```

```
  printf("The task is completed!\n");

  close(output1);
  close(output2);
  close(output3);
  close(output4);
  close(input);
  return(0);
}


void init()
{
  int l,p;
  double x1s,x1f,x2s,x2f;//set intervals for x-values
  x1s=-0.5;
  x1f=1.5;
  x2s=-0.5;
  x2f=1.5;

  delX1=(x1f-x1s)/N;
  delX2=(x2f-x2s)/N;
  delT=tmax/M;
  deltaT=tmax/nt;

  X1[0]=x1s;
  X2[0]=x2s;
  for(l=1; l<=N; l++) {
    X1[l]=X1[0]+l*delX1;
    X2[l]=X2[0]+l*delX2;
  }
  for(l=0; l<=N; l++)
    for(p=0; p<=N; p++) {
      I=p+l*(N+1);
      Theta[I]=0.;
      THETA[(nt-1)*(N+1)*(N+1)+I]=0.;
    }


}

void beesig(i,j)
{
  b[1]=a*X1[i]*(1-X1[i])-Ep*X1[i]+Ef*(1-X1[i])*X2[j];
  b[2]=a*X2[j]*(1-X2[j])-Ef*X2[j]+Ep*(1-X2[j])*X1[i];
  sigma[1]=c*X1[i]*(1-X1[i]);
  sigma[2]=c*X2[j]*(1-X2[j]);
}

void thetaeval()  //This function updates all  Theta1 Theta2 and Theta3
{
  int l,p;
  for(l=1; l<N; l++)
    for(p=1; p<N; p++) {
      beesig(l,p);
      I=p+l*(N+1);
      NTheta[I]=
          (+sigma[1]*sigma[1]/delX1/delX1+sigma[2]*sigma[2]/delX2/delX2)*Theta[I]
          +(-0.5*sigma[2]*sigma[2]/delX2/delX2+0.5*b[2]/delX2)*Theta[I-1]
          +(-0.5*sigma[2]*sigma[2]/delX2/delX2-0.5*b[2]/delX2)*Theta[I+1]
          +(-0.5*sigma[1]*sigma[1]/delX1/delX1+0.5*b[1]/delX1)*Theta[I-(N+1)]
          +(-0.5*sigma[1]*sigma[1]/delX1/delX1-0.5*b[1]/delX1)*Theta[I+(N+1)]
          +(0.25*sigma[1]*sigma[2]/delX1/delX2)*Theta[I-1+(N+1)]
          +(-0.25*sigma[1]*sigma[2]/delX1/delX2)*Theta[I-1-(N+1)]
          +(-0.25*sigma[1]*sigma[2]/delX1/delX2)*Theta[I+1+(N+1)]
          +(0.25*sigma[1]*sigma[2]/delX1/delX2)*Theta[I+1-(N+1)]
          +0.125*X2[p]*X2[p]/delX2/delX2*(Theta[I+1]-Theta[I-1])*(Theta[I+1]-Theta[I-1])
          -(Ef+Ep)*X1[l]*X2[p];
    }

  for(l=1; l<N; l++)
```

```
    for(p=1; p<N; p++) {
      I=p+l*(N+1);
      Theta[I]=Theta[I]-delT*NTheta[I];
    }

  Theta[0]=Theta[(N+1)+1];
  Theta[N]=Theta[(N+1)+N-1];
  Theta[N*(N+1)]=Theta[(N-1)*(N+1)+1];
  Theta[N*(N+1)+N]=Theta[(N-1)*(N+1)+N-1];

  for(p=1; p<N; p++) {
    Theta[p]=Theta[p+(N+1)];
    Theta[N*(N+1)+p]=Theta[(N-1)*(N+1)+p];
    Theta[p*(N+1)]=Theta[p*(N+1)+1];
    Theta[p*(N+1)+N]=Theta[p*(N+1)+N-1];
  }

}

void weiner()  //This proceedure generates Brownian Motion Process
{
  int l;
  gsl_rng_env_setup();

  gettimeofday(&t_snap,&tz_dummy);
  seed=t_snap.tv_sec+t_snap.tv_usec;
  gsl_rng_set(r,seed);

  for(l=0; l<=nt; l++) {
    K[l]=gsl_ran_gaussian(r,sigma1);
  }

  dW[0]=K[0]*sqrt(deltaT);
  W[0]=dW[0];
  for(l=0; l<=nt; l++) {
    dW[l]=K[l]*sqrt(deltaT);
    W[l]=W[l-1]+dW[l];
  }

}


void rand_euler()
{
  int l;
  double PY1,PY2; //Temporary Y for updates

  double overflow=0.;

  for(l=1; l<=nt; l++) {
    //this calculates interpolated values for Theta1, Theta2 and Theta3
    theta_interp(Y1[l-1],Y2[l-1],l-1);

    PY1=Y1[l-1]+(a*Y1[l-1]*(1-Y1[l-1])-Ep*Y1[l-1]+
        Ef*(1-Y1[l-1])*Y2[l-1])*deltaT+(c*Y1[l-1]*(1-Y1[l-1]))*dW[l];
    PY2=Y2[l-1]+(a*Y2[l-1]*(1-Y2[l-1])-Ef*Y2[l-1]+Ep*(1-Y2[l-1])*Y1[l-1]-
        Y2[l-1]*Y2[l-1]*dTHETA)*deltaT+(c*Y2[l-1]*(1-Y2[l-1]))*dW[l-1];

    Y1[l]=PY1;
    Y2[l]=PY2;
    overflow=overflow+(Ep+Ef)*Y1[l]*Y2[l]*deltaT;

    fprintf(output1,"%lf %lf\n",l*deltaT,Y1[l]);
    fprintf(output2,"%lf %lf\n",l*deltaT,Y2[l]);
    fprintf(output3,"%lf %lf\n",l*deltaT,overflow);
    fprintf(output4,"%lf %lf\n",l*deltaT,Y2[l]*dTHETA);
  }

}
```

```
void theta_interp(double y1,double y2,int K)
{
  int p,l;
  double THETA0;
  double THETA1,THETA2;

  double U,V;
  double delta2;

  l=(int) floor((y1+.5)/delX1);
  p=(int) floor((y2+.5)/delX2);

  I=p+l*(N+1)+K*(N+1)*(N+1);

  U=max(((y1-X1[l])/delX1),0);
  V=max(((y2-X2[p])/delX2),0);
  delta2=y2-X2[p];

  if(U==0 && V==0) {
    dTHETA=(THETA[I+1]-THETA[I])/delX2;
  } else {
    THETA0=(1-U)*(1-V)*THETA[I]+(1-U)*V*THETA[I+1]+
        U*V*THETA[I+1+(N+1)]+U*(1-V)*THETA[I+(N+1)];
    THETA1=(1-U)*THETA[I]+U*THETA[I+(N+1)];
    THETA2=(1-U)*THETA[I+1]+U*THETA[I+(N+1)+1];
    dTHETA=(1-V)*(THETA0-THETA1)/delta2+V*(THETA2-THETA0)/(delX2-delta2);
  }
}
```

Solution to Forward-Backward SDE using a Four-Step Scheme (R code Euler)

