The Pennsylvania State University
The Graduate School
College of Engineering

OPTIMAL CONTROL, DIFFERENTIAL VARIATIONAL INEQUALITIES, AND THEIR APPLICATION TO TRAFFIC SCIENCE, REVENUE MANAGEMENT AND SUPPLY CHAINS

A Dissertation in
Industrial and Manufacturing Engineering
by
Yiou Wang

Submitted in Partial Fulfillment
of the Requirements
for the Degree of

Doctor of Philosophy

May 2018

The dissertation of Yiou Wang was reviewed and approved* by the following:

Terry L. Friesz
Harold & Inge Marcus Chaired Professor of Industrial and Manufacturing Engineering
Dissertation Co-Advisor, Co-Chair of Committee

Tao Yao
Associate Professor of Industrial and Manufacturing Engineering
Dissertation Co-Advisor, Co-Chair of Committee

Uday V. Shanbhag
Gray and Sheila Bello Chair Professor of Industrial and Manufacturing Engineering
Committee Member

Alberto Bressan
Eberly Family Chair Professor of Mathematics
Committee Member

Janis Terpenny
Professor and Peter and Angela Dal Pezzo Department Head of Industrial and Manufacturing Engineering

*Signatures are on file in the Graduate School.

Abstract

Optimal control problems and differential Nash games have been employed by many scholars in the study of dynamic pricing, supply chain management and transportation network flow problems. This dissertation emphasizes the extension of frequently employed deterministic, open-loop modeling paradigms into feedback and stochastic cases respectively with a focus on the computational perspective.

For the feedback differential Nash games, this dissertation briefly reviews the classical theory of Hamilton-Jacobi-Bellman equation and the general technique to synthesis feedback optimal control from its solution. Such techniques are then applied to the investigation of a dynamic competitive pricing problem of perishable products with fixed initial inventories (DPFI). Other qualitative analysis and numerical extensions of the DPFI model are also provided.

In the study of differential Nash games with Ito-type of stochastic dynamics, this dissertation starts from reviewing the stochastic maximum principle. It then proposes stochastic differential variational inequality (S-DVI) as the necessary condition for stochastic differential Nash games. As an application, this dissertation provides formulation, qualitative analysis and algorithm for a stochastic differential oligopsony problem where multiple agents compete in the procurement of key raw material which follows Ito-type of stochastic price dynamics.

Table of Contents

List of Figures

List of Tables

Acknowledgments

I would like to express my special appreciation and thanks to my research advisor Dr. Terry L. Friesz for his guidance, motivation and patience during my graduate studies at Penn State. He has provided me with the opportunity to work on interesting research projects in an engaging learning environment in his laboratory. I could not have imagined having a better advisor and mentor for my Ph.D study.

I would also like to express my thanks to my research co-advisor Dr. Tao Yao for his insightful comments on my research and career developments as well as daily life. He has always been the role model to me.

I thank my supervisory committee members: Dr. Uday V. Shanbhag, Dr. Alberto Bressan and Dr. Paul M. Griffin for their support and guidance on my graduate research.

In addition, special thanks from me to Dr. Ke Han for all his support, guidance and all collaborative research opportunities he has been providing. Another special thanks to my life-long friend Simin Feng, and congratulations on his achievements.

My thanks are extended to the Friesz-Yao research group members who have been part of my experience in this lab over the past years: Amir Meimand, Sai Zhang, Xiaohuang "Hunter" Wu, Andreas Thoreson, Pedro Neto, Hongcheng Liu, Yuqi Sun, Wenjing Song, Zhaohu "Jonathan" Fan, Afrooz Ansaripour, Haipeng Wang and Xue Wang. Those discussions and debates with you have become my best memories.

Last but not least, I would like to thank my family for their unconditional love, encouragement and support through this process.

Dedication

To my grandfathers: Zhongtian Sun and Xueshun Wang

Chapter 1 Introduction

This chapter provides a general description of the motivation of the problems, the background on methodology, and the applications proposed by this dissertation. It also introduces the contribution and structure of each chapter.

Since the seminal work of von Neumann and Morgenstern in 1944 [143], game theory has become one of the most recognized research topic with "central importance" (Kuhn et al.) [85]. Especially, the concept of a Nash equilibrium, in which agents are assumed to know the each others' equilibrium strategies, and no agents could benefit from unilateral actions, has become one of the most widely studied and applied concepts in mathematics, economics, operations research and engineering, leading to a host of insightful models into real-world phenomena from warfare to policy choices to salary negotiations.

The notion of differential games is brought up towards the modeling and analysis for games taking place over an entire decision horizon instead of only one instant of time, and typically, the strategy applied by each player can be described by a function of time. For the earliest models in this domain, please refer to Issacs [74]. For a general dynamic game, the time interval can also take a discrete set of values, but in this dissertation, we use the term "differential games" to refer to those models with time variable t in a continuous interval $[t_0, t_f]$. In this context, each agent's decision will be dependent on the solution of a dynamic optimization/optimal control problem parameterized by the other agents strategies. Specifically, in this dissertation we summarize, analyze and extend one model of differential game with the application in revenue management.

It is well known that Pontryagin's Maximum Principle (PMP) and Bellman's Principle Of Optimality (POO) are two major approaches towards optimal control problems both in deterministic and stochastic settings. Interestingly, PMP and POO are two theories that have been developed separately and independently to investigate the same class of problem. More specifically, either a system of ordinary differential equations (in PMP) or a system of partial differential equations (in POO) can be used towards the solution of optimal/equilibrium strategies. In fact, such phenomenon is not alone in the history of scientific discoveries; other examples are: Hamilton's canonical system (ODE) and Hamilton-Jacobi equation (PDE) in classical mechanics; and different representations of quantum mechanics. Please see an insightful high-level discussion of this type of queries in the preface of Yong and Zhou [148].

1.1 About This Dissertation

In this dissertation, we review the relevant mathematical results on mathematical programming, optimal control and differential games, paying special attention to comparing the PMP and POO perspectives. From a computational view point, we try to investigate three topics that are less visited by existing literature, each with interesting and unique applications: dynamic competitive pricing and revenue management computed with differential variational inequalities, feedback generalization of the last model computed with HJB-PDEs and finally, stochastic dynamic oligopolies computed with stochastic maximum principle.

1.2 Organization and Contributions

In this subsection we describe the organization of this dissertation and introduce the contributions of each part. The first part of this dissertation consists of Chapter 2, 3 and 4, in which existing results and theoretical methodologies are reviewed to prepare the application in later chapters. Each of these individual chapters can also be regarded as a tutorial of the specific topic:

- Chapter 2 presents preliminary concepts and results necessary for the presentation and understanding of the rest of this dissertation: background from real and functional analysis, optimization and variational inequalities. It then introduces the concept of differential Nash games in deterministic, open loop settings and reviews how a differential variational inequality could be used as a characterization of such differential Nash equilibrium.

- Chapter 3 presents a tutorial on solving deterministic optimal control and differential games with state feedbacks. It also includes numerical examples and case studies to illustrate the effectiveness of the methodologies reviewed. The main results from the famous linear-quadratic problem are also presented.

- Chapter 4 discusses another important case of optimal control and differential games when the state dynamics are stochastically driven by Brownian motions. Both PMP and POO based methodologies are reviewed with an emphasize on stochastic PMP and the dynamic dual variables. Finally, in parallel to the notion of differential variational inequalities (DVI) we propose the stochastic DVI and establish such notion as an effective characterization of stochastic Nash equilibrium.

The second part is of this dissertation consists of Chapter 5, 6 and 7; and there is a correspondence of each of these chapters with one of the previous theoretical chapters:

- Chapter 5 first presents Dynamic Pricing with Fixed Inventory (DPFI) model and then provides insights into its connections with DVI. The DPFI model is designed to capture the essence of revenue management: opportunities lost cannot be saved, and costs are fixed so that profit maximization is the same as revenue maximization. Additionally, by controlling both price and sales, demand may be rejected. Later in this chapter, the DPFI model is extended to a case with dual time scale where the inventory decisions and pricing/demand allocation decisions are made on different time scales.

- Chapter 6 further extends the DPFI problem by first introducing its equivalent state dynamics. The agent's best response problem is then formulated and analyzed as optimal control problem with state feedback. Furthermore, the feedback equilibrium is presented and solved with numerical examples given.

- Chapter 7 is application of stochastic DVI developed in Chapter 4 in the study of supply chain problems. Especially this chapter considers the lesser explored problem of stochastic differential monopsony/oligopsony, where the dynamic procurement and inventory decisions are made. Qualitative analysis are provided.

Chapter 2 Mathematical Preliminaries

In this chapter, we summarize the basic results in mathematical programming, open-loop optimal control theory, along with Nash games and variational inequalities in various forms. Direct application of these methodologies could be found in Chapter 5, where a dynamic competitive revenue management model in continuous time is reviewed and extended. These results also serve as starting points for further development of this dissertation in feedback problems (Chapter 3) and stochastic optimal control and differential games (Chapter 4).

2.1 Elements of Analysis

We start with listing some necessary definitions and theorems without proof, in order to facilitate the formulation and analysis of dynamic optimization problem in continuous time. Most of these definitions and theorems are reproduced from Chapter 10 of Minoux(1986) [101]. Other references on real and functional analysis includes: Royden and Fitzpatrick (1988) [125] and Bressan (2012) [30].

Definition 2.1. *(Hilbert Space) A vector space V with a scalar product $\langle \cdot, \cdot \rangle$ is called a Hilbert space if V is complete for the topology associated with the norm*

$$\|v\| = [\langle v, v \rangle]^{1/2}.$$

A Hilbert space is a Banach space in which the norm derives from a scalar product.

This definition is from Minoux (1986) [101] Definition 10.7, we also refer to Theorem 10.2 of this textbook about the commonly used Riesz representation theorem:

Theorem 2.1. *(Riesz Rrepresentation Theorem) Let V be a Hilbert space and $L \in V^*$ a continuous linear form on V. Then there exists a unique element $u_L \in V$ such that*

$$L(v) = \langle u_L, v \rangle \qquad \forall v \in V$$

and

$$\|L\|_{V^*} = \|u_L\|_V$$

Conversely, we can associate with each $u \in V$ the continuous linear form L_u defined by

$$L_u(v) = \langle u, v \rangle \qquad \forall v \in V$$

In addition, following Definition 10.10 and 10.11 of Minoux (1986) [101] we list:

Definition 2.2. *(G-derivative) J has a directional derivative (or a differential in the sense of Gateaux) at $v \in V$ in the direction $\varphi \in V$ if*

$$\lim_{\theta \to 0} \frac{J(v + \theta\varphi) - J(v)}{\theta}$$

exists. Denote this limite by $\delta J(v, \varphi)$. If for all $\varphi \in V, \delta J(v, \varphi)$ exists, then J is said to be G-differentiable at $v \in V$.

Definition 2.3. *(Gradient) Let V be a Hilbert space with the scalar product $\langle \cdot, \cdot \rangle$. If J is G-differentiable at $v \in V$, and if $\delta J(v, \varphi)$ is a continuous linear form with respect to φ, then by Theorem 2.1, there exists an element $J'(v) \in V$, called the gradient of J at v, such that*

$$\delta J(v, \varphi) = \langle J'(v), \varphi \rangle .$$

2.2 Optimization Problems

2.2.1 Nonlinear Programming

Nonlinear programming (NLP) is the process of solving a minimization or maximization problem defined by a system of equalities and inequalities, over a set of real variables. In particular, it treats nonlinear objective functions and constraints.

In this section we list a general formulation of the NLP problem and its necessary conditions following Bazaraa et al. (2013). For more comprehensive review on the theory and application of NLP, see for example: Bazaraa et al. (2013) [11], Fiacco and McCormick (1990) [51], Luenberger and Ye (1984) [93]. Throughout the rest of this dissertation, we consider the following nonlinear programming problem from Bazaraa et al. (2013) [11] as a "standard form":

Problem 2.1. *(Nonlinear Programming)*

$$\begin{aligned} & \min f(x) \\ s.t.\ g_i(x) &\leq 0 \quad \text{for } i = 1, \cdots, m \\ h_i(x) &= 0 \quad \text{for } i = 1, \cdots, l \\ x &\in X \end{aligned}$$

Three types of approaches are commonly used as necessary conditions to characterize the local optimal solution to the nonlinear programming problem 2.1: the necessary conditions expressed in conic sets, the Fritz John (FJ) conditions, and the Karush-Kuhn-Tucker (KKT) conditions. In order to validate the KKT conditions, the so-called constraint qualifications must be assumed. Here we list an example of KKT necessary conditions. Please refer to Chapter 4 and 5 of Bazaraa et al. (2013) [11] for further details.

Theorem 2.2. *(KKT Necessary Condition) Let X be a nonempty open set in $\mathbb{R}^n$, and let $f : \mathbb{R}^n \to \mathbb{R}, g_i : \mathbb{R}^n \to \mathbb{R}$ for $i = 1, \cdots, m$, and $h_i : \mathbb{R}^n \to \mathbb{R}$ for $i = 1, \cdots, l$. Consider Problem 2.1, let x^* be a feasible solution and let $I = \{i : g_i(x^*) = 0\}$. Suppose that f and g_i are differentiable at x^* and that each h_i is continuously differentiable at x^* for $i = 1, \cdots, l$. Further, suppose that $\nabla g_i(x^*)$ for $i \in I$ and $\nabla h_i(x^*)$ for $i = 1, \cdots, l$ are linearly independent. If x^* solves Problem 2.1 locally, there exist unique scalars μ_i for $i \in I$ and ν_i for $i = 1, \cdots, l$ such that:*

$$\begin{aligned} \nabla f(x^*) + \sum_{i=1}^{m} \mu_i \nabla g_i(x^*) + \sum_{i=1}^{l} \nu_i \nabla h_i(x^*) &= 0 \\ \mu_i g_i(x^*) &= 0 \quad \text{for } i = 1, \cdots, m \\ \mu_i &\geq 0 \quad \text{for } i = 1, \cdots, m \end{aligned}$$

It is a well known fact that convexity plays an important role in nonlinear optimization, and monotonicity properties are linked with the gradients of convex functions. See Mas-Colell (1989) [98], Hadjisavvas et al. (2006) [67] and Facchinei and Pang (2007) [49] for detailed reviews on different definitions of (generalized) convexity, (generalized) monotonicity, and the relationships among different concepts. Here we here list some definitions and results related to (generalized) monotonicity of a vector function, which will be employed for our further analysis. Also see Rockafellar (1970) [122] for a collection of theory of convex sets and functions with a focus on applications on optimization problems.

Definition 2.4. *(Convexity) Let X be a nonempty convex set in $\mathbb{R}^n$, and let $f : X \subseteq \mathbb{R}^n \to \mathbb{R}$, then f is said to be*
(i) pseudo convex at $\overline{x} \in X$ if for $\forall x \in X$

$$\nabla f(\overline{x})^T (x - \overline{x}) \geq 0 \Rightarrow f(x) \geq f(\overline{x});$$

(ii) convex at $\overline{x} \in X$ if for $\forall \lambda \in (0, 1)$, and $\forall x \in X$

$$f[\lambda \overline{x} + (1 - \lambda)x] \leq \lambda f(\overline{x}) + (1 - \lambda) f(x)$$

(iii) strictly convex at $\overline{x} \in X$if for $\forall \lambda \in (0, 1)$, and $\forall x \in X$, $x \neq \overline{x}$

$$f[\lambda \overline{x} + (1 - \lambda)x] < \lambda f(\overline{x}) + (1 - \lambda) f(x)$$

Definition 2.5. *(Monotonicity) Let $F : \mathbb{R}^n \to \mathbb{R}^n$ be a vector function defined on $X \subseteq \mathbb{R}^n$, then F is said to be*
(i) pseudo monotone on X if for $\forall x, y \in X$

$$(x - y)^T F(y) \geq 0 \Rightarrow (x - y)^T F(x) \geq 0;$$

(ii) monotone on X if for $\forall x, y \in X$

$$(F(x) - F(y))^T (x - y) \geq 0$$

(iii) strictly monotone on X if for $\forall x, y \in X$ and $x \neq y$

$$(F(x) - F(y))^T (x - y) > 0$$

(iv) strongly monotone on X if there exist a positive constant C such that for $\forall x, y \in X$

$$(F(x) - F(y))^T (x - y) \geq C \|x - y\|^2$$

The characterization of (generalized) monotonicities of a vector function, which will be employed for our further analysis are provided by the following theorem. The intuition is similar to the case of a 1-dimensional function: the Jacobian of a continuously differentiable, (generalized) monotone function should satisfy some regularities such as positive semi-definiteness.

Theorem 2.3. *(Jacobian of Monotone Functions) Let $F : \mathbb{R}^n \to \mathbb{R}^n$ be continuously differentiable function defined on a nonempty open, convex set $X \subseteq \mathbb{R}^n$, and denote its Jacobian by $JF(x)$. We have the following regarding to its (generalized) monotonicity:*
(i) F is monotone on X if and only if $JF(x)$ is positive semidefinite for all $x \in X$;
(ii) F is strictly monotone on X if $JF(x)$ is positive definite for all $x \in X$;
(iii) F is strongly monotone on X if and only if $JF(x)$ is uniformly positive definite for all $x \in X$, which means there exist positive constant C such that for $\forall x \in X$:

$$y^T JF(x) y \geq C \|y\|^2, \quad \forall y \in \mathbb{R}^n$$

Proof. See details from Facchinei and Pang (2007) [49], Proposition 2.3.2 and Hadjisavvas et al. (2006) [67], Theorem 9.6. □

The condition for a differentiable function to be pseudo-monotone is more complicated, we list without proof the following theorem based on Brighi et al. (2003) [33] and Hadjisavvas et al. (2006) [67], Theorem 2.8;

Theorem 2.4. *(Jacobian of Pseudomonotone Functions) Assume that $F : \mathbb{R}^n \to \mathbb{R}^n$ be continuously differentiable function defined on a nonempty open, convex set $X \subseteq \mathbb{R}^n$, and denote its Jacobian by $JF(x)$. Then F is pseudomonotone on X if and only if the following two conditions hold:*
(i)$JF(x)$ is positive semi-definite on a simplex for all $x \in X$:

$$F(x)^T y = 0 \Rightarrow y^T JF(x) y \geq 0, \quad \forall y \in \mathbb{R}^n$$

(ii) for all $x \in X$

$$F(x) = 0, JF(x)y = 0$$
$$\Rightarrow \forall \bar{t} > 0 \exists t \in (0, \bar{t}] \text{ such that } F(x + ty)^T y \geq 0$$

2.2.2 Infinite Dimensional Mathematical Program

In applications there are many optimization problems in which the optimization problem is not presented in a finite dimensional real vector space, but on functional spaces defined on, for example, continuous time $[t_0, t_f]$. The review of such problems starts from the minimization of a functional:

Problem 2.2. *(Minimization of a Functional) Let $J(v)$ be a functional on V. Consider the following minimization problem:*

$$\min_{v \in U \subset V} J(v)$$

Still, a good reference of existence (Theorem 10.3) and necessary condition (Theorem 10.6) for the problem above is Minoux (1986) [101]:

Theorem 2.5. *(Weierstrass existence) If the subset $U \subset V$ is strongly (weakly) compact, and if $J(v)$ is strongly (weakly) continuous on U, then Problem 2.2 has an optimal solution in U.*

Theorem 2.6. *(Necessary condition) Let $J(v)$ be G-differentiable at v^*, and let $U \subset V$ be convex. A necessary condition for v^* to be a minimum of problem 2.2 is the following variational inequality:*

$$\delta J(v^*, v - v^*) \geq 0 \qquad \forall v \in U$$

Notice that this necessary condition is presented in the form of a Variational Inequality (VI). Further on the minimization of functionals, we give another problem formulation in the form of an infinite dimensional mathematical program reported by Ritter (1967) [119] and re-casted by Friesz (2010) [54].

Problem 2.3. *(Infinite dimensional mathematical program)*

$$\min J(v)$$

$$s.t.\ g_i(v) \leq 0 \quad \forall i \in [1, m]$$

Necessary conditions similar to the nonlinear programming KKT conditions hold for infinite dimensional mathematical programs given appropriate constraint qualifications, we list the following definition and theorem as examples:

Definition 2.6. *(CQ for infinite dimensional mathematical program) Consider Problem 2.3, we say Kuhn-Tucker constraint qualification holds if there exists a differential mapping* $h(t) : [0,1] \to V$ *such that:*
1.$h(t) \in U$ *for* $t \in [0, \alpha] \subset \mathbb{R}^1_+$*;*
2.$h(0) = v^*$ *and* $\frac{dh(0)}{dt} = \beta\phi$ *for* $\beta \in \mathbb{R}^1_{++}$
for any point $v^* \in U$ *and* $\phi \in V$ *such that* $\delta g_i(v^*, \phi) < 0$ *for all* $i \in I = \{i : g(v^*) = 0\}$.

Theorem 2.7. *(Infinite dimensional KKT necessary condition) If* $J(v)$ *and* $g_i(v)$ *are differentiable and their G-derivatives are continuous linear forms on the Hilbert space* V*, then there exist scalar multipliers* $\eta_i, \forall i \in [1, m]$ *such that the following conditions are necessary for* v^* *to be a local minimum of Problem 2.3:*

$$\begin{aligned} \nabla J(v^*) + \sum_{i=1}^{m} \eta_i g_i(v^*) &= 0 \\ \eta_i g_i(v^*) &= 0, i \in [1, m] \\ \eta_i &\geq 0, i \in [1, m] \end{aligned}$$

Proof. Ritter (1967) [119], Friesz (2010) [54] □

For most of the models considered later in this dissertation, we will use the interval $[t_0, t_f]$ to represent continuous time, and seek to build optimization problems and games based on functions defined on this time interval. For example, the best response problem in Chapter 5 is an embodiment of the infinite dimensional mathematical programs reviewed above. The next class of dynamic optimization problem in continuous time is optimal control problem.

2.2.3 Optimal Control Problem

In this part, we follow the notation employed by Friesz (2010) [54] of an optimal control problem of interest. It consists of an objective functional to minimize subject to state dynamics, initial condition, terminal condition and pure control constraints. After presenting the formulation we analyze the problem starting from regularity conditions employed, then different versions of its necessary conditions for optimal solution. A large body of literature exists for optimal control problems, for further review of this topic please see: Bryson and Ho (1975) [35], Sethi and Thompson (1981) [132], Bressan and Piccoli (2007) [32].

Problem 2.4. *(Optimal Control Problem, Minimization) The following minimization problem is the optimal control problem considered in this chapter:*

$$\min J(u) = \int_{t_0}^{t_f} f_0(x, u, t)dt + K[x(t_f), t_f]$$

subject to

$$\begin{aligned} \frac{dx}{dt} &= f(x(t), u(t), t) \\ x(t_0) &= x_0 \in \mathbb{R}^n \\ \Psi[x(t_f), t_f] &= 0 \\ u(t) &\in U \subset \mathbb{R}^m \end{aligned}$$

Here the control mapping $u(t)$ is in the m-fold product of the space of square-integrable functions:

$$u \in \mathcal{U} \subset \left(\mathcal{L}^2[t_0, t_f]\right)^m$$

with the inner product defined by

$$\langle u, v \rangle = \int_{t_0}^{t_f} u(t)^T v(t)dt.$$

We now study the existence, continuity and G-differentiability of the state operator following Bressan and Piccoli (2007) [32]. Here the right hand side of initial condition $x_0 \in \mathbb{R}^n$ is supposed to be fixed, and the right hand side of the dynamics is a mapping $f : \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^1 \to \mathbb{R}^n$. The state operator tells us the state vector x given each control $u(\cdot)$ for all $t \in [t_0, t_f]$, and can be interpreted as a mapping from $u(\cdot) \in (\mathcal{L}^2[t_0, t_f])^m$ to its image $x(u, \cdot) \in (C^0[t_0, t_f])^n$

$$x(u, t) = \arg\left\{\frac{dx}{dt} = f(x, u, t), x(t_0) = x_0\right\} \in \left(C^0[t_0, t_f]\right)^n,$$

where $(C^0[t_0, t_f])^n$ is in the n-fold product of the space of absolutely continuous real valued function defined on $[t_0, t_f]$. At the same time, we need the following regularity conditions for the study of the state operator from Bressan and Piccoli (2007) [32], Assumption 3.H:

Definition 2.7. *(Regularities of the optimal control problem) Assume the following for Problem 2.4:*

1. *The set $U \subset \mathbb{R}^m$ is compact.*
2. *The function f is continuous on $\mathbb{R}^n \times U \times \mathbb{R}^1$, continuously differentiable with respect to x, and continuously differentiable in an open neighborhood V of U.*

f satisfies the following for all (x, u, t) and some constant C, L :

$$\begin{aligned} |f(x,u,t)| &\leq C \\ \|D_x f(x,u,t)\| &\leq L \end{aligned}$$

The existence of the state operator could be interpreted as the existence of a solution to an ODE given the control variable. We have the following theorem:

Theorem 2.8. *(Existence of state operator) Consider the initial value problem:*

$$\begin{aligned} \frac{dx}{dt} &= f(x,u,t) \\ x(t_0) &= x_0 \end{aligned}$$

for $t \in [t_0, t_f]$, and suppose $f(x, u, t)$ is Lipschitz continuous in x for all $t \in [t_0, t_f]$. Then the initial value problem has a unique solution.

Proof. See Bressan and Piccoli (2007) [32] Theorem 3.2.1. □

Also we have the following results concerning the continuity and G-differentiability of the state operator.

Theorem 2.9. *(Continuity of state operator) Let the regularity conditions in Definition 2.7 hold. Then the state operator is continuous mapping from $(\mathcal{L}^2[t_0, t_f])^m$ to $(C^0[t_0, t_f])^n$.*

Proof. See Bressan and Piccoli (2007) [32] Theorem 3.2.1. □

Theorem 2.10. *(G-differentiability of state operator) Let the regularity conditions in Definition 2.7 hold. Then for every bounded and measurable* $\Delta u(\cdot)$, $x(u,\cdot)$ *is G-differentiable with respect to* u*, namely the derivative*

$$\delta x(u, \Delta u) = \lim_{\theta \to 0} \frac{x(u + \theta \Delta u, t) - x(u, t)}{\theta}$$

exists for every such Δu*.*

Proof. See Bressan and Piccoli (2007) [32] Theorem 3.2.6. □

A famous necessary condition for optimal control problems is known as Pontryagin's maximum/minimum principle. We start from defining the Hamiltonian for all $t \in [t_0, t_f]$ as:

$$H(x(t), u(t), \lambda(t), t) = f_0(x(t), u(t), t) + \lambda(t)^T f(x(t), u(t), t) \tag{2.1}$$

Here the adjoint variable $\lambda(t) \in \mathbb{R}^n$ is introduced. The PMP states that the triplet of optimal state, control and adjoint variable $(x^*(t), u^*(t), \lambda^*(t))$ must be such that the Hamiltonian is minimized with respect to u:

$$H(x^*(t), u^*(t), \lambda^*(t), t) \leq H(x^*(t), u(t), \lambda^*(t), t) \quad \forall u \in \mathcal{U}, \forall t \in [t_0, t_f],$$

along with the state dynamics, the adjoint equation:

$$(-)\frac{d\lambda^*(t)}{dt} = \lambda^*(t)^T \nabla_x f(x^*(t), u^*(t), t) + \nabla_x f_0(x^*(t), u^*(t)),$$

and the transversality condition:

$$\lambda^*(t_f)^T = \nabla_x \left\{ K[x^*(t_f), t_f] + \nu^T \Psi[x^*(t_f), t_f] \right\}.$$

In addition, variational inequalities could also be applied as part of the necessary conditions for optimal control, starting from analyzing the G-derivative of the objective functional J. Here we reproduce a theorem presented by Friesz (2010) [54] as an example, and notice that some of the regularities concerning the RHS of the dynamics is already listed previously in Definition 2.7:

Theorem 2.11. *(Necessary condition for optimal control) Consider Problem 2.4, and assume the following regularity conditions:*

1. *$u \in U \subseteq (\mathcal{L}^2[t_0, t_f])^m$, U is convex;*

2. *the operator $x(u,t) : (\mathcal{L}^2[t_0, t_f])^m \times \mathbb{R}^1_+ \to (\mathcal{H}^1[t_0, t_f])^n$ is regular in the sense of Definition 2.7;*

3. *$K : \mathbb{R}^n \times \mathbb{R}_+ \to \mathbb{R}$ is continuously differentiable w.r.t. x and t;*

4. *$\Psi : \mathbb{R}^n \times \mathbb{R}^1_+ \to \mathbb{R}^r$ is continuously differentiable w.r.t. x and t;*

5. *$f_0 : \mathbb{R}^n \times U \times [t_0, t_f] \to \mathbb{R}$ is continuously differentiable w.r.t. x and u with bounded partials;*

6. *$f : \mathbb{R}^n \times U \times [t_0, t_f] \to \mathbb{R}^n$ is continuously differentiable w.r.t. x and u with bounded partials;*

7. *$x_0 \in \mathbb{R}^n, t_0 \in \mathbb{R}^1_+$ and $t_f \in \mathbb{R}^1_{++}$ are known and fixed, $\nu \in \mathbb{R}^r$*

Then any solution $u^* \in U$ must obey the following:

1. *the variational inequality:*

$$\sum_{i=1}^{m} \frac{\partial H(x^*, u^*, \lambda^*, t)}{\partial u_i}(u - u^*) \geq 0, \quad \forall u \in U;$$

2. *the state dynamics:*

$$\frac{dx^*}{dt} = f(x^*, u, t), x^*(t_0) = x_0;$$

3. *the adjoint dynamics:*

$$(-)\frac{d\lambda^*(t)}{dt} = \nabla_x H^*;$$

4. *the transversality conditions:*

$$\lambda^*(t_f)^T = \nabla_x \left\{ K[x^*(t_f), t_f] + \nu^T \Psi[x^*(t_f), t_f] \right\}.$$

Proof. See Friesz (2010) [54], Theorem 4.16. □

Remark 2.1. *It is also possible to recast the theorem above into one with sufficient condition, however, as noted by Friesz (2010) [54], convexity of Hamiltonian $H(\cdot)$*

in control only may not be enough in general. On the other hand, by imposing joint convexity of the Hamiltonian in (x, u), it is possible to recast the existing sufficiency theorems using the VI necessary condition. Please refer to Seierstad and Sydaeter (1999) [130] for the Arrow sufficiency theorem and the Mangasarian sufficiency theorem.

So far we have discussed the PMP-type of necessary conditions for the deterministic, open-loop optimal control problem. At the same time, we are aware of Bellman's Principle of Optimality (POO) and the application of Dynamic Programming Principle (DPP). The application of POO, in continuous time, will lead to the Hamilton-Jacobi-Bellman partial differential equation (HJB-PDE). In this dissertation, we will use the HJB-PDE as a starting point in Chapter 3, and discuss how such techniques could be used to obtain the optimal control with feedback information structure. Also, note that similar variational principles have been used towards the derivation of stochastic maximum principle and stochastic variational inequalities as optimality conditions when the underlying problems are stochastic. This will be the main topic in Chapter 4.

In terms of numerical computation, we have the following gradient projection algorithm on Hilbert space:

Algorithm 2.1. *(Gradient Projection Algorithm) Consider an optimal control problem on a finite time horizon $t \in [t_0, t_f]$ as Problem 2.4:*

Step 0. *Initialization* *Set $k = 0$ and pick $u^0(t) \in (L^2[t_0, t_f])^m$.*

Step 1. *Find state trajectory* *Using $u^k(t)$ solve the state initial-value problem:*

$$\begin{aligned} \frac{dx}{dt} &= f(x, u^k, t) \\ x(t_0) &= x_0 \end{aligned}$$

and obtain the solution $x^k(t)$.

Step 2. *Find adjoint trajectory* *Using $u^k(t)$ and $x^k(t)$ solve the adjoint final value problem:*

$$(-)\frac{d\lambda(t)}{dt} = \frac{\partial H(x^k, u^k, \lambda, t)}{\partial x}$$

$$\lambda(t_f) \quad = \quad \frac{\partial K(x(t_f), t_f)}{\partial x}$$

and call the solution $\lambda^k(t)$

Step 3. *Find gradient* *Use* $u^k(t)$, $x^k(t)$, *and* $\lambda^k(t)$ *to calculate*

$$\nabla_u J(u^k) = \frac{\partial H(x^k, u^k, \lambda, t)}{\partial u}$$

Step 4. *Update and apply stopping test* *For a suitably small step size* θ_k, *update* $u(\cdot)$:

$$u^{k+1} = P_U[u^k - \theta_k \nabla_u J(u^k)]$$

If some stopping criterion is satisfied, declare $u^*(t) = u^{k+1}(t)$. *Otherwise, set* $k = k + 1$ *and go to Step 1.*

There are existing literature on the convergence of gradient projection type of algorithms, examples are: Tian and Dunn (1994) [137] and Nikol′skii (2007) [105], which provide detailed convergence proofs under different regularity conditions. In this dissertation, we refer to the following high-level theorem from Friesz (2010) [54]. The main idea is that coerciveness is important in providing the desired convergence.

Theorem 2.12. *Assume objective functional* $J : V \leftarrow \mathbb{R}$ *is coercive with* $\alpha > 0$, *which means* $\exists \alpha > 0$ *such that:*

$$J[(1-\theta)u + \theta v] \leq (1-\theta)J(u) + \theta J(v) - \frac{\alpha}{2}\theta(1-\theta)||u - v||^2$$

$\forall u, v \in V$ *and* $\theta \in (0, 1)$; *also assume that* $\nabla J(u)$ *is well defined and Lipschitz:*

$$||\nabla J(u) - \nabla J(v)|| \leq \beta ||u - v||,$$

then the projection algorithm converges to the minimum u^* *of* J *for fixed step size choices in* $\theta \in \left(0, \frac{2\alpha}{\beta^2}\right)$.

2.3 Variational Inequality

Variational inequalities (VIs) are another important class of mathematical problem in which a functional inequality has to be solved for all possible values within a given (convex) feasible region. It has been widely used in the literature of economics and operation research, often as a characterization of equilibrium or optimality conditions. We here list some formulations of VIs that are most commonly applied. For detailed reference on finite dimensional VIs and their applications, see for example Facchinei and Pang (2007) [49] and Nagurney (2013) [103]. Here we follow the notation in Friesz (2010) [54] Definition 5.6.

Problem 2.5. *(Variational Inequality, finite-dimensional) Let $\Lambda \subseteq \mathbb{R}^n$ be a non-empty subset of $\mathbb{R}^n$ and given a function $F : \Lambda \to \mathbb{R}^n$. The problem $VI(X, F)$ is to find vector y such that:*

$$\begin{aligned} y &\in \Lambda \\ [F(y)]^T (x - y) &\geq 0 \qquad \forall x \in \Lambda \end{aligned}$$

This formulation can be generalized into infinite dimensional.

Problem 2.6. *(Quasi-Variational Inequality, finite-dimensional, Friesz (2010) [54], Definition 5.8) Given a function $F : \mathbb{R}^n \to \mathbb{R}^n$ and a point to set mapping $K(y) \subset \mathbb{R}^n$, the quasi-variational inequality problem $QVI(F, K(y))$ is to find vector y such that:*

$$\begin{aligned} y &\in K(y) \\ [F(y)]^T (x - y) &\geq 0 \qquad \forall x \in K(y) \end{aligned}$$

Similarly, the variational inequality problem could also be defined on Hilbert space.

Problem 2.7. *(Infinite Dimensional VI) Let $X \subseteq V$ be a non-empty subset of a Hilbert space V and let $F : X \subseteq V \to V$ be a mapping from X to itself. The problem $VI(F, \Lambda)$ is to find $x^* \in X$ such that the following conditions hold:*

$$F(x^*)^T (x - x^*) \geq 0 \qquad \forall x \in X \subseteq \mathbb{R}^n$$

Problem 2.8. *(DVI with state dynamics, Friesz (2010) [54]) Consider the control vector $u \in (\mathcal{L}^2[t_0, t_f])^m$ and associated state operator $x(u, t) : (\mathcal{L}^2[t_0, t_f])^m \times \mathbb{R}^1_+ \to$*

$(\mathcal{H}^1[t_0, t_f])^n$ *where*

$$x(u,t) = arg\{\frac{dy}{dt} = f(y,u,t), y(t_0) = y_0, \Psi[y_f, t_f] = 0\} \tag{2.2}$$

Furthermore, assume every control vector is constrained in a set U, we then denote the following variational inequality as $DVI(F, f, \Psi, U.x_0)$: find $u^* \in U$ such that

$$\int_{t_0}^{t_f} [F(x(u^*), u^*, t)]^T (u - u^*) \geq 0 \tag{2.3}$$

Still, the DVI problem here is closely related to differential Nash games and optimal control problem. We list the following necessary conditions of the DVI problem and the regularities assumed.