```
#====================================================================
# ver3.R
#
# Implementation of version2.c from Kirill's Dissertation
#  (with some minor mods)
#
# Tried to maintain as close to original C code as I could.
#
# Note that some arrray indices will be offset in the positive direction
# since C arrays start with index 0 and R arrays start with index 1
#
# K. Yakovlev - 2/24/2011
#
# based on work done by M. Saum - 1/16/2011
#====================================================================
rm(list=ls())

# Declare some functions
#========================= Start of Functions ======================
#-------------------------------------------------------------------
beesig <- function(i,j) {
  b[1] <<- a*X1[i] * (1-X1[i])-Ep*X1[i] + Ef*(1-X1[i])*X2[j]
  b[2] <<- a*X2[j] * (1-X2[j])-Ef*X2[j] + Ep*(1-X2[j])*X1[i]
  sigma[1] <<- c*X1[i]*(1-X1[i])
  sigma[2] <<- c*X2[j]*(1-X2[j])

}
#-------------------------------------------------------------------
thetaeval <- function() {
  for(i in (2:N)) {
    for(j in (2:N)) {
      beesig(i,j)
      NTheta[i,j] <<- (sigma[1]*sigma[1]/delX1/delX1+sigma[2]*sigma[2]/delX2/delX2)*Theta[i,j]
                      +(-0.5*sigma[2]*sigma[2]/delX2/delX2+0.5*b[2]/delX2)*Theta[i,j-1]
                      +(-0.5*sigma[2]*sigma[2]/delX2/delX2-0.5*b[2]/delX2)*Theta[i,j+1]
                      +(-0.5*sigma[1]*sigma[1]/delX1/delX1+0.5*b[1]/delX1)*Theta[i-1,j]
                      +(-0.5*sigma[1]*sigma[1]/delX1/delX1-0.5*b[1]/delX1)*Theta[i+1,j]
                      +(0.25*sigma[1]*sigma[2]/delX1/delX2)*Theta[i+1,j-1]
                      +(-0.25*sigma[1]*sigma[2]/delX1/delX2)*Theta[i-1,j-1]
                      +(-0.25*sigma[1]*sigma[2]/delX1/delX2)*Theta[i+1,j+1]
                      +(0.25*sigma[1]*sigma[2]/delX1/delX2)*Theta[i-1,j+1]
                      +0.125*X2[j]*X2[j]/delX2/delX2*(Theta[i,j+1]-Theta[i,j-1])
                      * (Theta[i,j+1]-Theta[i,j-1]) -(Ef+Ep)*X1[i]*X2[j]
    }
  }
  Theta <<- Theta - delT*NTheta

  # BC's
  Theta[1,1] <<- Theta[2,2]
  Theta[1,NN] <<- Theta[2,N]
  Theta[NN,1] <<- Theta[N,2]
  Theta[NN,NN] <<- Theta[N,N]

  Theta[1,2:N] <<- Theta[2,2:N]
  Theta[NN,2:N] <<- Theta[N,2:N]
  Theta[2:N,1] <<- Theta[2:N,2]
  Theta[2:N,NN] <<- Theta[2:N,N]

}
#-------------------------------------------------------------------
rand_euler <- function() {

  overflow <- 0.0

  for(t in (2:nnt)) {
    t1 <- t-1
    dTHETA <- theta_interp(Y1[t-1],Y2[t-1],t-1)

    Y1[t] <<- Y1[t-1]+(a*Y1[t-1]*(1-Y1[t-1])-Ep*Y1[t-1]
```

```
              + Ef*(1-Y1[t-1])*Y2[t-1])*deltaT+(c*Y1[t-1]*(1-Y1[t-1]))*dW[t];
    Y2[t] <<- Y2[t-1]+(a*Y2[t-1]*(1-Y2[t-1])-Ef*Y2[t-1]
              +Ep*(1-Y2[t-1])*Y1[t-1]- Y2[t-1]*Y2[t-1]*dTHETA)*deltaT+(c*Y2[t-1]*(1-Y2[t-1]))*dW[t-1];
    overflow <- overflow + (Ep + Ef)*Y1[t]*Y2[t]*deltaT

    #Write out to somewhere
    pstr <- sprintf("%f %f\n",t1*deltaT,Y1[t])
    cat(pstr,file=output1)
    pstr <- sprintf("%f %f\n",t1*deltaT,Y2[t])
    cat(pstr,file=output2)
    pstr <- sprintf("%f %f\n",t1*deltaT,overflow)
    cat(pstr,file=output3)
    pstr <- sprintf("%f %f\n",t1*deltaT,Y2[t]*dTHETA)
    cat(pstr,file=output4)

  }

}
#----------------------------------------------------------------------
theta_interp <- function(y1,y2,yk) {

  ii <- as.integer(floor((y1+0.5)/delX1)) + 1
  jj <- as.integer(floor((y2+0.5)/delX2)) + 1

  U <- max((y1 - X1[ii])/delX1,0.)
  V <- max((y2 - X2[jj])/delX2,0.)
  delta2 <- y2 - X2[jj]

  if(U == 0. && V == 0.) {
     dTHETA <- (THETA[yk,ii,jj+1]-THETA[yk,ii,jj])/delX2;
  } else {
    THETA0 <- (1-U)*(1-V)*THETA[yk,ii,jj]+(1-U)*V*THETA[yk,ii,jj+1]+
               U*V*THETA[yk,ii+1,jj+1]+U*(1-V)*THETA[yk,ii+1,jj];
    THETA1 <-(1-U)*THETA[yk,ii,jj]+U*THETA[yk,ii+1,jj];
    THETA2 <-(1-U)*THETA[yk,ii,jj+1]+U*THETA[yk,ii+1,jj+1];
    dTHETA <-(1-V)*(THETA0-THETA1)/delta2+V*(THETA2-THETA0)/(delX2-delta2);
  }

  return(dTHETA)
}
#----------------------------------------------------------------------
#========================== End of Functions ========================

# Declare Global Dimensions and Variables
N <- 20 # Space
NN <- N + 1
NN2 <- NN*NN
M <- 1000 # Time
nt <- M
nnt <- nt + 1
tmax <- 5 # Max t
x1s <- -0.5
x2s <- -0.5
x1f <- 1.5
x2f <- 1.5
delX1 <- (x1f - x1s)/N
delX2 <- (x2f - x2s)/N
delT <- tmax/M
deltaT <- delT
zind <- rev(1:(nt-1)) # used to index backwards in time
sigma1 <- 1.0

# Initial Allocation
X1 <- numeric(NN)
X2 <- numeric(NN)
Y1 <- numeric(nnt)
Y2 <- numeric(nnt)
NTheta <- matrix(0,nrow=NN,ncol=NN)
K <- numeric(nnt)
dW <- numeric(nnt)
W <- numeric(nnt)
```

```
THETA <- array(0,dim=c(nnt,NN,NN))
sigma <- numeric(2)
b <- numeric(2)

# read in data for run
inpdf <- read.table("./input.dat",header=TRUE)
print(inpdf)
NP <- length(inpdf$a)

#Loop over number of paths
for(k in (1:NP)) {
  a <- inpdf$a[k]
  c <- inpdf$c[k]
  Ep <- inpdf$Ep[k]
  Ef <- inpdf$Ef[k]
  Y1[1] <- inpdf$Y10[k]
  Y2[1] <- inpdf$Y20[k]
  case <- inpdf$case[k]
  seed <- inpdf$seed[k]

  # open up output file connections
  output1 <- file(sprintf("%s_1.dat",case),"w")
  output2 <- file(sprintf("%s_2.dat",case),"w")
  output3 <- file(sprintf("%s_3.dat",case),"w")
  output4 <- file(sprintf("%s_4.dat",case),"w")
  output5 <- file(sprintf("%s_5.dat",case),"w")

  # initialize values (formerly init())
  X1[1] <- x1s
  X1[2:NN] <- x1s + (1:N)*delX1
  X2[1] <- x2s
  X2[2:NN] <- x2s + (1:N)*delX2
  Theta <- matrix(0,nrow=NN,ncol=NN)
  THETA[nt,,] <- matrix(0,nrow=NN,ncol=NN)

  #Go backwards in time
  for(z in zind) {
    #Theta <- thetaeval(NTheta,Theta,delT)
    thetaeval()
    THETA[z,,] <- Theta
    if(z == 1) {
      cat("Case: ",case," Reached beginning of time\n")
      for(i in (1:NN)) {
        for(j in (1:NN)) {
          if(j == NN) {
            pstr <- sprintf("%f %f %f\n\n",X1[i],X2[j],THETA[z,i,j])
          } else {
            pstr <- sprintf("%f %f %f\n",X1[i],X2[j],THETA[z,i,j])
          }
          cat(pstr,file=output5)
        }
      }
    }
  }

  #apply stochasticity
  # replaces weiner()
  # May want to reinitialize seed for different paths, althought not sure necessary
  rngpr <- RNGkind(normal.kind="Box")
  print(rngpr)
  if(seed != 0) {
    set.seed(seed)
  }

  K <- rnorm(nnt,0,sigma1)
  dW <- sqrt(deltaT)*K
  W <- cumsum(dW)

  #apply SDE solver
  rand_euler()
```

```
  # close output files
  close(output1)
  close(output2)
  close(output3)
  close(output4)
  close(output5)

}

rm(list=ls())
```

Solution to Forward-Backward SDE using a Four-Step Scheme (R code Strong Order 1.0)

```
#====================================================================
# Explicit Order 1.0 strong scheme.R
#
# Modification of Stochastic Euler scheme to Stochastic Strong Order 1.0
#
# K. Yakovlev - 3/28/2011
#
# M. Saum - 1/16/2011
#====================================================================
rm(list=ls())

# Declare some functions
#========================== Start of Functions =======================
#-------------------------------------------------------------------
beesig <- function(i,j) {
  b[1] <<- a*X1[i] * (1-X1[i])-Ep*X1[i] + Ef*(1-X1[i])*X2[j]
  b[2] <<- a*X2[j] * (1-X2[j])-Ef*X2[j] + Ep*(1-X2[j])*X1[i]
  sigma[1] <<- c*X1[i]*(1-X1[i])
  sigma[2] <<- c*X2[j]*(1-X2[j])

}
#-------------------------------------------------------------------
thetaeval <- function() {
  for(i in (2:N)) {
    for(j in (2:N)) {
      beesig(i,j)
      NTheta[i,j] <<- (sigma[1]*sigma[1]/delX1/delX1+sigma[2]*sigma[2]/delX2/delX2)*Theta[i,j]
          +(-0.5*sigma[2]*sigma[2]/delX2/delX2+0.5*b[2]/delX2)*Theta[i,j-1] +(-0.5*sigma[2]*sigma[2]/delX2/delX2
          -0.5*b[2]/delX2)*Theta[i,j+1] +(-0.5*sigma[1]*sigma[1]/delX1/delX1+0.5*b[1]/delX1)*Theta[i-1,j]
          +(-0.5*sigma[1]*sigma[1]/delX1/delX1-0.5*b[1]/delX1)*Theta[i+1,j]
          +(0.25*sigma[1]*sigma[2]/delX1/delX2)*Theta[i+1,j-1] +(-0.25*sigma[1]*sigma[2]/delX1/delX2)*Theta[i-1,j-1]
          +(-0.25*sigma[1]*sigma[2]/delX1/delX2)*Theta[i+1,j+1] +(0.25*sigma[1]*sigma[2]/delX1/delX2)*Theta[i-1,j+1]
          +0.125*X2[j]*X2[j]/delX2/delX2*(Theta[i,j+1]-Theta[i,j-1])* (Theta[i,j+1]-Theta[i,j-1]) -(Ef+Ep)*X1[i]*X2[j]
    }
  }
  Theta <<- Theta - delT*NTheta

  # BC's
  Theta[1,1] <<- Theta[2,2]
  Theta[1,NN] <<- Theta[2,N]
  Theta[NN,1] <<- Theta[N,2]
  Theta[NN,NN] <<- Theta[N,N]

  Theta[1,2:N] <<- Theta[2,2:N]
  Theta[NN,2:N] <<- Theta[N,2:N]
  Theta[2:N,1] <<- Theta[2:N,2]
  Theta[2:N,NN] <<- Theta[2:N,N]

}
#-------------------------------------------------------------------
rand_RK <- function() {

  overflow <- 0.0

  for(t in (2:nnt)) {
    t1 <- t-1
    dTHETA <- theta_interp(Y1[t-1],Y2[t-1],t-1)

        a1=a*Y1[t-1]*(1-Y1[t-1])-Ep*Y1[t-1]+ Ef*(1-Y1[t-1])*Y2[t-1];
        a2=a*Y2[t-1]*(1-Y2[t-1])-Ef*Y2[t-1]+Ep*(1-Y2[t-1])*Y1[t-1]- Y2[t-1]*Y2[t-1]*dTHETA;
        a1a=Y1[t-1]+a1*deltaT+(c*Y1[t-1]*(1-Y1[t-1]))*sqrt(deltaT);
        a2a=Y2[t-1]+a2*deltaT+(c*Y2[t-1]*(1-Y2[t-1]))*sqrt(deltaT);
        b11=c*Y1[t-1]*(1-Y1[t-1]);
        b22=c*Y2[t-1]*(1-Y2[t-1]);
        ba11=c*a1a*(1-a1a);
        ba22=c*a2a*(1-a2a);
        db1=ba11-b11;
        db2=ba22-b22;
```