Definition 2.8. *(DVI regularity conditions) Consider $DVI(F, f, \Psi, U.x_0)$ defined in Problem 2.8, we call $DVI(F, f, \Psi, U.x_0)$ regular if:*

1. *$u \in U \subseteq (\mathcal{L}^2[t_0, t_f])^m$, U is convex;*
2. *the operator $x(u,t) : (\mathcal{L}^2[t_0, t_f])^m \times \mathbb{R}^1_+ \to (\mathcal{H}^1[t_0, t_f])^n$ exists and is unique, strongly continuous and G-differentiable for all admissible u;*
3. *$K : \mathbb{R}^n \times \mathbb{R}^1_+ \to \mathbb{R}^1$ is continuously differentiable w.r.t. x and t;*
4. *$\Psi : \mathbb{R}^n \times \mathbb{R}^1_+ \to \mathbb{R}^r$ is continuously differentiable w.r.t. x and t;*
5. *$F : \mathbb{R}^n \times U \times [t_0, t_f] \to \mathbb{R}$ is continuously differentiable w.r.t. x and u with bounded partials;*
6. *$f : \mathbb{R}^n \times U \times [t_0, t_f] \to \mathbb{R}^n$ is continuously differentiable w.r.t. x and u with bounded partials;*
7. *$x_0 \in \mathbb{R}^n, t_0 \in \mathbb{R}^1_+$ and $t_f \in \mathbb{R}^1_{++}$ are known and fixed, $\nu \in \mathbb{R}^r$ for the terminal constraints $\Psi[x(t_f), t_f] = 0$.*

Remark 2.2. *In the definition above and in Theorem 2.11, note that all the regularities are listed w. r. t. the functions instead of referring to the "substitution operators" (also known as "Nemytskii operators"). This is due to the fact that regularities on the F, f and f_0 functions are much more practical to check than, the continuous Gateaux differentiability of mappings from a function space to another function space.*

The DVI necessary conditions are obtained from the following observations: assume the knowledge of u^* is known, then Problem 2.8 could be re-stated as a optimal control problem parameterized by u^*. The necessary conditions of this auxiliary optimal control problem then become the necessary conditions for the original DVI. Hence we have the following theorem:

Theorem 2.13. *(DVI necessary condition, Friesz (2010) [54], Theorem 6.4) Consider $DVI(F, f, \Psi, U.x_0)$ defined in Problem 2.8, when the regularities stated in Definition 2.8 holds, necessary conditions for $u^* \in U$ to be a solution of Problem 2.8 are:*

1. *the variational inequality:*

$$\left[F(x^*, u^*, t) + \nabla_u(\lambda^*)^T f(x^*, u^*, t)\right]^T (u - u^*) \geq 0, \quad \forall u \in U;$$

2. *the state dynamics:*

$$\frac{dx^*}{dt} = f(x^*, u^*, t), x^*(t_0) = x_0;$$

3. *the adjoint dynamics:*

$$(-)\frac{d\lambda^*(t)}{dt} = \nabla_x(\lambda^*)^T f(x^*, u^*, t);$$

4. *the transversality conditions*

$$\lambda^*(t_f)^T = \nabla_x \left\{\nu^T \Psi[x^*(t_f), t_f]\right\}.$$

Remark 2.3. *At this point it is interesting to look back and make some bibliographical notes on the term 'variational inequality' and 'differential variational inequality'. Lions and Stampacchia (1967) [90] can be considered as the seminal paper where the mathematical theory of variational inequalities started formally. Before this, there are different papers considering special cases of the variational inequality problem, see the Wikipedia page on the term 'variational inequality' for a brief historical remark.*

There have other problems bearing the name 'differential variational inequality' which are different than what is considered in this dissertation. For example, Aubin *and Cellina (1984) [5] used this term to describe a different problem which is later categorized as 'variational inequality of evolution' (VIE) by Pang and Stewart (2008) [108].*

2.4 Nash Games

In game theory, the Nash equilibrium is a solution of a non-cooperative multi-player game in which no player has anything to gain by changing only his own strategy. In this section we list definitions of different forms of Nash equilibrium solutions and give their equivalent variational inequalities without proof. For references on game theory in general, please refer to Fudenberg and Tirole (1991) [60]; for in-depth discussions on dynamic and differential games, there are multiple good references such as Basar and Ozder (1999) [10], Dockner et al. (2000) [41]. Below is a reproduction of Friesz (2010) [54], Definition 5.1:

Definition 2.9. *(Nash Equilibrium) Suppose there are N agents/players, each of which chooses a feasible strategy x^i from the strategy set Λ_i which is independent of other players' strategies. Furthermore, every agent i has a cost/disutility function $\Theta_i(x) : \Lambda \to \mathbb{R}^1$ that depends on all agents' strategies where*

$$\begin{aligned} \Lambda &= \prod_{i=1}^{N} \Lambda_i \\ x &= (x^i : i = 1, \cdots, N) \end{aligned}$$

Every agent $i \in [1, N]$ seeks to minimize his own disutility:

$$\min \Theta_i(x^i; x^{-i}) \qquad s.t. x^i \in \Lambda_i \tag{2.4}$$

for each arbitrary fixed non-own tuple $x^{-i} = (x^j : j \neq i)$. A Nash equilibrium is a tuple of strategies x^*, one for each agent, such that each $x^{i,*}$ solves the mathematical program (2.4), and is denoted as $NE(\Theta, \Lambda)$.

Notice that this in fact means no agent could lower his disutility by unilaterally changing his own strategy. The following theorem from Friesz (2010) [54] states that a Nash equilibrium is equivalent to a variational inequality given appropriate regularity conditions.

Theorem 2.14. *(Nash equilibrium and VI) Assume the following regularity conditions: (1) each $\Theta_i(x^i) : \Lambda_i \to \mathbb{R}^1$ is convex and continuously differentiable in x^i; (2) each Λ_i is a closed convex subset of $\mathbb{R}^{n_i}$. The Nash equilibrium $NE(\Theta, \Lambda)$ in Definition 2.9 is equivalent to the following variational inequality $VI(\nabla\Theta, \Lambda)$ to find a vector x^* such that*

$$
\begin{aligned}
x^* &\in \Lambda \\
\left[\nabla\Theta(x^*)\right]^T (x - x^*) &\geq 0 \quad x \in \Lambda
\end{aligned}
$$

Proof. See Friesz (2010) [54], Theorem 5.5. □

When the feasible strategy set of any agent depends on other agents' strategies, the concept of a Nash equilibrium is extended to the following definition of a generalized Nash equilibrium. Below is from Friesz (2010) [54], Definition 5.2:

Definition 2.10. *(Generalized Nash Equilibrium) Suppose there are N agents/-players, each of which chooses a feasible strategy x^i from the strategy set $\Lambda_i(x)$ which is dependent of other players' strategies. Furthermore, every agent i has a cost/disutility function $\Theta_i(x) : \Lambda \to \mathbb{R}^1$ that depends on all agents' strategies where*

$$
\begin{aligned}
\Lambda(x) &= \prod_{i=1}^{N} \Lambda_i(x) \\
x &= (x^i : i = 1, \cdots, N)
\end{aligned}
$$

Every agent $i \in [1, N]$ seeks to minimize his own disutility:

$$
\min \Theta_i(x^i; x^{-i}) \qquad s.t. x^i \in \Lambda_i(x) \tag{2.5}
$$

for each arbitrary fixed non-own tuple $x^{-i} = (x^j : j \neq i)$. A generalized Nash equilibrium is a tuple of strategies x^*, one for each agent, such that each $x^{i,*}$ solves the mathematical program (2.5), and is denoted as $GNE(\Theta, \Lambda(x^*))$.

Similar to the case of a Nash equilibrium, a generalized Nash equilibrium could be characterized by a QVI given appropriate regularities.

Theorem 2.15. *(Generalized NE and Quasi VI) Assume the following regularity conditions: (1) each $\Lambda_i \subseteq X_i$, and X_i is a compact and convex subset of $\mathbb{R}^{n_i}$; (2)each*

$\Theta_i(x^i) : \Lambda_i \to \mathbb{R}^1$ is convex and continuously differentiable in x^i. The generalized Nash equilibrium $GNE(\Theta, \Lambda(x))$ in Definition 2.10 is equivalent to the following quasivariational inequality $QVI(\nabla\Theta, \Lambda(x))$ to find a vector x^* such that

$$\begin{aligned} x^* &\in \Lambda(x^*) \\ \left[\nabla\Theta(x^*)\right]^T(x - x^*) &\geq 0 \quad x \in \Lambda(x^*) \end{aligned}$$

Proof. See Harker (1991) [68]. □

For the case when each player is facing a dynamic optimization problem in continuous time, similar definitions could be made on differential (generalized) Nash equilibrium. Here as an example, we follow Friesz (2010) [54] and define the differential Nash game with each player's ration is based on an optimal control problem and give without proof its equivalence to a differential variational inequality. We have the following definition:

Definition 2.11. *(Differential Nash equilibrium) Suppose there are N agents, each of which chooses a feasible strategy vector u^i from the strategy set Λ_i which is independent of the other players' strategies. Furthermore, every agen i has a cost/disutility functional $J_i(u) : \Lambda \to \mathbb{R}^1$ that depends on all agents' strategies where*

$$\begin{aligned} \Lambda &= \prod_{i=1}^{N} \Lambda_i \\ u &= (u^i : i = 1, \cdots, N) \end{aligned}$$

Every agent seeks to solve the optimal control problem

$$\min J(u^i; u^{-i}) = \int_{t_0}^{t_f} \Theta_i(x^i, u^i, x^{-i}, u^{-i}, t)dt + K_i[x^i(t_f), t_f]$$

subject to

$$\begin{aligned} \frac{dx^i}{dt} &= f^i(x^i(t), u^i(t), t) \\ x^i(t_0) &= x_0^i \\ \Psi^i[x^i(t_f), t_f] &= 0 \\ u^i(t) &\in \Lambda_i. \end{aligned}$$

A differential Nash equilibrium is a tuple of strategies u such that u^i solves the optimal control problem.

Again, no agent may lower his disutility unilaterally. Still, a DVI could be used to characterize the equilibrium of a differential Nash game.

Theorem 2.16. *(DVI equivalent to differential Nash equilibrium) The differential Nash equilibrium in Definition 2.11, is equivalent to the following DVI: find u^* such that:*

$$\begin{aligned} u^* &\in U \\ \langle F((u^*,t),u^*,t), u^* - u\rangle &\geq 0 \qquad \forall u \in U \end{aligned}$$

where

$$\begin{aligned} u &\in U \subseteq \mathcal{L}^2[t_0,t_f] \\ x(u,t) &= \arg\left\{ \frac{dx}{dt} = f(x,u,t), x(t_0) = x_0, \Psi[x(t_f),t_f] = 0 \right\} \subseteq \mathcal{H}^1[t_0,t_f] \end{aligned}$$

when $f^i(x^i(t),u^i(t),t)$ and $\Theta_i(x^i,u^i,x^{-i},u^{-i},t)$ are convex and continuously differentiable with respect to (x^i,u^i) for all fixed non-own tuples (x^{-i},u^{-i}) for all $i \in [1,N]$.

Proof. See Friesz (2010) [54], Theorem 5.5. □

In terms of numerical computation of DVIs and therefore DNEs, we have the following fixed point algorithm:

Algorithm 2.2. *(Fixed Point Algorithm, Friesz (2010) [54]) Consider a differential variational inequality on a finite time horizon $t \in [t_0,t_f]$ as Problem 2.8:*

Step 0. *Initialization* *Set $k=0$ and pick $u^0(t) \in U$.*

Step 1. *Solve the optimal control problem* *Solve the optimal control problem parameterized by $u^k(t)$:*

$$\min_v J^k(v) = \gamma^T \Psi[x(t_f),t_f] + \int_{t_0}^{t_f} \frac{1}{2}\left[u^k - \alpha F(x^k,u^k,t) - v\right]^2 dt$$

subject to

$$\begin{aligned} \frac{dx}{dt} &= f(x(t), v(t), t) \\ x(t_0) &= x_0 \\ v(t) &\in U \end{aligned}$$

and obtain the solution $u^{k+1}(t)$.

Step 2. $\boxed{\textit{Update and apply stopping test}}$ *If* $\left\|u^{k+1} - u^k\right\| \leq \epsilon$, *stop and declare* $u^* = u^{k+1}$. *Otherwise set* $k = k + 1$ *and go to Step 1.*

This algorithm relies on the repeated solution of an optimal control sub-problem, which could be handled by, for example, the functional gradient projection such as Algorithm 2.1. This type of algorithm is directly applied in the solution of dynamic Nash equilibrium in Chapter 5.

2.5 Summary

This chapter briefly reviews results in deterministic optimization and Nash games, it also serves as a starting point towards further development of this dissertation: optimal control problems and Nash games with feedback information structure are reviewed and applied in Chapter 3; stochastic optimal control and differential games are reviewed in Chapter 4 and then applied in Chapter .

Direct application of these methodologies could be found in Chapter 5, where we first review a dynamic competitive revenue management model, then extended the model towards more insightful results.

Chapter 3

Optimal Control Problems and Differential Games with Feedback

It is well known that Pontryagin's maximum principle (PMP) and Bellman's Principle of Optimality (POO) are two most widely applied methodologies in continuous time optimal control problems. The latter one is also known as Dynamic Programming Principle (DPP). In the PMP approach, the adjoint variables are introduced which result in a system of ODEs that are linked via the maximum principle. In the DPP approach, as we will review in this chapter, the Hamilton-Jacobbi-Bellmann partial differential equation (HJB-PDE) can be derived from the DPP.

At the same time, the DPP approach and the associated HJB-PDE are both centered around the value function representing the optimum objective attainable given a state-time tuple as a starting point. This gives the continuous time DPP approach natural advantage in dealing with optimal control with feedback information structure. In this chapter, several key results from the DPP approach are reviewed, including a comprehensive numerical scheme in the synthesis of optimal control in its feedback form.

The rest of this chapter is structured as follows. In Section 3.1 we review existing results on optimal control problem with feedback information structure based on a continuous time version of POO and the HJB-PDE thus induced. We also study the so-called verification theorem and the numerical procedure for the synthesis of feedback optimal control when the value function should be given numerically.

Such a procedure of feedback control synthesis is the main content of Section 3.3. The special case of linear-quadratic problems is also listed as a special case. Lastly, Section 3.4 covers the differential Nash games with agents' best response are based on feedback optimal control. Later in Chapter 6, we will propose applications of these methodologies in the area of revenue management when state information with respect to service capacities are taken advantage of by the decision makers.

3.1 Feedback Optimal Control

Comparing with the Pontryagin Maximum Principle (PMP), which solves for the open loop optimal control with a given initial state as a function of time, the techniques based on the Bellman principle provides insights on optimal control expressed as a function of both the state and time variable. Later on in this chapter, we will review how feedback controls could be synthesized based on the solution of Hamilton-Jaccobi-Bellman equation. Before that, we list the following definition from Basar and Olsder (1999) [10] regarding different information structures.

Definition 3.1. *(Information Structures) For finite horizon, continuous time optimal control problems and differential games, denote time as $t \in [t_0, t_f]$. We use the notation $\eta(t)$ to define information structures:*
(i) open-loop (OL): $\eta(t) = \{x_0\}$, $t \in [t_0, t_f]$;
(ii) feedback: $\eta(t) = \{x(t)\}$, $t \in [t_0, t_f]$;
(iii) closed-loop memoryless (CLM): $\eta(t) = \{x_0, x(t)\}$, $t \in [t_0, t_f]$;
(iv) closed-loop (CL): $\eta(t) = \{x(s),\ 0 \leq s \leq t\}$, $t \in [t_0, t_f]$.

In this chapter and throughout the rest of this dissertation, we will consider the feedback information structure as defined above. This means that the decision maker has knowledge of current state variable for all time t throughout the horizon, and will use this information toward his advantage. Therefore the control he considers will be a function of both state and time as $u(x(t), t)$, instead of just $u(t)$ as in open-loop optimal control. There are other interesting problems considering optimal control and differential games with imperfect or incomplete information, we will make further comments in the context of stochastic optimal control problems in Chapter 4. Following the spirit of our problem setting, we restate the optimal control problem into its feedback form:

Problem 3.1. *(Optimal Control in State Feedback) Consider a compact set of admissible control values* $U \subset \mathbb{R}^m$*, then* $f : \mathbb{R}^n \times U \to \mathbb{R}^n$*,* $f_0 : \mathbb{R}^n \times U \to \mathbb{R}$*,* $K : \mathbb{R}^n \to \mathbb{R}$*. A feedback optimal control problem in finite time horizon* $t \in [t_0, t_f]$ *is formulated as the following:*

$$\min_{u \in U} J(u) = \int_{t_0}^{t_f} f_0(x(t), u(x,t), t)dt + K[x(t_f)]$$

such that:

$$\begin{aligned} \frac{dx}{dt} &= f(x(t), u(x,t), t) \\ x(t_0) &= x_0 \end{aligned}$$

Note that with a proper discount factor, Problem 3.1 can also be formulated and solved in infinite time horizon, please refer to Bressan and Piccoli (2007) [32] for details. In the rest of this section we review the application of Dynamic Programming Principle (DPP) in continuous time, and the famous result of value function as a viscosity solution of Hamilton-Jaccobi-Bellman partial differential equation (HJB-PDE). For a comprehensive review of (discrete-time) Dynamic Programming please refer to Bertsekas (1995) [19], and Puterman (2014) [116]. The key towards this study is the value function, as defined by the following:

$$V(x,t) = \inf_{u \in \mathcal{U}} J(x,t,u) \tag{3.1}$$

where $\mathcal{U}$ is the family of admissible controls acting on the same dynamics and cost functionals as in Problem 3.1:

$$\mathcal{U} = \{u : \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}^m \text{ measurable, } u(x,t) \in U \;\; a.e. \; t\}$$

and functional $J(x,t,u)$ is the cost functional of Problem 3.1 from time t, state x with an admissible control u:

$$J(x,t,u) = \int_{t}^{t_f} f_0(x(t), u(x,t), t)dt + K[x(t_f)]$$

We list without proof the DPP in continuous time:

Theorem 3.1. *(Dynamic Programming Principle) For every* $\tau \in [t, t_f]$ *and* $x \in \mathbb{R}^n$*,* *we have:*

$$V(x,t) = \inf_u \left\{ \int_t^\tau f_0(x(t), u(x,t), t)dt + V(x(\tau), \tau) \right\}$$

Proof. For detailed proof, see Bressan and Piccoli (2007) [32] □

Now we give more details about regularity conditions on the optimal control problem to give insights to the regularity of $V(x,t)$.

Definition 3.2. *Problem 3.1 is regular if:*
(i)the functions $f(\cdot)$, $f_0(\cdot)$, $K(\cdot)$ satisfy the following for some constant C:

$$\begin{aligned} \|f_0(x,u)\| &\leq C, \ \|f_0(x,u) - f_0(y,u)\| \leq C\|x-y\| & (3.2)\\ \|f(x,u)\| &\leq C, \ \|f(x,u) - f(y,u)\| \leq C\|x-y\| & (3.3)\\ \|K(x)\| &\leq C, \ \|K(x) - K(y)\| \leq C\|x-y\| & (3.4) \end{aligned}$$

(ii) the set U is compact.

The value function enjoys Lipschitz continuity if the regularities in Definition 3.2 is met, as in the following lemma:

Lemma 3.1. *(Analytical properties of value function) If the feedback optimal control Problem 3.1 is regular in the sense of Definition 3.2, then the value function $V(x,t)$ is bounded and Lipschitz continuous. For some constant C_1 and C_2:*

$$\begin{aligned} \|V(x,t)\| &\leq C_1 & (3.5)\\ \|V(x,t) - V(y,s)\| &\leq C_2(\|x-y\| + \|t-s\|) & (3.6) \end{aligned}$$

Proof. For details, see Bressan and Piccoli (2007) [32], Lemma 8.6.2. □

We are ready to review the following theorem, which states that the value function of Problem 3.1 could be characterized by viscosity solution (Evans (2010) [47]) of the HJB-PDE:

Theorem 3.2. *(Value function as viscosity solution of HJB-PDE) For Problem 3.1 which is regular in the sense of Definition 3.2, consider the value function V defined by (3.1). Then V is the unique viscosity solution of the HJB-PDE:*

$$\frac{\partial V(x,t)}{\partial t} + \min_u \left[\nabla_x V(x,t) \cdot f(t,x,u) + f_0(t,x,u)\right] = 0 \qquad (3.7)$$

for $\forall(x,t) \in \mathbb{R}^n \times (t_0, t_f)$, *with the boundary condition:*

$$V(t_f, x) = K(x) \tag{3.8}$$

Proof. Please refer to Bressan and Piccoli (2007) [32] for a detailed proof. □

This type of technique also works for stochastic optimal control problems with Brownian motion type of stochastic dynamics, we will review such techniques briefly in Chapter 4.

3.1.1 Linear Quadratic Problem as a Special Case

Here we take a detour to briefly review one of the most widely used spacial case of the above discussions, which is the so-called Linear (dynamics) Quadratic (objective) problem. Please refer to Engwerda (2005) [46] for more details. In this subsection, we first introduce the formulation of an LQ optimal control problem as the following:

Problem 3.2. *(Linear Quadratic Optimal Control) A LQ problem in finite time horizon* $t \in [t_0, t_f]$ *is the following:*

$$\min_u J = \int_{t_0}^{t_f} \{x^T(t)Kx(t) + u^T(t)Ru(t)\}dt + x^T(T)K_{t_f}x(T) \tag{3.9}$$

such that:

$$\frac{dx}{dt} = Ax(t) + Bu(t), \tag{3.10}$$

$$x(t_0) = x_0 \tag{3.11}$$

The feedback optimal control of Problem 3.2 enjoys a particularly neat form in state feedback. This is due to the fact that the HJB-PDE equation (3.7) corresponding to LQ problems admits a separation of variables. We use the following theorem from Engwerda (2005) [46] to summarize this result:

Theorem 3.3. *(Riccati System) The LQ Problem 3.2 has, for* $\forall x_0$, *a solution if and only if the following Riccati system of ODE has a solution* $Q(\cdot)$ *on* $[t_0, t_f]$:

$$\dot{Q}(t) = -A^TQ(t) - Q(t)A + Q(t)SQ(t) - Q, Q(t_f) = K_{t_f} \tag{3.12}$$

where $S \triangleq BR^{-1}B^T$, and then it is unique and the optimal control in feedback form is:

$$u^*(t) = -R^{-1}B^TQ(t)x(t) \tag{3.13}$$

Proof. See Engwerda (2005) [46], Theorem 5.1. □

Nevertheless, we will use the LQ case as a benchmark when testing our numerical procedure which starts from the solution of HJB-PDE (3.7). In next section we will review such numerical procedure and provide a small numerical example demonstrating its effectiveness towards the LQ problem, and compare it to the solution based on Theorem 3.3.

3.2 Synthesis of Feedback Optimal Control

Now given that we could obtain a solution $V(x,t)$ of the HJB-PDE, we want to know how one could utilize this information and generate feedback optimal control. In other words, the main motivation of this section is to present how continuous time dynamic programming can be applied to generate optimal feedback control. The following theorem states that the HJB-PDE is a way of testing the optimality of a given admissible state-control pair. In the optimal control literature, this theorem is the so called "verification theorem". In this section we review such verification theorem carefully, and then present the numerical procedure of feedback control construction justified by it.

Theorem 3.4. *(Verification Theorem) If a continuously differentiable function $V(t,x)$ can be found that satisfies the HJB equation (3.7) subject to the boundary condition (3.8), then it generates the optimal feedback strategy through the static (pointwise in time) minimization problem defined by RHS of (3.7).*

Proof. In this proof we take the state equation as 1-dimensional, bearing in mind that such analysis is easily extendable into multi-dimensional case. This proof is casted from Basar Olsder (1999) [10], Theorem 5.3. If we compare a point-wise optimal trajectory (x^*, u^*) with an arbitrary one $(\tilde{x}, \tilde{u})$, the two of them starts from the same initial state x_0, then equation (3.7) will turn out to be the following:

$$\frac{\partial V(\tilde{x},t)}{\partial x} f(\tilde{x},\tilde{u},t) + f_0(\tilde{x},\tilde{u},t) + \frac{\partial V(\tilde{x},t)}{\partial t} \geq 0.$$

This is due to the fact that the minimum on RHS of (3.7) is not necessarily true for arbitrary trajectory and

$$\frac{\partial V(x^*,t)}{\partial x} f(x^*,u^*,t) + f_0(x^*,u^*,t) + \frac{\partial V(x^*,t)}{\partial t} \equiv 0.$$

Integrate the above with corresponding time horizons $[t_0, t_f]$, we have:

$$\int_{t_0}^{t_f} f_0(\widetilde{x}, \widetilde{u}, t)dt + V(\widetilde{x}(t_f), t_f) - V(x_0, t_0) \geq 0$$

and

$$\int_{t_0}^{t_f} f_0(x^*, u^*, t)dt + V(x^*(t_f), t_f) - V(x_0, t_0) = 0.$$

The first term of the integration is immediate. The $V(x(t_f), t_f) - V(x_0, t_0)$ term comes from the following:

$$\begin{aligned} & \frac{\partial V(x,t)}{\partial x} f(x,u,t) + \frac{\partial V(x,t)}{\partial t} \\ = \quad & \frac{\partial V(x,t)}{\partial x}\frac{dx}{dt} + \frac{\partial V(x,t)}{\partial t} \\ = \quad & \frac{dV(x,t)}{dt} \end{aligned}$$

Therefore,

$$\begin{aligned} & \int_{t_0}^{t_f} \left[\frac{\partial V(x,t)}{\partial x} f(x,u,t) + \frac{\partial V(x,t)}{\partial t} \right] dt \\ = \quad & \int_{t_0}^{t_f} \frac{dV(x,t)}{dt} dt \\ = \quad & V(x(t_f), t_f) - V(x_0, t_0) \end{aligned}$$

The elimination of $V(x_0, t_0)$ yields

$$\int_{t_0}^{t_f} f_0(\widetilde{x}, \widetilde{u}, t)dt + V(\widetilde{x}(t_f), t_f) \geq \int_{t_0}^{t_f} f_0(x^*, u^*, t)dt + V(x^*(t_f), t_f). \tag{3.14}$$

Also due to the boundary condition (3.8) this means

$$\int_{t_0}^{t_f} f_0(\widetilde{x}, \widetilde{u}, t)dt + K(\widetilde{x}(t_f), t_f) \geq \int_{t_0}^{t_f} f_0(x^*, u^*, t)dt + K(x^*(t_f), t_f) \tag{3.15}$$

(x^*, u^*) is in fact the optimal feedback strategy. □

Notice that one can interpret the above theorem as a sufficient condition of optimal control Problem 3.1. Further more, this theorem can serve as a justification of the numerical procedure to retrieve optimal feedback control. The following algorithm, which has been recorded by Evans (2010, Section 10.3.3) [47], Bressan and Piccoli (2007, Remark 7.3) [32], Yong and Zhou (1999) [148], takes advantage of the fact that the Hamiltonian should be minimized by the feedback optimal control.

Algorithm 3.1. *(Feedback Optimal Control Synthesis) Consider the feedback optimal control in Problem 3.1 and a finite time horizon* $t \in [t_0, t_f]$*:*

Step 1. *Solve the HJB-PDE* *Solve (3.7) subject to the boundary condition (3.8), and thereby numerically compute the value function* $V(x, t)$*.*

Step 2. *Establish optimal control* *At each time instance* t*, use the value function* $V(x, t)$ *from Step 1 to design an optimal feedback control* $u^*(\cdot)$*, with the following strategy: for each point* x *and each time* $t \leq t_f$*, solve*

$$\frac{\partial V(x,t)}{\partial x} f(x, u^*, t) + f_0(x, u^*, t) + \frac{\partial V(x,t)}{\partial t} = 0 \tag{3.16}$$

for u^**, and define* $u^*(x, t)$ *to be the root of the above equation. Note that with* $V(x, t)$ *known (numerically from Step 1),* $\frac{\partial V(t,x)}{\partial x}, \frac{\partial V(t,x)}{\partial t}$ *could be evaluated (by techniques such as finite difference) and the solution of the above equation is in fact a root-finding procedure given a fixed state-time tuple* (t, x)*.*

Step 3. *Recover optimal state-control pair* *Solve the following ODE, assuming* $u(\cdot, t)$ *is sufficiently regular to let us do so:*

$$\frac{dx^*(t)}{dt} = f(x^*(t), u(x^*(t), t)), t \in (s, t_f] \tag{3.17}$$

$$x(s) = x_0 \tag{3.18}$$

Then define the feedback control-state-time trajectory:

$$u^* \triangleq u(x^*(t), t) \tag{3.19}$$

Remark 3.1. *Remember in open loop problems what we solve is a mapping from time to control, whereas in feedback problems we seek a mapping from state-time to control. Step 2 of the Algorithm above defines such mapping. The numerical solution of HJB-PDE in Step 1 is to be carried out by implementing some existing numerical packages, which relies on so-called semi-Lagrangian schemes. We shall review such numerical procedure in Appendix A.*

We will use the rest of this section to present an illustrative numerical example for the procedures above, we consider the following LQ problem from Engwerda (2005) [46], and solve the same problem twice: firstly by Riccati system following Theorem 3.3, then by Algorithm 3.1:

Example 3.1. *(LQ feedback optimal control) The problem to consider, with a background in investment decisions is:*

$$\max \int_0^T [px(t)^2 - ru(t)^2]dt \tag{3.20}$$

$$\frac{dx}{dt} = -\alpha x(t) + u(t) \tag{3.21}$$

$$x(0) = x_0 \tag{3.22}$$

$$u \in U \tag{3.23}$$

We first follow Theorem 3.3 and match up this example with each component of the "standard form": $A \sim -\alpha, B \sim 1, K \sim -p, R \sim r, K_T \sim 0$. *This lead to the following Riccati system:*

$$\frac{dk}{dt} = 2\alpha k(t) + \frac{1}{r}k^2(t) + p, k(T) = 0;$$

hence the optimal control in feedback form is:

$$u(x,t) = -R^{-1}BK(t)x(t) = -\frac{1}{r}k(t)x(t).$$

On the other hand, we follow Algorithm 3.1, by first solving the HJB-PDE and then synthesis the feedback optimal control. We let $\alpha = 0.1, r = 0.1, p = 0.001, x_0 = 1$, $U = [0,1]$ *in the HJB-PDE approach, we discretize the feasible control set with* $N_c = 1000$, 2000, 5000. *In Figure 3.1 we compare the optimal feedback control with* *the Riccati system based feedback law. And Figure 3.2 compares the corresponding state trajectories. In addition, from Step 2 of Algorithm 3.1, we can numerically compute the optimal feedback control mapping as in Figure 3.1.*

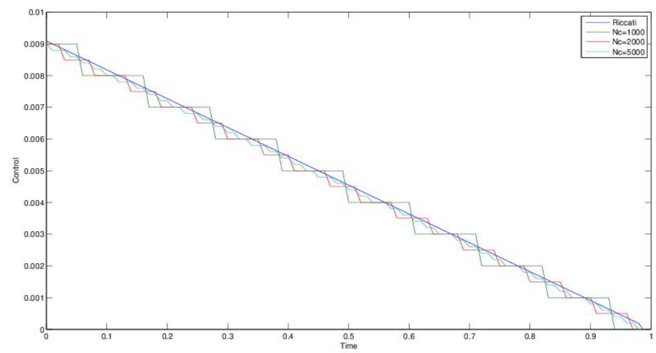

Figure 3.1. Comparing optimal feedback control calculated by Riccati system and by directly solving HJB-PDE with different control discretization (N_c = 1000, 2000, 5000)

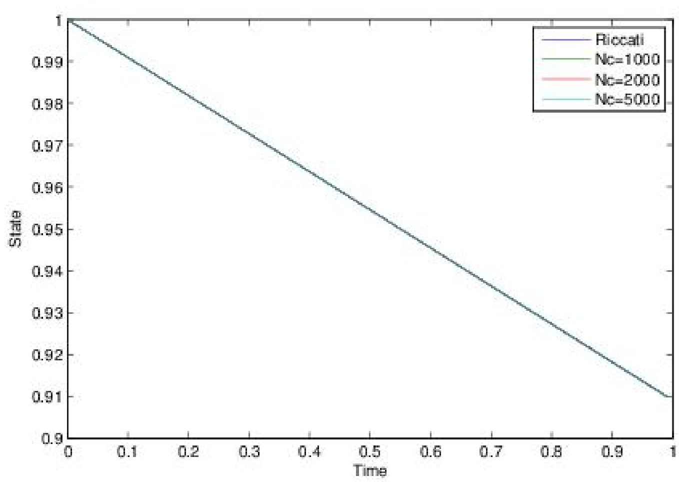

Figure 3.2. Comparing state trajectories corresponding to different feedback optimal control trajectories

From these figures we see that for LQ problems, optimal feedback solutions can be obtained by either solving the Riccati system (ODE), or by solving the HJB-PDE first then synthesize optimal solution via Algorithm 3.1.

3.3 Feedback SO-DTA: a Numerical Study

This section is devoted to the development of system optimal DTA with feedback information structure. Notice that this problem is not necessarily LQ, making this a case study to apply the methodologies developed in this chapter, especially, Algorithm 3.1.

Optimal control has been one of the major mathematical models towards the study of DTA problems (Peeta (2001) [111]), others include mathematical programming (e.g. Waller and Ziliaskopoulos (2006) [144]), variational equality/Nash-like games (e.g. Friesz et al. (1993) [55]), and simulation (e.g. Ben-Akiva et al. (2012) [12]). Examples of the application of optimal control methodologies are Friesz et al. (1989) [58], Ran et al. (1993) [118]. These papers feature traffic demands modeled in continuous time, use traffic flow conditions as state variables, and commuters' routing decisions as control variables. Optimal control analysis techniques such as Pontryagin's necessary conditions are used for the analysis of the system/user optimal conditions to gain insights about the dynamic traffic flow on networks. There have been very few papers on the study of dynamic traffic network with feedback information structure, and most of these works focus on the design of traffic control and routing systems instead of DTA models. For example, Gartner and Stamatiadis (1998) [63] uses feedback control techniques towards the design of a real-time traffic adaptive signal control system. Peeta and Yang (2003) [110] studies the route guidance strategies with feedback control and the effects on the stability of the associated DTA problems. In addition, there is another group of work from the automated control literature that proposes different versions of feedback regulators for closed loop DTA. Examples of this group of works are: and Papageorgiou (1990) [109], Kachroo et el. (1998) [75], Kachroo and Ozbay (2005) [76]. Please see Kachroo and Ozbay (2012) [77] for an in-depth review of this group of literature. Here we must point out that, despite of the similarities in the general methodologies employed, which are different versions of HJB-PDEs, the models in this group of work are very different from the model in this section in the following aspects:

- in this section we would like to propose another version of the DTA model following system optimal user behavior given the state information, whereas in Kachroo and Ozbay (2012) [77], the feedback models are developed as route guidance system;

- our results in this section are based on the solution of an HJB-PDE with time derivative of the value function, or the "evolutionary" HJB-PDE, whereas in Kachroo and Ozbay (2012) [77] the technique of H_∞ controller design are applied where the feedback control is in fact built on the solution of a steady-state HJB-PDE, even if the problem they consider is also defined on a finite time horizon.

Now we layout the feedback optimal control formulation of dynamic traffic assignment with feedback information structure under system optimal condition. These feedback optimal control problems are analyzed by HJB-PDE and solved by numerical techniques provided by for example, Altarovici et al. (2013) [3]. We will first review DTA formulated as an open-loop optimal control problem by Friesz et al. (1989) [58], and borrow its notations for our development of feedback DTA. For further analysis such as necessary conditions, etc. please refer to the full paper.

Assume that a transportation network could be represented by a directed graph $\mathcal{G}(\mathcal{N}, \mathcal{A})$: $\mathcal{A}$ denotes the set of all arcs; $\mathcal{N}$ is the set of all nodes. $\mathcal{W}$ is the set of all origin-destination pairs, for each $(o, d) \in \mathcal{W}$, $\mathcal{P}_{od}$ is the set of utilized paths that connects origin-destination pair $p = \{a_1, a_2, \ldots, a_{m(p)}\} \in \mathcal{P}$, $a_i \in \mathcal{A}$ path represented by the set of arcs it uses. On the modeling of traffic flow over time, we consider a finite planning time horizon $t \in [t_0, t_f]$. We use $x_a(t)$ as the volume of traffic on arc a at time t, which is the state variable; $g_a[x_a(t)]$ is the exit flow for arc a at time t; and $u_a(t)$ is the flow entering arc a at time t, which serves as the control variable. For a node k, $A(k)$ is the set of arcs leaving k; $B(k)$ is the set of arcs entering k. $s_k(t)$ is the traffic flow generated at node k. Given the traffic flow on arcs, $C_a(x_a)$ is used to describe the network traffic cost on arc a in system optimal (SO) problem.

Friesz et al. (1989) [58] defines the set of feasible arc flows as the following state-control tuple:

1. For all arcs, the following state dynamics is used to describe the change of traffic volume:

$$\frac{dx_a(t)}{dt} = u_a(t) - g_a[x_a(t)] \quad \forall a \in \mathcal{A}, t \in [0, T] \tag{3.24}$$

2. For all nodes, flow conservation should be satisfied:

$$s_k(t) + \sum_{a \in B(k)} g_a[x_a(t)] = \sum_{a \in A(k)} u_a(t) \tag{3.25}$$

3. Initial condition:

$$x_a(0) = x_a^0 \geq 0 \tag{3.26}$$

4. Non-negativity:

$$u_a(t) \geq 0, x_a(t) \geq 0 \tag{3.27}$$

Together, we define the following set of feasible solutions:

$$\Lambda_o = \{(x, u) : \text{(3.24),(3.25),(3.26),(3.27) are satisfied}\}$$

Definition 3.3. *(Dynamic System Optimal) If, over the fixed planning horizon* $[t_0, t_f]$*, the total cost of transportation is minimized, the corresponding flow pattern is said to be Dynamic System Optimal.*

Furthermore, Friesz et al. (1989) [58] formulates the dynamic UE and dynamic SO problem in the form of optimal control.

Problem 3.3. *(System Optimal DTA) In system optimal DTA, total transportation cost is minimized, resulting in the following optimal control problem:*

$$\begin{aligned} \min J &= \sum_{a \in \mathcal{A}} \int_{t_0}^{t_f} C_a(x_a(t))dt \\ \textit{subject to } (x, u) &\in \Lambda_o \end{aligned}$$

In these formulations, the functions $g_a(\cdot), C_a(\cdot), c_a(\cdot)$ are used to describe the dynamic network delay, or the so-called dynamic network loading (DNL) procedure. We are aware that such functions, especially with those strong regularities needed for a well-behaved DTA problem, may not be readily available to explicitly capture complicated and realistic traffic physics, such as spill-back, which has been included in state-of-the-art DNL modules (Garavello et al. 2016, [62]).

Before we begin to recast the open-loop Problem 3.3 in to their feedback counterparts, let us provide more details about the information structure following

Definition 3.1 in the context of traffic assignment, and explain what we mean by "feedback information structure" in this subsection.

In general, a feedback solution is a relationship between control variables and observable state variables. In this section, it is reasonable to assume the following: the network traffic flow $x_a(t)$ of any time $t \in [t_0, t_f]$, observable link-by-link, for each link a is observable in real time and taken into travelers' consideration when their routing decisions are made. This means that we assume the following:

$$u_a = u_a(x_1, \cdots, x_a, \cdots; t) = u_a(x, t), \quad \forall a \in \mathcal{A}, t \in [t_0, t_f]$$

With the above information structure assumed, we have the revised set of feasible solutions defined for the feedback optimal control problem:

1. For all arcs, the following state dynamics is used to describe the change of traffic volume:

$$\frac{dx_a(t)}{dt} = u_a(x, t) - g_a[x_a(t)] \quad \forall a \in \mathcal{A}, t \in [t_0, t_f] \tag{3.28}$$

2. For all nodes, flow conservation should be satisfied:

$$s_k(t) + \sum_{a \in B(k)} g_a[x_a(t)] = \sum_{a \in A(k)} u_a(x, t) \tag{3.29}$$

3. Initial condition:

$$x_a(0) = x_a^0 \geq 0 \tag{3.30}$$

4. Non-negativity:

$$u_a(x, t) \geq 0, x_a(t) \geq 0 \tag{3.31}$$

In addition, we let $f_a(x, u) = u_a(x, t) - g_a[x_a(t)]$, then $f = [\cdots, f_a, \cdots]$, and (3.28) becomes

$$\frac{dx(t)}{dt} = f(x, u)$$

Also define the set of feasible control as:

$$U(t, x(t)) = \{u(x, t) : (3.29) \text{ is satisfied}, u_a(x, t) \geq 0\}$$

We see that the user behavior principles still hold for the feedback information structure, leading to the reformulation of the DTA problems into feedback optimal control. And let us start with the feedback SO-DTA problem.