```
    Y1[t] <<- Y1[t-1]+a1*deltaT+b11*dW[t]+db1/(2*sqrt(deltaT))*(dW[t]^2-deltaT);
    Y2[t] <<- Y2[t-1]+a2*deltaT+b22*dW[t]+db2/(2*sqrt(deltaT))*(dW[t]^2-deltaT);
    overflow <- overflow + (Ep + Ef)*Y1[t]*Y2[t]*deltaT

    #Write out to somewhere
    pstr <- sprintf("%f %f\n",t1*deltaT,Y1[t])
    cat(pstr,file=output1)
    pstr <- sprintf("%f %f\n",t1*deltaT,Y2[t])
    cat(pstr,file=output2)
    pstr <- sprintf("%f %f\n",t1*deltaT,overflow)
    cat(pstr,file=output3)
    pstr <- sprintf("%f %f\n",t1*deltaT,Y2[t]*dTHETA)
    cat(pstr,file=output4)

  }

}
#---------------------------------------------------------------------
theta_interp <- function(y1,y2,yk) {

  ii <- as.integer(floor((y1+0.5)/delX1)) + 1
  jj <- as.integer(floor((y2+0.5)/delX2)) + 1

  U <- max((y1 - X1[ii])/delX1,0.)
  V <- max((y2 - X2[jj])/delX2,0.)
  delta2 <- y2 - X2[jj]

  if(U == 0. && V == 0.) {
     dTHETA <- (THETA[yk,ii,jj+1]-THETA[yk,ii,jj])/delX2;
  } else {
    THETA0 <- (1-U)*(1-V)*THETA[yk,ii,jj]+(1-U)*V*THETA[yk,ii,jj+1]+ U*V*THETA[yk,ii+1,jj+1]
                  +U*(1-V)*THETA[yk,ii+1,jj];
    THETA1 <-(1-U)*THETA[yk,ii,jj]+U*THETA[yk,ii+1,jj];
    THETA2 <-(1-U)*THETA[yk,ii,jj+1]+U*THETA[yk,ii+1,jj+1];
    dTHETA <-(1-V)*(THETA0-THETA1)/delta2+V*(THETA2-THETA0)/(delX2-delta2);
  }

  return(dTHETA)
}
#---------------------------------------------------------------------
#========================= End of Functions ========================

# Declare Global Dimensions and Variables
N <- 20 # Space
NN <- N + 1
NN2 <- NN*NN
M <- 1000 # Time
nt <- M
nnt <- nt + 1
tmax <- 5 # Max t
x1s <- -0.5
x2s <- -0.5
x1f <- 1.5
x2f <- 1.5
delX1 <- (x1f - x1s)/N
delX2 <- (x2f - x2s)/N
delT <- tmax/M
deltaT <- delT
zind <- rev(1:(nt-1)) # used to index backwards in time
sigma1 <- 1.0

# Initial Allocation
X1 <- numeric(NN)
X2 <- numeric(NN)
Y1 <- numeric(nnt)
Y2 <- numeric(nnt)
NTheta <- matrix(0,nrow=NN,ncol=NN)
K <- numeric(nnt)
dW <- numeric(nnt)
W <- numeric(nnt)
THETA <- array(0,dim=c(nnt,NN,NN))
```

```
sigma <- numeric(2)
b <- numeric(2)

# read in data for run
inpdf <- read.table("./Rinput.dat",header=TRUE)
print(inpdf)
NP <- length(inpdf$a)

#Loop over number of paths
for(k in (1:NP)) {
  a <- inpdf$a[k]
  c <- inpdf$c[k]
  Ep <- inpdf$Ep[k]
  Ef <- inpdf$Ef[k]
  Y1[1] <- inpdf$Y10[k]
  Y2[1] <- inpdf$Y20[k]
  case <- inpdf$case[k]
  seed <- inpdf$seed[k]

  # open up output file connections
  output1 <- file(sprintf("%s_1.dat",case),"w")
  output2 <- file(sprintf("%s_2.dat",case),"w")
  output3 <- file(sprintf("%s_3.dat",case),"w")
  output4 <- file(sprintf("%s_4.dat",case),"w")
  output5 <- file(sprintf("%s_5.dat",case),"w")

  # initialize values (formerly init())
  X1[1] <- x1s
  X1[2:NN] <- x1s + (1:N)*delX1
  X2[1] <- x2s
  X2[2:NN] <- x2s + (1:N)*delX2
  Theta <- matrix(0,nrow=NN,ncol=NN)
  THETA[nt,,] <- matrix(0,nrow=NN,ncol=NN)

  #Go backwards in time
  for(z in zind) {
    #Theta <- thetaeval(NTheta,Theta,delT)
    thetaeval()
    THETA[z,,] <- Theta
    if(z == 1) {
      cat("Case: ",case," Reached beginning of time\n")
      for(i in (1:NN)) {
        for(j in (1:NN)) {
          if(j == NN) {
            pstr <- sprintf("%f %f %f\n\n",X1[i],X2[j],THETA[z,i,j])
          } else {
            pstr <- sprintf("%f %f %f\n",X1[i],X2[j],THETA[z,i,j])
          }
          cat(pstr,file=output5)
        }
      }
    }
  }

  #apply stochasticity
  # replaces weiner()
  rngpr <- RNGkind(normal.kind="Box")
  print(rngpr)
  if(seed != 0) {
    set.seed(seed)
  }

  K <- rnorm(nnt,0,sigma1)
  dW <- sqrt(deltaT)*K
  W <- cumsum(dW)

  #apply SDE solver
  rand_RK()

  # close output files
  close(output1)
```

```
  close(output2)
  close(output3)
  close(output4)
  close(output5)

}

rm(list=ls())
```

Solution to Forward-Backward SDE using Markovian Iterations (R code)

```
rm(list=ls())

#options(error=recover)

#library(lattice)

# Declare the update functions
update_x <- function() {

  # Loop forward in time
  for(j in (2:(n+1))) {
    xx1 <- X[,j-1,1]
    xx2 <- X[,j-1,2]
    if( MI > 1 && (j <= n)) { upd <- predict(U2[[j]],data.frame(xx1,xx2)) } else { upd <- 0.0 }
    X[,j,1] <<-  xx1+(a*xx1* (1.0-xx1)-ep*xx1+ ef*(1.0-xx1)*xx2 )*DeltaT + cc*xx1*(1.0-xx1)*dW[,j]
    X[,j,2] <<- xx2+(a*xx2* (1.0-xx2)-ef*xx2 + ep*(1.0-xx2)*xx1 - upd*xx2*xx2/m)*DeltaT
                   + cc*xx2*(1.0-xx2)*dW[,j]
  }
}

update_y <- function() {

  # Loop backwards in time
  for(j in rev(2:n)) {
    xx1 <- X[,j+1,1]
    xx2 <- X[,j+1,2]

    if(j == n) {
      p[,n,1] <<- 0.0
      p[,n,2] <<- 0.0
    } else {
      # Predict p_{j+1}
      p[,j+1,1] <<- predict(U1[[j+1]],data.frame(xx1,xx2))
      p[,j+1,2] <<- predict(U2[[j+1]],data.frame(xx1,xx2))
    }

    # Update Y (temporary storage)
    Y[,j+1,1] <<- n * p[,j+1,1] * dW[,j+1]
    Y[,j+1,2] <<- n * p[,j+1,2] * dW[,j+1]

    # Projections on subspaces (V)
    xx1 <- X[,j,1]
    xx2 <- X[,j,2]
    V1.lm <- lm(Y[,j+1,1] ~ xx1 * xx2)
    V2.lm <- lm(Y[,j+1,2] ~ xx1 * xx2)

    # predict q_{j}
    q[,j,1] <<- predict(V1.lm,data.frame(xx1,xx2))
    q[,j,2] <<- predict(V2.lm,data.frame(xx1,xx2))

    # Update Y (temporary storage)
    pp1 <- p[,j+1,1]
    pp2 <- p[,j+1,2]
    qq1 <- q[,j,1]
    qq2 <- q[,j,2]
    Y[,j,1] <<- pp1 -DeltaT*((a*(1.0-2.0*xx1)-ep-ef*xx2)*pp1+ep*(1.0-xx2)*pp2 +cc*(1.0-2.0*xx1)*qq1)
    Y[,j,2] <<- pp2 -DeltaT*(ef*(1.0-xx1)*pp1 +(a*(1.0-2.0*xx2)-ef-ep*xx1 -2.0/m*pp2*xx2)*pp2
                     +cc*(1.0-2.0*xx2)*qq2 -(ep+ef)*xx1 +pp2*pp2*xx2/m)

    # Projections on subspaces (U)
    U1[[j]] <<- lm(Y[,j,1] ~ xx1 * xx2)
    U2[[j]] <<- lm(Y[,j,2] ~ xx1 * xx2)

  }

  # Handle the updates at 0
  xx1 <- X[,2,1]
  xx2 <- X[,2,2]
```

```
  p[,2,1] <<- predict(U1[[2]],data.frame(xx1,xx2))
  p[,2,2] <<- predict(U2[[2]],data.frame(xx1,xx2))

  q1_0 <<- sum(p[,2,1]*dW[,2])/L * n
  q2_0 <<- sum(p[,2,1]*dW[,2])/L * n

  q[,1,1] <<- q1_0
  q[,1,2] <<- q2_0

  pp1 <- p[,2,1]
  pp2 <- p[,2,2]

  p1_0 <<- sum( pp1 -((a*(1.0-2.0*X1_0)-ep -ef*X2_0)*pp1 +ep*(1.0-X2_0)*pp2 +cc*(1.0-2.0*X1_0)*q1_0
                       -(ep+ef)*X2_0)*DeltaT)/L
  p2_0 <<- sum( pp2 -(ef*(1.0-X1_0)*pp1 +(a*(1.0-2.0*X2_0)-ef-ep*X1_0 -2.0/m*pp2*X2_0)*pp2
                       + cc*(1.0-2.0*X2_0)*q2_0 -(ep+ef)*X1_0 +pp2*pp2*X2_0/m)*DeltaT)/L

  p[,1,1] <<- p1_0
  p[,1,2] <<- p2_0

}

######################## Main Program #########################


# Input Parms
inp <- read.table("input.dat",header=FALSE,
            col.names=c("Case","n","Tmax","L","Mmax","Ef","Ep","a","c","X10","X20","seed"))
print(inp)

case.f <- factor(inp$Case)
case.n <- nlevels(case.f)
case.c <- levels(case.f)

for (k in (1:case.n)) {
  Case <- as.character(inp$Case[k])
  L <- inp$L[k]
  n <- inp$n[k]
  Tmax <- inp$Tmax[k]
  MMAX <- inp$Mmax[k]
  ef <- inp$Ef[k]
  ep <- inp$Ep[k]
  a <- inp$a[k]
  cc <- inp$c[k]
  X1_0 <- inp$X10[k]
  X2_0 <- inp$X20[k]
  seed <- inp$seed[k]

  # Parameters
  DeltaT <- Tmax/n;
  MI <- 1
  m <- 1.0
  M <- 1.0
  sigma1=1.0

  q1_0 <- 0.0
  q2_0 <- 0.0
  p1_0 <- 0.0
  p2_0 <- 0.0

  X <- array(0,dim=c(L,n+1,2))
  Y <- array(0,dim=c(L,n+1,2))
  p <- array(0,dim=c(L,n+1,2))
  q <- array(0,dim=c(L,n+1,2))
  U1 <- vector(mode="list", length=n+1)
  U2 <- vector(mode="list", length=n+1)

  dW <- array(0,dim=c(L,n+1))
  W <- array(0,dim=c(L,n+1))
  rtime <- array(0,dim=(n+1))

```

```
# The Random Walks
if ( seed > 0) {
  set.seed(seed)
}

for (i in (1:L)) {
  dW[i,] <- rnorm(n+1, sd=sqrt(sigma1))*sqrt(DeltaT)
  W[i,2:(n+1)] <- cumsum(dW[i,])[1:n]
}

for (i in 2:(n+1)) {
  rtime[i] <- (i-1)*DeltaT
}

# Initialize
X[,1,1] <- X1_0
X[,1,2] <- X2_0
p1_00 <- 0.0
p2_00 <- 0.0
q1_00 <- 0.0
q2_00 <- 0.0
X1_nn <- 0.0
X2_nn <- 0.0

# Open output files
fstr <- sprintf("p01_X_%s.dat",Case)
dumpX <- file(fstr,"wt")
pstr <-  sprintf("Case MI ntime rtime X1 X2 p1 p2 q1 q2 u X3\n");
cat(pstr,file=dumpX)
fstr <- sprintf("p01_N_%s.dat",Case)
dumpN <- file(fstr,"wt")
pstr <- sprintf("Case MI X1n X2n Xn p1n p2n pn\n");
cat(pstr,file=dumpN)