Problem 3.4. *(Feedback SO-DTA) In system optimal DTA, total transportation cost is minimized, resulting in the following optimal control problem:*

$$\begin{aligned} \min J &= \sum_{a\in\mathcal{A}} \int_{t_0}^{t_f} C_a(x_a)dt \\ \textit{subject to } \frac{dx(t)}{dt} &= f(x,u) \\ x_a(0) &= x_a^0 \geq 0, x_a(t) \geq 0 \\ u &\in U(t, x(t)) \end{aligned}$$

Since Problem 3.4 does not have terminal cost, the value function is defined as the minimum attainable cumulated traffic cost starting from a given state-time tuple, we have:

Definition 3.4. *(Value Function for Feedback-SO-DTA) The value function of Problem 3.4 is :*

$$V_{SO}(x,t) = \inf_{u(s), t\leq s\leq t_f} \left[\sum_{a\in\mathcal{A}} \int_{t}^{t_f} C_a(x_a)dt \right]$$

Still, we need the following regularities in the spirit of Definition 3.2 to proceed with further analysis:

Definition 3.5. *(Regularities for Feedback-SO-DTA) Problem (3.4) is said to be regular if the following holds:*
(i) the cost rate functions $C_a(x_a)$ are non-negative, increasing, Lipschitz continuous and bounded for all $x_a(t) \geq 0, a \in \mathcal{A}, t \in [t_0, t_f]$;
(ii) the exit function $g_a(x_a)$ are non-negative, increasing, Lipschitz continuous and bounded for all arc flow $x_a(t)$; in addition, for all $a \in \mathcal{A}$, $g_a(0) = 0$.
(iii) $C_a(x_a)$ is convex, $g_a(x_a)$ is concave for all $x_a(t) \geq 0, a \in \mathcal{A}, t \in [t_0, t_f]$, in addition, the set of feasible controls U should be compact.

In fact, this regularity condition is built by combining Definition 3.2 and Definition 1 from Friesz et al. (1989) [58]. Following such regularities we can easily get bounded and Lipschitz value function, which lead us to the following theorem:

Theorem 3.5. *(HJB-PDE of Feedback SO-DTA) Assume Problem 3.4 is regular as in Definition 3.5, the value function* $V_{SO}(x,t)$ *is a viscosity solution of HJB-PDE (3.32) and boundary condition (3.33):*

$$\frac{\partial V_{SO}}{\partial t} + \min_{u \in U} \left[f(x,u)^T \nabla V_{SO} + \sum_a C_a(x_a) \right] = 0 \tag{3.32}$$

$$V_{SO}(x, t_f) = 0 \tag{3.33}$$

Remark 3.2. *In the framework we introduced so far, we have introduced a simple link-delay model to serve as the network loading module so that the model together can be computationally tractable. Notice that possible changes of the traffic dynamics may not impact our ability to apply the optimal control synthesis techniques. Especially, a path based optimal control formulation with the SO traffic assignment principle may also be introduced. Meanwhile, with the rapid development of advanced transportation information system (ATIS) and spatial-temporal machine-learning techniques, it is possible to develop statistical DNL models that also enjoy easy-to-handle analytical properties. (Song et al. (2017) [135])*

For the preparation of numerical examples, we start from a simple two-node two-arc network as in Figure 3.3. There is only 1 OD pair on this network (1,2). We employ ROC-HJ software for the numerical computation of HJB-PDE solutions, please see details in Appendix. All other parts of the numerical example, including the feedback control synthesis, and the open-loop traffic assignment in this section are prepared in Matlab 2015b on a laptop computer equipped with Intel (R) Core i5 (TM) processor and 8GB RAM.

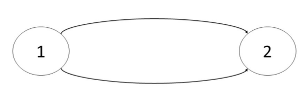

Figure 3.3. A small network with 2 arcs and 1 OD pair (1,2)

First we formulate the Feedback SO-DTA as in Problem 3.4:

$$\min J_{SO}(u_1) = \int_{t_0}^{t_f} [C_1(x_1) + C_2(x_2)] dt$$

$$
\begin{aligned}
\text{s.t.} \frac{dx_1(t)}{dt} &= u_1(x,t) - g_1[x_1(t)] \\
\frac{dx_2(t)}{dt} &= s(t) - u_1(x,t) - g_2[x_2(t)] \\
x_1(0) &= x_1^0, x_2(0) = x_2^0 \\
0 &\leq u_1(x,t) \leq s_1(t) \\
x_1(t) &\geq 0, x_2(t) \geq 0
\end{aligned}
$$

Note that in this problem, flow conservation constraint (3.29) only holds for node 1 as $s(t) = u_1(x,t) + u_2(x,t)$. To simplify the above formulation, this flow conservation constraint is immediately applied to eliminate u_2. The value function becomes:

$$
V_{SO}(x,t) = \min_{u(s), t_0 \leq s \leq t_f} \left[\int_t^{t_f} [C_1(x_1) + C_2(x_2)] dt \right]
$$

Furthermore, we formulate the HJB-PDE for this problem following Theorem 3.5:

$$
H_{SO}(x,p) = \min_{u \in U} \left[\begin{pmatrix} u - g_1(x_1) \\ s - u - g_2(x_2) \end{pmatrix} \cdot p + C_1(x_1) + C_2(x_2) \right]
$$

This gives us the HJB-PDE with boundary condition

$$
\frac{\partial V_{SO}}{\partial t} + H_{SO}(x, \nabla V_{SO}) = 0 \quad (3.34)
$$

$$
V_{SO}(x, t_f) = 0 \quad (3.35)
$$

Following Friesz et al. (1989) [58] we define the arc total cost rate function $C_a(x_a) \triangleq c_a(x_a) g_a(x_a)$. We can check that those network cost functions meet the regularities specified by Definition 3.5. Table 3.1 gives the detailed parameterization of different functions. We use quadratic function for $c_a(x_a)$ and use an exit flow function based on exponential functions similar to Papageorgiou (1990) [109]. Other parameters are:

$$
\begin{aligned}
t_0 &= 0, t_f = 2, s(t) = 1, t \in [0,2] \\
x^0 &= (1,1)
\end{aligned}
$$

As we know now, the value function in this example takes the form $V(x_1, x_2, t)$,

Arc	1	2
$c_a(x_a)$	$0.2x_1^2+0.1$	$0.1x_2^2+0.2$
$g_a(x_a)$	$0.6(1-\exp(\frac{-x_1}{2}))$	$0.8(1-\exp(\frac{-x_1}{3}))$
$C_a(x_a)$	$c_1(x_1)g_1(x_1)$	$c_2(x_2)g_2(x_2)$

Table 3.1. Parameters for the Two-Arc Network

and the numerical solution of the value function is shown in Figure 3.4, where we plot the function V_{SO} in (x_1, x_2) at different instant of time. In applying the ROC-HJ package, we assume that the state space to consider is $0 \leq x_1 \leq 2$, with $N_x = 40$ mesh points per dimension. Time is discretized into $N_t = 40$ mesh points also. In the numerical approximation of Hamiltonian function, a control mesh of $N_c = 5000$ points is taken.

As a benchmark we compare the feedback solution with the open loop solution based on time discretization and reformulation into Nonlinear Programming (NLP). From our definition of the value function we can see that the system optimal network traffic cost is $J^*_{SO} = V_{SO}(x_1^0, x_2^0, 0) = V_{SO}(1,1,0) = 0.5595$; whereas the open loop system optimal total traffic cost is $J^*_{SO,Open-Loop} = 0.7161$

Algorithm 3.1 is employed to carry out the synthesis of feedback optimal control. To be more specific, in Step 2 of Algorithm 3.1 we use the following finite difference scheme to prepare $\frac{\partial V(t,x)}{\partial x}, \frac{\partial V(t,x)}{\partial t}$ at each state-time mesh point: first order forward scheme in time, first order central scheme in space. As a result of this step, an optimal control $u_1^*(\cdot)$ can be obtained by solving (3.16) for all possible mesh points. Here Figure 3.5 presents the optimal control mapping at time $t = 0$ for all points on the state mesh grid.

On the other hand, in order to speed up the optimal control synthesis for the traffic assignment, there are two things we should pay attention to:

- It is not necessary to have $u_1^*(\cdot)$ solved for all mesh points, especially when the state dimension is high. This means that Step 2 through 4 of Algorithm 3.1 could be carried out by time forwarding, which is the faster way to obtain feedback DTA (control) along with the state trajectories.

- Since the state dynamics of this problem is linear in control and the running cost does not include control explicitly, (3.16) is in fact linear in control, which means the following control rule holds when (x,t), $V(x,t)$ and its derivatives

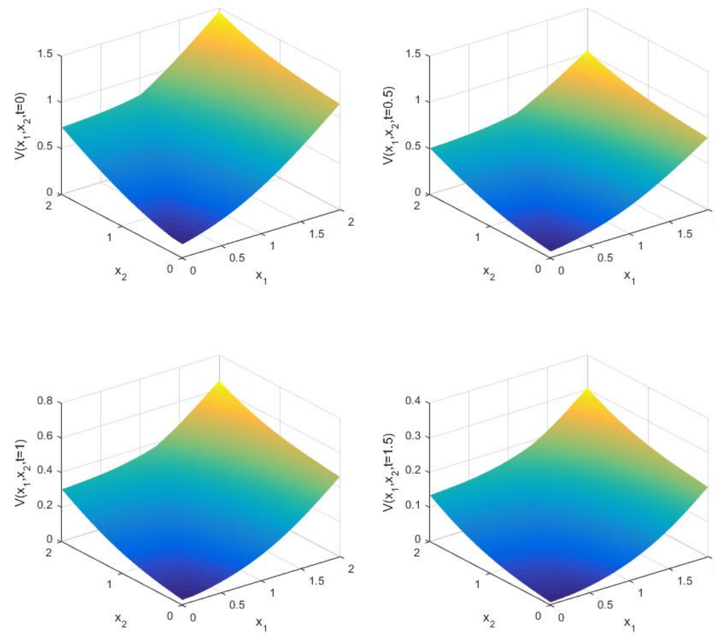

Figure 3.4. Numerical Solution of V_{SO}, the value function of Feedback SO-DTA problem; plotted in (x_1, x_2) at time $t = 0, 0.5, 1, 1.5$

are known:

$$u^*(x_1, x_2, t) = \frac{g_1(x_1)V_{x1} + g_2(x_2)V_{x2} - V_t - s(t)V_{x2} - C_1(x_1) - C_2(x_2)}{V_{x1} - V_{x2}}$$

where a subscript in V means its corresponding derivative. Such control rule is implemented for faster feedback control synthesis.

In fact, Figure 3.5 is generated by evaluating the above control law over all (x_1, x_2) at time $t = 0$.

Finally, we employ the optimal control rule and compare the feedback state-control trajectory with the solution from open-loop problem as a benchmark, Figure 3.6 reflects such comparison.

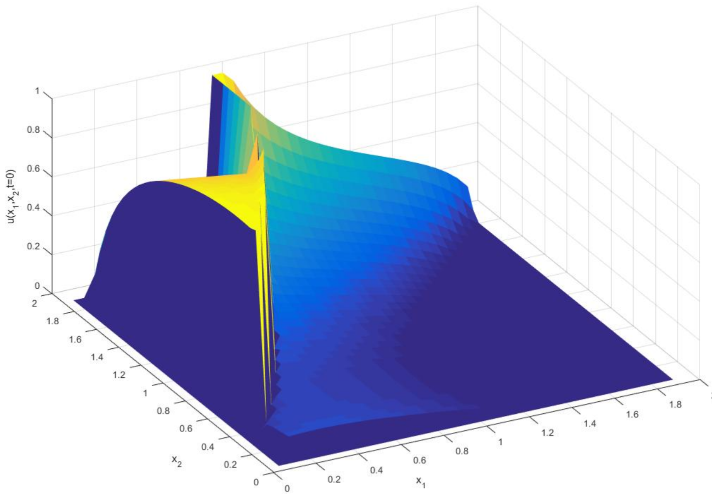

Figure 3.5. Feedback optimal control mapping at $t = 0$ plotting over x_1 and x_2

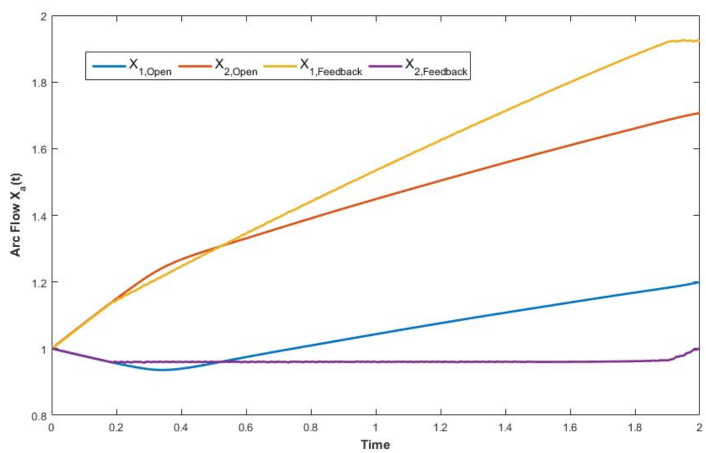

Figure 3.6. Compare the state trajectories of SO DTA with feedback and open loop policies

3.4 Feedback Differential Games

With the preparation of the feedback control problem, we are ready to review key results in differential Nash games with feedback. Similar to the open-loop case as in Definition 2.11, we consider N players simultaneously optimizing their own objective functional subject to a system dynamics affected by strategies of all players. Here, the state information, which is available to all agents in real time, is also used towards their decision, leading to the following definition:

Definition 3.6. *(Differential Nash equilibrium with Feedback) Suppose there are N agents, each of which chooses a feasible strategy vector u^i from the strategy set Λ_i which is independent of the other players' strategies. Furthermore, every agent i has a cost/disutility functional $J_i(u) : \Lambda \to \mathbb{R}^1$ that depends on all agents' strategies where*

$$\begin{aligned} \Lambda &= \prod_{i=1}^{N} \Lambda_i \\ u &= (u^i : i = 1, \cdots, N) \end{aligned}$$

Every agent seeks to solve the optimal control problem

$$\min J^i(u^i; u^{-i}) = \int_{t_0}^{t_f} f_0^i(x, u^i, u^{-i}, t)dt + K^i[x(t_f), t_f]$$

subject to

$$\begin{aligned} \frac{dx}{dt} &= f(x(t), u^i(t), u^{-i}(t), t) \\ x(t_0) &= x_0 \\ u^i(x, t) &\in \Lambda_i. \end{aligned}$$

A differential Nash equilibrium is a tuple of strategies u such that u^i solves the optimal control problem.

Intuitively, it is natural to pose the necessary condition of feedback Nash equilibrium as the solution of the corresponding system of HJB-PDEs. We list the following theorem:

Theorem 3.6. *(Feedback differential Nash equilibrium as system of HJBs) Let the N-tuple $\left(U^1, U^2, \cdots, U^N\right)$ of functions $U^i : [0, T] \times X \to \mathbb{R}$ be feedback strategies of each player, and make the following assumptions:*
(i)there exists a unique absolutely continuous solution $x : [0, T] \to X$ to the initial value problem:

$$\dot{x}(t) = f(t, x(t), U^1(t, x), U^2(t, x), \cdots, U^N(t, x)) \tag{3.36}$$

$$x(0) = x_0 \tag{3.37}$$

(ii)$\forall i, \exists$ a continuously differentiable function $V^i(x,t)$ such that the HJB equations: (treating u^{-i} as parameters)

$$-\frac{\partial V^i(t,x)}{\partial t} = \min_{u^i}\left[\frac{\partial V^i(t,x)}{\partial x} f(t,x,u^i,u^{-i}) + f_0(t,x,u^i,u^{-i})\right] \tag{3.38}$$

are satisfied for all $(t,x) \in [0,T] \times X$
(iii)Boundary conditions, for $\forall i, \forall x \in X$:

$$V^i(x,T) = q^i(x,T) \tag{3.39}$$

Denote by $\Phi^i(x,t)$ the set of all $u^i \in U^i(t,x(t))$ which maximize the right hand side of (3.38). If $u^i \in \Phi^i(x,t)$ holds the above assumptions for almost all t,then $\left(U^1, U^2, \cdots, U^N\right)$ is a feedback Nash equilibrium.

Proof. For detailed proof, see e.g. Dockner (2000) [41], Theorem 4.1. □

Note that this is the characterization based on the system of HJB-PDEs. There is an alternative approach based on the PMP (with the $\frac{\partial H}{\partial x}$ term). We refer the reader to Theorem 4.2 of Dockner [41].

3.4.1 The LQ Special Case for Feedback DNEs

Again, we list the LQ problem as a special case. With linear dynamics and quadratic objectives, the optimal control for Player i from Definition 3.6 becomes:

$$\min J^i(u^i;u^{-i}) = \int_{t_0}^{t_f}\left[x^T(t)K^i x(t) + \sum_{j=1}^{N} u_j^T(t)R_{ij}u_j(t)\right]dt + x^T(t_f)K_f^i x(t_f)$$

subject to

$$\begin{aligned} \frac{dx}{dt} &= Ax(t) + \sum_{i=1}^{N} B_i u_i(t) \\ x(t_0) &= x_0 \end{aligned}$$

For the case with 2-players, we have:

Definition 3.7. *(Differential Nash equilibrium with linear feedback) A set of controls* $u^{i,*} = F^{i,*}(t)x(t)$, $i = 1, 2$ *is a linear feedback differential Nash equilibrium if for all* $u_i \in \Lambda^{i,lfb}$, *where*

$$\Lambda^{i,lfb} = \left\{ u^i | u^i(x,t) = F^i(t)x(t) \right\}$$

Note that this definition is a special case of the general LQ DNE with feedback, later Engwerda (2005) argues that in the case of LQ problems the two cases coincide. Nevertheless, the linear feedback DNE also admits the solution of Riccati system as its necessary and sufficient condition. Similar to the case of LQ optimal control, define the intermediate variable

$$\begin{aligned} S_i &= B_i R_{ii}^{-1} B_i^T \\ S_{ij} &= B_i R_{ii}^{-1} R_{ji} R_{ii}^{-1} B_i^T \end{aligned}$$

Theorem 3.7. *(Two-person Feedback LQ games) The two-person LQ differential game has, for every initial state, a Nash equilibrium with linear feedback if and only if the following Riccati system has a set of symmetric solutions* $Q_{1,2}$*:*

$$\begin{aligned} \dot{Q}_1(t) &= -(A - S_2 Q_2(t))^T Q_1(t) - Q_1(t)(A - S_2 Q_2(t)) \\ &\quad + Q_1(t) S_1 Q_1(t) - K_1 - Q_2(t) S_{21} Q_2(t) \\ Q_1(t_f) &= K_f^1 \\ \dot{Q}_2(t) &= -(A - S_1 Q_1(t))^T Q_2(t) - Q_2(t)(A - S_1 Q_1(t)) \\ &\quad + Q_2(t) S_2 Q_2(t) - K_2 - Q_1(t) S_{12} Q_1(t) \\ Q_2(t_f) &= K_f^2 \end{aligned}$$

and then it is unique and the optimal control in feedback form is:

$$u^{i,*}(t) = -R_i^{-1} B_i^T Q_i(t) x(t) \quad i = 1, 2 \tag{3.40}$$

Proof. See Engwerda (2005) [46], Theorem 8.3. □

3.5 Summary

In this chapter we provided a tutorial on solving deterministic optimal control and differential games with state feedbacks. Central to this approach is the value function and its solution via the HJB-PDEs. Proper synthesis procedure should be applied to (numerically) obtain optimal control in feedback form. We also briefly reviewed the LQ problem as a special case, and noted the differential Riccati system as a consequence of separation of variables when solving the HJB-PDE. Later in Chapter 6 we will present the application of this techniques under the context of pricing and revenue management.

As a numerical case study, we take advantage of existing open-loop optimal control formulations of the SO-DTA and extended them into case with feedback information structure which could be effectively solved through the numerical procedure reviewed in this chapter.

Chapter 4
Stochastic Maximum Principle and Stochastic DVIs

It is well known that Pontryagin's maximum principle and Bellman's dynamic programming are the two most popular approaches in solving optimal control problems and differential Nash games. On the other hand, in the domain of stochastic problems, the latter is more popular, and a Markovian information structure is often assumed. In this chapter, we aim at investigating the solution strategy of stochastic differential Nash games with adapted open-loop information structure with stochastic differential variational inequalities (S-DVIs). S-DVI is a stochastic generalization of the deterministic DVI. Due to the non-anticipating nature of the stochastic problem considered, the computation of S-DVIs will require the solution of so-called forward-backward stochastic differential equations (FBSDEs), which is just the stochastic generalization of two-point boundary value problems. The methodologies summarized and developed in this chapter will be later applied in Chapter 7 in the computation of stochastic differential monopsony and oligopsony.

The rest of this chapter is structured as follows: we will start from the analysis on stochastic optimal control with Brownian motion in Section 4.1, in which several necessary or sufficient conditions for a stochastic optimal control problem are reviewed and presented. Based on these results, in Section 4.2 we propose the definition and necessary conditions of the S-DVI problem. Later in Section 4.3 we discuss the relationship between S-DVIs and stochastic dynamic Nash equilibriums. Finally, we discuss the solution of FBSDEs in Section 4.4.

4.1 Stochastic Optimal Control and Its Necessary Conditions

In this section we look at the stochastic extension of optimal control problem in Chapter 2. Starting form the stochastic dynamics, we look at formulation, existence and necessary/sufficient conditions of such problem. Some good references of stochastic optimal control theory in continuous time are: Malliaris and Brock (1982) [96], Yong and Zhou (1999) [148]. These theories have been widely applied in different aspects of operations research/management science (Barsar and Olzder (1999) [10], Dockner (2010) [41]) and financial economics (Duffie (2010) [43]).

4.1.1 Brownian Motions and Stochastic Optimal Control

We start with a very brief review of Brownian motion, for a comprehensive review on related topics please refer to, for example, Karatzas and Shreve (2012) [79].

Definition 4.1. *(Brownian Motion) A standard m-dimensional Wiener process B with time domain $[t_0, t_f]$ is a continuous time stochastic process with values in $\mathbb{R}^m$, that is, $B : [t_0, t_f] \times \Xi \to \mathbb{R}^m$ with the following properties:*
(i) $B(0, \xi) = B_0$ for all ξ in a set of probability 1 where $B_0 \in \mathbb{R}^m$ in an arbitrary initial value;
(ii) for any finite sequence of real numbers $(t_1, t_2, ..., t_l)$ with $t_0 \leq t_1 < t_2 < ... < t_l \leq t_f$ it holds that the random variables $B(t_1, \cdot)$ and $B(t_{i+1}, \cdot) - B(t_i, \cdot)$, $i \in \{1, 2, ..., l-1\}$ are stochastically independent;
(iii) for all pairs (s, t) of real numbers such that $t_0 \leq s < t \leq t_f$, the random variable $B(t, \cdot) - B(s, \cdot)$ has a normal distribution with mean vector $0 \in \mathbb{R}^m$ and covariance matrix $(t-s)I$, where $I \in \mathbb{R}^{m\times m}$.

Given a filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq t_0}, \mathbf{P})$ such that a standard m-dimensional Wiener process B_t is well-defined and consider the following controlled stochastic differential equation (SDE) to denote the evolution of the state variable x:

$$dx(t) = f(x(t), u(t), t)dt + \sigma(x(t), u(t), t)dB_t, \quad x(t_0) = x_0 \tag{4.1}$$

We assume that $x(t) \in \mathbb{R}^n$ and $u(t) \in U$ for all t, therefore:

$$\begin{aligned} f(x,u,t) &: \mathbb{R}^n \times U \times [t_0, t_f] \to \mathbb{R}^n \\ \sigma(x,u,t) &: \mathbb{R}^n \times U \times [t_0, t_f] \to \mathbb{R}^{n\times m} \end{aligned}$$

with $\sigma_{ij}(x,u,t)$ measures the direct influence of the j-th component of the m-dimensional Brownian motion on the evolution of the i-th component of the n-dimensional state vector. The interpretation of the above differential equation should be that $x(\cdot)$ satisfies the integral equation:

$$x(t) = x_0 + \int_{t_0}^{t} f(x(s), u(s), s)ds + \int_{t_0}^{t} \sigma(x(s), u(s), s)dB_s \tag{4.2}$$

for all ξ in a set of probability 1. The interpretation of the second integration is the following limit:

$$\lim_{\delta\to 0} \sum_{l=1}^{L-1} \sigma(x(t_l), u(t_l), t_l)[B(t_{l+1}) - B(t_l)] \tag{4.3}$$

where $t_0 = t_1 < t_2 < ... < t_L = t_f$ and $\delta = \max\{|t_{l+1} - t_l|, 1 \leq l \leq L-1\}$.

The next important result is Itô's lemma, which on a high level could be interpreted as a "chain rule" for stochastic differential equations.

Lemma 4.1. *(Itô's) Suppose that $x(\cdot)$ solves the SDE (4.1) and let $G : X \times [0,T] \to \mathbb{R}$ be a (deterministic) function with continuous partial derivatives G_t, G_x, G_{xx}. Then the function $g(t) = G(x(t), t)$ satisfies the following SDE:*

$$\begin{aligned} dg(t) = \{&G_t(x(t),t) + G_x(x(t),t) f(x(t),u(t),t) \\ &+ (1/2)tr[G_{xx}(x(t),t)\sigma(x(t),u(t),t)\sigma(x(t),u(t),t)^T]\}dt \\ &+ G_x(x(t),t)\sigma(x(t),u(t),t)dB_t \end{aligned}$$

Proof. See Dockner (2010) [41], Lemma 8.2. □

One of the most important application of Itô's lemma is the derivation of the SDE for the so-called Geometric Brownian Motion (GBM).

Example 4.1. *(Geometric Brownian Motion) The SDE describing GBM is the* *following: (with* $f(x(t), u(t), t) = \mu x, \sigma(x(t), u(t), t) = \sigma x$*)*

$$dx = \mu x dt + \sigma x dB_t$$

this is equivalent to:

$$\frac{dx}{x} = \mu dt + \sigma dB_t$$

means $G(x(t), t) = \ln(x), G_x = 1/x, G_t = 0, G_{xx} = -1/x^2$. *Hence*

$$d \ln(x(t)/x(0)) = (\mu - \frac{\sigma^2}{2})dt + \sigma dB_t$$

The stochastic optimal control problem considered in this chapter is given as follows, which is a natural extension of the deterministic Problem 2.4:

Problem 4.1. *(Stochastic Optimal Control Problem, Minimization) We consider the following stochastic optimal control problem with* $f_0(x, u, t) : \mathbb{R}^n \times U \times [t_0, t_f] \to \mathbb{R}$, $f(x, u, t) : \mathbb{R}^n \times U \times [t_0, t_f] \to \mathbb{R}^n$, $\sigma(x, u, t) : \mathbb{R}^n \times U \times [t_0, t_f] \to \mathbb{R}^{n\times m}$*:*

$$\begin{aligned} \min_u J &= E_{u(\cdot)}[\int_{t_0}^{t_f} f_0(x(t), u(t), t)dt + K(x(t_f))] && (4.4) \\ \textit{such that } dx &= f(x(t), u(t), t)dt + \sigma(x(t), u(t), t)dB_t, \quad x(t_0) = x_0 && (4.5) \\ u(t) &\in U \end{aligned}$$

Similar to the deterministic case, we first discuss the issue of how a feasible and an optimal control should be defined in the stochastic case. And this will eventually lead to the formulation of stochastic optimal control problem.

Definition 4.2. *(Feasible and Optimal Control)*
(i)A control path $u(\cdot)$ *is non-anticipating if its value at time* t *does not depend on any uncertainty revealed after time* t*. That is,* $u(t, \xi)$ *must not depend on realizations of the random variables* $B(t + \tau, \xi)$ *for any positive* τ*.*
(ii)A control path $u : [t_0, t_f] \times \Xi \to \mathbb{R}^m$ *is feasible for the stochastic optimal control problem stated above if it is non-anticipating, if there exists a unique solution* $x(\cdot)$ *to the SDE 4.5, if any constraints* $x(t) \in X$ *and* $u(t) \in U$ *are satisfied w.p. 1 for all t, and if the integral in (4.4) is well defined.*
(iii)(for finite t_f*) A control path* $u(\cdot)$ *is optimal if it is feasible and if* $J(u(\cdot)) \leq J(\tilde{u}(\cdot))$ *holds for all feasible control paths* $\tilde{u}(\cdot)$*.*

Remark 4.1. *The definitions and formulations listed above are minimal as a preparation for the application we need for the rest of this dissertation. There are other topics of great mathematical value and interest that we did not cover, such as reachable sets/controllability, existence of optimal control, and comparisons between the so-called strong and weak formulations. Please see Yong and Zhou (1999) [148], Chapter 2: Section 4, 5, 6 and 7 for detailed review and proof of these results.*

We will first look at the necessary and sufficient conditions for the stochastic optimal control problem based on the POO and the Hamilton-Jacobi-Bellman equation, the idea is very similar to what is reviewed in Chapter 3. We have the following theorem from Dockner (2010) [41] on a HJB-PDE in one dimension:

Theorem 4.1. *(Sufficient condition for stochastic optimal control) Let $V : X \times [t_0, t_f] \to \mathbb{R}$ be a function with continuous partial derivatives V_t, V_x, V_{xx} and assume that V satisfies the HJB equation:*

$$-V_t(x,t) = \max_{u(t)\in U(x,t)} \left\{ f_0(x,u) + V_x(x,u,t)f(x,u,t) + \frac{1}{2}tr[V_{xx}(x,t)\sigma(x,t)\sigma(x,t)^T] \right\} \tag{4.6}$$

for all $(x,t) \in X \times [t_0, t_f]$. Here $tr[\cdot]$ denote the trace of a square matrix, which is the sum of all elements on its diagonal. Now let $U^0(x,t)$ denote the set of controls maximizing the right hand side of (4.6) and let $u(\cdot)$ be a feasible control path with the state trajectory $x(\cdot)$ such that $u(t) \in U^0(x,t)$ holds w.p. 1 for almost all $t \in [t_0, t_f]$. If t_f is finite and if the boundary condition $V(x, t_f) = K(x)$ holds for all $x \in X$, then $u(\cdot)$ is an optimal control path.

Proof. See Dockner (2010) [41], pp. 229 for a full proof. □

At the same time, we are more interested in the characterization of optimal control via necessary conditions consisting of stochastic maximum principle associated with a series of stochastic state and adjoint dynamics. Malliaris and Brock (1982) [96] has a series of such necessary conditions. Again, we consider the one dimensional special case. And we will see later, the proof given by Malliaris and Brock utilizes some very strong assumptions on the value function and its Taylor expansions, despite of this, their result is useful in giving intuitions.

Theorem 4.2. *(Stochastic Pontryagin Maximum Principle) Suppose that $x(t)$ and $u^0(t)$ is a solution for Problem 4.1 in one dimension. And assume that a value*

function $V(x(t), t, t_f)$ in the sense of Theorem 4.1 is well defined with continuous partial derivatives in state and time. Then there is a co-state variable $\lambda(t)$ such that for each $t \in [t_0, t_f]$ we have, in addition to the state equation (4.5) the following:
(i) $u^0(t)$ maximizes $H(x, u, \lambda, \partial\lambda/\partial x)$ where

$$H(x, u, \lambda, \partial\lambda/\partial x) = f_0(x, u) + \lambda f(x(t), u(t), t) + \frac{1}{2}\sigma^2 \frac{\partial \lambda}{\partial x} \tag{4.7}$$

(ii) the co-state function $\lambda(t)$ satisfies the stochastic differential equation, here $H^0 = H(x, u^0, \lambda, \partial\lambda/\partial x)$

$$d\lambda = -H_x^0 dt + \sigma(x, u^0) V_{xx} dB_t \tag{4.8}$$

(iii) the transversality condition:

$$\lambda(x(t_f), t_f) = K_x(x(t_f), t_f) \tag{4.9}$$

Proof. Here we summarize the proof recorded in Malliaris and Brock (1982) [96], Proposition 10.1. Consider the value function $V(x(t), t, t_f)$, which is similar to the deterministic case, and defined as the optimal value one can get starting from state $x(t)$ at time t to the end of horizon t_f. The proof is in the following four steps:

[Step 1] use Bellman's Principle of Optimality on the value function:

$$\begin{aligned} V(x(t), t, t_f) &= \max_u E \int_t^{t_f} f_0(x(t), u(t)) dt \\ &= \max_u E \int_t^{t+\Delta t} f_0 x(t), u(t)) dt + \max_u E \int_{t+\Delta t}^{t_f} F(x(t), u(t)) dt \\ &= \max_u E \left[\int_t^{t+\Delta t} f_0(x(t), u(t)) dt + V(x(t+\Delta t), t+\Delta t, t_f) \right] \end{aligned} \tag{4.10}$$

Note that (4.10) is true due to the POO and the definition of value function at next instant of time $t + \Delta t$.

[Step 2] use Taylor expansion and get the HJB: Assume V has continuous partial derivatives of all orders less than 3. Then Taylor expansion could be applied towards (4.10):

$$V(x(t), t, t_f) = \max_u E[f_0(x(t), u(t))\Delta t + V(x(t), t, t_f)$$

$$+ V_x\Delta x + V_t\Delta t + \frac{1}{2}V_{xx}(\Delta x)^2 + V_{xt}(\Delta x)(\Delta t) + \frac{1}{2}V_{tt}(\Delta t)^2 + o(\Delta t)] \tag{4.11}$$

Also let the state equation to be approximated by the following:

$$\Delta x = f(x,u)\Delta t + \sigma(x,u)\Delta B_t + o(\Delta t) \tag{4.12}$$

Plug 4.12 back into 4.11, and note the following rule of multiplication due to the properties of Brownian motion:

$$(\Delta t)^2 = 0;\ (\Delta B_t)^2 = \Delta t;\ (\Delta t)(\Delta B_t) = 0.$$

We have

$$0 = \max_u E[f_0(x,t)\Delta t + (V_x f + V_t + \frac{1}{2}V_{xx}\sigma^2)\Delta t + V_x\sigma\Delta B_t + o(\Delta t)]$$

Let $\Delta J = [V_x f + V_t + \frac{1}{2}V_{xx}\sigma^2] + V_x\sigma\Delta B_t$ the above becomes:

$$0 = \max_u E[f_0(x,t)\Delta t + \Delta J + o(\Delta t)]$$

Also assume that we can pass expectation E onto each term. We then divide both sides by Δt, and let $\Delta t \to 0$, which will lead to the following:

$$-V_t = \max_u [F(x,u) + V_x f(x,u) + \frac{1}{2}V_{xx}\sigma^2]$$

with boundary condition

$$\frac{\partial V(x(t_f), t_f, t_f)}{\partial x} = K_x(x(t_f), t_f)$$

[Step 3] introduce the co-state variable: Now define the co-state variable $\lambda(t)$

$$\lambda(t) \doteq V_x(x, t, t_f)$$

Then $\lambda_x = V_{xx}$ and the HJB could be re-written as

$$-V_t = \max_u H(x, u, \lambda, \frac{\partial\lambda}{\partial x})$$

Assume that a function u^0 exists that solves the maximization problem of the RHS of the above, and denote this function by:

$$u^0 = u^0(x, \lambda, \frac{\partial \lambda}{\partial x})$$

And the HJB could be again re-written as:

$$-V_t = H^0(x, \lambda, \frac{\partial \lambda}{\partial x})$$

where

$$H^0(x, \lambda, \frac{\partial \lambda}{\partial x}) = f_0(x, u^0) + \lambda f(x, u^0) + \frac{1}{2}\frac{\partial \lambda}{\partial x}\sigma^2(x, u^0)$$

[Step 4] form the state and adjoint dynamics as the necessary condition. We first consider the state dynamics:

$$\begin{aligned} dx &= f(x, u^0)dt + \sigma(x, u^0)dB_t \\ &= H^0_\lambda dt + \sigma dB_t \end{aligned}$$

Use Itô's lemma, we have:

$$\begin{aligned} dp &= V_{xt}dt + V_{xx}dx + \frac{1}{2}V_{xxx}(dx)^2 \\ &= [V_{xt} + V_{xx}f + \frac{1}{2}V_{xxx}(\sigma)^2]dt + V_{xx}\sigma dB_t \end{aligned} \tag{4.13}$$

Since the HJB holds, we can compute V_{tx} based on it:

$$-V_{tx} = H^0_x + fV_{xx} + \frac{1}{2}V_{xxx}(\sigma)^2 \tag{4.14}$$

plug 4.14 back into 4.13 we have the desired result of co-state dynamics:

$$d\lambda = -H^0_x dt + \sigma V_{xx} dB_t$$

□

4.1.2 Spike Variation Method for Stochastic Maximum Principle

Notice that Theorem 4.2 above relies heavily on the existence and smoothness of the value function $V(\cdot)$. In fact, a series of more rigorous approaches have been proposed by, for example, Bensoussan (1983) [13] and Peng (1990) [112]. In this group of literature, a "spike variation method" has been considered. The main idea is to examine the relationship between the objective functional of the optimal-control-trajectory and a "neighboring" trajectory: let $u^*(\cdot)$ be an optimal control, this approach will consider $u^*(\cdot) + \theta(u(\cdot) - u^*(\cdot))$ where $u(\cdot)$ is another admissible control, hence by definition:

$$J(u^*(\cdot) + \theta(u(\cdot) - u^*(\cdot))) \geq J(u^*(\cdot))$$

and thus formatively we can write:

$$\langle J'(u^*(\cdot)), u(\cdot) - u^*(\cdot)\rangle \geq 0, \quad \forall u(\cdot) \in U$$

provided the Gateaux differential of $J(u(\cdot))$ as a functional on the Hilbert space $L^2_F(0,T)$ is well defined. In the rest of this section we review Bensoussan (1983) [13], Section 4 and Yong and Zhou (1999) [148] Chapter 3. Under reasonable regularity conditions, a variational inequality along with the stochastic state and adjoint dynamics will be employed to characterize the optimal control of Problem 4.1. We start from the regularity conditions:

Definition 4.3. *(Regularity Conditions of Stochastic Optimal Control Problem) Problem 4.1 is regular if the following conditions hold:*
(i) $f(x,u,t) : \mathbb{R}^n \times U \times [t_0, t_f] \to \mathbb{R}^n$ *is continuously differentiable w.r.t.* x *and* u*. Also the first order partials* f_x, f_u *are bounded. This means:*

$$|f(x,u,t)| \leq C_1(1 + |x| + |u|);$$

(ii) $\sigma(x,u,t) : \mathbb{R}^n\times U \times [t_0, t_f] \to \mathbb{R}^n\times \mathbb{R}^n$ *is continuously differentiable w.r.t.* x *and* u*. Also the first order partials* σ_x, σ_v *are bounded. This means:*

$$|\sigma(x,u,t)| \leq C_2(1 + |x| + |u|);$$

(iii) the set of feasible control $u \in U \subseteq \mathbb{R}^m$ is closed and convex;
(iv) $K(x)$ is continuously differentiable, moreover

$$K_x(x) \leq C_3(1 + |x|);$$

(v) $f_0(x, u, t) : \mathbb{R}^n \times U \times [t_0, t_f] \to \mathbb{R}$ is Borel, continuously differentiable w.r.t. x and u :

$$\begin{aligned} |f_{0,x}(x, u, t)| &\leq C_4(1 + |x| + |u|) \\ |f_{0,u}(x, u, t)| &\leq C_5(1 + |x| + |u|) \end{aligned}$$

The stochastic Hamiltonian is defined as

$$H(x, u, t; p, q) = p^T f(x, u, t) + \operatorname{tr}\left[q^T \sigma(x, u, t)\right] - f_0(x, u, t) \tag{4.15}$$

with $(x, u, p, q) \in \mathbb{R}^n \times U \times \mathbb{R}^n \times \mathbb{R}^{n\times n}$, and $\operatorname{tr}[\cdot]$ means the trace of a matrix which is the sum of all its diagonals. Notice that this time the stochastic Hamiltonian includes two adjoint variables: (i) variable p associated with the drift term of the state dynamics, (ii) variable q associated with volatility. Still following Yong and Zhou (1999) [148], we have:

Theorem 4.3. *(Stochastic Maximum Principle) We assume the regularities listed as in Definition 4.3. Let $u^*(t)$ be an optimal control and $x^*(t)$ the corresponding trajectory which solves Problem 4.1, then they solve the following stochastic Hamiltonian system consisting of three parts:*
(i) Stochastic Maximum Principle in the form of a variational inequality:

$$H_u(x^*(t), u^*(t), t; p(t), q(t))(u(t) - u^*(t)) \leq 0 \tag{4.16}$$

a.e. $t \in [t_0, t_f]$, $\mathbf{P}-$a.s. $\forall u \in U$
(ii) State dynamics as in (4.5)

$$dx^* = H_p(x^*(t), u^*(t), t; p(t), q(t))dt + H_q(x^*(t), u^*(t), t; p(t), q(t))dB_t \tag{4.17}$$

$$x^*(0) = x_0 \tag{4.18}$$

(iii)The stochastic adjoint dynamics as a terminal value SDE:

$$dp = -H_x(x^*(t), u^*(t), t; p(t), q(t))dt + q(t)dB_t \quad (4.19)$$

$$p(t_f) = -K_x(x^*(t_f)) \quad (4.20)$$

The full proof of this theorem, especially when the control enters the diffusion term $\sigma = \sigma(x, u, t)$ is lengthy and technical. Please see Yong and Zhou (1999) [148] for a full proof. We here review sketched proof of a simplified version of this theorem when: (i) the control does not enter diffusion term, namely $\sigma = \sigma(x, t)$; (ii) all quantities are scalars; and (iii) there are neither state nor control constraints for the problem.