# Markovian iteration loop
for (i in (1:MMAX)) {
  update_x()
  update_y()

  normp1 <- abs(p1_0-p1_00)
  normp2 <- abs(p2_0-p2_00)
  normp <- normp1 + normp2

  X1_n <- mean(X[,n+1,1])
  X2_n <- mean(X[,n+1,2])
  normx1 <- abs(X1_n-X1_nn)
  normx2 <- abs(X2_n-X2_nn)
  normx <- normx1 + normx2

  p1_00 <- p1_0
  p2_00 <- p2_0
  X1_nn <- X1_n
  X2_nn <- X2_n

  # Print out summary
  cat("Case= ",Case," MI= ",MI," normx= ",normx," normp= ",normp,"\n")
  pstr <- sprintf("%s %d %20.14e %20.14e %20.14e %20.14e %20.14e %20.14e\n",
                Case,MI, normx1, normx2, normx, normp1, normp2, normp )
  cat(pstr,file=dumpN)

  for(j in (1:(n+1))) {
    pstr <- sprintf("%s %d %d %20.14e %20.14e %20.14e %20.14e %20.14e %20.14e %20.14e %20.14e %20.14e\n",
           Case,MI, 1, (j-1)*DeltaT, X[1,j,1], X[1,j,2],p[1,j,1],p[1,j,2],q[1,j,1],q[1,j,2],
           -X[1,j,2]*p[1,j,2]/m,(ef+ep)*X[1,j,1]*X[1,j,2])
    cat(pstr,file=dumpX)
  }