Proof. (sketch proof of stochastic maximum principe) As mentioned at the beginning of this section, the method of spike variation will be employed throughout this proof. Without loss of generality, let u^* be the optimal control and u be some feasible control, we can then construct a perturbed control by letting the :

$$u^\epsilon(t) = \begin{cases} u^*(t) & t \in [t_0, t_f] \backslash T_\epsilon \\ u(t) & t \in T_\epsilon \end{cases}$$

where T_ϵ is a measurable subset of the planning horizon with measure ϵ. Also it is convenient to define the following indicator function

$$\chi_{T_\epsilon}(t) = \begin{cases} 0 & t \in [t_0, t_f] \backslash T_\epsilon \\ 1 & t \in T_\epsilon \end{cases}$$

In this proof we will use the superscript ϵ to represent the quantities corresponding to the perturbed control, especially, we have the perturbed state dynamics with the same initial condition:

$$\begin{aligned} dx^\epsilon &= f(x^\epsilon, u^\epsilon, t)dt + \sigma(x^\epsilon, t)dB_t \\ x^\epsilon(0) &= x_0 \end{aligned}$$

And we use the notation $\delta f(\cdot)$ to represent the difference of the drift term with optimal and feasible control mentioned above, i.e.

$$\delta f(t) = f(x^*, u, t) - f(x^*, u^*, t)$$

and similar definition holds for $\delta\sigma(\cdot)$ and $\delta f_0(\cdot)$. In addition, we introduce the following auxiliary processes:

$$\begin{aligned} dy^\epsilon &= f_x(t)y^\epsilon(t)dt + \sigma_x(t)y^\epsilon(t)dB_t \\ y^\epsilon(0) &= 0 \end{aligned}$$

$$\begin{aligned} dz^\epsilon &= \left\{f_x(t)z^\epsilon(t) + \delta f(t)\chi_{T_\epsilon}(t) + \frac{1}{2}f_{xx}(t)y^\epsilon(t)^2\right\}dt \\ &\quad + \left\{\sigma_x(t)z^\epsilon(t) + \frac{1}{2}\sigma_{xx}(t)y^\epsilon(t)^2\right\}dB_t \\ z^\epsilon(0) &= 0 \end{aligned}$$

The first step is to estimate the deviation of the perturbed state dynamics from the optimal one, and the difference between the perturbed objective functional with the optimal one, to do so, Yong and Zhou (1999) [148] gives the following estimation of the difference between the perturbed and the optimal objective functional as an intermediate theorem:

$$\begin{aligned} J(u^\epsilon) - J(u^*) &= E\left\langle K_x(x^*(t_f)), y^\epsilon(t_f) + z^\epsilon(t_f)\right\rangle + \frac{1}{2}E\left\langle K_{xx}(x^*(t_f))y^\epsilon(t_f), y^\epsilon(t_f)\right\rangle \\ &\quad + E\int_{t_0}^{t_f}\left\{\left\langle f_{0,x}(t), y^\epsilon(t) + z^\epsilon(t)\right\rangle + \frac{1}{2}\left\langle f_{0,xx}(t)y^\epsilon(t), y^\epsilon(t)\right\rangle\right\}dt \\ &\quad + E\int_{t_0}^{t_f}\left\{\delta f_0(t)\chi_{T_\epsilon}(t)\right\}dt + o(\epsilon) \end{aligned} \tag{4.21}$$

The detailed proof is very lengthy, and is carried out mainly by proving intermediate moment estimates for the expected value of the absolute deviation of perturbed and optimal trajectories. Notice that this means a necessary condition for optimality is to have the RHS of the above equation greater or equal to zero for any feasible $u(t)$ and positive measure $\epsilon > 0$.

The next step is to introduce the stochastic adjoint variables $p(t)$ and $q(t)$, note that by definition they solve the backward stochastic adjoint dynamics as in part

(iii) of our Theorem 4.3. The goal of this step is to cancel out the first order terms of auxiliary processes $y^\epsilon(t)$ and $z^\epsilon(t)$ with representation theorems. Firstly, with the application of Ito's lemma (Yong and Zhou (1999) [148], Chapter 1 Collorrary 5.6) and Taylor expansion ([148], Chapter 3, Lemma 4.3) we have:

$$\begin{aligned} & -E\left\langle K_x(x^*(t_f)), y^\epsilon(t_f)+z^\epsilon(t_f)\right\rangle \\ = \; & E\int_{t_0}^{t_f}\left\{\langle f_x(t), y^\epsilon(t)+z^\epsilon(t)\rangle + \frac{1}{2}\left\langle p(t), f_{xx}y^\epsilon(t)^2\right\rangle + \frac{1}{2}\left\langle q(t), \sigma_{xx}y^\epsilon(t)^2\right\rangle\right\} dt \\ & +E\int_{t_0}^{t_f}\left\{p(t)\delta f(t)+q(t)\delta\sigma(t)\right\}\chi_{T_\epsilon}(t)dt + o(\epsilon) \end{aligned}$$

This term is plugged back into (4.21) and we have a revised term of our target difference:

$$\begin{aligned} 0 \;\geq\; & J(u^*) - J(u^\epsilon) \\ = \; & \frac{1}{2}E[P(t_f)y^\epsilon(t_f)^2 + \int_{t_0}^{t_f} H_{xx}(t)y^\epsilon(t)^2 dt] \\ & +E[\int_{t_0}^{t_f}\delta H(t)\chi_{T_\epsilon}(t)dt] + o(\epsilon) \end{aligned}$$

We see that the first oder term of $y^\epsilon(t)$ and $z^\epsilon(t)$ have disappeared here. There are still terms with respect to $y^\epsilon(t)^2$, in Chapter 3, Lemma 4.6 of [148]. These terms are taken care of with the introduction of second order adjoint variables such as $P(\cdot)$. However, for our special case when control does not enter diffusion term, these second order terms vanish automatically. We can then take the limit with $\epsilon \to 0$ and pass it with the operator $\delta(\cdot)$ and obtain the maximum principle. Specifically, for the case with no control or state constraints, we have:

$$\left.\frac{\partial H}{\partial u}\right|_{u=u^*} = 0, \quad a.e., P - a.s.$$

□

Remark 4.2. *The construction of perturbation with T_ϵ as a general measurable set is due to the fact that U is just a metric space which does not hold a linear structure in general (Yong and Zhou (1999) [148]). At the same time, for the prototype problem considered and casted into applications in this dissertation, taking T_ϵ as a small time interval is sufficient, in fact the proof by Bensoussan (1983) [13] uses* *such construction of perturbation.*

The proof above, even re-casted in sketch with a reduced version of the theorem, is still too technical, especially, it lacks an explanation towards the intuitive/economical meaning of the dual variables. In the following example, we re-cast a summary of Bismut's approach by Malliaris and Brock (1982) [96], in which the primal-dual relationship of different quantities is specifically listed with emphasize of their economic meanings. Especially, the two concepts they try to emphasize here are "risk taking" and "information processing". This approach is originally reported by Bismut (1973) [24] and (1975) [25] and is inspired by Rockafellar (1970) [121]. With the verification theorem established, Yong and Zhou (1999) [148] also has a similar interpretation, we will review such interpretations later.

Example 4.2. *(Bismut's Approach) The problem to consider is:*

$$\max E\left[\int_{t_0}^{t_f} \pi(x,u,t)dt\right] \text{ or equivalently, } \min E\left[-\int_{t_0}^{t_f} \pi(x,u,t)dt\right]$$

subject to:

$$\begin{aligned} dx &= f(x,u,t)dt + \sigma(x,u,t)dB_t \\ x(t_0) &= x_0 \end{aligned}$$

we can interpret π as the instantaneous revenue and x as the capital stock and u is the investment decision. Bisumut (1973) [24] defined the first adjoint variable similar to that of the deterministic dual variable: let $p(t)$ denote the marginal value of capital at time t: (assuming u here is already the optimal policy)

$$p(t) = \frac{\partial}{\partial x} E\left[-\int_{t}^{t_f} \pi(x,u,t)dt\right]$$

The key assumption (see e.g. [24], pp387) is that $p(t)$ may be written as:

$$p(t) = p(t_0) + \int_{t_0}^{t} \dot{p}(t)dt + \int_{t_0}^{t} q(t)dB_t$$

here $\dot{p}(t)$ is the infinitesimal expected rate of growth of p, while q is the infinitesimal conditional covariance of p with B_t. This means that the marginal value of capital *at time t consists of three terms: the first one is the sum of the marginal value of capital at time 0; the second term is the expected value of cumulated infinitesimal increment $p(t)$; and the third term integrates uncertainties in the accumulation process.*

If we follow the same definition of Hamiltonian as in (4.15), the maximum principal that Bisumt gives is then identical to that derived from a direct application of Theorem 4.3 to this problem, we have:

$$\begin{aligned} \frac{\partial H}{\partial u} &= 0 \\ -dp &= \frac{\partial H}{\partial x} dt - q dB_t \\ p(t_f) &= 0 \end{aligned}$$

which means the following primal-dual correspondence:

$$\begin{aligned} f(x, u, t) &\rightarrow p(t) \\ \sigma(x, u, t) &\rightarrow q(t) \end{aligned}$$

Now let us discuss the interpretation of such Maximum principle. We will start with Hamiltonian H, which is the sum of the following three terms: (i) instantaneous profit: $\pi(t)$; (ii) the expected infinitesimal increment of capital valued at its marginal expected value: $p(t) \cdot f(t)$; and (iii) the risk associated with a given investment policy valued at its cost: $q(t) \cdot \sigma(t)$. Here we gain better insights into the economic meanings of the second adjoint variable $q(t)$, which measures the instantaneous attitude towards risk, and is perhaps unfamiliar to readers that are used to deterministic optimal control. $q(t)$ positive means the decision maker is risk-taking, and negative if he is risk-averting.

With the above discussion let us look at the intuitions behind the adjoint dynamics: on the left hand side, $-dp(t)$ is the conditional expected rate of depreciation in the marginal value of capital; and the right hand side consists of two terms:

1. The first term of RHS is the partial derivative of Hamiltonian w.r.t. capital, and according to our definition of Hamiltonian, it consists of three terms: (i) the capital's marginal contribution to profits; (ii) capital's contribution to enhancing the expected value of the increment of the capital stock; and (iii)

capital's contribution to increasing the conditional standard deviation of the increment of the capital stock valued at the cost of risk.

2. Consider the second term, starting with the following heuristic resemblance of the (primal) state dynamics:

$$dB_t = \frac{1}{\sigma} \cdot (dx - f dt)$$

multiply both sides with $q(t)$

$$q(t) dB_t = \frac{q(t)}{\sigma} \cdot (dx - f dt)$$

Clearly, this means that $q(t)dB_t$ measures the correlation between the following: (i) the evolution of the marginal value of capital (evaluated in terms of p) and (ii) the difference between the infinitesimal movement of state variable dx and its expected value $E(dx) = f dt$.

Remark 4.3. *Still, Theorem 4.3 is not the most general result of this type. In the case when the convexity assumption in Definition 4.3 is relaxed, the stochastic maximum principle should be substituted by the corresponding Hamiltonian maximization problem. More generally, if the control could enter the volatility term, which means* $\sigma = \sigma(x, u, t)$*, then a second order adjoint dynamics would be absolutely necessary, which makes the solution of the stochastic Hamiltonian system a 6-tuple instead of a 4-tuple. See for example, Peng (1990) [112], Yong and Zhou (1999) [148] for more results. A short summary of this type of results and recent development on this topic is Hu (2014) [71].*

Finally, we list the following existence result:

Theorem 4.4. *(S-OCP, existence result) We assume the regularities listed as in Definition 4.3, Problem 4.1 admits an optimal control.*

Proof. See Yong and Zhou (1999) [148]. □

4.1.3 Stochastic Verification Theorem

For the rest of this section, we study the relationship between the two version of stochastic maximum principle that we have reviewed. Intuitively, these two theorems should be equivalent given appropriate, and possibly stronger, regularities. Towards this end, Yong and Zhou (1999) [148] gives a presentation and discussion of so-called verification theorems. The main idea is: the stochastic maximum principle derived according to the two techniques reviewed thus far should be equivalent given the value function and its appropriate partials exist. We start from the regularity conditions:

Definition 4.4. *(Stochastic OCP Regularity Revisited) Problem 4.1 is regular if the following conditions hold:*
(i) define the value function of Problem 4.1 for all $(x,t) \in \mathbb{R}^n \times [t_0, t_f]$ *as follows:*

$$\begin{cases} V(x,t) = \inf_{u \in U} J(x,t;u(\cdot)) \\ V(x(t_f),t_f) = K(x(t_f)) \end{cases} \tag{4.22}$$

and the value function is three times continuously differentiable in state and continuously differentiable in time, V_{xt} *is also continuous;*
(ii) $f(x,u,t) : \mathbb{R}^n \times U \times [t_0,t_f] \to \mathbb{R}^n, f_0(x,u,t) : \mathbb{R}^n \times U \times [t_0,t_f] \to \mathbb{R}, \sigma(x,u,t) : \mathbb{R}^n \times U \times [t_0,t_f] \to \mathbb{R}^n \times \mathbb{R}^n$ *are all uniformly continuous differentiable w.r.t. x and u., and twice continuously differentiable w.r.t x. Let* φ *stand for any one of these mappings, then:*

$$\begin{aligned} |\varphi(x,u,t) - \varphi(y,u,t)| &\leq L_1(|x-y|) \\ |\varphi_x(x,u,t) - \varphi_x(y,v,t)| &\leq L_2(|x-y|) + L_3(|u-v|) \end{aligned}$$

(iii) the set of feasible control $u \in U \subseteq \mathbb{R}^m$ *is closed and convex;*
(iv) $K(x)$ *is continuously differentiable, moreover*

$$K_x(x) \leq C_3(1+|x|);$$

With such regularities we provide the following revision of Theorem 4.3 involving the value function. Notice that this is in fact a generalization of the result from Malliaris and Brock.

Theorem 4.5. *(Stochastic maximum principle with value functions) We assume the regularities listed as in Definition 4.4. Let* $u^*(t)$ *be an optimal control and* $x^*(t)$ *the corresponding trajectory which solves Problem 4.1, also let* $(p(t), q(t))$ *be the* *adjoint variables corresponding to the optimal state-control pair. Then* $p(t)$ *and* $q(t)$ *could be expressed in terms of the value function as follows:*

$$\begin{cases} p(t) = -V_x(x^*(t), t) \\ q(t) = -V_{xx}(x^*(t), t)\sigma(x^*(t), u^*(t), t) \end{cases} \tag{4.23}$$

a.e. $t \in [t, t_f]$, $\mathbf{P}-a.s.$.

Proof. See Yong and Zhou (1999) [148]. □

With the introduction of a well-defined value function, it is possible to discuss the economic interpretation of both adjoint variables, see the discussions regarding the notion of a dynamic stochastic "shadow price" in Yong and Zhou (1999) [148]. Later in Chapter 7, we will present such an aspect under a specific problem setting called dynamic stochastic monopsony. Moreover, following the virtue of convexity/concavity, the same authors have the following sufficiency result, which is important later this chapter in the analysis of stochastic Nash games:

Theorem 4.6. *(Sufficiency of stochastic maximum principle) Assume that, in addition to Definition 4.10,* $K(x)$ *is convex,* $H(x, u, t; p, q)$ *is concave in* (x, u) *for all* $t \in [t_0, t_f]$, $\mathbf{P}-a.s.$. *Let* $(x^*(\cdot), u^*(\cdot), p(\cdot), q(\cdot))$ *be a 4-tuple that satisfies the maximum principle as well as solves the system of equation presented in Theorem 4.3, then* $(x^*(\cdot), u^*(\cdot))$ *is an optimal state-control pair for Problem 4.1.*

Proof. See Yong and Zhou (1999) [148]. □

4.2 Stochastic DVI and Analysis

In this section we give list the definition of stochastic DVI and based on the results from last section, we propose its necessary conditions as well as the regularity conditions required. A S-DVI theory based on a sample-path treatment is reported by Mookherjee (2006) [102], then further developed by Meimand (2013) [100]. Later in Section 4.3 we will see how the S-DVIs could be used as an characterization of the stochastic Nash games.

Problem 4.2. *(Stochastic DVI) We refer to the following problem as a stochastic differential variational inequality (SDVI) and denote it by* $SDVI(F, f, \sigma, U, x^0)$:

find $u^* \in U$*, such that*

$$E\left\{\int_{t_0}^{t_f} F[x(u^*,t),u^*,t]^T(u-u^*)dt\right\} \geq 0, \forall u \in U, \quad \mathbf{P}-a.s. \tag{4.24}$$

here $u \in U \subseteq (L^2[t_0,t_f])^m$,$x(u,t) = \arg\{dx = f(x,u,t)dt + \sigma(x,u,t)dB_t, x(t_0) = x^0\}$.

This definition is the generalization of the deterministic DVI recorded in Chapter 2, which is originally from Friesz (2010) [54]. We assume the following regularity conditions for the analysis of the problem:

Definition 4.5. *(Stochastic DVI regularity conditions) We call* $SDVI(F,f,\sigma,U,x^0)$ *regular if:*
(i) $f(x,u,t) : \mathbb{R}^n \times U \times [t_0,t_f] \to \mathbb{R}^n$ *is continuously differentiable w.r.t.* x *and* u*. Also the first order partials* f_x, f_v *are bounded. This means:*

$$|f(x,u,t)| \leq M_1(1+|x|+|u|);$$

(ii) $\sigma(x,u,t) : \mathbb{R}^n \times U \times [t_0,t_f] \to \mathbb{R}^n \times \mathbb{R}^n$ *is continuously differentiable w.r.t.* x *and* u*. Also the first order partials* σ_x, σ_v *are bounded. This means:*

$$|\sigma(x,u,t)| \leq M_2(1+|x|+|u|);$$

(iii) the set of feasible control $U \subseteq \mathbb{R}^k$ *is convex; also the set* $\{f_i, (\sigma\sigma^T)^{ij}, F|u \in U, i = 1,\cdots,n, j = 1,\cdots,n\}$ *is a convex set in* $\mathbb{R}^{n+n^2+1}$*;*
(iv) $F(x,u,t) : \mathbb{R}^n \times U \times [t_0,t_f] \to \mathbb{R}$ *is Borel, continuously differentiable w.r.t.* x *and* u*:*

$$\begin{aligned} |F_x(x,u,t)| &\leq M_3(1+|x|+|u|) \\ |F_u(x,u,t)| &\leq M_4(1+|x|+|u|) \end{aligned}$$

Theorem 4.7. *(Necessary Conditions) When regularity conditions as in Definition 4.5 hold, a solution* (u^*,x^*) *of SDVI Problem 4.2 must obey the following stochastic Hamiltonian system along with adjoint variables* (p,q)*:*
(i) the variational inequality, for all $\forall u \in U, t \in [t_0,t_f]$ *a.e.,* $P-a.s.$*:*

$$\left\langle\left\{p^T f_u(x^*,u^*,t) + \nabla_u[trq^T\sigma(x^*,u^*,t)] - F(x^*,u^*,t)\right\}, u(t)-u^*(t)\right\rangle \leq 0 \tag{4.25}$$

(ii) the state dynamics:

$$dx^* = f(x^*, u^*, t)dt + \sigma(x^*, u^*, t)dB_t \tag{4.26}$$

$$x^*(t_0) = x^0 \tag{4.27}$$

(iii) the adjoint dynamics:

$$dp = -\left\{\nabla_x[p^T f(x^*, u^*, t)] + \nabla_x[trq^T \sigma(x^*, u^*, t)]\right\} dt + q(t)dB_t \tag{4.28}$$

$$p(t_f) = 0 \tag{4.29}$$

Proof. Inspired by Friesz (2010) [54], this proof is done by rewriting the SDVI into a stochastic optimal control problem and apply the corresponding necessary conditions from last section. More specifically, consider the following optimal control problem

$$\begin{aligned} &\min_u E_{u(\cdot)} \left\{ \int_{t_0}^{t_f} [F(x(u^*, t), u^*, t)]^T u dt \right\} \\ \text{such that } dx &= f(x(t), u(t), t)dt + \sigma(x(t), u(t), t)dB_t, \quad x(t_0) = x_0 \\ u(t) &\in U(x, t) \end{aligned}$$

it is not hard to list its necessary condition following Theorem 4.3. First write down its Hamiltonian:

$$H = p^T f(x, u, t) + \text{tr}q^T \sigma(x, u, t) - [F(x^*, u^*, t)]^T u$$

Then the stochastic maximum principle for this problem is:

$$H_u(x^*(t), u^*(t), t; p(t), q(t))(u(t) - u^*(t)) \leq 0$$

which is

$$\left\{p^T f_u(x^*, u^*, t) + \nabla_u[\text{tr}q^T \sigma(x^*, u^*, t)] - F(x^*, u^*, t)\right\} (u(t) - u^*(t)) \leq 0$$

the state dynamics still hold:

$$dx^* = f(x^*, u^*, t)dt + \sigma(x^*, u^*, t)dB_t$$

$$x^*(t_0) \quad = \quad x^0$$

the adjoint equation is then

$$dp = -H_x(x^*(t), u^*(t), t; p(t), q(t))dt + q(t)dB_t$$

which is

$$dp = -\left\{\nabla_x[p^T f(x^*, u^*, t)] + \nabla_x[\text{tr} q^T \sigma(x^*, u^*, t)]\right\} dt + q(t)dB_t$$

with terminal condition

$$p(t_f) = 0$$

□

Theorem 4.8. *(S-DVI, existence result) We assume the regularities listed as in Definition 4.5, Problem 4.2 admits a solution.*

4.3 Stochastic Differential Nash Games

In this section, we discuss the stochastic differential Nash games in which the best response problem of each player takes the form of stochastic optimal control problem as defined by Problem 4.1. The stochastic differential Nash game we list here is a natural extension of the deterministic differential games reviewed in Definition 2.11. In this section, we will formulate stochastic differential Nash games and study the necessary conditions of its equilibrium. Especially, we will see how stochastic DVI problem from Section 4.2 could be applied as one way to characterize the necessary conditions of stochastic differential Nash equilibrium (S-DNE).

4.3.1 Discussions on Information Structures

Before we proceed to the mathematical formulations, it is helpful to discuss the role information plays under the context of stochastic competition in continuous time. In this part, we will compare different assumptions existing in (stochastic) dynamic game theory in terms of what information is available on current situation or of the past, to a player himself or to his rivals. For detailed discussion on these topics, see Docker et al. (2000) [41], Fudenberg and Tirole [60] (1991).

The first pair of related assumption to compare is the classification of games into complete and incomplete information games. As introduced by von Neumann and Morgenstern (1944) [143], in a game of complete information, all players know all relevant information expressed in the rules of the game. We say Player i has private information if his knowledge includes something that other players do not know, yet other players are aware of the fact that Player i has such knowledge they do not know, so on so forth. Hence we list the following definition of a game of complete information from Fuduenberg and Tirole (1999) [60].

Definition 4.6. *(Complete Information) A game of complete information is such that no players have private information.*

And we list the following example of incomplete information game from Dockner et al. (2000) [41]:

Example 4.3. *(Incomplete Information Game) Consider a duopoly and assume that one firm is uncertain about the cost structure of the other firm and vice versa. What one firm knows is that each player must belong to one within a set of possible types, and the knowledge of specific types is not sure, or subject to an a priori distribution. If the game goes on repeatedly, each firm could use observations of their rivals to make inferences about the other firm's "type" that they initially were not clear about.*

The second topic of interest is the classification of games into ones with perfect and imperfect information. (Fuduenberg and Tirole (1999) [60])

Definition 4.7. *(Perfect Information) A game of perfect information is one with which all players' actions are observable by all players.*

In the context of differential games, this means that a player, when acting at time t, has perfect knowledge of all previous actions. And in the setting of stochastic games in this chapter, it also means that any exogenous uncertainties, or "acts of nature", are also included in the player's knowledge. In the last topic on information structure of this chapter, our discussion is about the extent to which the state variable will be used by all players in the competition. From Section 4.1 we know that when the state equation is driven by a Brownian motion as (4.5), it is natural to require that the players' decisions are adapted to the filtration $\{\mathcal{F}_t\}_{t_0 \le t \le t_f}$. Here, similar to Section 3.1 we list a series of information structures that are frequently employed towards the studies stochastic differential games following Bensoussan et al. (2015) [14]:

Definition 4.8. *(Information Structures) For finite horizon, continuous time stochastic optimal control problems and differential games, denote time as $t \in [t_0, t_f]$. We use the notation $\eta(t)$ to define information structures:*
(i) adapted open-loop (AOL): $\eta(t) = \{x_0, \mathcal{F}_t\}$, $t \in [t_0, t_f]$;
(ii) adapted feedback (AF): $\eta(t) = \{x(t), \mathcal{F}_t\}$, $t \in [t_0, t_f]$;
(iii) adapted closed-loop memoryless (ACLM): $\eta(t) = \{x_0, x(t), \mathcal{F}_t\}$, $t \in [t_0, t_f]$;
(iv) adapted closed-loop (ACL): $\eta(t) = \{x(s),\ 0 \le s \le t, \mathcal{F}_t\}$, $t \in [t_0, t_f]$.

In this chapter, however, we emphasize on the study of stochastic differential Nash games with the following information structure:

- perfect information: which leads to perfect observable control history for all players;
- complete information: this means no players could take advantage of any information that other players do not know;
- adapted open-loop (AOL): which means players use only the initial state towards their decisions, but not the (realtime) state information.

As a matter of fact, there has been a large literature on the study of games with incomplete and/or imperfect information. Under the settings of incomplete information, Chatterjee and Samuelson (1983) [37] studies a two-sided bargain with each player knows his own reservation price, but is uncertain about, hence relying on a subjective distribution of the other player's reservation price. On the study of imperfect information games, for example, Kaitala (1993) [78] considers stochastic resource management game where the choice of cooperation and competition among players are based on imperfect estimates of other players' efforts; Querou (2010) [117] analyzes a resource extraction game in which each player has no information on other players' payoff functions. Note that case (iii) in the above definition is also often referred to as Markovian information structure. A collection of results on the equilibrium of this information structure are summarized by Barsar and Olsder (1999) [10] and Dockner et al. (2000) [41], which relies on the solution of stochastic version of Hamilton-Jacobi Bellman equations similar to Theorem 4.1.

4.3.2 Stochastic Differential Nash Games and Equivalent S-DVIs

Firstly, following the information structure we have just specified, we define the following Nash equilibrium of interest:

Definition 4.9. *(Stochastic Differential Nash equilibrium, Adapted Open-Loop) Assume that there are N players, and use $u^i(\cdot)$ to denote the control trajectory of of the i-th player, in response of the following stochastic optimal control problem*

$$\begin{aligned} \min_{u^i} J_i &= E\left[\int_{t_0}^{t_f} f_0(x, u^i, u^{-i}, t)dt + K^i(x(t_f))\right] \\ s.t. dx &= f(x, u^1, u^2, \cdots, u^N, t)dt + \sigma(x,t)dB_t \\ x(t_0) &= x_0 \\ u^i(t) &\in U^i(u^{-i}, t) \end{aligned}$$

where every agent i has a cost/disutility functional $J_i(u) : U \to \mathbb{R}^1$ that depends on all agents' strategies. Furthermore, apply the following notations where

$$\begin{aligned} U &= \prod_{i=1}^{N} U_i \\ u(\cdot) &= (u^i(\cdot) : i = 1, \cdots, N) \end{aligned}$$

Then $u^*(\cdot)$ is defined as the stochastic differential Nash equilibrium if each of the inequalities

$$J_i(u^*(\cdot)) \leq J_i(u^{*,1}(\cdot), \cdots, u^{*,i-1}(\cdot), u^i(\cdot), u^{*,i+1}(\cdot), \cdots, u^{*,N}(\cdot))$$

hold for all other $u(\cdot) \in U$.

The formulation in this section consists of the stochastic dynamics where the control $u(\cdot)$ does not enter the diffusion coefficient. For the discussion on a more general case, please refer to Tun et al. (2008) [140]. Follow the spirit of the definition above, at the equilibrium point $u^*(\cdot)$, the stochastic optimal control problem of

Player i, which is parameterized by other players' control strategies, must reach its minimum. We can therefore use Theorem 4.3 as the starting point of our analysis. First we propose a series of regularity conditions to facilitate the analysis. Consider the Hamiltonian of Player i's problem:

$$H^i = -f_0^i(x,u,t) + (p^i)^T f(x,u,t) + \mathrm{tr}\left[(q^i)^T \sigma(x,t)\right], \tag{4.30}$$

with $p^i(\cdot)$ and $q^i(\cdot)$ being the adjoint variables, we have the following:

Definition 4.10. *(Regularity Conditions for Stochastic Differential Nash Game) The stochastic differential Nash game problem as in Definition 4.9 is called regular if the following hold:*
(i) $f(\cdot), \sigma(\cdot)$ are measurable in all their arguments and twice continuously differentiable in x. Both the functions and their derivatives $f_x, f_{xx}, \sigma_x, \sigma_{xx}$ are continuous in (x,u) and bounded.
(ii) in objective functionals, mapping $f_0^i(\cdot)$ and $K^i(\cdot)$ are also twice continuously differentiable in x, with continuous derivatives; moreover their first order derivatives are bounded by the following,

$$|f_{0,x}^i(x,u,t)| + |K_x^i(x)| \leq C(1+|x|+|u|).$$

(iii) each player's problem have a concave Hamiltonian $H^i(x,u,t;p,q)$ in (x,u^i); also $K^i(x)$ is convex, and the set of feasible control U^i is convex.

Now we present the main theorem of this chapter, in which the stochastic dynamic Nash equilibrium admits a characterization in the form of S-DVI problems similar to Problem 4.2.

Theorem 4.9. *(Stochastic DVI representation) When the stochastic DNE is regular in the sense of Definition 4.10, it admits a stochastic differential variational inequality.*

Proof. We shall apply Theorem 4.3 and obtain the optimality conditions for Player i's problem, starting with variational inequality:

$$\nabla_{u^i} H^i(x^*(t), u^{*,i}(t), t; p^i(t), q^i(t))(u^i(t) - u^{*,i}(t)) \leq 0 \tag{4.31}$$

a.e. $t \in [t_0, t_f]$, $\mathbf{P}$−a.s. , $\forall u^i \in U^i$ and the corresponding adjoint dynamics:

$$dp^i = -\nabla_x H^i(x^*(t), u^{*,i}(t), t; p^i(t), q^i(t))dt + q^i(t)dB_t \tag{4.32}$$

$$p^i(t_f) = -\nabla_x K^i(x^*(t_f)) \tag{4.33}$$

while the state dynamics still hold:

$$dx^* = f(x^*(t), u^*(t), t)dt + \sigma(x^*, t)dB_t \tag{4.34}$$

$$x^*(0) = x_0. \tag{4.35}$$

Notice that by Definition 4.10, the DNE is regular such that each of Player i's stochastic optimal control problem meet the regularities both in Definition 4.3 and in Theorem 4.6, we know that these conditions are also sufficient for all Player $i \in \{1, 2, \cdots, N\}$. In addition, we know that in order for the Nash equilibrium condition to hold, all these conditions must hold simultaneously. At the same time, let's introduce the following tuples:

$$F^i(y, u, t) = \nabla_{u^i} H^i(y^*(t), u^{*,i}(t), t)$$

where y is defined as:

$$y^i(u, t) = \begin{pmatrix} x(u, t) \\ p^i(u, t) \\ q^i(u, t) \end{pmatrix};$$

with augmented state-adjoint dynamics:

$$g^i = \begin{pmatrix} f(x, u, t) \\ -\nabla_x H^i(x^*(t), u^{*,i}(t), t; p^i(t), q^i(t)) \\ 0 \end{pmatrix}, \quad \pi^i = \begin{pmatrix} \sigma(x, t) \\ 0 \\ q^i(t) \end{pmatrix};$$

such that the following dynamics holds:

$$dy^i = g^i(y^i, u, t)dt + \pi^i(y^i, u, t)dB_t$$

and initial-terminal conditions

$$
\begin{aligned}
y^i(t_0) &= y^{i,0} = \begin{pmatrix} x^0 \\ p^i(0) \text{ free} \\ q^i(0) \text{ free} \end{pmatrix}; \\
y^i(t_f) &= y^{i,f} = \begin{pmatrix} x(t_f) \text{ free} \\ -\nabla_x K^i(x^*(t_f)) \\ q^i(t_f) \end{pmatrix}.
\end{aligned}
$$

Finally, let $y = (y^i : i = 1, \cdots, N), g = (g^i : i = 1, \cdots, N), \pi = (\pi^i : i = 1, \cdots, N)$, and $G = (G^i : i = 1, \cdots, N)$, we shall see that the following S-DVI is equivalent to (4.31) throughout (4.35) holds for all the players: find $u^* \in U$, and the corresponding $y^* = y(u^*, t)$ such that

$$
E\left\{\int_{t_0}^{t_f} F[y^*, u^*, t]^T (u - u^*) dt \leq 0\right\}, \forall u \in U, \quad \mathbf{P} - a.s. \tag{4.36}
$$

here

$$
\begin{aligned}
u &\in U \subseteq (L^2[t_0, t_f])^m, \\
y(u,t) &= \arg\left\{dy = g(y,u,t)dt + \pi(y,u,t)dB_t, y(t_0) = y^0, y(t_f) = y^f\right\}.
\end{aligned}
$$

□

Remark 4.4. *Notice that instead of relying on the solution of HJB-PDEs, analysis and solution of dynamic stochastic Nash equilibrium via S-DVIs will rely on the solution of the maximum principle and the coupled stochastic Hamiltonian system. We will later review the solution of such systems of FBSDEs. In Chapter 7, we will illustrate this type of necessary conditions by applying it towards a problem under the setting of dynamic stochastic oligopsony.*

4.4 Solution of FBSDE

The last issue remaining in this chapter is to solve the system of forward-backward stochastic differential equations (FBSDEs) induced by the stochastic maximum principle with the state and adjoint equations. The solution of such a system is crucial to the computation of stochastic optimal controls and Nash equilibriums. In this section, we review the 4-step treatment by Ma et al. (1994) [94] (later reviewed by Yong and Zhou (1999) [148]). We start by taking the following abstract formulation of the FBSDE system concerning a triple of processes (X, Y, Z) with $X \in L^2_{\mathcal{F}}(\Omega; C([t_0, t_f]; \mathbb{R}^n))$, $Y \in L^2_{\mathcal{F}}(\Omega; C([t_0, t_f]; \mathbb{R}^k))$, $Z \in L^2_{\mathcal{F}}([t_0, t_f]; \mathbb{R}^{k\times m}))$.

Problem 4.3. *(FBSDE) Given a filtered probability space* $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq t_0}, \mathbf{P})$ *with* $m-$*dimensional standard Brownian motion* B_t *defined, consider the following system of forward-backward stochastic differential equations (FBSDEs):*

$$\begin{cases} dX(t) = f(t, X(t), Y(t), Z(t))dt + \sigma(t, X(t), Y(t), Z(t))dB_t \\ dY(t) = h(t, X(t), Y(t), Z(t))dt + Z(t)dB_t \\ X(t_0) = x_0, Y(t_f) = g(X(t_f)) \end{cases}$$

with functions f, σ, h, g *given as deterministic. Then a triple of processes* (X, Y, Z) *is called a solution of the system above if the following holds:*

$$\begin{aligned} X(t) &= x_0 + \int_{t_0}^{t} f(s, X(s), Y(s), Z(s))ds + \int_{t_0}^{t} \sigma(s, X(s), Y(s), Z(s))dB_s \\ Y(t) &= g(X(t_f)) - \int_{t}^{t_f} h(s, X(s), Y(s), Z(s))ds - \int_{t}^{t_f} Z(s)ds \end{aligned}$$

Let us make the following observations:

1. the FBSDE system in Problem 4.3 is a stochastic generalization of the two-point boundary value problem;

2. the stochastic Hamiltonian system in Theorem 4.3 is a special case of Problem 4.3: (i) the state dynamics is the forward component $X(\cdot)$, (ii) the two adjoint variables $(p(\cdot), q(\cdot))$ correspond to the backward components $(X(\cdot), Y(\cdot))$, (iii)the control variable could be coupled through the stochastic maximum principle.

Now we suppose that (X, Y, Z) is a set of adapted solution of Problem 4.3, in addition, assume that $Y(\cdot)$ and $X(\cdot)$ follows the following implicit relationship $\theta(t, x)$ which is continuously differentiable in t and twice continuously differentiable in x:

$$Y(t) \triangleq \theta(t, X(t)), \quad \forall t \in [t_0, t_f], \mathbf{P} - a.s.$$

Then per Itô's Lemma, consider the $l-$th dimension of backward dynamics we have:

$$\begin{aligned} dY^l(t) &= d\theta^l(t, X(t)) \\ &= \Big\{\theta^l_t(t, X(t)) + \Big\langle \theta^l_x(t, X(t)), f(t, X(t), \theta(t, X(t)), Z(t)) \Big\rangle \\ &\quad + \frac{1}{2}\mathrm{tr}\Big[\theta^l_{xx}(t, X(t))(\sigma\sigma^T)\sigma(t, X(t), \theta(t, X(t)), Z(t))\Big]\Big\} dt \\ &\quad + \Big\langle \theta^l_x(t, X(t)), \sigma(t, X(t), \theta(t, X(t)), Z(t))dB_t \Big\rangle . \end{aligned} \tag{4.37}$$

Hence by comparing the above with the original system listed in Problem 4.3 we see that in order to obtain the right θ it suffices to have for $l = 1, \cdots, k$:

$$\begin{aligned} h^l(t, X(t), \theta(t, X(t)), Z(t)) &= \theta^l_t(t, X(t)) + \Big\langle \theta^l_x(t, X(t)), f(t, X(t), \theta(t, X(t)), Z(t)) \Big\rangle \\ &\quad + \frac{1}{2}\mathrm{tr}\Big[\theta^l_{xx}(t, X(t))(\sigma\sigma^T)(t, X(t), \theta(t, X(t)), Z(t))\Big] \\ Z(t) &= \theta_x(t, X(t))\sigma(t, X(t)), Z(t)) \\ \theta(t_f, X(t_f)) &= g(X(t_f)) \end{aligned}$$

The heuristic argument given above follows from Yong and Zhou (1999) [148], and still, there are many issues concerning the mathematical rigorousness. For instance, the solvability of Problem 4.3 might not hold in general. Here we argue that such proofs are beyond the scope of this dissertation and please refer to Ma et al. (1994) [94], Yong and Zhou (1999) [148] for richer mathematical details. Nevertheless, we following the heuristic argument above and summarize the procedure:

Algorithm 4.1. *(4-Step Algorithm for FBSDE) Consider the FBSDE system in Problem 4.3 and a finite time horizon $t \in [t_0, t_f]$,*

Step 1. $\boxed{\textit{Find } z(t, x, y, p)}$ *Find $z(t, x, y, p)$ such that $\forall t \in [t_0, t_f]$, $(x, y, p) \in \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^n$:*

$$z(t, x, y, p) = p\sigma(t, x, y, z(t, x, y, p)) \tag{4.38}$$

Step 2. *Solve PDE* *Solve the following system of PDEs:*

$$\begin{cases} \theta_t^l + \frac{1}{2} tr\left[\theta_{xx}^l(\sigma\sigma^T)t, x, \theta, z(t, x, \theta, \theta_x)\right] + \left\langle f(t, x, \theta, z(t, x, \theta, \theta_x)), \theta_x^l \right\rangle \\ -h^l(t, x, \theta, z(t, x, \theta, \theta_x)) = 0, \quad (t, x) \in (t_0, t_f) \times \mathbb{R}^n, \quad 1 \le l \le k \\ \theta(t_f, x) = g(x), x \in \mathbb{R}^n \end{cases} \tag{4.39}$$

Step 3. *Recover $X(t)$* *Use the θ and z obtained to define*

$$\begin{aligned} \widetilde{f}(t, x) &= f(t, x, \theta(t, x), z(t, x, \theta(t, x), \theta_x(t, x)) \\ \widetilde{\sigma}(t, x) &= \sigma(t, x, \theta(t, x), z(t, x, \theta(t, x), \theta_x(t, x)) \end{aligned}$$

and solve the forward SDE:

$$X(t) = x_0 + \int_{t_0}^{t} \widetilde{f}(t, X(s))ds + \int_{t_0}^{t} \widetilde{\sigma}(t, X(s))dB_s$$

Step 4. *Recover $Y(t)$ and $Z(t)$* *Finally, use the relationship defined previously and set*

$$\begin{cases} Y(t) = \theta(t, X(t)) \\ Z(t) = z(t, X(t), \theta(t, X(t)), \theta_x(t, X(t))) \end{cases} \tag{4.40}$$

Later in Chapter 7, we will apply Algorithm 4.1 towards the solution of an embodiment of Problem 4.3 under the context of S-DVI applied to stochastic dynamic monopsony and oligopsony. For other worked-out example of Algorithm 4.1 with application in option pricing, see Yong and Zhou (1999) [148].

4.5 Summary

In this chapter we start from the concept of a Brownian motion and reviewed the techniques necessary for solving a problem of stochastic optimal control. We then propose stochastic DVI as a characterization of a stochastic differential Nash equilibrium. Later in Chapter 7 we will apply such techniques towards the analysis of stochastic dynamic monopsony and oligopsony.

Chapter 5
DVIs in Revenue Management: the Fixed Inventory Model

In this chapter we explore the dynamic pricing of an abstract commodity whose inventory is specified at an initial time but never subsequently replenished. The discrete time version of this model is proposed by Perakis and Sood (2006) [114] and is as *dynamic pricing with fixed inventory* (DPFI) and studied by using a discrete time, finite dimensional quasi-variational inequality perspective. The context of this chapter is close to Perakis and Sood (2006), but we study the DPFI problem using a continuous time perspective. Friesz (2010) [54] provides the analysis that re-casts the problem as a infinite dimensional quasi-variational inequality. In this chapter, we build upon such analysis and prove that the DPFI model admits an equivalent variational inequality, which allows us to effectively compute the DPFI problem.

Furthermore, using the DPFI model as a basic model for the sellers' short term pricing strategies, we continue in this chapter to explore the interaction of long-term inventory replenishment policies and the short term pricing strategies of a seller, which is the focus of Section 5.3. Then in Section 5.4, we provide numerical examples and further discussions for all these models.

5.1 The Fixed Inventory Model

5.1.1 Introduction and Literature Review

Nowadays with the fast growth of electronic commerce, online shopping/price catalog and mobile payment, it is almost costless to a retailer to implement changes on pricing and product availability in real-time. With this in mind, the research covered in this section proposes a continuous time model to best capture the frequent changes in price and product availability/service capacity in a short term planning horizon. At the same time, this gives rise to the fixed-inventory assumption of our model since the planning horizon we are considering is too short to allow for restocking.

There is a huge literature on the topic of revenue management and pricing. Talluri and van Ryzin (2006) [136] provides a review of a series of basic models, where monopoly is usually assumed. On the other hand, oligopoly pricing is still a challenging topic and a growing literature. Vives (2001) [142] gives details of classical models in quantity competition (Cournot) and price competition (Bertrand and Edgeworth). Recently, with the emergence of e-commerece and fast growing information systems, the role of dynamic pricing is becoming increasingly important. Elmaghraby and Keskinocak (2003) [44] gives a review of existing models categorized by replenishment or fixed inventory, dependent or independent demand over time, and myopic or strategic customers. In competitive pricing models in the service industry, Netessine and Shumsky (2005) [104] provides insights into pricing competitions in single-leg and multi-leg airline services.

In the study of dynamic oligopoly pricing with fixed inventories for a homogeneous, perishable product, or a set of complementary products, different scholars takes on different models and assumptions. Perakis and Sood (2006) studies the General Nash Equilibrium (GNE) of sellers using robust optimization in their response with prices and protection inventory levels as decisions. Gallego and Hu (2014) [61] studies the competition with both deterministic and stochastic demand, and with both open-loop and feedback information structure. The decision variables for the sellers are prices over time, and if a seller runs out-of-stock, a choke price, generating zero demand, is posted. Martinez-de-Albeniz and Talluri (2011) [97] considers sellers facing fixed number of units to sell over a fixed number of periods, in which customers arrive following stochastic processes and each customer purchase one unit of product from the seller offering lowest price.

5.1.2 Best Response Problem, Generalized Nash Equilibrium and QVI

In this part we follow Chapter 10 of Friesz (2010) [54] and Friesz et al. (2012) [57] to quickly summarize the settings of the fixed inventory model, starting from some basic assumptions. (1) *Perfect and complete information:* We assume that perfect information is obtained by each seller about the structure of demand, the impact of price changes on demand, and the initial inventory; (2) *Consumer choice and demand:* We assume that demand for the output of each seller is a function only of current prices and that prices are the only factor that distinguishes products from one another. That is, we assume there is a deterministic demand function faced by each seller that depends on own-period prices; (3) *Product:* We assume there is a single commodity and that inventory is salable for all $t \in [t_0, t_f]$ and is perishable meaning worthless at t_f; (4) *Objectives:* We further assume that sellers maximize the present value of their respective revenues by setting prices and do not employ any other type of strategies. Accordingly a differential Nash equilibrium will describe the market of interest.

Now we introduce the notations. Let us use $\pi_s(t)$ to denote the price charged by seller $s \in \mathcal{S}$ at time $t \in [t_0, t_f]$, where $\mathcal{S}$ denotes the set of all sellers and $\pi_s \in L^2[t_0, t_f]$, $t_0 \in \mathbb{R}^1_+$, $t_f \in \mathbb{R}^1_{++}$, $t_f > t_0$ where $L^2[t_0, t_f]$ is the space of square-integrable functions. For any subscript $s \in \mathcal{S}$ and any given instant of time $t_1 \in [t_0, t_f]$, we stress that $\pi(t_1) \in \mathbb{R}^1_+$ is a scalar. We also use the notation $\pi = (\pi_s : s \in \mathcal{S}) \in (L^2[t_0, t_f])^{|\mathcal{S}|}$ to represent the column vector of prices. We use

$$h_s[\pi(t); \xi_s] : \left(L^2[t_0, t_f]\right)^{|\mathcal{S}|} \times \mathbb{R}^1_+ \longrightarrow \mathcal{H}^1[t_0, t_f] \tag{5.1}$$

to represent the *observed demand* for the output of each seller $s \in \mathcal{S}$ corresponding to a specific realization of the vector of parameters $\xi_s \in (L^2[t_0, t_f])^m$. We let $d_s(t)$ represent the *realized demand* served by seller $s \in \mathcal{S}$, and define the column vectors $d = (d_s : s \in \mathcal{S}) \in (L^2[t_0, t_f])^{|\mathcal{S}|}$. This is the other set of decision variables for the sellers. At any time for any seller s, the realized demand must be less than or equal to its observed demand. Furthermore, $\pi_{\min}$ is a lower bound on prices, and $\pi_{\max}$ is the upper bound on prices; $D_{\min}$ is the lower bound on realized demand. These quantities are introduced to reflect the possible regulations existing on the market. Following the tradition of game-theoretic literature we take $\pi = (\pi_s, \pi_{-s})$ to be the complete column vector of prices, with π_{-s} being price of all other sellers' from seller s's perspective.

Problem 5.1. *(Best Response Problem) With each seller s aiming at maximizing its own revenue given other sellers' price π^*_{-s}, we have the following formulation:*

$$\max_{\pi_s, d_S} J_s = \int_{t_0}^{t_f} \exp(-\rho t)\pi_s\left(t\right) d_s(t)dt \tag{5.2}$$

subject to

$$\begin{aligned} d_s &\leq h_s\left(\pi_s, \pi^*_{-s}; \xi_s\right) \\ K_s &\geq \int_{t_0}^{t_f} d_s\left(t\right) dt \\ \pi_{\min} &\leq \pi_s \leq \pi_{\max} \\ D_{\min} &\leq d_s \end{aligned}$$

Note that this problem falls in the general category of an infinite dimensional mathematical program. In addition, we assume the following regularity conditions for Problem 5.1.

Definition 5.1. *(Regularities for the Best Response Problem) Problem 5.1 has the following regularity conditions:*
1. Prices are bounded from above and below, especially for all $s \in \mathcal{S}$ we have:

$$ess \sup_{\xi_s, \pi_{-s}} h_s\left(\pi_{\max}, \pi_{-s}; \xi_s\right) = 0$$

2. Every demand function $h_s(\cdot, \pi_{-s}; \xi_s)$ is concave in π_s for all feasible prices; and $h_s(\cdot, \pi_{-s}; \xi_s)$ is strictly monotonically decreasing in its own price, that is

$$\int_{t_0}^{t_f} \exp(-\rho t)[h_s\left(\pi^1_s, \pi^1_{-s}; \xi_s\right) - h_s\left(\pi^2_s, \pi^2_{-s}; \xi_s\right)](\pi^1_s - \pi^2_s)dt < 0$$

for all $s \in \mathcal{S}$ and feasible prices $\pi^1_s \neq \pi^2_s$, and all parameters ξ_s.

3. The bounds $\pi_{\min}, \pi_{\max}, D_{\min} \in \mathbb{R}^1_{++}$

Especially we consider the following affine function of observed demand that satisfies all the regularities above:

$$h(\pi_s, \pi_{-s}; \xi_s) = \alpha_s(t) - \beta_s(t)\pi_s + \sum_{r \neq s} \gamma_{sr}(t)\pi_r.$$

For detailed comprehensive review on different demand functions applied in revenue management along with their foundations from choice theory, see for example, Talluri and van Ryzin (2006) [136], Bodea and Ferguson (2014) [27] and Ledvnia and Sircar (2011) [89]. Chapter 10 of Friesz (2010) [54] and Friesz et al. (2012) [57] study the analytical properties of this problem including convexity of the feasible region, existence and uniqueness of the solution, etc. We here will not repeat such analysis. Nevertheless, it is helpful to re-state seller s's feasible region as:

$$\Lambda_s(\pi^*_{-s}, \xi_s) \equiv \Big\{(\pi_s, D_s) : \pi_{\min} - \pi_s \leq 0,\, \pi_s - \pi_{\max} \leq 0,\, \int_{t_0}^{t_f} d_s\,(t)\, dt - K_s \leq 0 \quad D_{\min} - d_s \leq 0 \quad d_s - h_s\left[\pi_s, \pi^*_{-s}, \xi_s\right] \leq 0\Big\}$$

This helps us realize that each seller's feasible region is not only parameterized by ξ_s, but also by other sellers' price, which indicates that the price equilibrium to solve for is a generalized Nash equilibrium.

Definition 5.2. *(Generalized NE for DPFI) With the set of sellers* $\mathcal{S}$ *and with each seller's problem defined as in Problem 5.1, we further define*

$$\Lambda(\pi, d) = \prod_{s=1}^{|\mathcal{S}|} \Lambda_s(\pi_{-s}, \xi_s)$$

A generalized Nash equilibrium of dynamic pricing with fixed inventories is a tuple of strategies (π^*, d^*) *such that each* (π^*_s, d^*_s) *solves the infinite dimensional mathematical program in Problem 5.1.*

Based on the analysis of the Best Response Problem, Chapter 10 of Friesz (2010) [54] gives the following theorem that the DPFI generalized Nash equilibrium is equivalent to a differential quasi-variational inequality.

Theorem 5.1. *(Generalized Nash equilibrium among sellers expressible as a* ***quasi-variational inequality****) The generalized Nash equilibrium among sellers $s \in \mathcal{S}$ that is described by Definition 5.2, is equivalent to the following quasi-variational inequality:*

$$\left.\begin{array}{c} \textit{find } (\pi^*, d^*) \in \Lambda(\pi^*, \xi) \ \textit{ such that} \\ \sum_{s\in\mathcal{S}} \int_{t_0}^{t_f} \exp(-\rho t) \left[d_s^* \cdot (\pi_s - \pi_s^*) + \pi_s^* \cdot (d_s - d_s^*)\right] dt \leq 0 \quad \forall\ (\pi, d) \in \Lambda(\pi^*, \xi) \end{array}\right\} \tag{5.3}$$

where

$$\Lambda(\pi^*, \xi) \triangleq \prod_{s\in\mathcal{S}} \Lambda_s\left(\pi_{-s}^*, \xi_s\right)$$

when the regularity conditions in Definition 5.1 hold.

5.2 Analysis of DPFI model by DVI

5.2.1 From GNE to DVI

In this section, we give the DVI that is equivalent to the generalized Nash equilibrium. This result is important to the qualitative and numerical analysis of this model, since there is a series of results such as necessary conditions and equivalence to the fixed point problem that could be built upon an equivalent differential variational inequality. In order to facilitate further analysis, we introduce the shorthand $y \triangleq (\pi, D)$ to denote the collection of decision variables of all the players on the market. We let

$$\Lambda(\xi) \triangleq \prod_{s\in\mathcal{S}} \Lambda_s(\pi_{-s}, \xi_s)$$

be the feasible set of the market. For a known and given ξ, let $\Lambda(\xi) \triangleq \Lambda$; and the mapping $F(y) : \Lambda \to \left(L^2[t_0, t_f]\right)^2$ expressed as

$$F(y,t) \triangleq \sum_{s\in\mathcal{S}} \left(\nabla_{y^s} J_s(y)\right) = \sum_{s\in\mathcal{S}} e^{-\rho t} \begin{pmatrix} d_s \\ \pi_s \end{pmatrix} \tag{5.4}$$

Theorem 5.2. *(Generalized Nash equilibrium among sellers expressible as a* ***variational inequality****) The generalized Nash equilibrium among sellers $s \in \mathcal{S}$ that is described by Definition 5.2, is equivalent to the solution of the following variational* *inequality:*

$$\left.\begin{array}{c}\textit{find } (\pi^*, d^*) \in \Lambda(\xi) \ \textit{ such that} \\ \sum_{s\in\mathcal{S}} \int_{t_0}^{t_f} \exp(-\rho t) \left[d_s^* \cdot (\pi_s - \pi_s^*) + \pi_s^* \cdot (d_s - d_s^*)\right] dt \le 0 \quad \forall\ (\pi, d) \in \Lambda\end{array}\right\} \tag{5.5}$$

when the regularity conditions in Definition 5.1 hold. In addition, the set Λ *is weakly compact under the above regularity conditions, thus a solution of* $VI(F, \Lambda)$ *exists, and the market admits a general equilibrium. (5.5) in compact form could be expressed as:*

$$\left.\begin{array}{c}\textit{find } y^* \in \Lambda \textit{ such that} \\ \int_{t_0}^{t_f} F(y,t)^T (y(t) - y^*(t)) dt \le 0 \quad \forall\ y \in \Lambda\end{array}\right\} \tag{5.6}$$

In proving the theorem above we used the conclusion from Facchinei et al. (2007) [48], the notation follows from the classical results of existence and uniqueness of equilibrium points for n-person Nash games by Rosen (1965) [123].

Lemma 5.1. *(Finite dimensional GNE to VI, Facchinei et al.) For a finite dimensional GNE with player* $\nu, (\nu = 1, \cdots, n)$*'s problem*

$$\max_{x^\nu} J_\nu \;=\; \theta_\nu(x^\nu, x^{-\nu}) \tag{5.7}$$

$$s.t. x^\nu \;\in\; X_\nu(x^{-\nu}) \triangleq \left\{x^\nu : (x^\nu, x^{-\nu}) \in X\right\} \tag{5.8}$$

and with regularities: for all player ν *(1)* θ_ν *is continuously differentiable in (the whole vector)* x*;(2)* $\theta_\nu(x^\nu, x^{-\nu})$ *is pseudo-concave in* x^ν*; also (3) the set* X *is closed and convex. Then every solution of the* $VI(X, \mathcal{F})$*, is a solution of the GNEP. With*

$$\mathcal{F}(x) \triangleq \begin{pmatrix} \nabla_{x^1}\theta_1(x) \\ \vdots \\ \nabla_{x^N}\theta_N(x) \end{pmatrix}. \tag{5.9}$$

If, in addition,(4) the set X *is compact, then the* $VI(X, \mathcal{F})$ *admits a solution therefore the original GNEP has a solution.*

We also notice other similar results in the literature of finite dimensional quasi-variational inequality that gives conditions when an equivalent VI to a QVI could be obtained, for example, Harker (1991) [68] gives one such result based on regularity conditions imposed onto the feasible region of a QVI. The proof we give here for Theorem 5.2 is also inspired by Bressan and Han (2013) [31].

Proof. (Generalized Nash equilibrium among sellers expressible as a **variational inequality**)

Step1: We first consider a finite time discretization of the generalized Nash game. Let $\Delta t = \frac{t_f - t_0}{n}, t_i = t_0 + i\Delta t \; \forall i$, then $t_n = t_f$. We then define a modification of the decision variable y_n and feasible region accordingly. We also remind ourselves with the definition of the operator in this problem:

$$F(y,t) \triangleq \sum_{s\in\mathcal{S}} \Big(\nabla_{y^s} J_s(y)\Big) = \sum_{s\in\mathcal{S}} e^{-\rho t} \begin{pmatrix} d_s \\ \pi_s \end{pmatrix}. \tag{5.10}$$

For each player, the decision variable $y_{s,n}$ is a $2-$tuple: $(\pi_{s,n}, d_{s,n})$ which is in $\left(L^2\left[t_0, t_f\right]\right)^2$ and also piecewise constant. Similarly, we restrict the coefficients $\alpha_s(t), \beta_s(t), \gamma_{rs}(t)$ to be also piecewise constant functions with the mesh defined above, and name the feasible region induced in such a way $\Lambda_{p,n}$. Therefore the restricted definition of feasible region Λ becomes the following:

$$\Lambda_n \triangleq \{y_n : y_n \in \Lambda_{p,n}, y_n \text{ piecewise constant with given mesh}\}.$$

On the other hand, we formulate the VI that is targeted to be equivalent to the restricted GNE. Here we define the mapping $F_n(y_n) : \Lambda_n \to \mathbb{R}^{2n}$

$$F_n(y_n) \triangleq \begin{pmatrix} \vdots \\ \frac{1}{t_{i+1}-t_i}\int_{t_i}^{t_{i+1}} F(y_n,t)dt \\ \vdots \end{pmatrix}; \tag{5.11}$$

we also notice that $y_n(t) = \sum_{i=1}^{n} y_{s,i} e_i(t)$ where $e_i(t)$ are the natural basis:

$$e_i(t) = \begin{cases} 1 & \text{if } t \in [t_i, t_{i+1}) \\ 0 & \text{otherwise} \end{cases}$$

Then there is a one-to-one mapping that maps $y_n(t)$ to a finite dimensional vector $\psi \in \mathbb{R}^{2n|s|}$. Thus, $F_n(\cdot)$ could be viewed as a function from $\mathbb{R}^{2n|s|}$ to $\mathbb{R}^{2n}$. Therefore if we consider the following variational inequality:

$$\left.\begin{array}{c} \text{find } \psi^* \in \mathbb{R}^{2n|s|} \text{ such that} \\ \int_{t_0}^{t_f} F^n(\psi^*,t)^T(\psi(t) - \psi^*(t))dt \leq 0 \quad \forall\, \psi \in \Lambda_n \end{array}\right\} \tag{5.12}$$

Here we use ψ and $y_n(t)$ interchangeably due to the natural isomorphism mentioned above.

Moreover, we point out that the constraint of observed and in the restricted in the finite dimensional GNE problem qualifies the regularity condition (5.8) of Lemma 5.1. Consider a case with seller s under a n-divided time discretization, by the virtue of her best response problem, the corresponding constraint of $d_s - h_s \left[\pi_s, \pi^*_{-s}, \xi_s\right] \le 0$ at time instance i in the time-discretized GNE reads:

$$d_{s,i} \le h_1(\pi_{s,i}, \pi_{-s,i}, i\Delta t) \tag{5.13}$$

note that $y_{s,i} = (\pi_{s,i}, d_{s,i}), \forall i = 1, \cdots, n$, this means $y_s \in \Lambda_s(y_{-s})$. To further illustrate this, just consider for a moment the special case with $|\mathcal{S}| = 2$ with affine $h(\cdot)$ time-invariant α, β and γ, we have for $s = 1$ and each $i = 1, \cdots, n$, constraint (5.13) becomes:

$$(\pi_{1,i}, d_{1,i})^T \cdot (\beta_1, 1) \le \alpha_1 + (\pi_{2,i}, d_{2,i})^T \cdot (\beta_2, 0)$$

Now are able to claim that Lemma 5.1 could be applied to such VI, which is equivalent to the original infinite dimensional VI restricted on Λ_n, such that the following is true: VI (5.12) admits a solution, and it is the solution to the GNE defined on the restrained feasible region.

To substantiate this claim we need to check other regularity conditions. First of all, the continuity and pseudoconcavity of the objective function of each player still holds. We then check that Λ_n is closed and convex, which holds true due to the regularities assumed, the restriction of parameters $\alpha_s(t), \beta_s(t), \gamma_{rs}(t)$ to piecewise constant functions with same mesh, and the linearity of the observed demand with respect to price. The above also leads to, finally, the compactness of Λ_n.

Step 2: Notice that the time meshing in the last step is valid for any given $n \ge 1, n \in \mathbb{N}$, it is clear that we now have a sequence of generalized Nash games whose solutions are given via equivalent VIs. Namely, we have that $\forall n$, $\exists y^*_n \in \Lambda_n$ such that

$$\langle F(y^*_n),\, y_n - y^*_n \rangle \;\le\; 0, \qquad \forall y_n \in \Lambda_n; \tag{5.14}$$

and with the y^*_n's being the solution of the generalized Nash games defined on a restricted (piecewise constant) function space. In this step, we will show that any

GNE matching the regularity conditions in Theorem 5.2 could be well approximated by the series of GNEs defined on the restrained feasible region. This is based on the fact that all the functions in $L^2[t_0, t_f]$, with the regularities imposed by Theorem 5.2 could be well approximated by the series of piecewise constant functions.

Step 3: The remaining step is to show that the sequence of equivalent VIs converge to the target VI. Due to the fact that such sequence $\{y_n^*\}$, $n \geq 1$ is uniformly bounded, there exists a subsequence $\{y_{n_k}^*\} \subset \{y_n^*\}$ that converges weakly to a point $y^* \in \Lambda$. It remains to show that y^* obtained in this way is the solution of the original infinite-dimensional VI as well as the generalized Nash equilibrium. We proceed as follows:

Given any $y \in \Lambda$, there exists a sequence of piecewise-constant functions $\{\bar{y}_n\}$, $n \geq 1$ with each $\bar{y}_n \in \Lambda_n$ such that $\bar{y}_n \rightharpoonup y$ weakly. According to (5.14), we have:

$$\langle F(y_n^*), \, \bar{y}_n - y_n^* \rangle \;\leq\; 0 \qquad \forall n \geq 1. \tag{5.15}$$

Due to the fact that $F(y,t)$ is linear in y, we have that

$$\lim_{n \to +\infty} \langle F(y_n^*), \, \bar{y}_n - y_n^* \rangle \;=\; \langle F(y^*), \, y - y^* \rangle$$

That is,

$$\langle F(y^*), \, y - y^* \rangle \;\leq\; 0$$

Finally, since y is arbitrarily chosen, we have proven the desired variational inequality.

□

Remark 5.1. *Notice that Lemma 5.1 comes as a combination of Theorem 2.1 and Proposition 2.2 from Facchinei's paper, which is applicable in the proof of necessary condition and existence condition. In the same paper however, there is one more result further discussing the solution of the corresponding KKT with the GNE solution. That result requires more regularity conditions and is not utilized here. Facchinei's paper also points out that Aubin (2007) [6] contains similar results. For further review and developments on this end, please refer to, for example Kulkarni and Shanbhag (2009) [86].*

5.2.2 DVI equivalent FPP and computation

It is possible to solve the VI problem (5.6) using a simple but effective fixed point algorithm. To understand the fixed point algorithm, it is helpful to consider the following abstract variational inequality: find $u^* \in U$ such that

$$\langle F(u^*, t), u - u^* \rangle \leq 0 \qquad \forall\, u \in U \tag{5.16}$$

where

$$F : \left(L^2[t_0, t_f]\right)^m \times \mathbb{R}^1_+ \longrightarrow \left(\mathcal{H}^1[t_0, t_f]\right)^m$$
$$U \subseteq \left(L^2[t_0, t_f]\right)^m$$

We have the following result:

Theorem 5.3. *Fixed point formulation of infinite dimensional variational inequality. If $F(u,t)$ is differentiable and pseudo-concave with respect to u while U is convex, the variational inequality (5.6) is equivalent to the following fixed point problem:*

$$u^* = P_U[u + \alpha F(u^*, t)] \tag{5.17}$$

where $P_U[.]$ is the minimum norm projection onto U and $\alpha \in \mathbb{R}^1_{++}$ is an arbitrary positive constant.

Naturally there is a fixed point algorithm associated with the above fixed point formulation and expressed as the following iterative scheme:

$$u^{k+1} = P_U\left[u^k - \alpha F\left(u^k, t\right)\right] \tag{5.18}$$

Algorithm 5.1. *(Fixed Point Algorithm for DPFI) The positive scalar $\alpha \in \mathbb{R}^1_{++}$ may be chosen empirically to assist convergence and may even be changed as the algorithm progresses:*

Step 0. *Initialization* *Identify an initial feasible solution u^0 such that $u^0 \in U$ and set $k = 0$.*

Step 1. *Solve optimal control sub-problem* *Solve the following optimal control prob-*

lem:

$$\min_{v} J^k(v) = \int_{t_0}^{t_f} \frac{1}{2}\left[u^k - \alpha F\left(u^k, t\right) - v\right]^2 dt \quad s.t. \quad v \in U \tag{5.19}$$

Call the solution u^{k+1}.

Step 2. *Stopping test* *If $\left\|u^{k+1} - u^k\right\| \leq \varepsilon_1$ where $\varepsilon_1 \in \mathbb{R}^1_{++}$ is a preset tolerance, stop and declare $u^* \approx u^{k+1}$. Otherwise set $k = k + 1$ and and go to Step 1.*

The convergence of the above algorithm is guaranteed under certain conditions by the following result:

Theorem 5.4. *Convergence of the fixed point algorithm. Let the variational inequality (5.16) be regular in the sense of Theorem 5.2.When $F(u, t)$ is strongly monotonically increasing with respect to u and satisfies the Lipschitz condition*

$$\|F(u, t) - F(v, t)\| \leq \rho_0 \|u - v\|$$

for some $\rho_0 \in \mathbb{R}^1_{++}$ and all $u, v \in \left(L^2[t_0, t_f]\right)^m$, while U is non-empty and convex, the fixed point algorithm presented above converges for appropriate α.

Proof. The strong monotonicity requires that for some $\beta \in \mathbb{R}^1_{++}$

$$\langle F(u) - F(v), u - v\rangle \geq \beta \|u - v\|^2 \tag{5.20}$$

holds for any $u, v \in \Gamma$, and since Γ is parameterized by ξ, we have $\beta = \beta(\xi)$. The α consistent with convergence shall satisfy

$$0 < \zeta = 1 - 2\alpha\beta(\xi) + \rho_0\alpha^2 < 1$$

we first consider

$$1 - 2\alpha\beta(\xi) + \rho_0\alpha^2 < 1 \tag{5.21}$$

which leads to

$$\alpha\left(-2\beta(\xi) + \rho_0\alpha\right) < 0 \Longrightarrow \alpha < \bar{\alpha} = \frac{2\beta(\xi)}{\rho_0}.$$

We should also have

$$1 - 2\alpha\beta(\xi) + \rho_0\alpha^2 > 0 \tag{5.22}$$

and since $\rho_0 > 0$ we only need to consider $\Delta = 4(\beta^2(\xi) - \rho_0)$. If we have $\Delta \geq 0$, then $\beta^2(\xi) - \rho_0 \geq 0, \beta^2(\xi) \geq \rho_0$

$$\alpha_{1,2} = (\frac{\beta(\xi)}{\rho_0}) \pm \sqrt{(\frac{\beta(\xi)}{\rho_0})^2 - \frac{1}{\rho_0}} \tag{5.23}$$

in this case we should have

$$0 < \alpha < \alpha_2 \text{or } \alpha_1 < \alpha < \bar{\alpha} \tag{5.24}$$

On the other hand, if $\Delta < 0, \beta^2(\xi) < \rho_0$ we then have

$$\zeta = 1 - 2\alpha\beta(\xi) + \rho_0\alpha^2 = (\alpha - \frac{\beta(\xi)}{\rho_0})^2 + \frac{\rho_0 - \beta^2(\xi)}{\rho_0} > 0 \tag{5.25}$$

and we only need

$$0 < \alpha < \bar{\alpha} \tag{5.26}$$

□

Theorem 5.5. *Let $\{u^k\}_{k=0}^{\infty}$ be the series of solutions generated by the fixed point algorithm and let u^* be the solution of the problem. Then the number of iterations required to reach a solution within ε distance from u^* is of order $O(\frac{\ln \frac{\varepsilon(1-\eta)}{\Upsilon}}{\ln \eta})$ where $\eta = \sqrt{\zeta}$, and Υ is the diameter of the feasible region Γ.*

Proof. With the same arguments in the convergence proof, the following relationship could be established between successive iterations:

$$\left\|u^{k+1} - u^k\right\|^2 \leq \zeta^k \left\|u^1 - u^0\right\|^2 \tag{5.27}$$

which means

$$\left\|u^{k+1} - u^k\right\| \leq \eta^k \left\|u^1 - u^0\right\| \tag{5.28}$$

thus by yielding triangular inequality to this countable number of iterations we have

$$\left\|u^{k+1} - u^*\right\| \leq \left\|u^1 - u^0\right\| \sum_{m=1}^{\infty} \eta^{k+m} \leq \Upsilon \sum_{m=1}^{\infty} \eta^{k+m} \tag{5.29}$$

$$= \Upsilon \left[\sum_{m=0}^{\infty} \eta^m - \sum_{m=0}^{k} \eta^m\right] \tag{5.30}$$

$$= \quad \Upsilon \frac{\eta^k}{1-\eta} \tag{5.31}$$

($\Upsilon = \sup_{u,v\in\Gamma} \|u - v\|$), thus if we want $\left\|u^{k+1} - u^*\right\| \leq \varepsilon$ we need

$$k \geq \frac{\ln \frac{\varepsilon(1-\eta)}{\Upsilon}}{\ln \eta} \tag{5.32}$$

This gives the result that the number of iterations required for the algorithm to converge to an ε-close solution to the equilibrium is of order $O(\frac{\ln \frac{\varepsilon(1-\eta)}{\Upsilon}}{\ln \eta})$. □

We could relate this complexity result to that by Perakis and Sood (2006) [114], in which a similar expression is given in terms of how many iterations would be required. The main difference is in how the expressions are parameterized. In 5.5, the parameter η should be consistent with the fixed point algorithm where as in Perakis and Sood, a parameter with similar function should be explicitly related to: (1) the Lipschitz constant and the strongly monotonicity of the observed demand function $h(\cdot, \cdot)$; (2) the Lipschitz constant and the strongly monotonicity of the robust inverse demand function. We will present the numerical examples of DPFI problem later in Section 5.4 along with other examples.

5.3 The Dual-time-scale Fixed Inventory Model

5.3.1 Introduction and Literature Review

From Section 5.1 and Section 5.2, we have reviewed and analyzed the DPFI model as a competitive revenue management model in continuous time. At the same time, in reality we know different decisions such as service capacity adjustment, inventory replenishment and adjustment of pricing would have to "take place on different time scales" (Bitran and Caldentey (2003) [26]). For example for as a retail service provider, "the number of shirts to purchase from an overseas supplier are decided long before demand is realized and price policies are implemented" [26]. Therefore in this section, in order to capture and study such retailers' competitive behavior in different time scales, we form the dual-time-scale dynamic pricing with fixed inventory model (DT-DPFI) by giving the DPFI model in the last section a discrete time scale for inventory replenishment.

More specifically, we try to give a mathematical statement of the following story: There are two time scales for the problem. The discrete time scale correspond to the selling seasons experienced by the retailers, over this time scale, the retailers build their inventory by deciding on the volume of product to order. At the end of every season, where a DPFI game as in Section 5.1 has been carried out by all sellers, the inventory decisions are made. Those decisions could be made based on all historical information on available to a retailer. And the inventory is fixed once such a decision has been made for a seller. The sellers get their commodity from sources with ample capacity and the cost-side factors such as wholesale price is taken as exogenous to our model. At the beginning of the next selling season, every retailer begins to make the most revenue by making tactical decisions on pricing and demand-to-fulfill in continuous time, which is exactly the case described by the DPFI model in Section 5.1. The next selling season ends and reflections on the historical inventory and pricing experience leads to the retailers' new ordering decisions. Also following the spirit of the DPFI model, the left-over inventory of a given season has no salvage value.