  # Update Markovian Iteration
  MI <- MI + 1
}
```

```
  # Close output files
  close(dumpX)
  close(dumpN)
}

rm(list=ls())
```

Appendix B

Graphs

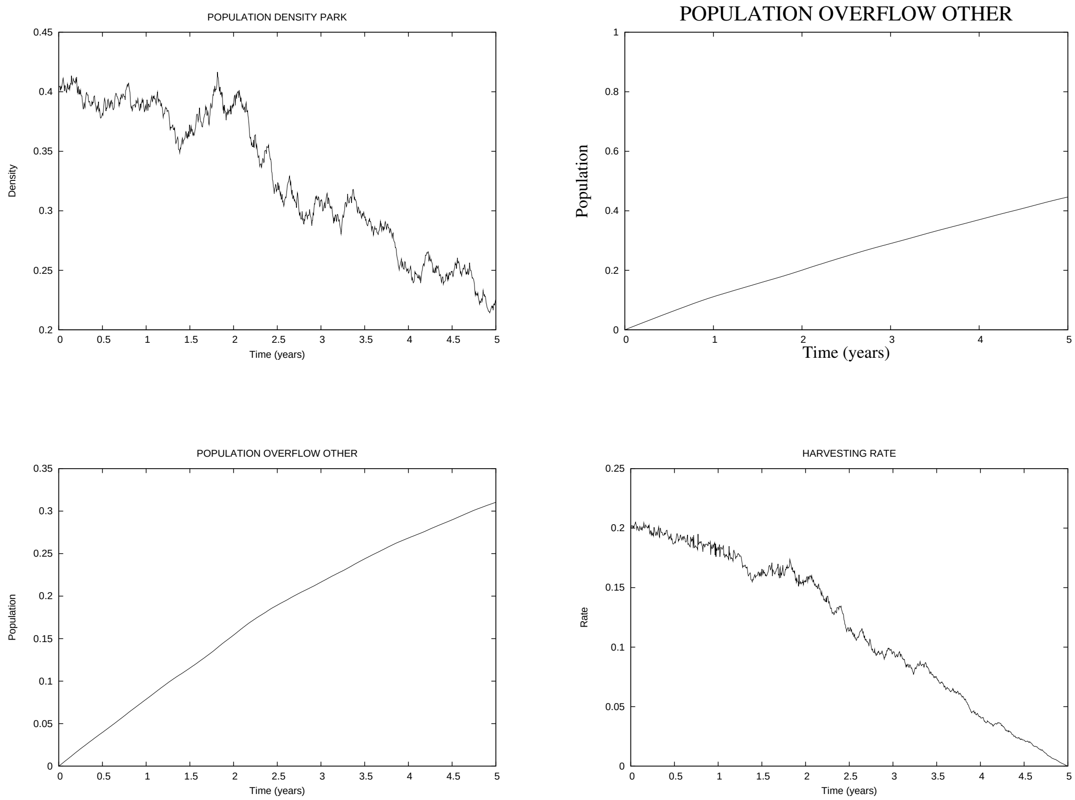

Figure B.1: Population results of the optimal control problem using one sample path and the variance of the population growth $c = 0.2$.

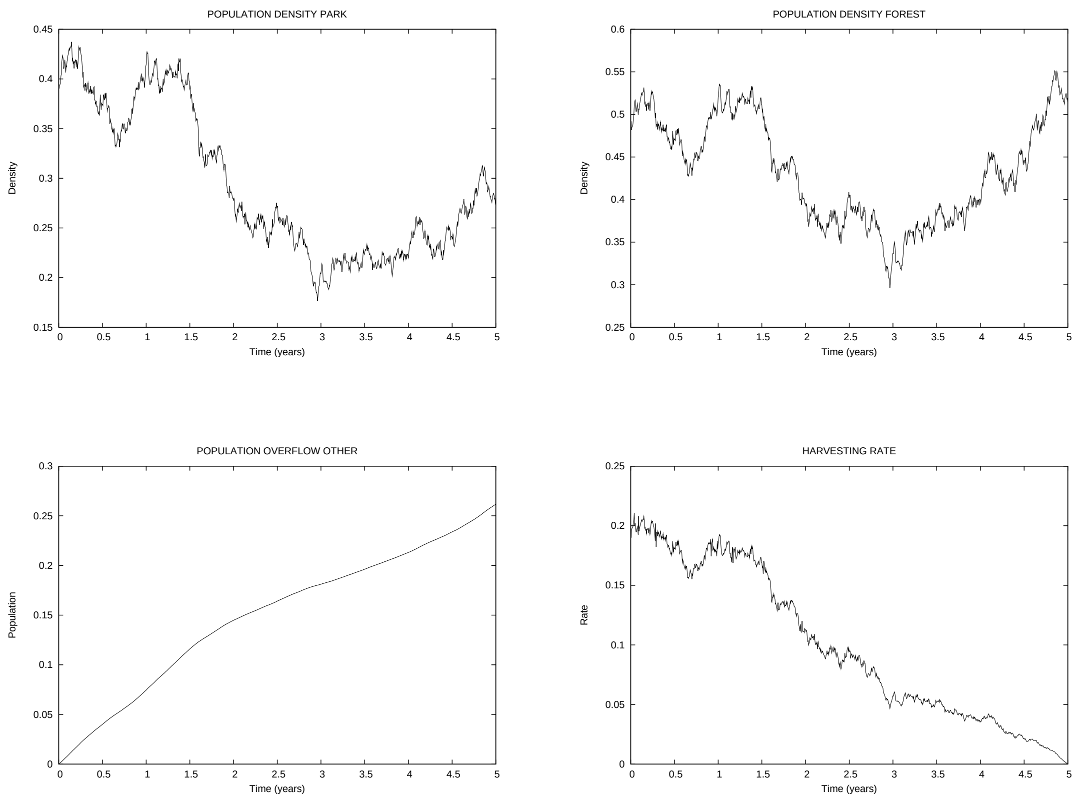

Figure B.2: Population results of the optimal control problem using one sample path and the variance of the population growth $c = 0.3$.

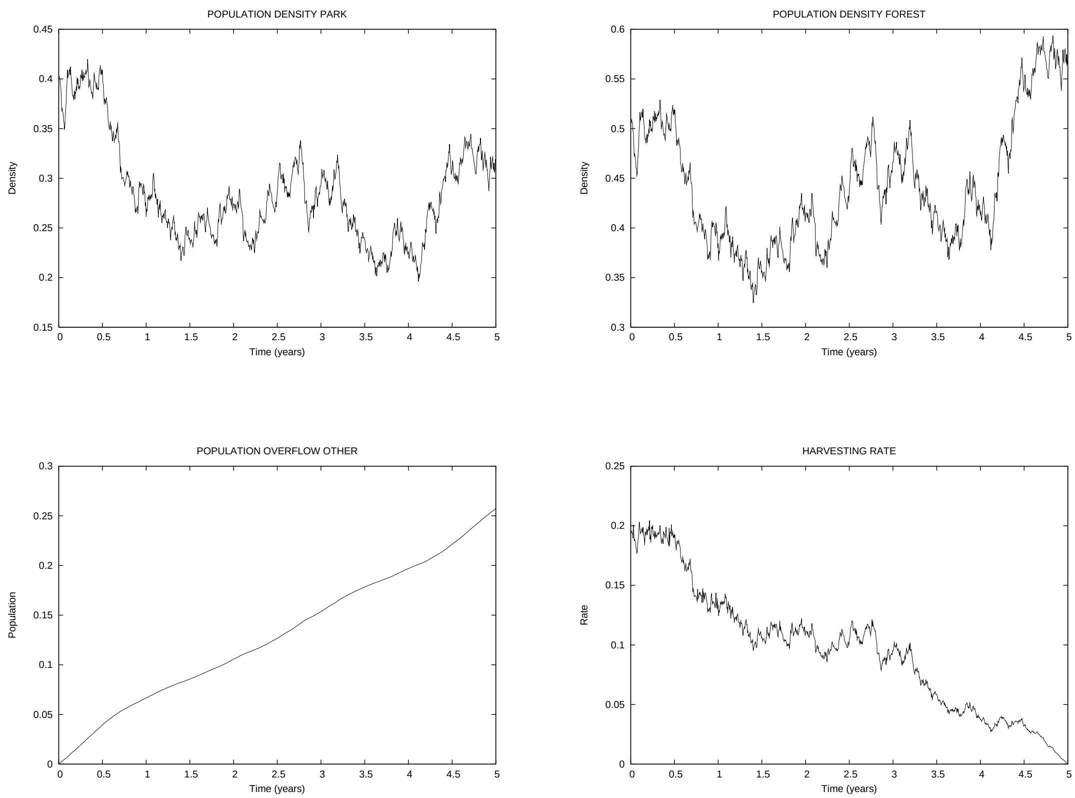

Figure B.3: Population results of the optimal control problem using one sample path and the variance of the population growth $c = 0.4$.

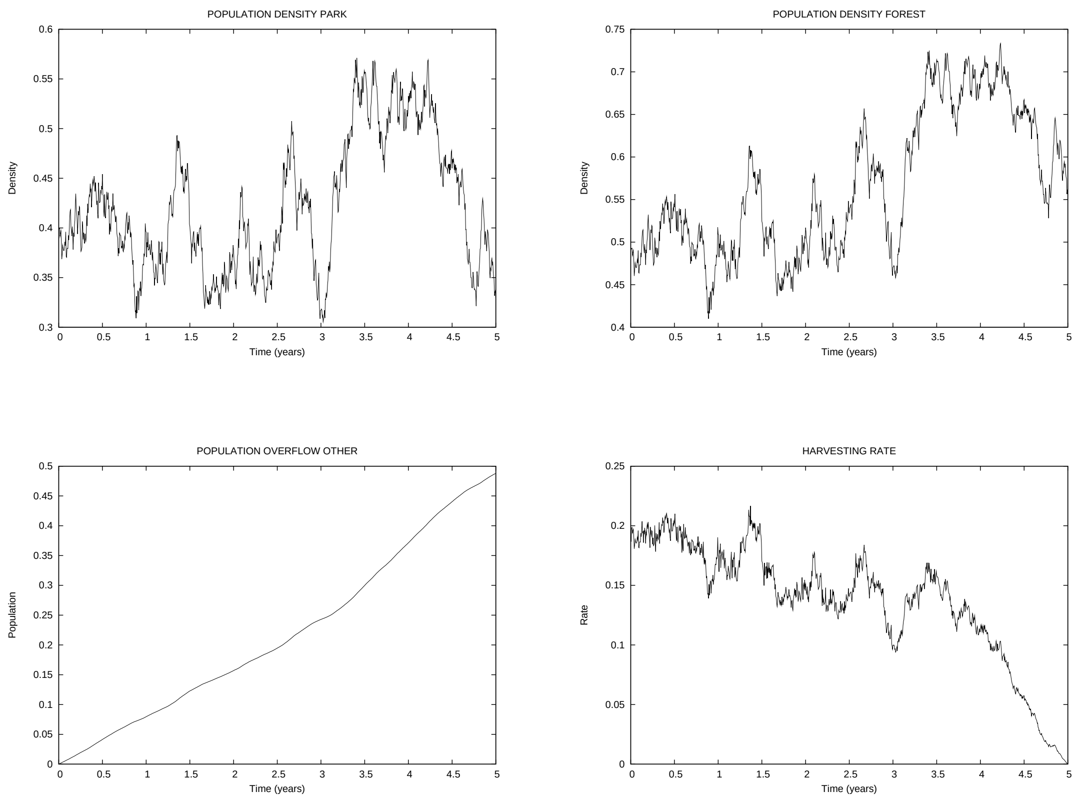

Figure B.4: Population results of the optimal control problem using one sample path and the variance of the population growth $c = 0.6$.

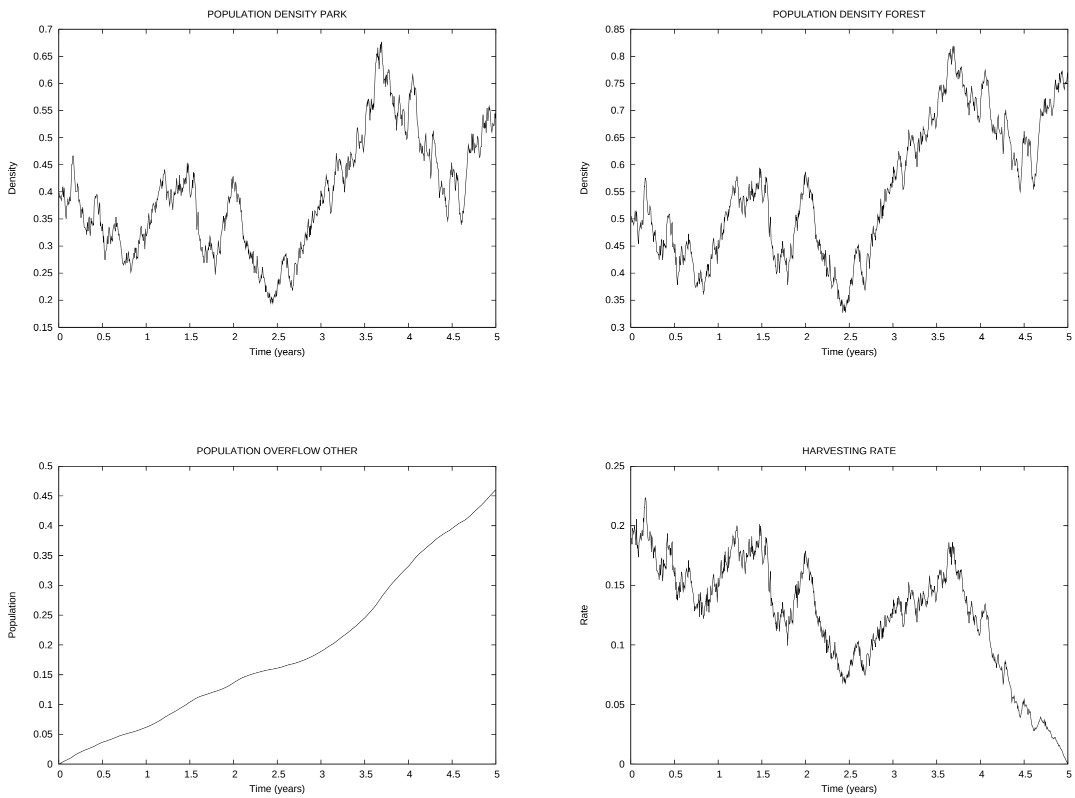

Figure B.5: Population results of the optimal control problem using one sample path and the variance of the population growth $c = 0.7$.

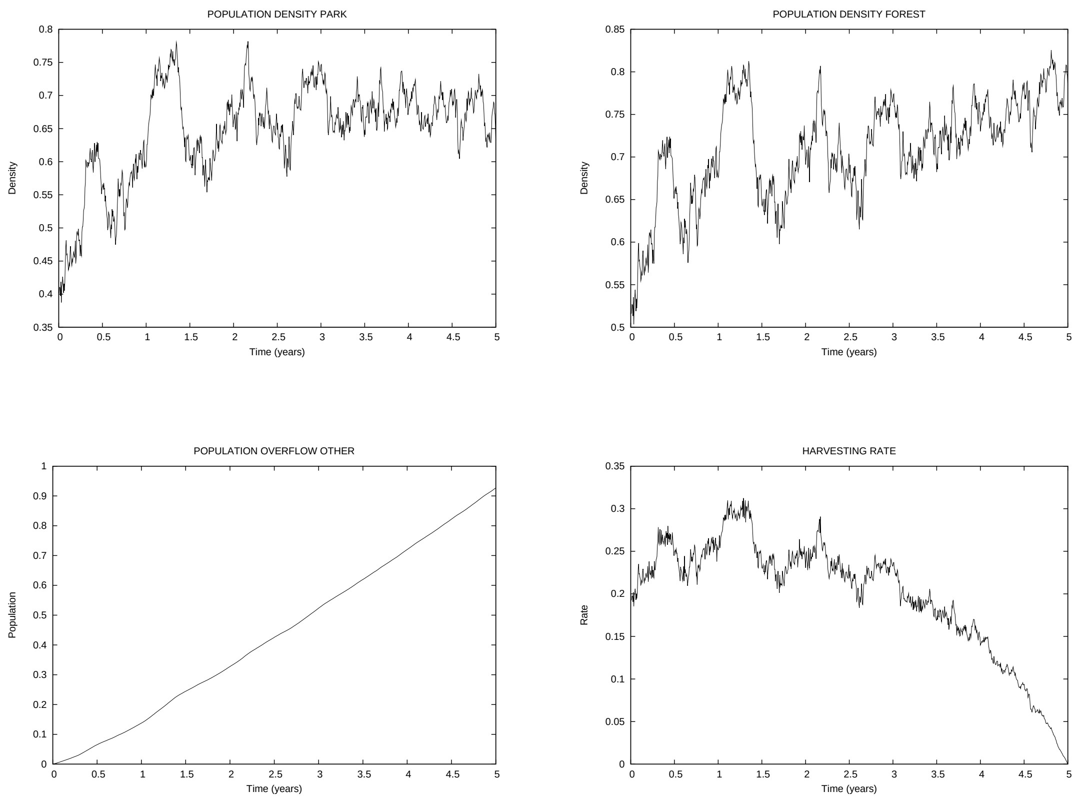

Figure B.6: Population results of the optimal control problem using one sample path and the variance of the population growth $c = 0.8$.

Vita

Kirill Yakovlev was born on January 24 in St. Petersburg, Russia where he attended school, played basketball, and enjoyed time with his parents and family. In 1997 he graduated from high school #281 and then attended Saint Petersburg State University for a brief period before moving to America in December 1998. He attended college in the United States and graduated from the University of the Cumberlands in Kentucky in summer of 2002. Subsequently, he attended graduate school at the University of Tennessee. He worked as a Graduate Teaching Assistant and was later promoted to Graduate Teaching Associate. He also enjoyed a partial fellowship with compensation and reduced teaching load provided by the UT Science Alliance. As a young researcher, Kirill has taken classes, attended conferences, and given presentations on the topic of probability. Kirill is as a Statistical Consultant since fall 2009.