For a literature review on dynamic pricing competition and game-theoretic models in revenue management, please refer to Section 5.1. In this section, we focus on reviewing dynamic inventory models in discrete and continuous time, especially, we are interested in competitive dynamic inventory models that also integrates pricing or production decisions. Examples of such research are: Chen et al. (2001) [38] first reviewed different dynamic joint pricing and inventory model, and then proposes near optimal strategies for a two-echelon distribution system. Bernstein and Federgruen (2004) [17] addresses an infinite-horizon oligopolistic model with periodic review on inventory and discusses the optimization of the aggregate profits in the supply chain by deciding on wholesale pricing schemes. Adida and Perakis (2010) [2] study a make-to-stock manufacturing system with a duopoly through dynamic pricing and inventory control. Here the inventory is modeled in continuous time with a fluid dynamic model. Friesz et al. (2006) [59] also employs this type of inventory dynamics. The problem to consider in this paper is an n-player oligopoly on a network with production, pricing and transportation decisions.

There is also a series of research that tries to build dynamic inventory models on a discrete time scale by using time-series methods. For example, Aviv (2003) [7] proposes a general linear state-space model to capture the adaptive inventory replenishment policies of different supply chain members that utilizes the Kalman Filter technique. Gilbert (2005) [65] also presents multi-stage dynamic supply chain model based on the autoregressive integrated moving average (ARIMA) models. Our discrete-time scale inventory model in this section has been inspired by those models.

Unfortunately, all of the dynamic supply chain and revenue management models reviewed thus far dwell on either the discrete or continuous time scale, making these models unable to capture the competitive behaviors of different supply chain members with both long and short term decision capabilities. Very few contributions have been made in the literature of supply chain management with multiple time scales being dynamically captured. In terms of time scale modeling, the dual index approach in Timpe and Kallrath (2000) [138] is the closest to the model in this section, where two time scales, one for production decisions and one for marketing decisions, are treated with different time resolutions. In this section, we propose a novel model utilizing the idea of dual-time-scales; and especially, we emphasize the dynamic pricing competition by giving it a continuous time. The idea of using a dual-time scale dynamic game-theoretic model to reflect both short-term and long-term competitive decisions has been applied in the area of transportation studies. For example, given the evolution of traffic demand, Friesz et al. (2011) [56] presents a dual-time-scale layout of dynamic traffic assignment (DTA) model, see Chapter 3 for a brief review on the topic of DTA.

In the remaining of this section we propose the season-to-season evolutionary dynamics of sellers' inventory policies and then build a dual-time-scale DVI model for the DT-DPFI problem. We then analyze the dual-time-scale DVI and propose effective algorithm for the solution of the DT-DPFI problem.

5.3.2 Problem Formulation

We extend the problem formulation from Section 5.1 and let $\tau \in \{1, 2, \cdots, N\}$ be one typical discrete selling season and assume the length of each selling season to be Υ, then the continuous time within each selling season would be $t \in [(\tau-1)\Upsilon, \tau\Upsilon]$ for all $\tau \in \{1, 2, \cdots, N\}$. On this discrete time scale, we have a inventory replenishment model with the following abstract form:

$$\begin{aligned} K_s^{\tau+1} &= \Xi(K^\tau; \pi^\tau, d^\tau; \theta_s^{\tau+1}),\ \forall s \in \mathcal{S}, \quad \tau \in \{1, 2, \cdots, (N-1)\} \\ K_s^\tau &\geq 0, \quad \forall s \in \mathcal{S}, \quad \tau \in \{2, 3, \cdots, N\} \\ K_s^1 &= \kappa_s > 0, \quad \forall s \in \mathcal{S} \end{aligned}$$

here the notations are expanded from Section 5.1: K_s^τ is the inventory for seller s at the start of selling season τ; thus $K^\tau = (K_s^\tau : s \in \mathcal{S})$ and $K = (K^\tau : \tau \in \{1, 2, \cdots, N\})$. The initial inventory at the beginning of the first selling season is taken as given by $K_s^1 = \kappa_s$. The update of the inventory decision is captured by an abstract mapping that takes into account not only the within-season decisions of all sellers (π^τ, D^τ), but also the information on discrete time scale such as the inventory K^τ and $\theta_s^{\tau+1}$ which is a set of exogenous parameters to capture the (projected) market condition in selling season $(\tau + 1)$. This includes, for example, the projected wholesale price of the commodity in the coming selling season. Here we assume that as soon as one selling season ends, the information on all sellers in this season are available to be used by seller s to help its inventory decision.

As an example, in this section we assume that $K_s^{\tau+1}$ is to be a weighted average of:

1. K_s^τ which is the the inventory built with current selling season τ; and
2. a linear model taking $\bar{\pi}^\tau$, which is the scaled average sold price in season τ and $\theta_s^{\tau+1}$, which is the projected wholesale price, as input.

And the above set of evolutionary dynamics become: for $\forall s \in \mathcal{S}$:

$$K_s^{\tau+1} = \alpha_s K_s^\tau + (1-\alpha_s)(k_{0,s}^\tau + k_{1,s}^\tau \bar{\pi}_s^\tau + k_{2,s}^\tau \theta_s^{\tau+1}), \tau \in \{1, 2, \cdots, (N-1)\} \quad (5.33)$$

$$K_s^\tau \geq 0, \quad \tau \in \{2, 3, \cdots, N\} \quad (5.34)$$

$$K_s^1 = \kappa_s > 0, \quad (5.35)$$

where the scaled average sold price $\bar{\pi}^\tau$ could be computed solely by information available by the end of selling season τ:

$$\bar{\pi}_s^\tau = \frac{\int_{(\tau-1)\Upsilon}^{\tau\Upsilon} \pi_s(t)\, d_s(t) dt}{\int_{(\tau-1)\Upsilon}^{\tau\Upsilon} d_s(t) dt} \left[\frac{\sum_{s \in \mathcal{S}} \int_{(\tau-1)\Upsilon}^{\tau\Upsilon} \pi_s(t)\, d_s(t) dt}{\sum_{s \in \mathcal{S}} \int_{(\tau-1)\Upsilon}^{\tau\Upsilon} d_s(t) dt} \right]^{-1},$$

which is the average sold price of seller s in season τ scaled by the overall sold price among all sellers. In this section we use such a quantity in Equation (5.33) to capture seller s's perspective on potential profitability. One ad-hoc specification of the quantity $k_{0,s}^{\tau}$ is:

$$k_{0,s}^{\tau} = \frac{1}{2}\left[\int_{(\tau-1)\Upsilon}^{\tau\Upsilon} d_s(t)dt + K_s^{\tau}\right]$$

which is the average of the inventory built and the actual volume sold in the current season. We use this quantity to reflect the seller's effort to try to lower the redundant inventory, which has zero salvage value according to our model. Such a weighted average approach is inspired by the time series models mentioned in the literature review of this section.

Note here that equation (5.33) is just an illustrative example of many alternative models to implement. Furthermore, with the rapid development of information systems, data mining and machine learning techniques (see, for review, Bishop (2007) [23], and for example, Oroojlooyjadid et al. (2016) [107]) and popular industrial application of those practice (see, for example, Williams et al. (2013) [146]), it is likely that equation (5.33) should be replaced by very sophisticated statistical/deep learning models based on very high dimensional historical data. Direct implementation of those models are beyond the scope of this dissertation. Please refer to Bertsimas and Perakis (2006) [20] for an example of how a seller could refine its forecasting on market demand and, at the same time, optimize the pricing strategy in a competitive setting. And notice that this learning-while-optimizing model is built using a discrete-time, multi-dimensional dynamic programming approach.

Based on the analysis about equivalent DVI of the within season DPFI problem and the above evolutionary dynamics to describe inventory replenishment policies, we are ready to propose the dual-time-scale dynamic pricing with fixed inventory model (DT-DPFI):

Problem 5.2. *(Dual-Time-Scale dynamic pricing with fixed inventory as a dual-time-scale DVI) Given a set of initial inventory at the beginning of selling season* $\tau = 1$, *a dual-time-scale dynamic pricing equilibrium with fixed inventory is to find*

$K^* \geq 0$ *and* $(\pi^*(t), d^*(t)) \in \Lambda^\tau$ *such that:*

$$\sum_{s \in \mathcal{S}} \int_{t_0}^{t_f} \exp(-\rho t) \left[d_s^* \cdot (\pi_s - \pi_s^*) + \pi_s^* \cdot (d_s - d_s^*) \right] dt \leq 0 \quad \forall \, (\pi, D) \in \Lambda$$

$$K_s^{\tau+1} = \alpha_s K_s^\tau + (1 - \alpha_s)(k_{0,s}^\tau + k_{1,s}^\tau \bar{\pi}^\tau + k_{2,s}^\tau \theta_s^{\tau+1}), \quad \tau \in \{1, 2, \cdots, (N-1)\}$$

$$K_s^\tau \geq 0, \quad \tau \in \{2, 3, \cdots, N\}$$

$$K_s^1 = \kappa_s > 0$$

5.3.3 Analysis and Algorithm

The within-season DPFI has been analyzed in detail in Section 5.1 and Section 5.2. In this part we focus on the discussion of the dual time scale model as described by Problem 5.2.

First of all, similar to the techniques applied in Friesz et al. (2011) [56] and reviewed by Pang and Stewart (2008) [108], we claim that a time stepping technique could be applied to solve DVI as in Problem 5.2. To see this, we first notice that at the end of season $\tau = 1$, according to our assumptions each seller knows about each other's initial fixed inventories represented by the full vector $K^{\tau=1}$, and each other's pricing and fulfilled demand strategies by the full vector $(\pi^1(t), d^1(t))$. Then following our discussions about the inventory replenishment decision process, the inventory decision for the next selling season $\tau = 2$ is determined solely by those information obtained from the current season. This means that another within-season DPFI equilibrium as in Definition 5.2 or in DVI form as in Theorem 5.2 is well defined for season $\tau = 2$ provided the new full vector $K^{\tau=2}$ is determined by (5.33). This gives rise to the following time stepping algorithm:

Algorithm 5.2. *(Time stepping for DT-DPFI) We give the following detailed statement of the time stepping algorithm for Problem 5.2:*

Step 0. *Initialization* *Set the initial fixed inventory at the beginning of season as* $K_s^1 = \kappa_s > 0$ *for all* $s \in \mathcal{S}$ *and fix model parameters* α_s, $\theta_s^{\tau+1}$, *etc. Set* $\tau = 1$.

Step 1. *Solve DPFI model in season* τ *Solve the DPFI model in season* τ *with Algorithm 5.1 and denote the equilibrium solution by* $(\pi^{\tau,*}(t), d^{\tau,*}(t))$. *Specifically, the within-season DPFI model takes fixed inventory* K^τ, *and the planning horizon is* $t \in [(\tau-1)\Upsilon, \tau\Upsilon]$.

Step 2. Inventory Replenishment *Compute the fixed inventory of the next selling season with* $(\pi^{\tau,*}(t), d^{\tau,*}(t))$ *by evolutionary dynamics (5.33).*

Step 3. Time Stepping *If* $\tau = N$*, stop. Otherwise* $\tau = \tau + 1$*, go to Step 1.*

5.4 Numerical Examples

In this section we present a series of numerical examples to illustrate different models and algorithms we proposed throughout this chapter. It is important to realize that in Algorithm 5.1, each iteration requires the solution of a fixed point subproblem in continuous time. These subproblems are linearly constrained quadratic programs and could be solved by many functional space algorithms, see for example Mioux (1986) [101] for more details. In this chapter, we use time-discretization approximation of each subproblem. All the numerical examples in this section are prepared in Matlab 2015b on a laptop computer equipped with Intel dual-core i5 processors and 8GB RAM.

5.4.1 Numerical Examples for the DPFI Model

We first consider a small example concerning the basic DPFI model. There are 2 sellers on the market with $s = 1, 2$, the time horizon to consider is $[t_0, t_f] = [0, 10]$. the initial fixed inventory endowments are $K_1 = 6000$ and $K_2 = 4000$, and we take the following parameterization:

$$\begin{aligned}
\alpha_1(t) &= 300, \alpha_2(t) = 200 \\
\beta_1(t) &= 1 - 0.04t, \beta_2(t) = 0.8 - 0.02t \\
\gamma_1(t) &= 0.6 + 0.03t, \gamma_2(t) = 0.6 + 0.01t
\end{aligned}$$

Other parameters include:

$$\begin{aligned}
p_{\min} &= (50, 35)^T, p_{\max} = (300, 300)^T, \\
d_{\min} &= (1, 1)^T, \rho = 0.01.
\end{aligned}$$

Algorithm 5.1 is employed for the solution of the DPFI equilibrium, with the fixed point subproblem solved by time discretization. Figure 5.1 gives the pricing and demand strategy with both players.

Here we need to make note of one important when operating Algorithm 5.1: a continuous time framework is used based on the fact that we represent the optimal control sub-problem in continuous time. This also applies to other algorithms that requires solving sub-problems in continuous time. On the other hand, since closed-form solutions are not always obtainable for these subproblems, time discretization is often applied in solving the sub problems. In this case, some curve fitting method is usually used onto the discrete-time solution in continuous time t to obtain an approximated continuous time solution and passed into the next iteration.

Remark 5.2. *In this numerical example, we use a discrete time approximation of each subproblem that is solved using the optimization toolbox of Matlab. A forth order polynomial in time is then fit to the discrete time solution of each subproblem, and the next fixed point iteration is carried out in continuous time.*

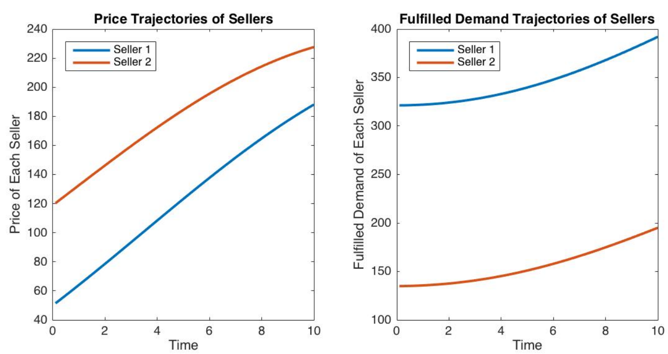

Figure 5.1. Dynamic pricing and demand fulfillment policies for both sellers in DPFI equilibrium

Remark 5.3. *Many other schemes may be invoked for solving the subproblem, and in fact, even for this problem the design of time discretization may not be perfect: a universal time meshing of* $N = 1000$ *is used through out each iteration of the numerical example, however it is possible that the error due to time discretization may build up, therefore, a varying scheme with* N_k *goes up in each iteration may lead to better numerical outcome. Certainly, a study on numerical errors and how they build up with each iteration is a promising direction of future research.*

Another interesting quantity to discuss is the efficiency/deficiency of the price equilibrium. In supply chain literature, there has been a series of research considering quantifying the efficiency of a supply chain when different coordination mechanisms are to be applied. For example, in Perakis and Roels (2007) [113], such efficiency is studied and bounded in the case of price-only contracts (where arbitrary quantity could be traded with a constant transaction price that is specified by the contract). At the same time, in the transportation and computer science literature, a quantity named the Price of Anarchy (PoA) is proposed in comparing the overall network congestion following either the user equilibrium (UE) or system optimal (SO) principle, see Roughgarden (2005) [124] for a comprehensive review on bounding the PoA with static traffic assignment. Another example is Cominetti et. al. (2009) [40], which gives PoA results on an atomic network game with a finite number of competitors, each might control a fraction of the total flow. However, there has been relatively few papers considering PoA type of quantities when the underlying games are in continuous time, dynamic setting.

In this section, we propose a scheme to systematically explore the deficiency of the DPFI game, and as a more interesting numerical example. First of all, we start with the formulation of the system optimal problem, which would capture the following scenario: all the sellers $s \in \mathcal{S}$ are in fact brunches of a larger organization, which seeks the maximal total profit with given inventory. There could be two cases under such scenario: (1) the organization could not re-allocate the inventory thus fixed inventory constraints should be satisfied locally (by each branch s); and (2) the organization has the power to re-allocate inventory, (and suppose doing so would not cause additional cost), thus there should be one fixed inventory constraint on the total inventory. For simplicity our definition follows the latter.

Problem 5.3. *(DPFI, System Optimal) Letting* $\pi = (\pi_1, \cdots, \pi_s, \cdots, \pi_{|s|})$ *and* $d = (d_1, \cdots, d_s, \cdots, d_{|s|})$*we define the system optimal DPFI problem with a given set of parameters* $\{\xi_s\}$*:*

$$J^{SO} = \max_{(\pi,D)} \sum_{s\in s} \int_{t_0}^{t_f} \exp(-\rho t)\pi_s(t)\, d_s(t)dt \tag{5.36}$$

$$s.t. d_s \leq h_s(\pi_s, \pi_{-s}; \xi_s) \quad \forall s \tag{5.37}$$

$$K = \sum_{s\in s} K_s \geq \sum_{s\in s} \int_{t_0}^{t_f} d_s(t)dt \tag{5.38}$$

$$\pi_{\max} \leq \pi_s \leq \pi_{\min} \quad \forall s \tag{5.39}$$

$$D_{\min} \leq D_s \forall s \tag{5.40}$$

At the same time, denote J_s^{GNE} as the total revenue for a seller s as a result of the pricing competition per Definition 5.2. With the total revenue in both the SO and the NE case, we are ready to define the PoA of the DPFI game as the supremum of the deficiency ratio:

Definition 5.3. *(Price of Anarchy, DPFI) The price of anarchy is the supreme of the ratio between the profit generated by the system optimal problem and the total profit generated by the generalized Nash equilibrium problem*

$$PoA = \sup \frac{J^{SO}}{\sum_{s \in s} J_s^{NE}} \tag{5.41}$$

Note that given a set of parameters, we can compute the solution of Problem 5.3 as well as the solution of the equivalent VI as in Theorem 5.2, thus numerically generate the deficiency of the competition as a ratio between different revenues. Furthermore, inspired by Gallego and Hu (2014) [61], and Maglaras and Meissner (2006) [95], we study the asymptotic behavior of the deficiency of the competition when the initial fixed inventory and the observed demand function are subject to tuning. Specifically, there are two important classes of parameters in our model, the first one is K_s, which stands for the inventory endowment of each seller, the second one is α_s, which is the zeroth order term in the observed demand function and measures the seller's ability to generate demand. As a special case, we study duopoly ($s = 1, 2$) with $\frac{K_1}{K_2} = r_1$, and without loss of generality, let $r_1 \geq 1$; at the same time, fix $\frac{\alpha_1}{\alpha_2} = \frac{1}{r_1}$. On the other hand, we define $r_2 = \frac{K_1+K_2}{\alpha_1+\alpha_2} > 0$, which is a measure of the ratio of total inventory over the level of demand that could be generated.

Figure 5.2 is an example when r_1 and r_2 are tuned each time by 2 times of the previous experiment:

Remark 5.4. *The concept of PoA originated in the literature of theoretical computer science (Koutsoupias and Papadimitriou [84]), see Section 1.5 of Roughgarden (2005) [124] for a historical remark. In this dissertation, we borrow the general idea of this type of metrics in bringing more interesting extension to the DPFI problem.*

Also note that with the combination of this model of the sellers' and some reasonable assumptions on static and dynamic change of consumers' utilities or

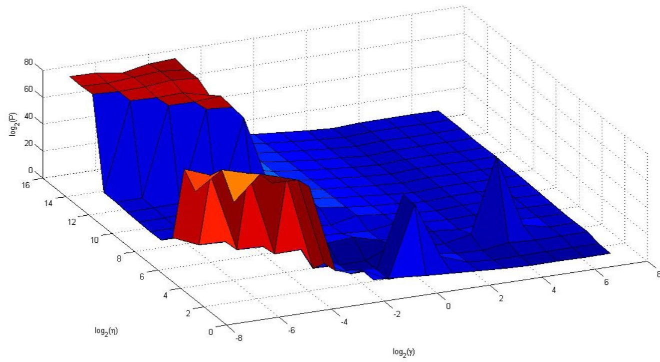

Figure 5.2. Plot of deficiency ratio for two-seller DPFI equilibriums

other metrics of benefits (see e.g. Ke et al. 2016 [83] and the literature review therein), it is possible to re-consider the definition of PoA here such as the ratio of total savings and utility increase for the consumer under different structure of competition.

5.4.2 Numerical Examples for the DT-DPFI Model

In this part, we provide the numerical example of Problem 5.2 and solved by the time stepping Algorithm 5.2. Our parameterization of the problem is an extension of the numerical example recorded by Friesz (2010) [54] into the case that explicitly incorporates season-to-season factors. There are 3 sellers on the market with $s = 1, 2, 3$. the time horizon to consider is $[t_{0,} t_f] = [0, 40]$ in 8 seasons, this means one season is with length $\Upsilon = 5$. The initial fixed inventory endowments are $K_1 = 3000$ and $K_2 = 2000$, and $K_3 = 2500$. Also, take the following parameterization of the function $h_s(\pi_s, \pi_{-s})$:

$$\begin{aligned}
\alpha_1(t) &= 350, \alpha_2(t) = 250 + 10\tau, \alpha_3(t) = 350 - 5\tau \\
\beta_{1,3}(t) &= 2 - 0.2(t - (\tau - 1)\Upsilon), \\
\beta_2(t) &= 2 - 0.2(t - (\tau - 1)\Upsilon) - 0.05\tau \\
\gamma_{1,2,3}(t) &= 1 + 0.1(t - (\tau - 1)\Upsilon)
\end{aligned}$$

Other parameters include:

$$p_{\min} = (50, 30, 45)^T, p_{\max} = (300, 300, 300)^T,$$

$$d_{\min} = (0,0,0)^T, \rho = 0.01.$$

and the parameters for the season-to-season evolution dynamics are:

$$\begin{aligned} \alpha_s &= 0.5; k_{1,s}^{\tau} = (150,200,250)^T, \forall \tau = 1,\cdots,8; \\ k_{2,s}^{\tau} &= (-9,-8,-7)^T, \forall \tau = 1,\cdots,8; \\ \theta^{\tau} &= (28,27,26,25,28,27,26,25)^T, \forall s = 1,2,3. \end{aligned}$$

Again, realizing sellers' decisions should be changing from season to season, it is interesting to plot the total profit, and the ordered inventory change from season to season. To measure the in-season performance of each seller, we also plot the average sold price and the percentage of the inventory that is sold compared to the initial fixed inventory. All these plots are gathered in Figure 5.3. We see that despite of the illustrative nature of our season-to-season inventory dynamics, each seller in this numerical example has the ability to, for example, adjust its inventory ordering decision to make sure that less and less goods are left at the end of each selling season.

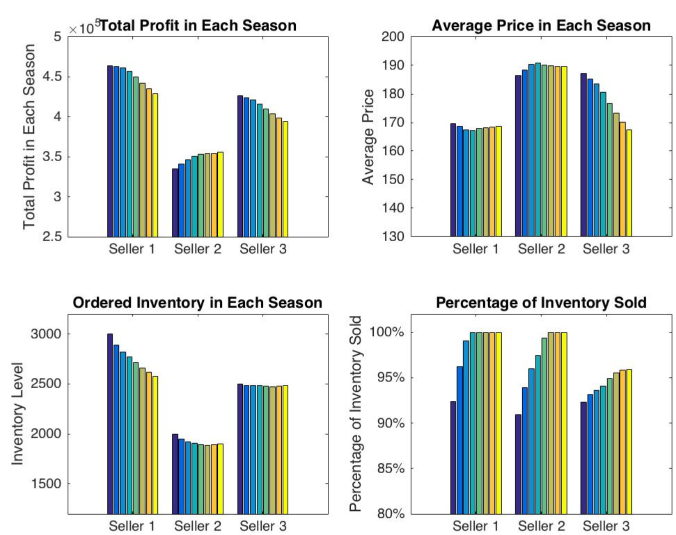

Figure 5.3. Season-to-season numerical outputs for DT-DPFI model

5.5 Summary and Future Work

In this chapter, we first summarize the continuous time DPFI model proposed by Friesz (2010) [54] and then provide in-depth analysis to the problem as a useful and important application of dynamic variational inequality. One of the importance of the DPFI model is that it captures the essence of the competition in dynamic pricing and demand allocation by simple assumptions, and could serve as a basic model for further analysis. This chapter justifies that also by extending the DPFI model into a dual-time-scale (DT-DPFI) model with service capacity replenishment decisions made on a different time scale.

In the future, given the increasing interest in investigating decision mechanisms in data-rich environments, we would like to extend this competitive model towards a data-driven model with which pricing, along with service quality and other product/service features could be incorporated. We already identified a series of such research that takes advantage of information regarding product features, and use such "big-data" to support supply chain management decisions. For example towards the (dynamic) newsvendor problem, Ban and Rudin (2016) [126] proposes Machine Learning and Kernel-weights Optimization, Oroojlooyjadid et al. (2007) [107] applies deep neural network.

Chapter 6
The Fixed Inventory Model with State Dynamics and Feedback Information

In this chapter, we continue the analysis of the Dynamic Pricing with Fixed Inventory (DPFI) model from Chapter 5 and study its equivalent formulation with explicit state dynamics. Following the spirit of Friesz (2010) [54] Section 10.1.2, it is possible to reformulate the fixed inventory constraint as a so-called isoperimetric constraint:

$$\begin{aligned} \frac{dy_s}{dt} &= d_s(t) \\ y_s(t_0) &= 0, y_s(t_f) \leq K_s \end{aligned}$$

note that such reformulation will give rise to an equivalent optimal-control formulation to Problem 5.1, with the state variable y_s. In this chapter, we will provide more qualitative analysis on the DPFI model with the aid of the optimal control formulation.

Remark 6.1. *Similar reformulations have also been noted by Gallego and Hu (2015) [61] with similar state variables proposed as "Inventory" and it is not hard to observe that these two types of dynamics are equivalent mathematically. However, we argue that under the setting of a service problem of this chapter, the physical meaning of this state variable is immediate: at time t, $y_s(t)$ is the cumulated service volume provided by seller s from the beginning of the time horizon t_0. Hence the final time value of such state variable shall be less than or equal to the initial fixed* *endowment of seller s.*

Note that all regularity assumptions of the DPFI model, especially in the sense of Definition 5.1, are inherited in this chapter, this means that $d_s(t) \geq D_{\min} > 0$ for $\forall t \in [t_0, t_f]$ and $y_s(t) \geq 0$ is automatically satisfied. In addition, in order to make the problem well posed, it is necessary to make sure that the cumulated service level is less than the initial endowment throughout the planning horizon:

$$y_s(t) \leq K_s \quad \forall t \in (t_0, t_f).$$

We will carry out more analysis on the state dynamics, it turns out that such analysis will greatly help the construction of necessary conditions.

In addition to the techniques reviewed in Chapter 2 and Chapter 3, special care must be taken towards the analysis since the state variable is constrained. Let us first review the literature on state-constrained optimal control problems. Towards the theory and applications of revised Maximum Principle, and solution of state constrained open loop OCPs, Hartl et al. (1995) [70] collects and reviews most previous results including existence, necessary conditions with the direct and indirect adjoining approaches, this review paper also provides worked-out examples. Later this chapter, and inspired by Gallego and Hu (2015) [61], we will apply the direct adjoint approach towards the analysis of the best response problem of DPFI. Examples of related literature using this method are: Girsanov (1972) [66], Norris (1973) [106], Maurer (1979) [99] and Ioffe and Tihomirov (1979) [73]. Especially, in Section 6.1 and Section 6.2, the optimality conditions will be derived following Chapter 5 of Seierstad and Sydsaeter (1986) [130]. For more details on the indirect adjoint approach, please also refer to Hartl et al. [70] for a description of the theory and more examples in textbooks such as Bryson and Ho (1975) [35] and Pontryagin (1987) [115].

The introduction of the state variable y_s also allows us to investigate the best response and Nash equilibrium of sellers in the DPFI model when each seller knows his and his competitors' cumulated service volume offered. Such analysis will be built on the solution of the Hamilton-Jaccobi-Bellman partial differential equation (HJB-PDE) with constrained state space (Soner (1986a) [133] (1986b) [134]). Those techniques, as well as the synthesis of feedback optimal control, are reviewed and summarized in Schättler (2009) [127]. Recently, Altrarovici et al. (2013) [3]

proposes a general framework for the solution of such problems, and we will employ the numerical package based on their framework for the computation of the best response problem later in this chapter. For more details of the numerical package and a brief review of the Semi-Lagrangian scheme, please refer to Appendix A of this dissertation and the references therein.

The rest of this chapter is structured as follows: Section 6.1 formulates and analyzes the equivalent OCP of Problem 5.1 in both open loop and state feedback case. After that, Section 6.2 analyzes the open-loop DPFI equilibrium utilizing the necessary conditions of DVI with state dynamics; Section 6.3 looks at the feedback equilibrium of the DPFI model and provides algorithm for the computation of dynamic feedback policy of pricing and demand fulfillment. Finally, numerical examples of feedback best response problem and feedback equilibrium will be provided in Section 6.4.

6.1 Best Response Problem as Optimal Control

Similar to Chapter 5, we follow the discussions above and start our analysis from the Best Response Problem (BRP). It is very helpful to re-formulate the BRP of Seller s from a infinite dimensional mathematical program into an optimal control problem with state constraints. And we shall recast this formulation following the general formulation of the state-constrained OCP proposed by Hartl et al. (1995) [70]:

Problem 6.1. *(Best Response Problem as OCP) With each seller aiming at maximizing its own revenue given other seller's price strategy π^*_{-s}, we have the following recast of Problem 5.1 as an OCP:*

$$\max_{\pi_s, d_s} J_s = \int_{t_0}^{t_f} \exp(-\rho t)\pi_s(t)d_s(t)dt$$

subject to

$$\frac{dy_s}{dt} = d_s(t), \quad y_s(t_0) = 0$$
$$n_s(y_s, t) = K_s - y_s(t) \geq 0$$
$$K_s - y_s(t_f) \geq 0$$

$$u_s(t) = (\pi_s(t), d_s(t)) \in U_{0,s}$$

with the set of feasible control defined by pure control constraints:

$$\begin{aligned} U_{0,s} \triangleq \{(\pi_s(t), d_s(t)) : \pi_s(t) - \pi_{\max} \leq 0 \\ \pi_{\min} - \pi_s(t) \leq 0 \\ D_{\min} - d_s(t) \leq 0 \\ d_s(t) - h_s(\pi_s, \pi^*_{-s}) \leq 0\} \end{aligned}$$

We can make two observations toward Problem 6.1: (1) this problem includes pure state constraints and pure control constraints, and does *not* include any mixed-state-control constraints; (2) the state dynamics is linearized, especially it is linear in control. Hence according to the references reviewed at the beginning of this chapter, Problem 6.1 belongs to the "normal case", which means its Hamiltonian and Lagrangian can be formed as follows:

$$\begin{aligned} H_s &= \exp(-\rho t)\pi_s(t)d_s(t) + \lambda_s d_s(t) \\ L_s &= H_s + \nu_s n_s = \exp(-\rho t)\pi_s(t)d_s(t) + \lambda d_s(t) + \nu_s(K_s - y_s(t)) \end{aligned}$$

And as an embodiment of Theorem 4.1 in Hartl et al. (1995) [70], in particular, see Chapter 5 of Seierstad and Sydsaeter (1986) [130], we have the following lemma as the necessary condition of Problem 6.1:

Lemma 6.1. *(Necessary condition of Problem 6.1 as a general state constrained problem) Assume Problem 6.1 is regular in the sense of Definition 5.1 and let* $(\pi^*_s(t), d^*_s(t), y^*_s(t))$ *be an optimal state-control pair such that* $u^*_s = (\pi^*_s(t), d^*_s(t))$ *is right continuous with left-hand limits. Assume* $y^*_s(t)$ *has finitely many junctions times. Then there exists an adjoint trajectory* $\lambda_s(t)$*, a piecewise continuous multiplier* $\nu_s(t)$*, a* $\eta_s(\tau)$ *for the point of discontinuity of* $\lambda_s(t)$*. In addition to state dynamics, state constraints and control constraints, the following necessary conditions should hold for every* $t \in [t_0, t_f]$*:*
(i) maximum principle:

$$u^*_s = (\pi^*_s(t), d^*_s(t)) = \arg\max_{u \in U_0} H_s(y^*_s, u_s, \lambda^*_s) \tag{6.1}$$

(ii) adjoint dynamics

$$\frac{d\lambda_s(t)}{dt} = -\frac{\partial L_s}{\partial y_s} = \nu_s(t) \tag{6.2}$$

with

$$\nu_s(t) \geq 0, \quad \nu_s(t)n_s(t) = 0 \tag{6.3}$$

(iii) transversality conditions:

$$\lambda_s(t_f^-) = -\gamma_s \tag{6.4}$$

where

$$\gamma_s \geq 0, \quad \gamma_s(K_s - y_s(t_f)) = 0 \tag{6.5}$$

(iv) contact time conditions: λ_s may contain jump discontinuities for (any) contact time τ:

$$\lambda_s(\tau^-) = \lambda_s(\tau^+) - \eta_s(\tau) \tag{6.6}$$

where

$$\eta_s(\tau) \geq 0, \quad \eta(\tau)(K_s - y_s(\tau)) = 0 \tag{6.7}$$

Notice that the above Lemma only takes in the knowledge of Problem 6.1 as a generic OCP with state constraints, and is only the starting point of our qualitative analysis. Firstly, for a given time t, the maximum principle becomes a finite-dimensional nonlinear program parameterized by $\lambda_s(t)$ and $\pi^*_{-s}(t)$, we write down its equivalent minimization problem:

$$\min_{\pi_s, d_s} - \exp(-\rho t)\pi_s d_s - \lambda_s d_s$$

such that:

$$\begin{aligned}
\pi_s(t) - \pi_{\max} &\leq 0 \quad (\alpha_1) \\
\pi_{\min} - \pi_s(t) &\leq 0 \quad (\alpha_2) \\
D_{\min} - d_s(t) &\leq 0 \quad (\beta_1) \\
d_s(t) - h_s(\pi_s, \pi^*_{-s}) &\leq 0 \quad (\beta_2)
\end{aligned}$$

Note that the last constraints, should be an equality constraint due to the following lemma:

Lemma 6.2. *(Demand constraint binds) For every competitor $s \in \mathcal{S}$, given her* *competitors' strategies π_{-s}, a solution (π_s^*, d_s^*) of a best response problem having control as $(\pi_s(t), d_s(t)) \in U_0$ and whose criterion is monotonically increasing in own price satisfies*

$$d_s^* - h_s\left[\pi_s^*, \pi_{-s}^*\right] = 0 \quad (\beta_2) \tag{6.8}$$

General nonlinear programming KKT condition as in Theorem 2.2 yields:

$$\begin{aligned}
-\exp(-\rho t)d_s + \alpha_{1s} - \alpha_{2s} - \beta_{2s}(\frac{\partial h_s}{\partial \pi_s}) &= 0 \\
-\exp(-\rho t)\pi_s - \lambda_s - \beta_{1s} + \beta_{2s} &= 0 \\
\alpha_{1s}(\pi_s - \pi_{\max}) &= 0, \quad \alpha_{1s} \geq 0 \\
\alpha_{2s}(\pi_{\min} - \pi_s) &= 0, \quad \alpha_{2s} \geq 0 \\
\beta_{1s}(D_{\min} - d_s) &= 0, \quad \beta_{1s} \geq 0 \\
\beta_{2s}(d_s - h_s(\pi_s, \pi_{-s}^*)) &= 0,
\end{aligned}$$

In fact, we can use such system to solve for the optimal price and demand fulfillment policy parameterized by $\lambda_s(t)$ and $\pi_{-s}^*(t)$, and the price and fulfilled demand policies of Seller s can be expressed as a combination of dynamic dual variables:

$$\begin{aligned}
d_s(t) &= \exp(\rho t)[\alpha_{1s} - \alpha_{2s} - \beta_{2s}(\frac{\partial h_s}{\partial \pi_s})] \\
\pi_s(t) &= \exp(\rho t)[-\lambda_s - \beta_{1s} + \beta_{2s}]
\end{aligned}$$

In general, different cases should be examined according to whether each control constraint is binding or not. Due to the limited space here, we take one special case:

Example 6.1. *(Maximum Principle with more details) Consider the case when (i) the observed demand function takes on the linear form as in Chapter 5:$h(\pi_s, \pi_{-s}) = a_s - b_s\pi_s + \sum_{r\neq s}\gamma_{sr}\pi_r^*$; (ii) observed demand are fulfilled; and (iii) price is strictly interior, which leads to:*

$$\alpha_{1s} = 0, \alpha_{2s} = 0, \beta_{1s} = 0, \frac{\partial h_s}{\partial \pi_s} = -b_s$$

hence the following system, parameterized by λ_s and π_{-s}^*, holds:

$$d_s(t) \quad = \quad \exp(\rho t)[\beta_{2s} b_s]$$

$$\begin{aligned}\pi_s(t) &= \exp(\rho t)[-\lambda_s + \beta_{2s}] \\ d_s(t) &= a_s - b_s \pi_s(t) + \sum_{r \neq s} \gamma_{sr} \pi_r^*\end{aligned}$$

which leads to:

$$\begin{aligned}d_s(t) &= \frac{1}{2}\left[a_s + \sum_{r \neq s} \gamma_{sr} \pi_r^* + b_s \lambda_s \exp(\rho t)\right] \\ \pi_s(t) &= \frac{1}{2}\left[\frac{a_s + \sum_{r \neq s} \gamma_{sr} \pi_r^*}{b_s} - \lambda_s \exp(\rho t)\right]\end{aligned}$$

Then, let us get back to the time-evolution of the state variable $y_s(t)$ and adjoint variable $\lambda_s(t)$. We begin by explaining what a contact time is, following Hartl et al. (1996), a contact time τ is defined as the time when the state trajectory "just touches the boundary" of its feasible region, and is "in the interior just before". Under the setting of Problem 6.1, it is the time when Seller s has just run out of capacity. The following lemma describes important observations regarding the contact time.

Lemma 6.3. *(Contact Time) Assume Problem 6.1 is regular in the sense of Definition 5.1, then it admits at most one contact time τ. And if τ exists, $\tau = t_f$, leading to $y_s(t) < K_s$, $\forall t \in [t_0, t_f)$.*

Proof. Assume there exists multiple contact times, τ_1 and τ_2, then $y_s(\tau_1) = y_s(\tau_2) = K_s$. This is contradictory to the following observation: since $d_s(t) \geq D_{\min} > 0$ for $\forall t \in [t_0, t_f]$, the state trajectory $y_s(t)$ must strictly increasing. Then Problem 6.1 admits at most one contact time τ. Similarly, if $\tau < t_f$, then it must be that $y_s(t) > y_s(\tau)$ for some $t \in (\tau, t_f]$, which contradicts the feasibility of such a state trajectory. Hence if τ exists, $\tau = t_f$.

Finally, this lead to $y_s(t) < K_s$, $\forall t \in [t_0, t_f)$. □

This lemma is illustrated by Figure 6.1, where only the red and blue lines are possible state trajectories. The purple trajectory is infeasible due to Definition 5.1. Note that this is one of the basic assumptions of this chapter that is significantly different from Gallego and Hu (2014) [61], in which the decision variables for seller s is only price over time, and if a seller runs out-of-stock, a choke price, generating zero observed demand, must be posted.

Now we analyze the two possible cases:

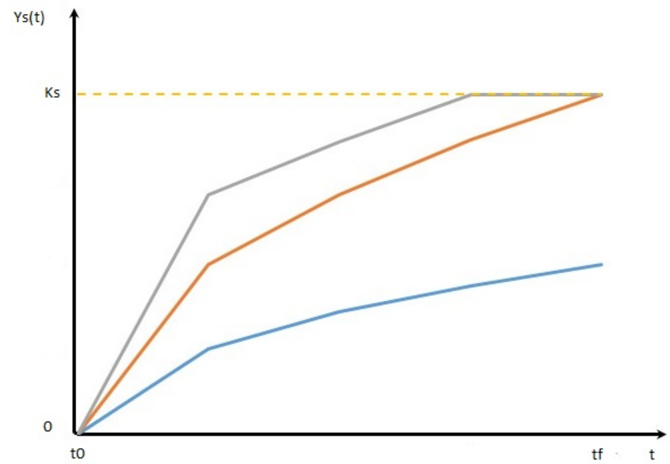

Figure 6.1. Illustration of $y_s(t)$. Only the cases with red and blue lines are possible

1. (illustrated with the blue line, $y_s(t) < K_s$, $t \in [t_0, t_f]$) In this case, $n_s(t) = K_s - y_s(t) > 0$, hence from (6.3) we know $\nu_s(t) = 0$. Moreover, from the transversality condition (6.4) and (6.5) we have $\gamma_s = 0$. Then back to (6.2) we know $\frac{d\lambda_s(t)}{dt} = 0$, $\lambda_s(t_f) = 0$, $\forall t \in [t_0, t_f]$.

2. (illustrated with the with red line, $y_s(t) < K_s$, $t \in [t_0, t_f), y_s(t_f) = K_s$) In this case $\tau = t_f$, which is the only discontinuity for adjoint variable λ_s. Similar to the previous case, $\frac{d\lambda_s(t)}{dt} = 0$ for $t \in [t_0, t_f)$. At time t_f, the transversality conditions (6.4) and (6.5) state that $\lambda_s(t) = -\gamma_s$ for $t \in [t_0, t_f^-)$. From the contact time condition (6.6) and (6.7) $\lambda_s(t_f) = -\gamma_s + \eta_s$.

The intuitive implication of the adjoint variable is clear: a non-trivial shadow price will be present if the initial fixed endowment is exhausted at the end of the selling horizon. Such shadow price is constant throughout the horizon except by the end. The following theorem summarizes the analysis above, and provides insights about a Seller's best response policies with a series of dual variables:

Theorem 6.1. *(Necessary Condition, More specific) Assume Problem 6.1 is regular in the sense of Definition 5.1 and let $(\pi_s^*(t), d_s^*(t), y_s^*(t))$ be an optimal state-control pair such that $u_s^* = (\pi_s^*(t), d_s^*(t))$ is right continuos with left-hand limits. Then there exists an adjoint trajectory $\lambda_s(t)$, a tuple of piecewise continuos multiplier $(\alpha_{1s}(t), \alpha_{2s}(t), \beta_{1s}(t), \beta_{2s}(t))$ and a couple of non-negative constants γ_s, η_s. The*

following necessary condition in addition to state dynamics, state constraints and control constraints should hold for every $t \in [t_0, t_f]$.
(i) maximum principle:

$$
\begin{aligned}
d_s(t) &= \exp(\rho t)[\alpha_{1s}(t) - \alpha_{2s}(t) - \beta_{2s}(t)(\frac{\partial h_s}{\partial \pi_s})] \\
\pi_s(t) &= \exp(\rho t)[-\lambda_s(t) - \beta_{1s}(t) + \beta_{2s}(t)]
\end{aligned}
$$

with

$$
\begin{aligned}
\alpha_{1s}(\pi_s - \pi_{\max}) &= 0, \quad \alpha_{1s} \geq 0 \\
\alpha_{2s}(\pi_{\min} - \pi_s) &= 0, \quad \alpha_{2s} \geq 0 \\
\beta_{1s}(D_{\min} - d_s) &= 0, \quad \beta_{1s} \geq 0 \\
\beta_{2s}(d_s - h_s(\pi_s, \pi^*_{-s})) &= 0,
\end{aligned}
$$

(ii) adjoint dynamics

$$
\frac{d\lambda_s(t)}{dt} = 0, \quad t \in [t_0, t_f) \tag{6.9}
$$

(iii) transversality and contact time conditions:

$$
\lambda_s(t_f^-) = -\gamma_s, \quad \gamma_s \geq 0, \quad \gamma_s(K_s - y_s(t_f)) = 0 \tag{6.10}
$$

$$
\lambda_s(t_f^-) = \lambda_s(t_f) - \eta_s \tag{6.11}
$$

where

$$
\eta_s \geq 0, \quad \eta_s(K_s - y_s(t_f)) = 0 \tag{6.12}
$$

Nevertheless, it is interesting to look at how the above theorem works with the special case recorded in Chapter 10 of Friesz (2010) [54]:

Corollary 6.1. *("Sold up" case) If, in addition to the assumptions in Theorem 6.4, we assume Seller s must sold out the fixed inventory, then the optimal state-control pair can be described by similar* $(\alpha_{1s}(t), \alpha_{2s}(t), \beta_{1s}(t), \beta_{2s}(t)), \gamma_s, \eta_s$ *and:*

$$
\begin{aligned}
d_s(t) &= \exp(\rho t)[\alpha_{1s}(t) - \alpha_{2s}(t) - \beta_{2s}(t)(\frac{\partial h_s}{\partial \pi_s})] \\
\pi_s(t) &= \exp(\rho t)[-\lambda_s(t) - \beta_{1s}(t) + \beta_{2s}(t)]
\end{aligned}
$$

with

$$\begin{aligned}
\alpha_{1s}(\pi_s - \pi_{\max}) &= 0, \quad \alpha_{1s} \geq 0 \\
\alpha_{2s}(\pi_{\min} - \pi_s) &= 0, \quad \alpha_{2s} \geq 0 \\
\beta_{1s}(D_{\min} - d_s) &= 0, \quad \beta_{1s} \geq 0 \\
\beta_{2s}(d_s - h_s(\pi_s, \pi^*_{-s})) &= 0,
\end{aligned}$$

and

$$\begin{aligned}
\lambda_s(t) &= -\gamma_s \quad t \in [t_0, t_f) \\
\lambda_s(t_f^-) &= -\gamma_s \\
\lambda_s(t_f) &= -\gamma_s + \eta_s
\end{aligned}$$

As discussed in the opening remarks of this chapter, the introduction of the state variable y_s also allows us to investigate the best response problem of sellers in the DPFI model when each seller knows his cumulated service volume offered. In such a case, we employ the methods reviewed in Chapter 3 for its analysis. Let us start by considering the value function of Seller s's BRP:

Definition 6.1. *(Value function of the BRP) The value function of each Seller's best response problem is defined as the maximum remaining revenue-to-gain at time t with cumulated service level y_s, hence $V(y_s, t) : [0, K_s] \times [t_0, t_f] \to \mathbb{R}$ with:*

$$V_s(y_s, t) = \sup_{u \in U_{0,s}} \left[\int_t^{t_f} \exp[-\rho(\chi - t)] \pi_s(\chi) d_s(\chi) d\chi \right]$$

Then as a direct application of Theorem 3.2 and Theorem 3.4, we take advantage of the following theorem to state the HJB-PDE for Seller s's best response problem:

Theorem 6.2. *(HJB-PDE of BRP) Let the value function of each seller's BRP follows Definition 6.1, and let the optimal control problem be regular in the sense of Definition 5.1, then the value function solves the following HJB-PDE:*

$$\frac{\partial V_s}{\partial t} + \max_{u \in U_{0,s}} \left[\frac{\partial V_s}{\partial y_s} d_s(t) + \pi_s(t) d_s(t) \right] = 0$$

with boundary condition:

$$V(y_s(t_f), t_f) = 0 \quad \textit{whenever} \;\; y_s(t_f) \leq K_s$$

and it generates Seller s's feedback best response strategy through the maximization of the RHS of this HJB-PDE.

Then similar to Algorithm 3.1, we list the algorithm for the synthesis of feedback optimal pricing and demand fulfillment strategies. In this algorithm, we take advantage of two properties of Problem 6.1: (1) Lemma 6.2 tells us the observed demand constraint should be binding; (2) we compute the optimal control and recover the optimal state at the same time instance, avoiding the solution of (6.13) on all mesh points.

Algorithm 6.1. *(Feedback Best Response Policy Synthesis) Consider Seller s's best response Problem 6.1 in feedback and a finite time horizon $t \in [t_0, t_f]$, on appropriate state and time mesh: $\Delta t = \frac{t_f - t_0}{N_{time}}, \Delta y = \frac{K_s}{N_{state}}$*

Step 1. *Solve HJB-PDE* *Solve the HJB-PDE subject to the boundary condition following Theorem 6.2, thereby compute the value function $V_s(y_s, t)$ on given mesh points. Set $k = 0$.*

Step 2. *Feedback control synthesis* *For $k \leq N_{time}$, $t_k = t_0 + k\Delta t$, use the value function $V(y_s, t)$ from step 1 to design an optimal feedback control $u^*(\cdot) = (\pi_s^*, d_s^*)$, as follows: for each point y and each time $t \leq t_f$, solve*

$$\frac{\partial V(y_s, t_k)}{\partial y_s} h_s(\pi_s^*) + \exp(-\rho t) h_s(\pi_s^*) \pi_s^* + \frac{\partial V(y_s, t_k)}{\partial t} = 0 \tag{6.13}$$

for π_s^*, and define $u^*(y_s, t) = (\pi_s^*, d_s^*)$ accordingly.

Step 3. *Define the feedback control-state-time trajectory* *Set the next state:*

$$y_s^*(t_{k+1}) = y_s^*(t_k) + \Delta t d_s^*(y_s^*(t), t_k) \tag{6.14}$$

and set $k = k + 1$, hence the optimal trajectory becomes

$$u^* \triangleq u(y_s^*(t), t) \tag{6.15}$$

Remark 6.2. *Note that in Step 2, the solution of (6.13) is in fact a 1-dimensional root finding problem. The reason is the following: we know that with $V(y_s, t)$ known on the given mesh points (from Step 1), $\frac{\partial V}{\partial y}, \frac{\partial V}{\partial t}$ could be approximated by standard numerical procedures such as Lagrangian interpolation, also the functional form of h_s is given, the RHS of (6.13) at this instance of time is only a function of π_s*

We will provide numerical examples utilizing this algorithm later in Section 6.4. For now, let us move on towards the analysis of the equilibrium conditions for the DPFI problem.

6.2 Open Loop Equilibrium Revisited

In Chapter 5, Theorem 5.2 establishes the equivalent DVI of the DPFI generalized Nash equilibrium, the DVI from there are based on infinite dimensional mathematical programming. In this section, we give another equivalent DVI with explicit state dynamics (and constraints) in the spirit of Problem 2.8. First consider the multi-dimensional state variable $y(t) = (y_1(t), \cdots, y_{|\mathcal{S}|}(t))^T$, and a vector of fixed endowments: $K = (K_1, \cdots, K_{|\mathcal{S}|})^T$, then similarly we denote $\pi(t) = (\pi_1(t), \cdots, \pi_{|\mathcal{S}|}(t))^T$, $d(t) = (d_1(t), \cdots, d_{|\mathcal{S}|}(t))^T$, for each Seller s, $(\pi_s, d_s) \in U_{0,s}(\pi_{-s}, \xi_s)$. Here we let

$$U(\xi) \triangleq \prod_{s \in \mathcal{S}} U_{0,s}(\pi_{-s}, \xi_s)$$

be the feasible set of the market. For a known and given set of parameters ξ, let $U(\xi) \triangleq U$. The following theorem is the equivalent reformulation of Theorem 5.2 with explicit state variables:

Theorem 6.3. *(DVI of the DPFI equilibrium with state dynamics) The generalized Nash equilibrium among sellers $s \in \mathcal{S}$ that is described by Definition 5.2, is equivalent to the solution of the following differential variational inequality:*

$$\left.\begin{array}{c} \text{find } (\pi^*, d^*) \in U \text{ such that} \\ \sum_{s \in \mathcal{S}} \int_{t_0}^{t_f} \exp(-\rho t) \left[d_s^* \cdot (\pi_s - \pi_s^*) + \pi_s^* \cdot (d_s - d_s^*) \right] dt \leq 0 \quad \forall\ (\pi, d) \in U \end{array}\right\} \quad (6.16)$$

when the regularity conditions in Definition 5.1 hold, with the following state *dynamics:*

$$\frac{dy}{dt} = d(t), \quad y(t_0) = 0$$
$$n(y,t) = K - y(t) \geq 0$$
$$K - y(t_f) \geq 0$$

Inspired by Friesz (2010) [54], in the pursuit of necessary conditions of DVI (6.16), it is helpful to formulate and examine the following auxiliary optimal control problem which is parameterized by (π^*, d^*):

$$\max_{\pi,d} J^* = \sum_{s\in\mathcal{S}} \int_{t_0}^{t_f} \exp(-\rho t)[\pi_s^*(t)d_s(t) + d_s^*(t)\pi_s(t)]dt$$

subject to

$$\frac{dy}{dt} = d(t), \quad y(t_0) = 0$$
$$n(y,t) = K - y(t) \geq 0$$
$$K - y(t_f) \geq 0$$
$$u(t) = (\pi(t), d(t)) \in U$$

with the set of feasible control defined by pure control constraints:

$$U = \Big\{(\pi_1(t), \cdots, \pi_{|\mathcal{S}|}(t); d_1(t), \cdots, d_{|\mathcal{S}|}(t)) :$$
$$\pi_s(t) - \pi_{\max} \leq 0$$
$$\pi_{\min} - \pi_s(t) \leq 0$$
$$D_{\min} - d_s(t) \leq 0$$
$$d_s(t) - h_s(\pi_1(t), \cdots, \pi_{|\mathcal{S}|}(t)) \leq 0\Big\}$$

This is a multi-dimensional optimal control problem with state constraints, and we will apply similar methodologies as in last section. Consider its Hamiltonian:

$$H = \sum_s \exp(-\rho t)[\pi_s^*(t)d_s(t) + d_s^*(t)\pi_s(t)] + \sum_s \lambda_s d_s(t)$$

and Lagrangian:

$$L = H + \sum_s \nu_s(K_s - y_s(t))$$

The maximum principle at each instant t can be analyzed by KKT, resulting in vector notation:

$$\begin{aligned} -\exp(-\rho t)\pi^* - \lambda - \beta_1 + \beta_2 &= 0 \\ -\exp(-\rho t)d^* + \alpha_1 - \alpha_2 - M\beta_2 &= 0 \end{aligned}$$

with M being a $|\mathcal{S}|$-by-$|\mathcal{S}|$ matrix containing partial derivatives of the observed demand functions, $M_{ij} \triangleq \frac{\partial h_i}{\partial \pi_j}$, following the regularities in Definition 5.1, we know that $M_{ii} = \frac{\partial h_i}{\partial \pi_i} < 0$, $M_{ij} = \frac{\partial h_i}{\partial \pi_j} > 0$ when $i \neq j$. Also apply Lemma 6.2, the CSCs are:

$$\begin{aligned} \alpha_{1,s}(\pi_s - \pi_{\max}) &= 0, \quad \alpha_{1,s} \geq 0 \\ \alpha_{2,s}(\pi_{\min} - \pi_s) &= 0, \quad \alpha_{2,s} \geq 0 \\ \beta_{1,s}(D_{\min} - d_s) &\leq 0, \quad \beta_{1,s} \geq 0 \\ d_s(t) - h(\pi_s, \pi_{-s}) &= 0, \quad \beta_{2,s} \text{ free} \end{aligned}$$

For the adjoint equation, initially we have:

$$\begin{aligned} \frac{d\lambda}{dt} &= \nu(t), \\ \nu(t)(K - y(t)) &\geq 0, \quad \nu(t) \geq 0 \end{aligned}$$

applying Lemma 6.3 for each seller s we know that $\nu_s(t) = 0$ for $t \in [t_0, t_f)$, and similar contact-time transversality conditions indicates:

$$\lambda_s(t_f^-) = -\gamma_s, \quad \gamma_s \geq 0, \quad \gamma_s(K_s - y_s(t_f)) = 0 \tag{6.17}$$

$$\lambda_s(t_f^-) = \lambda_s(t_f) - \eta_s \tag{6.18}$$

where

$$\eta_s \geq 0, \quad \eta_s(K_s - y_s(t_f)) = 0 \tag{6.19}$$

Summing up the analysis above, we have:

Theorem 6.4. *(DVI necessary conditions for DPFI) Assume the DPFI equilib-*

rium described by Definition 5.2 is regular in the sense of Definition 5.1, and let $(\pi^*(t), d^*(t), y^*(t))$ be an equilibrium state-control tuple. Then there exists an adjoint trajectory $\lambda(t)$, a tuple of piecewise continuous multiplier $(\alpha_1(t), \alpha_2(t), \beta_1(t), \beta_2(t))$ and vectors of non-negative constants γ, η. The following necessary condition in addition to state dynamics, state constraints and control constraints should hold for every $t \in [t_0, t_f]$.

(i)maximum principle:

$$\begin{aligned} d^*(t) &= \exp(\rho t)[\alpha_1(t) - \alpha_2(t) - M\beta_2(t)] \\ \pi^*(t) &= \exp(\rho t)[-\lambda(t) - \beta_1(t) + \beta_2(t)] \end{aligned}$$

with and for each $s \in \{1, 2, \cdots, |\mathcal{S}|\}$:

$$\begin{aligned} \alpha_{1,s}(\pi_s - \pi_{\max}) &= 0, \quad \alpha_{1,s} \geq 0 \\ \alpha_{2,s}(\pi_{\min} - \pi_s) &= 0, \quad \alpha_{2,s} \geq 0 \\ \beta_{1,s}(D_{\min} - d_s) &\leq 0, \quad \beta_{1,s} \geq 0 \\ d_s(t) - h(\pi_s, \pi_{-s}) &= 0, \quad \beta_{2,s} \textit{ free} \end{aligned}$$

(ii)adjoint dynamics

$$\frac{d\lambda(t)}{dt} = 0, \quad t \in [t_0, t_f) \tag{6.20}$$

(iii)transversality and contact time conditions:

$$\lambda(t_f^-) = -\gamma, \quad \gamma \geq 0, \quad \gamma^T(K - y(t_f)) = 0 \tag{6.21}$$

$$\lambda(t_f^-) = \lambda(t_f) - \eta \tag{6.22}$$

where

$$\eta \geq 0, \quad \eta(K - y(t_f)) = 0 \tag{6.23}$$

6.3 The Feedback Equilibrium

In the competitive case with feedback information structure, the feedback equilibrium is obtained when each player faces a feedback optimal control problem, still parameterized by other players' strategies. The following theorem is a direct application of Theorem 3.6 (Theorem 4.1 of Dockner (2000) [41]), which states the necessary condition of a differential Nash equilibrium with feedback can be expressed as a system of HJB-PDEs along with other conditions:

Theorem 6.5. *(Feedback DPFI Equilibrium) For all* $s \in \{1, 2, \cdots, |\mathcal{S}|\}$*, let* (π^*, d^*) *be a given tuple of functions with* $d_s : [0, K_s] \times [t_0, t_f] \to \mathbb{R}, \pi_s : [0, K_s] \times [t_0, t_f] \to \mathbb{R}$*, and assume the following holds:*
(i) the state equation as initial value problem has a unique solution:

$$\frac{dy}{dt} = d^*(t), \quad y(t_0) = 0, \; y(t) \leq K$$

(ii) There exists a tuple of continuously differentiable value function $V_s(y_s, t)$ *such that the HJB-PDEs are satisfied for all* $s \in \{1, 2, \cdots, |\mathcal{S}|\}$*:*

$$\frac{\partial V_s(y_s, t)}{\partial t} + \max_{u \in U} \left\{ d_s^*(t) \pi_s^*(t) + \frac{\partial V_s(y_s, t)}{\partial y_s} d_s^*(t) \right\} = 0 \tag{6.24}$$

with boundary condition:

$$V_s(y_s, t_f) = 0 \text{ whenever } y_s(t_f) \leq K_s$$

then (π^*, d^*) *is a feedback DPFI equilibrium.*

From a computational perspective, the solution of the system of HJB-PDEs is extremely difficult for a problem that is not linear quadratic. Hence this chapter, inspired by Perakis and Sood (2006) [114], we propose the following algorithm for the solution of Feedback DPFI equilibrium:

Algorithm 6.2. *(DPFI Feedback) Initialize a set of feasible policies for each seller: For all* $s \in \{1, 2, \cdots, |\mathcal{S}|\}, d_s^0 = d_{s,initial}\ \pi_s^0 = \pi_{s,initial}$. *Set counter* $k = 1$*, a maximum number of iterations* C *and a stopping criterion* ϵ*.*

Step 1. *Solve BRP with feedback* *For all* $s \in \{1, 2, \cdots, |\mathcal{S}|\}$*, use Algorithm 6.1 to obtain* (π_s^k, d_s^k) *by solving Seller s's Feedback BRP parameterized by* π_{-s}^{k-1}*.*

Step 2. *Convergence check* *If* $\left\| (\pi_s^k, d_s^k) - (\pi_s^{k-1}, d_s^{k-1}) \right\| \leq \epsilon$*, for all* $s \in \{1, 2, \cdots, |\mathcal{S}|\}$*, set* $(\pi_s^*, d_s^*) = (\pi_s^k, d_s^k)$*, break; otherwise set* $k = k + 1$ *and go back to Step 1.*

Remark 6.3. *We notice that in fact Algorithm 6.2 should be named as a heuristics. It simulates an iterative decision process of the Sellers which is very close to the notion of "fictitious play" (see Brown (1951) [34] and Robinson (1951) [120]). In fictitious play each agent's response is based on the opponents playing a stationary strategy. In the context of static and multi-stage Bertrand competition, the tâtonnement process described in Vives (2001) [142] also inspires Algorithm 6.2. In general, some regularities that lead to convergence will be an interesting research direction.*

On the other hand, in the theorem that follows, we see that despite of the fact that convergence is not guaranteed in general, a result from Algorithm 6.2 that is converging will lead to the NE with feedback.

Theorem 6.6. *(Convergence) In Algorithm 6.2, if policies of all sellers from two consecutive iterations are equal, namely for some k,* $(\pi_s^k, d_s^k) = (\pi_s^{k+1}, d_s^{k+1}),\ \forall s \in \mathcal{S}$, *then the set of policies* (π^k, d^k) *is satisfies the feedback Nash equilibrium conditions as in Theorem 6.5.*

Proof. If $(\pi_s^k, d_s^k) = (\pi_s^{k+1}, d_s^{k+1}),\ \forall s \in \mathcal{S}$, then for any Seller s, substituting (π_s^k, d_s^k) back into the HJB-PDE system as in equation (6.24) leads to:

$$\frac{\partial V_s(y_s^k, t)}{\partial t} + \max_{u \in U} \left\{ d_s^k(t)\pi_s^k(t) + \frac{\partial V_s(y_s^k, t)}{\partial y_s^k} d_s^k(t) \right\} = 0 \qquad (6.25)$$

This indicates that for each Seller s, the feedback optimal condition is met with his problem parameterized by other seller's policies, and in turn means that (π^k, d^k) is a set of equilibrium dynamic pricing and demand fulfillment policies. □

Note that this theorem ensures that an equilibrium must be attained once Algorithm 6.2 converges. In the numerical experiments in Section 6.4, we will see that such convergence is rapid. The technique employed in the proof above is one that commonly used in the literature of diagonalization methods for solving variational inequality problems derived from equilibrium models, see, for example, Theorem 2 of Abudulaal and Leblanc (1979) [1].

6.4 Numerical Examples

In addition to the computation towards open loop equilibrium strategies, let us consider computation towards feedback solutions. We employ ROC-HJ software for the numerical computation of HJB-PDE solutions, please see details in Appendix A. All other parts of the numerical example, including the feedback control synthesis, are prepared in Matlab 2015b on a laptop computer equipped with Intel (R) Core i5 (TM) processor and 8GB RAM.

6.4.1 Numerical Examples for Best Response Problem with Feedback

We start with the case of a best response problem, namely, we seek numerical solution of an embodiment of Problem 6.1 with feedback information structure.

Example 6.2. *(Feedback Best Response Problem with linear observed demand) Seller 1 on the market with two sellers* $s = 1, 2$, *the time horizon to consider is* $[t_0, t_f] = [0, 10]$. *The initial fixed inventory endowment is* $K_1 = 6000$, *and we take the following parameterization:*

$$\begin{aligned} \alpha_1(t) &= 300 + 12t,\ \beta_1(t) = 1 - 0.04t,\ \gamma_{21}(t) = 0.6 + 0.03t \\ p_{\min} &= 50,\ p_{\max} = 300,\ d_{\min} = 1,\ \rho = 0.01. \end{aligned}$$

Assume Seller 1 faces Seller 2's pricing strategy at $\pi_2^*(t) = 75,\ \forall t \in [t_0, t_f]$. *Figure 6.2 shows the value function of this BRP. The HJB-PDE is solved with* $N_t = 100$ *discrete times steps and the state space is discretized in* $N_x = 6000$.

Furthermore, following Algorithm 6.1 we can synthesis the feedback optimal control. Figure 6.3 shows the optimal feedback pricing, demand fulfillment strategies of Seller 1's and the corresponding state variable. Note that the initial fixed endowment is not exhausted at the end of time horizon.

In a second numerical example, we consider the case with the multinomial logit (MNL) function for observed demand. The MNL function is another widely applied demand function in the literature of revenue management. Please refer to the following literature for more discussions of this class of demand function: Talluri and Van Ryzin (2006) [136] on general review, Berry (1994) [18] for parameter

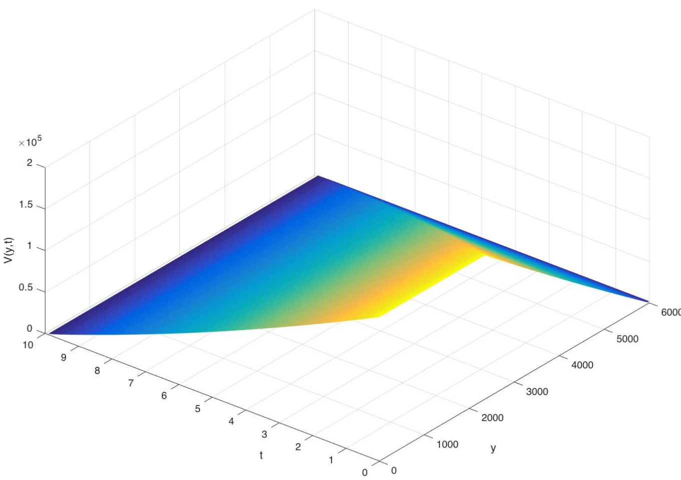

Figure 6.2. The value function $V_1(y_1, t)$ of Seller 1's best response problem facing a given pricing policy from Seller 2 $\pi_2(t) = 75$.

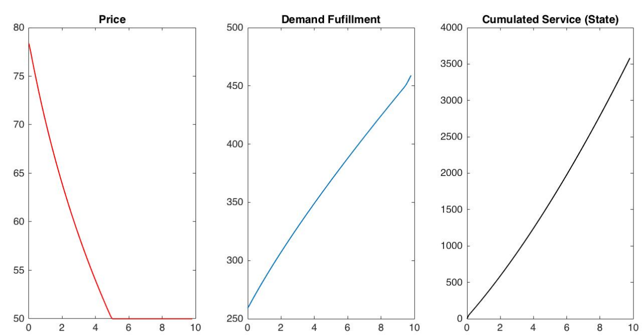

Figure 6.3. Optimal feedback control policy and corresponding state trajectory for Seller 1 with linear demand function, with Seller 2's pricing $\pi_2^*(t) = 75$

estimation under environment of oligopolistic competition. In practice, another way of building a MNL demand function is to take advantage of its relationship with customer-based measurements such as willingness-to-pay (WTP). See Bodea and Ferguson (2014) [27] for instructions on such parameter estimation with industry level datasets.

In this example, the MNL observed demand function for Seller s takes the form by Talluri and Van Ryzin (2006) [136]:

$$h_s(\pi_s, \pi_{-s}, t) = M(t)\frac{e^{-b_s(t)\pi_s(t)}}{1 + \sum_s e^{-b_s(t)\pi_s(t)}} \tag{6.26}$$

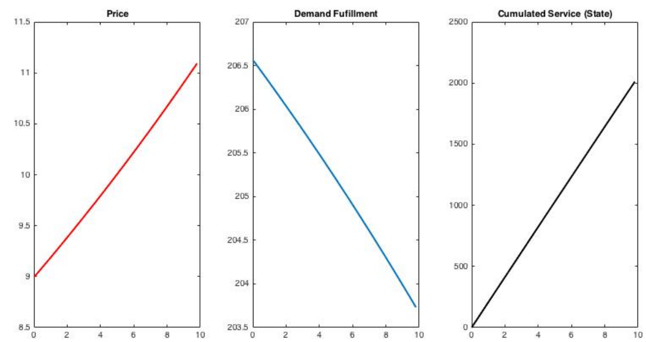

Figure 6.4. Optimal feedback control policy and corresponding state trajectory for Seller 1 with MNL demand function, with Seller 2's pricing $\pi_2^*(t) = 15$

As a BRP, still we take the perspective of Seller 1 and assume Seller 2, who posts a constant pricing strategy $\pi_2^*(t) = 15, \ \forall t \in [t_0, t_f]$, is his only competitor, the initial fixed inventory endowment is $K_1 = 3000$. Other parameters include:

$$
\begin{aligned}
b_1(t) &= 0.01, b_2(t) = 0.02, M(t) = 600 \\
p_{\min} &= 5, p_{\max} = 20, d_{\min} = 1, \rho = 0.01.
\end{aligned}
$$

Example 6.3. *(Feedback Best Response Problem with MNL observed demand) Similar to Figure 6.3, we can come up with the dynamic optimal feedback control policy as in Figure 6.4. In this example, however, we carry out one major iteration of Algorithm 6.1 to emphasize the details. Let us follow Algorithm 6.1 and now we are at time $t = 5$ and state $y = 1009$, remember from the HJB solver we are equipped with a table of the numerical value of the value function $V(y, t)$. Let us use the table below to show part of it: At this time, we use numerical*

Table 6.1. Value function on discrete state-time mesh for feedback BRP with MNL demand function around $(y, t) = (1009, 5)$

	$y = 1008.5$	$y = 1009$	$y = 1009.5$
$t = 4.9$	10403	10403	10403
$t = 5.0$	10208	10208	10208
$t = 5.1$	10013	10013	10013

interpolation to obtain the following partial derivatives of the value function can be *obtained:* $\frac{\partial V}{\partial y}(1009,5) = 0, \frac{\partial V}{\partial t}(1009,5) = -1951$*. This means we are ready for the numerical root-finding with equation (6.13). Here Lemma 6.2 is taken advantage of, which means there is only one variable* $\pi(5)$ *to solve for. And the solution yields:* $\pi_1^*(1009,5) = 9.9840$*, we can compute the demand fulfillment accordingly* $d_1^*(1009,5) = 205.2269$*.*

Finally, we update the state-time to compute for the next iteration: $t^{k+1} = t^k + \Delta t = 5+0.1 = 5.1$*, and* $y_1^{k+1} = y_1^k + d_1^{k,*}(y_1^k, t^k)\cdot\Delta t = 1009 + 205.2269 * 0.1 = 1029.5$*, which will be taken as a starting point for the next iteration.*

6.4.2 Numerical Examples for DPFI Equilibrium with Feedback

In this part we provide the numerical results of the feedback equilibrium following Algorithm 6.2, where the BRP of each seller is solved with following Algorithm 6.1. For simplicity in providing necessary data to the ROC-HJ solver, the optimal feedback best response of one previous iteration is passed on to the next with a fitted model in second order polynomial with respect to time. And the numerical discretization schemes with respect to state space, time and control are also inherited. The settings are: there are 2 sellers $s = 1,2$ on the market and the time horizon to consider is $[t_0, t_f] = [0,10]$ The initial fixed inventory endowments are $K_1 = 6000, K_2 = 3000$ and we take the following parameterization:

$$\begin{aligned}
\alpha_1(t) &= 500 + 15t,\ \beta_1(t) = 3 - 0.15t,\ \gamma_{21}(t) = 0.5 + 0.1t \\
\alpha_2(t) &= 300 + 5t,\ \beta_2(t) = 2 - 0.1t,\ \gamma_{12}(t) = 0.5 + 0.03t \\
p_{\min,1,2} &= 10,\ p_{\max,1,2} = 100,\ d_{\min,1,2} = 1,\ \rho = 0.01.
\end{aligned}$$

In order to save computational time, the parallel computation capability of the ROC-HJ package is invoked. In each round of iteration, the HJB-PDE is solved by 4 threads powered by OpenMP, which is one of the most popular toolbox in parallel computing. For more details on this end, please refer to Appendix.

Remark 6.4. *In this numerical example, we use a time mesh of* $N = 200$ *in the solution of optimal control subproblem, its HJB-PDE by the ROC-HJ toolbox, then optimal control synthesis is carried out in this discrete time mesh. Then, a third order polynomial in time is used to fit the discrete time solution of each subproblem into continuous time. The solution is then passed onto the next iteration.*

In this numerical example, the convergence of Algorithm 6.2 is presented in Figure 6.5, where each dot represents the 2-norm of the difference between two consecutive best responses. Comparing with the convergence rate analysis from Chapter 5 on the open-loop problem, we can see that the convergence rate towards the feedback equilibrium is fast. Intuitively, this is due to the feedback information structure of the game where state dynamics is endogenous. On the other hand, the open-loop best response strategies can be more sensitive.

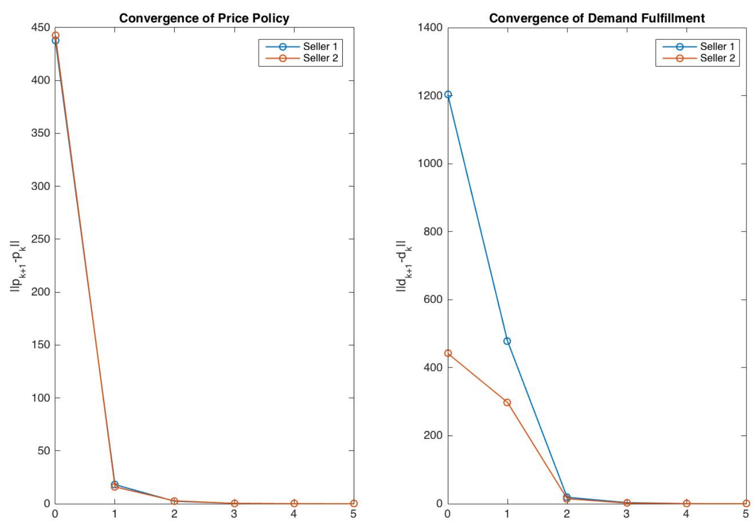

Figure 6.5. Control Convergence

Furthermore, in Figure 6.6 we present, for both sellers, the equilibrium pricing and demand fulfillment policies along with the equilibrium state trajectories.

6.5 Summary and Future Work

This chapter we introduce the state dynamics into the DPFI model and provide further analysis into the model with the aid of such state dynamics. In the context of open-loop best response and equilibrium, we utilize the necessary conditions for optimal control problems with state constraints and provide a new set of necessary conditions, especially for the DVI describing the open loop equilibrium. When assuming a feedback information structure, we propose the solution schemes the of the feedback BRP, and upon that, build (algorithm/heuristics) inspired by fictitious

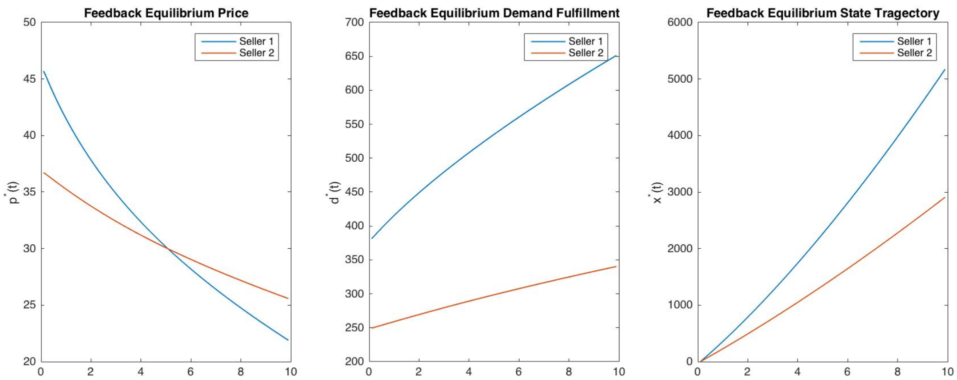

Figure 6.6. Feedback Equilibrium State and Control

play. Furthermore, we provide numerical examples utilizing an existing HJB-PDE solver and demonstrate the effectiveness of the proposed algorithm.

In the future, we would like to provide further qualitative analysis for this model such as stability (Seierstad and Sydsaeter (1986) [130]), time-consistency and sub-game perfect (Dockner (2000) [41]). We would also like to investigate the equilibrium with limited information structures such as conjecture variation (Vives (2001) [142]).

Chapter 7
Stochastic Dynamic Monopsony and Oligopsony

7.1 Introduction and Literature Review

In this chapter we investigate the problem of monopsony and oligopsony in a dynamic, stochastic setting, utilizing the theories investigated in Chapter 4. The term monopsony, meaning "single-purchase", comes from ancient Greek. It describes the market structure in which only one buyer interacts with many would-be sellers of a particular product or resource. Accordingly, the term oligopsony refers to the market structure where such interaction is carried out with only a few such buyers. Models based on such market structure have been applied in the study of labor and natural resource market. Recently, there has been observations that in industries such as consumer electronics, some dominating corporations could have become monopsonists/oligopsonists. See, for example Elmer-deWitt (2011) [45] which discusses the case of Apple company.

In fact, there has been a series of static and deterministically dynamic models built under the context of monopsony and oligopsony. In the study of agricultural/-food supply chains, Chen and Lent (1992) [39] studies the effect of a supply shift on the equilibrium of farm products purchased by food processors in oligopsonistic competition; Fofana and Jaffry (2008) [52] conducts empirical study on measuring the oligopsony power of UK salmon retailers; on dynamic modeling, Katchova et al. (2005) [81] proposes a linear-quadratic dynamic oligopsony model with a case study on the US potato processing industry; recently, Yu and Bouamra-Mechemache

(2016) [149] compares the oligopsony models with oligopoly models in vertical food supply chains under the context of food production standards.

Monopsony and oligopsony competition models are also widely applied in the research of labor markets. Examples are: Baily (1975) [8] studies the dependency of labor supply on the wage paid by the firm modeled as a dynamic monopsonist using the formulation of an infinite-horizon optimal control problem; Bhaskar et al. (2002,2003) [21], [22] propose comprehensive and tractable oligopsony models that capture several essence of wage dispersion in labor markets.

Another important application of monopsony models lies in the energy and natural resource management literature. For instance, Schworm (1983) [129] studies the steady state policy of dynamic monopsonistic control of a common property renewable resource, then generalize this model to the case of nonlinear lump-sum pricing; Vargas and Schreiner (1999) [141] uses monopsonistic computable general equilibrium (CGE) model to form a case study in the forest product industry.

At the same time, we notice that to the best of our knowledge, there has been no model addressing both the dynamic and stochastic nature of those markets, namely, taking a point of view with stochastic optimal control/stochastic Nash games. Whereas in reality, the dynamics of the (future) price of large commodities often encounters stochastically in addition to the supply-demand relationship that is driving the market (Hull (2006) [72]). For example, there has been research concerning the (future) price of large commodities such as crude oil in energy and agricultural industry, and justifying such price could be modeled by different stochastic processes (see Schwartz (1997) [128] for a review and Du et al. (2011) [42] as an example of recent development of such models). Hence in this chapter, we chose the price of some resource as the state variable, and use a controlled stochastic differential equation (SDE) to model the dynamic change of such price.

In fact, this chapter is also inspired by the recent development of dynamic stochastic monopolistic/oligopolisitc models. In these models, either demand or capacity of competing firms are modeled as diffusion process induced by Brownian motion, and different versions of the stochastic Hamilton-Jacobi-Bellman PDE are employed to study the optimal policies of competitors. Solutions in the form of special functions are often produced to give analytical insights. These examples are: Ledvina and Sircar (2011) [89] in dynamic Bertrand duopoly; Harris et al. (2010) [69], Ludkovski and Sircar (2012) [91], Ludkovski and Yang (2015) [92] in the case of dynamic Cournot duopoly/oligopoly; Chan and Sircar (2015) [36] in the case of dynamic mean field games (MFGs, see Larsy and Lions (2007) [88] for a comprehensive review). Note that instead of solving the stochastic HJB-PDEs, the analysis we propose in this chapter relies on the stochastic maximum principle and stochastic dynamic variational inequality (S-DVI) that has been summarized in Chapter 4.

The rest of this chapter is structured as follows: in Section 7.2 we layout the formulation of aspatial monopsony as optimal control problem and analyze its necessary condition. In Section 7.3 we introduce multiple firms and discuss the aspatial oligopsony problem. And as a start, we will spend the rest of this introductory section to review examples of commonly employed (competitive) market structures under static, deterministic settings.

7.1.1 The Deterministic, Static Case

In this part we take a small detour and review the following concepts by examples: perfect competition, monopoly-monopsony, Bertrand versus Cournot oligopoly, and finally Cournot-oligopoly-oligopsony. We illustrate these concepts by simple example problems in static, deterministic settings, and then use them as a benchmark for our study with the dynamic, stochastic case. For reference on more details of these terms and their applications, we refer the readers to Tirole (1988) [139], Talluri and Van Ryzin (2006) [136]. Also see Friesz (2010) [54] for a comprehensive review on deterministic dynamic aspatial monopolies and oligopolies. We start with the case of perfect competition, in which some commodity is produced in a market of interest and all firms are price takers in selling their product and buying resources.

Example 7.1. *(Perfect Competition) Under perfect competition, firms can produce and sell as much quantity x as they want. And a firm's problem can be formulated as the following:*

$$\max_{x\geq 0} V(x) = px - c(x) - wx$$

here $p, w > 0$ are final product price and resource price. The first order condition leads to:

$$p = \frac{\partial c(x)}{\partial x} + w$$

If a firm is the only one to purchase resources, carry-out the production and then sell its final product, the firm is then a monopolist and a monopsonist at the same time, we have the following example:

Example 7.2. *(Monopsony-Monopoly) Following the previous example, but now assume that firm is a monopsonist in procurement of resource and a monopolist in production and selling of the final product, assume the firm uses production quantity x as the decision variable, its optimization problem is the following:*

$$\max_{x \geq 0} V(x) = p(x)x - c(x) - w(x)x.$$

Here, in addition, $p(x)$ is the price function, $w(x)$ is the resource cost function, and the first order condition leads to

$$\frac{\partial p(x)}{\partial x} + p(x) = \frac{\partial c(x)}{\partial x} + \frac{\partial w(x)}{\partial x} + w(x).$$

Furthermore, if the monopsonist is a price taker when selling the final product, the above becomes

$$p(x) = \frac{\partial c(x)}{\partial x} + \frac{\partial w(x)}{\partial x} + w(x) \tag{7.1}$$

Example 7.3. *(Cournot duopoly, linear demand) We here look at the Cournot duopoly. Let $x = (x_1, x_2)$ be the production of two firms producing and selling the same product. and let $X = x_1 + x_2$ be the aggregate supply. The market determines the price that clears this output, and we denote this price by the inverse-demand function $p(X) = \frac{\alpha - X}{\beta}$. Also assume a constant marginal cost c for each firm. We then have firm i's objective which is to maximize profit*

$$\max_{x_i} V_i(x) = p(X)x_i - cx_i.$$

The necessary condition for this problem:

$$\frac{\alpha - (x_1 + x_2)}{\beta} + x_1(-\frac{1}{\beta}) = c, \frac{\alpha - (x_1 + x_2)}{\beta} + x_2(-\frac{1}{\beta}) = c$$

And the Cournot equilibrium is given by the simultaneous solution of these two first order conditions. This means production level is $x_{Cournot,1} = x_{Cournot,2} = \frac{1}{3}(\alpha - \beta c)$ and $X = x_1 + x_2 = \frac{2}{3}(\alpha - \beta c)$. The market price is then $p_{Cournot}(X) = \frac{1}{3}\frac{\alpha}{\beta} + \frac{2}{3}c$.

The profit for either firm is:

$$V_{Cournot,1,2} = (\frac{1}{3}\frac{\alpha}{\beta} - \frac{1}{3}c)x_{1,2} = \frac{1}{9\beta}(\alpha - \beta c)^2$$

This means that in order to generate non-negative net profit we should assume,

$$\frac{\alpha}{\beta} \geq c, \ \text{or} \ \alpha - \beta c \geq 0 \tag{7.2}$$

We can argue that this is as reasonable as assuming the market price should not be lower than the marginal cost.

Example 7.4. *(Bertrand duopoly, linear demand) In Bertrand competition we assume that all customers buy only from firms offering the lowest price, and if the firms post the same price the demand will be shared evenly. Firms compete by posting their prices and producing in quantities just sufficient to satisfy all the demand they face. With two firms 1 and 2, and the same marginal cost c, let the market-demand function be denoted by $d(p)$. The demand for firm 1 at price p_1 is given by:*

$$d_1(p_1, p_2) = \begin{cases} d(p_1), & \text{if } p_1 < p_2 \\ d(p_1)/2, & \text{if } p_1 = p_2 \\ 0, & \text{if } p_1 > p_2 \end{cases}$$

Firm i's profit is then:

$$V_i = (p_i - c)d_i(p_1, p_2)$$

And it is easy to see that the Bertrand-Nash equilibrium is unique and it is for both firms to price at the marginal cost c, therefore making zero net profits. This leads to $p_{Bertrand,1} = p_{Bertrand,2} = c$, and the demand faced are equal. Let $d = d_{Bertrand,1} = d_{Bertrand,2}$, thus $X = 2d$ while $p(X) = c$ and solve for d:

$$c = \frac{\alpha - 2d}{\beta}, \ \text{or} \ d = \frac{1}{2}(\alpha - \beta c)$$

And since firms are pricing at marginal cost, the total profit should be zero.

$$V_{Bertrand,1,2} = 0$$

Remark 7.1. *The above examples show that to the firms, the Cournot competition will be more beneficial than the Bertrand competition under (7.2). Intuitively, the firms in the Cournot game are taking advantage of their power to control the scarcity of the commodity, in fact, they produce less to gain better market price.*

We list a static oligopsony of Cournot type as a final example:

Example 7.5. *(Duopsony) Let $x = (x_1, x_2)$ be the procurement quantity of two firms producing and selling the same product. and let $X = x_1 + x_2$ be the aggregate procurement quantity. The market determines the procurement price $w(X) = \frac{\alpha+X}{\beta}$. Firm i's problem:*

$$\max_{x_i} V_i(x_i; x_{-i}) = px_i - c_i(x_i) - w(X)x_i$$

And the first order conditions are:

$$w(X) + x_i \frac{\partial w(X)}{\partial X} = p - \frac{\partial c_i(x_i)}{\partial x_i}$$

with the assumption of $w(X) = \frac{\alpha+X}{\beta}, c_i(x_i) = cx_i$ we have:

$$\frac{\alpha + (x_1 + x_2)}{\beta} + x_1(\frac{1}{\beta}) = p - c, \frac{\alpha + (x_1 + x_2)}{\beta} + x_2(\frac{1}{\beta}) = p - c$$

and the equilibrium solution is:

$$x_{1,2} = \frac{\beta(p - c)}{3} - \frac{\alpha}{3}$$

7.2 Dynamic Stochastic Monopsony

In this section, we consider the dynamic and stochastic generalization of the monopsony model from Example 7.2. The time horizon to consider is $[t_0, t_f]$, the decision environment is adaptive open-loop. The firm of interest is the only buyer of some input on the market. Such input resource from procurement is immediately turned into production which leads to the final product. We assume that the firm is a price taker for its final product, resulting in our model an exogenous, deterministic price function $p(t)$. As mentioned earlier, here we use the purchasing cost of the input resource $w(t)$ as the state variable and assume it follows a controlled diffusion process expressed by:

$$\begin{aligned} dw(t) &= f(q(t), w(t))dt + \sigma(q(t), w(t))dB_t \\ w(t_0) &= w_0. \end{aligned}$$

Here the drift term reflects both the demand-supply relationship due to the quantity purchased and other deterministic dynamic/seasonal effects, and the diffusion term is added to reflect stochasticity.

Here we assume that one unit of final product can be produced from one unit of raw material, and similar to the problem settings of Section 8.1 of Friesz (2010) [54], we also allow the monopsonist firm to build its inventory, leading to the following dynamics, where $d(t)$ is the other part of the firm's control, meaning the rate of quantity sold or demand fulfillment:

$$\begin{aligned} \frac{dI(t)}{dt} &= q(t) - d(t) \\ I(t_0) &= I_0 \end{aligned}$$

At each time instant t, the profit for the firm is

$$\pi(t) = p(t)d(t) - w(t)q(t) - c(q(t)) - \Psi(I(t))$$

where $c(\cdot)$ is production cost function and $\Psi(\cdot)$ is the inventory cost. In addition, let p_f be the liquidated value of the firm's final inventory at time t_f, we know that the firm's objective is:

$$\max_{q,d} E\left[p_f I(t_f) + \int_{t_0}^{t_f} [p(t)d(t) - w(t)q(t) - c(q(t)) - \Psi(I(t))]dt\right]$$

subject to the dynamics mentioned above. Here denote the set of feasible control as Λ and summarizing the discussions above, we propose the monopsonist firm's problem:

Problem 7.1. *(Dynamic Stochastic Monopsony) The monopsonist firm seeks the* *dynamic procurement and selling policy to maximize total profit:*

$$\max_{(q,d)\in\Lambda} E\left[p_f I(t_f) + \int_{t_0}^{t_f} [p(t)d(t) - w(t)q(t) - C(q(t)) - \Psi(I(t))]dt\right]$$

such that

$$\begin{aligned} dw(t) &= f(q(t), w(t), t)dt + \sigma(q(t), w(t), t)dB_t \\ \frac{dI(t)}{dt} &= q(t) - d(t) \\ w(t_0) &= w_0, I(t_0) = I_0 \end{aligned}$$

Clearly, Problem 7.1 is a special case of the general Problem from Chapter 4, where the deterministic inventory dynamics can be interpreted as a stochastic dynamics with zero diffusion term. We here propose its regularity conditions:

Definition 7.1. *(Regularities of Monopsony Problem) Problem 7.1 is called regular if the following conditions hold:*
(i) $f(w, q, t) : \mathbb{R}\times\ \Lambda \times [t_0, t_f] \to \mathbb{R}$ *is continuously differentiable w.r.t.* w *and* q*. Also the first order partials* f_w, f_q *are bounded. This means:*

$$|f(w, q, t)| \leq L_1(1 + |w| + |q|);$$

(ii)$\sigma(w, q, t) : \mathbb{R} \times [t_0, t_f] \to \mathbb{R}$ *is continuously differentiable w.r.t.* w *and* q*. Also the first order partials* σ_w, σ_q *are bounded. This means:*

$$|\sigma(w, q, t)| \leq L_2(1 + |w| + |q|);$$

(iii) the set of feasible control $\Lambda \subseteq \mathbb{R}^2$ *is convex;*
(iv) production cost $C(\cdot) : \mathbb{R} \to \mathbb{R}$ *is convex, continuously differentiable w.r.t.* q*, inventory cost* $\Psi(\cdot) : \mathbb{R} \to \mathbb{R}$ *is convex, continuously differentiable w.r.t.* I*:*

$$|C_q(q)| \leq L_3(1 + |q|), \quad |\Psi_I(I)| \leq L_3(1 + |I|)$$

We start the analysis by direct application of the necessary conditions in Theorem 4.3 from Chapter 4. Firstly, notice that our stochastic optimal control problem has an equivalent objective of minimization:

$$\min_{q,d} J = E\left[-p_f I(t_f) + \int_{t_0}^{t_f} [w(t)q(t) + C(q(t)) + \Psi(I(t)) - p(t)d(t)]dt\right]$$

The Hamiltonian to employ to this minimization problem is with adjoint variable (λ, μ, η):

$$H = pd - \Psi(I) - C(q) - wq + \lambda f(q, w) + \mu\sigma(q, w) + \eta(q - d) \tag{7.3}$$

along with a list of corresponding partial derivatives:

$$\begin{aligned} H_q &= \frac{\partial H}{\partial q} = -\frac{\partial C}{\partial q} - w + \lambda\frac{\partial f}{\partial q} + \mu\frac{\partial \sigma}{\partial q} + \eta; \\ H_w &= \frac{\partial H}{\partial d} = -q + \lambda\frac{\partial f}{\partial w} + \mu\frac{\partial \sigma}{\partial w}; \\ H_I &= \frac{\partial H}{\partial I} = -\frac{\partial \Psi}{\partial I}; \quad H_d = \frac{\partial H}{\partial d} = p - \eta. \end{aligned}$$

We are ready to apply Theorem 4.3. Firstly, observing the set of feasible controls Λ is convex, the stochastic maximal principle is in the form of VI. That is, for all $(q, d) \in \Lambda$, a.e. $t \in [t_0, t_f]$, $P-$a.s., we have:

$$\langle H_q(t, w^*, q^*, d^*; \lambda, \mu, \eta), q - q^*\rangle + \langle H_d(t, w^*, q^*; \lambda, \mu, \eta), d - d^*\rangle \leq 0$$

The adjoint variables of the stochastic dynamics should follow:

$$d\lambda(t) = -H_w dt + \mu dB_t$$

Let us also write down the equivalent form of the inventory dynamics:

$$dI(t) = (q - d)dt + \tilde{\sigma} d\tilde{B}_t$$

with $\tilde{\sigma} \equiv 0$ due to the deterministic nature of the inventory dynamics. This leads to the remaining part of the adjoint dynamics:

$$\frac{d\eta(t)}{dt} = -H_I$$

Both adjoint equations should be paired with appropriate transversality conditions:

$$\lambda(t_f) = -\frac{\partial(-p_f I(t_f))}{\partial w} = 0, \quad \eta(t_f) = -\frac{\partial(-p_f I(t_f))}{\partial I} = p_f.$$

Let us summarize the analysis above into the following theorem:

Theorem 7.1. *(Necessary Condition of Aspatial Monopsony Problem) Let Problem 7.1 be regular in the sense of Definition 7.1 and admit an optimal state-control tuple* $(q^*, d^*,\ w^*,\ I^*)$*. Then such optimal tuple along with the adjoint variables* $(q^*, d^*, w^*, I^*; \lambda, \mu, \eta)$ *of Problem 7.1 solve the following stochastic Hamiltonian system consisting of state dynamics:*

$$\begin{aligned} dw^*(t) &= f(q^*(t), w^*(t))dt + \sigma(w^*(t))dB_t, \quad w^*(t_0) = w_0 \\ \frac{dI^*(t)}{dt} &= q^*(t) - d^*(t), \quad I(t_0) = I_0 \end{aligned}$$

adjoint dynamics:

$$d\lambda(t) = -H_w dt + \mu dB_t,\ \frac{d\eta(t)}{dt} = -H_I$$

transversality condisions:

$$\lambda(t_f) = 0, \quad \eta(t_f) = \frac{\partial(-p_f I(t_f))}{\partial I} = p_f$$

along with the following variational inequality to reflect the maximum principle:

$$\begin{gathered} \text{for all } (q,d) \in \Lambda,\ a.e.t \in [t_0, t_f],\ P - a.s. \\ \langle H_q(t, w^*, q^*, d^*; \lambda, \mu, \eta), q - q^* \rangle + \langle H_d(t, w^*, q^*; \lambda, \mu, \eta), d - d^* \rangle \leq 0 \end{gathered}$$

As a direct application of the existence result in Theorem (CrossRef) from Chapter 4, we have the following existence result:

Theorem 7.2. *(Existence of Monopsony Policy) Let Problem 7.1 be regular in the sense of Definition 7.1, then it admits an optimal control and state tuple* $(q^*, d^*;\ w^*, I^*)$*.*

In order to facilitate further analysis and computation, we introduce a series of additional assumptions. Firstly, following Schwartz (1997) [128] and Gibson

(1990) [64], we assume input commodity spot price follows the stochastic process that takes into account the mean reversion effect. More specifically, in the rest of this chapter, we assume the input price dynamics take the following form:

$$dw(t) = \kappa(m(q(t)) - \ln w(t))w(t)dt + \sigma w(t)dB_t. \tag{7.4}$$

Remark 7.2. *Note that one way to interpret Equation (7.4) is that, according to [128], if we define $X = \ln w(t)$, Itô's lemma suggests the log price should be in the form of the famous Ornstein-Uhlenbeck stochastic process:*

$$dX = \kappa(m(q(t)) - \frac{\sigma^2}{2\kappa} - X)dt + \sigma dB_t.$$

Here $\kappa > 0$ represents the magnitude of price adjustment speed; $\sigma > 0$ measures the magnitude of diffusion. Both are constants that could be fitted from historical data. The term $m(\cdot)$ represents the "dynamic mean", which reflects the "true" supply-demand relationship of the raw input. Specifically, if $m(t) = m$ for all $t \in [t_0, t_f]$, then it is to represent "the long-run mean log price [128]". In this chapter, we assume that the monopsonist's dynamic procurement policy affects such dynamic mean, which means $m(t) = m(q(t))$, and we will discuss regularity conditions for the mapping $m(\cdot)$ later.

We can see the state dynamics (7.4) as an immediate generalization of previously employed static and discrete-time dynamic models. For example, Karp (1983) [80] uses deterministic linear price dynamics, coupled with quadratic objective in a model for the corn industry, our model is a continuous time dynamic and stochastic extension of their linear dynamics.

Remark 7.3. *There exists a significant group of literature which studies the so-called theory of storage, for example Brennan (1958) [29]. A key feature of such theory, which is also noted by Schwartz (1997) [128], is the establishment of commodity price volatility with the inventory. For example, there are real-world data driven models reflecting the correlation between sudden changes in commodity demand and spikes in volatility.*

Another notion related to dynamic inventory decisions in the presence of stochasticity is safety stock. In literature and in industry practice, this issue of safety stock is often investigated by utilizing multi-stage stochastic linear and integer program-

ming methods. See, for example, Sen and Higle (1999) [131] for a review of basic methodologies, also see Wieland et al. (2012) [145] and Zanjani et al. (2010) [82] for case studies and industry implementation of these models. As shown by these examples, most existing literature focus on introducing uncertainties on the demand side and production itself, whereas in this Chapter we discuss stochasticities related to supply of raw materials.

Also note that in this chapter, inventories of powerful buyers are explicitly modeled as part of the state dynamics, but we are not considering its contribution to the diffusion term, namely, we assume only a constant σ will be involved for the simplicity of this model. And we will leave the study towards the state dynamics involving theory of storage to a further research.

Remark 7.4. *In addition, it is a well-known fact that in the real world, from a operational perspective, downstream buyers will use different financial instruments such as futures and forward contracts to hedge as much risk as possible in the procurement of raw input. Examples of such procurement models are Benth et al. (2014) [15] and Arnold (2010) [4]. In this chapter, we use the rate of purchase (and production) $q(t)$ as a description of such policies in the spirit of fluid limits. In the OR/OM literature, previous research utilizing such fluid models in the study of production and inventory polices includes: Yan and Kulkarni (2008) [147], Kulkarni and Yan (2012) [87], Berman and Perry (2006) [16].*

Along with the embodiment of commodity price dynamics, we further assume that the monopsonist is free to make purchase with a positive rate with simple bounds in the planning horizon, which leads to a convex set of feasible control. Similarly we enforce a simple upper bound to the demand fulfillment rate. We argue that such assumption is close enough to reality in a lot of industries where the availability of raw materials are limited as well as the production capacity of the monopsonist. This means the feasible region Λ is the following:

$$\begin{aligned} \Lambda \;\triangleq\; & \{(q(t), d(t)) : -q(t) \le Q_{min}, \quad q(t) - Q_{max} \le 0 \\ & -d(t) \le D_{min}, \quad d(t) - D_{max} \le 0, t \in [t_0, t_f]\} \end{aligned} \tag{7.5}$$

Finally, another set of assumption is needed in the spirit of Definition 4.4, which means the value function of Problem 7.1 should be well defined and should have continuous partial derivatives of all orders necessary. Combining those stronger assumptions we have the following corollary of Theorem 7.1:

Corollary 7.1. *(Monopsony with Mean Reversion Stochastic Commodity Price) If, in addition to Definition 7.1, define the value function of Problem 7.1 for all* $(w,t) \in \mathbb{R} \times [t_0, t_f]$ *as follows:*

$$\begin{cases} V(w,I,t) = \inf_{q\in\Lambda} J(w,I,t;q(\cdot),d(\cdot)) \\ V(w(t_f),I(t_f),t_f) = 0 \end{cases}$$

Assume the value function exists and is three times continuously differentiable in state and continuously differentiable in time, V_{xt} *is also continuous. Furthermore, assume the commodity price process follow (7.4) with* $m(q(t))$ *convex and continuously differentiable, and with the set of feasible controls defined by (7.5). Then Problem 7.1 admits an optimal tuple* (q^*, d^*, w^*, I^*)*, and that the following conditions must hold:*

$$\begin{aligned} dw^*(t) &= \kappa(m(q^*) - \ln w^*)w^* dt + \sigma w^* dB_t, \frac{dI^*(t)}{dt} = q^*(t) - d^*(t) \\ d\lambda(t) &= \left[q^* - \lambda\kappa(m(q^*) - 1 - \ln w^*) - \sigma^2 w^* \frac{\partial \lambda}{\partial w}\right] dt + \sigma w^* \frac{\partial \lambda}{\partial w} dB_t \\ \frac{d\eta(t)}{dt} &= \frac{\partial \Psi}{\partial I} \\ w^*(t_0) &= w_0, I^*(t_0) = I_0, \lambda(t_f) = 0, \eta(t_f) = p_f \end{aligned}$$

along with the following variational inequality:

$$\begin{cases} \text{for all } (q,d) \in \Lambda, \text{ a.e.} t \in [t_0, t_f], \ P - a.s. \\ \left\langle \left[\lambda\kappa w^* \frac{\partial m(q^*)}{\partial q} - \frac{\partial c(q^*)}{\partial q} + \eta - w^*\right], q - q^* \right\rangle + \langle p - \eta, d - d^* \rangle \le 0 \end{cases} \tag{7.6}$$

Proof. We first apply the specific form of mapping $f(w,q,t) : \mathbb{R}\times \Lambda \times [t_0, t_f] \to \mathbb{R}$ and $\sigma(w,t) : \mathbb{R} \times [t_0, t_f] \to \mathbb{R}$

$$\begin{aligned} \frac{\partial f(w,q,t)}{\partial w} &= \kappa(m(q) - 1 - \ln w), \frac{\partial f}{\partial q} = \kappa w \frac{\partial m}{\partial q} \\ \frac{\partial \sigma(w,t)}{\partial w} &= \sigma \end{aligned}$$

combine this with the Hamiltonian (7.3) we have:

$$\begin{aligned} H_q &= \lambda\kappa w\frac{\partial m}{\partial q} - \frac{\partial c(q)}{\partial q} + \eta - w \\ -H_w &= q - \lambda\kappa(m(q) - 1 - \ln w) - \mu\sigma. \end{aligned}$$

In addition, we know per Definition 7.1 and the additional regularities assumed, Theorem 4.5 holds as well as the assumptions therein, we have

$$\begin{cases} \lambda(t) = -V_w(w^*(t), t) \\ \mu(t) = -V_{ww}(w^*(t), t)\sigma w^* \end{cases} \tag{7.7}$$

which leads to

$$\mu(t) = \sigma w^* \frac{\partial \lambda}{\partial w}$$

and we have the revised adjoint dynamics:

$$\begin{aligned} d\lambda(t) &= \left[q^* - \lambda\kappa(m(q^*) - 1 - \ln w^*) - \sigma^2 w^* \frac{\partial \lambda}{\partial w}\right] dt + \sigma w^* \frac{\partial \lambda}{\partial w} dB_t \\ \frac{d\eta(t)}{dt} &= \frac{\partial \Psi}{\partial I} \end{aligned}$$

□

Notice that for this special case, the necessary conditions we derived would be the same should we apply Theorem 4.2 from Malliaris and Brock (1982) [96] directly.

Remark 7.5. *Corollary 7.1 allows us to further investigate the economic insights into Problem 7.1. In fact by virtue of concavity of the Hamiltonian and due to the simple constrained feasible region we assumed, variational inequality problem (7.6) is transformed into:*

$$\begin{aligned} q^* &= [\arg(H_q = 0)]_0^Q, \quad a.e.t \in [t_0, t_f],\ P - a.s. \\ d^* &= [\arg(H_d = 0)]_0^D, \quad a.e.t \in [t_0, t_f],\ P - a.s. \end{aligned}$$

assuming the control constraints are non-binding we have for a.e.$t \in [t_0, t_f]$, $P-$a.s.:

$$\eta(t) - \kappa w^*(t)\frac{\partial V(w^*,t)}{\partial w}\frac{\partial m(q^*(t))}{\partial q} = w^*(t) + \frac{\partial c(q^*)}{\partial q} \tag{7.8}$$

$$p(t) = \eta(t) \tag{7.9}$$

which is exactly the dynamic, and stochastic generalization of the necessary condition for the static problem as in equation (7.1). This condition indicates that at the optimal state-control pair, for a.e. $t \in [t_0, t_f]$, the procurement rate $q^*(t)$ should be such that marginal revenue $p(t)$, less a special term $\kappa w^*(t)\frac{\partial V(w^*,t)}{\partial w}\frac{\partial m(q^*(t))}{\partial q}$ which equals marginal cost. And from the right hand side of (7.8) we see that as in (7.1), the marginal cost consist of procurement price $w^*(t)$ and marginal production cost at the optimal production rate $\frac{\partial c(q^*)}{\partial q}$.

Remark 7.6. *In the case where control enters the diffusion term, a more general stochastic maximum principle still hold, with (7.6) replaced by:*

$$\begin{cases} \text{for a.e.} t \in [t_0, t_f],\ P - a.s. \\ (q^*, d^*) = \arg\max_{(q,d)\in\Lambda} H(t, w^*, q, d, \lambda(t), \mu(t), \eta(t)) \end{cases} \tag{7.10}$$

7.2.1 Algorithm for Stochastic Monopsony

In this part we list the algorithm used for the stochastic monopsony problem. We must point out that the algorithm developed here is based on the stochastic maximum principle instead of dynamic programming. And it is a generalization of the gradient projection algorithm for deterministic optimal control problems as in Friesz (2010) [54]. The algorithm proposed in this section is also inspired by Mookherjee [102] and Meimand (2013) [100]. We start by considering the major iteration given some control (q^k, d^k). The following system is an intermediate step of Corollary 7.1, and is a special case of the more general FBSDE system in Problem 4.3 since we can take the diffusion term of the dynamics related to inventory $I(t)$ as 0:

$$
\begin{cases}
dw(t) = \kappa(m(q^k) - \ln w)wdt + \sigma w dB_t \\
dI(t) = [q^k - d^k]dt \\
d\lambda(t) = \left[q^k - \lambda\kappa(m(q^k) - 1 - \ln w) - \mu\sigma\right] dt + \mu dB_t \\
d\eta(t) = \frac{\partial \Phi}{\partial I} dt \\
w(t_0) = w_0, I(t_0) = I_0 \\
\lambda(t_f) = 0, \eta(t_f) = p_f.
\end{cases}
\tag{7.11}
$$

Moreover, in each iteration with given (q^k, d^k), the deterministic part of system (7.11) can be treated with the same method reviewed in Chapter 2 where the state and adjoint trajectories are computed after one another: first find the state trajectory $I^k(t)$ with initial time problem and then compute the adjoint trajectory using both the state and control information. On the other hand, such technique could not be directly applied to the stochastic part of system (7.11) due to the non-anticipation property required by the stochastic problem. Instead, let us take a closer look at the following sub-system:

$$
\begin{cases}
dw(t) = \kappa(m(q^k) - \ln w)wdt + \sigma w dB_t \\
d\lambda(t) = \left[q^k - \lambda\kappa(m(q^k) - 1 - \ln w) - \mu\sigma\right] dt + \mu dB_t \\
w(t_0) = w_0, \lambda(t_f) = 0
\end{cases}
\tag{7.12}
$$

It is easy to see that system (7.12) have the following features: (i) the control $q(\cdot)$ does not enter the diffusion term of the forward SDE; (ii) the state and adjoint variables are of the same dimension (both are of 1-dimension). These features make system (7.12) a special, therefore easier to solve than the general FBSDE system. We shall refer to Yong and Zhou (1999) [148] for a rigorous proof on the solvability of system (7.12) by the 4-step Algorithm 4.1, in the sense that the 4-step approach should provide a unique adapted solution to FBSDE system (7.12). Therefore, taking the solution of system (7.12) as a subproblem, we propose the following stochastic gradient projection algorithm:

Algorithm 7.1. *(Stochastic Gradient Projection for Monopsony) Consider a finite time horizon $t \in [t_0, t_f]$, and Problem 7.1:*

Step 0 Initialization *Set $k = 0$ and pick $q^0(t) \in \Lambda$.*

Step 1 FBSDE Solution *Solve system (7.11) parameterized by (q^k, d^k) and call solution $(w^k, I^k; \lambda^k, \mu^k, \eta^k)$:*

Step 1a Find $z(t, w, p)$ such that $\forall t \in [t_0, t_f]$, $(w, p) \in \mathbb{R} \times \mathbb{R}$:

$$z(t, w, p) = \sigma w p$$

Step 1b Solve the following system of PDEs:

$$\left\{ \begin{array}{c} \theta_t + \frac{\sigma^2 w^2}{2}\theta_{ww} + \left[\kappa m(q^k) - \kappa \ln w + \sigma^2\right] w\theta_w \\ -\kappa(m(q^k) - 1 - \ln w)\theta - q^k = 0, \quad (t, w) \in (t_0, t_f) \times \mathbb{R} \\ \theta(t_f, w) = 0, \quad w \in \mathbb{R} \end{array} \right. \tag{7.13}$$

Step 1c Use the θ and z obtained to define

$$\begin{aligned} \widetilde{f}(t, w) &= \kappa(m(q^k) - \ln w)w \\ \widetilde{\sigma}(t, w) &= \sigma w \end{aligned}$$

and solve the forward SDE:

$$w^k(t) = w_0 + \int_{t_0}^{t} \widetilde{f}(t, w(s))ds + \int_{t_0}^{t} \widetilde{\sigma}(t, w)dB_s$$

Step 1d Finally, use the relationship defined previously and set

$$\left\{ \begin{array}{c} \lambda^k(t) = \theta(t, w^k(t)) \\ \mu^k(t) = \sigma w^k \theta_w(t, w^k(t)) \end{array} \right.$$

Step 2 Find gradient *Use $q^k(t)$, $w^k(t)$, $\lambda^k(t)$, $\mu^k(t)$ to calculate:*

$$\begin{aligned} \nabla_q J &= \frac{\partial H(q^k, w^k, \lambda^k, \mu^k, t)}{\partial q} \\ &= \lambda^k \kappa w^k \frac{\partial m(q^k)}{\partial q} - \frac{\partial c(q^k)}{\partial q} + p - w^k \end{aligned}$$

Step 3 *Update and apply stopping test* *For a suitably small step size φ^k, update (q, d):*

$$q^{k+1} = P_\Lambda[q^k - \varphi^k \nabla_q J(q^k, d^k)]$$
$$d^{k+1} = P_\Lambda[d^k - \varphi^k \nabla_d J(q^k, d^k)]$$

If some stopping criterion is satisfied, declare $u^*(t) = (q^{k+1}, d^{k+1})$. Otherwise, set $k = k + 1$ and go to Step 1.

We can easily see that the structure of the algorithm above is very similar compared to Algorithm 2.1 in the deterministic case. The key difference is that due to the stochastic nature of Problem 7.1, the forward state dynamics and the backward adjoint dynamics have to be solved together as one FBSDE system. And according to the 4-step Algorithm 4.1, such solution will be relying on the solution of a parabolic PDE system.

Remark 7.7. *One special case of interest is the case of deterministic problem, for brevity we will not formulate and analyze the full problem in detail since this will be a straightforward application of the methodologies reviewed in Chapter 2.*

7.3 Dynamic Stochastic Oligopsony

In this section we introduce Nash competition and consider the oligopsony problem by introducing N firms with each firm $i = 1, \cdots, N$ applying procurement policy $q^i(\cdot)$, making the production/sales decision $d^i(\cdot)$, and setting the final product with exogenously given price $p(t)$, we have:

Problem 7.2. *(Stochastic Differential Oligopsony) Assume that there are N firms, firm i use $(q^i(\cdot), d^i(\cdot))$ as purchasing and sales quantity per unit of time. Firms seek to maximize total profit:*

$$\max_{(q^i,d^i)\in\Lambda^i} \Pi_i = E\left[p_f I^i(t_f) + \int_{t_0}^{t_f} [p(t)d^i(t) - w^i(t)q^i(t) - C^i(q^i(t)) - \Psi^i(I^i(t))]dt\right]$$

such that:

$$\begin{aligned} dw(t) &= f(w, q^1, q^2, \cdots, q^N, t)dt + \sigma w dB_t, w(t_0) = w_0 \\ \frac{dI^i(t)}{dt} &= q^i(t) - d^i(t), I(t_0) = I_0^i,\ i = 1, \cdots, N \end{aligned}$$

where every agent i has a cost/disutility functional $\Pi_i(q^i, d^i) : \Lambda^i \to \mathbb{R}$ *that depends on all agents' strategies. Furthermore, apply the following notiations where*

$$\begin{aligned} \Lambda &= \prod_{i=1}^{N} \Lambda^i \\ (q, d) &= ((q^i(\cdot), d^i(\cdot)) : i = 1, \cdots, N) \end{aligned}$$

Then to find the equilibrium of such aspatial stochastic differential oligopsony is to find $q^*(\cdot)$ *such that the following*

$$\Pi_i(q^*, d^*) \geq \Pi_i(q^{*,1}, \cdots, q^i, \cdots, q^{*,N}; d^{*,1}, \cdots, d^i, \cdots, d^{*,N})$$

hold for all firm i *and for any other* $(q, d) \in \Lambda$.

We will first provide analysis to this oligopsony problem following the spirit of Theorem 4.9, notice that similar to the monopsony case, we can write down each firm's equivalent objective in minimization:

$$\min_{(q^i,d^i)\in\Lambda^i} E\left[-p_f I^i(t_f) + \int_{t_0}^{t_f} [p(t)d^i(t) - w(t)q^i(t) - C^i(q^i(t)) - \Psi^i(I^i(t))]dt\right]$$

and Hamiltonian:

$$H^i = pd^i - \Psi^i(I^i) - C^i(q^i) - wq^i + \lambda f(w, q^1, \cdots, q^i, \cdots, q^N) + \mu\sigma(w) + \eta^i(q^i - d^i) \tag{7.14}$$

The partial derivatives of the Hamiltonian with respect to state and control follows immediately from previous Sections. In addition, we introduce some shorthand notations $u^i \triangleq (q^i, d^i)^T$ to represent firm i's control, and let $u = (u^1, \cdots, u^N)$ be the full vector.

Then as an embodiment of Definition 4.10 we propose a series of regularity conditions to facilitate the analysis:

Definition 7.2. *(Regularity Conditions for Stochastic Dynamic Oligopsony) (0)Information structure: adaptive open-loop (1)Each firm i's problem is regular in the sense of Definition 7.1. (2)each player's problem have a concave Hamiltonian* $H^i(x, u, t; p, q)$ *in* (x, u^i)*; also* $K^i(x)$ *is convex, and the set of feasible control* U^i *is convex*

Now we present the main theorem of this chapter, in which the stochastic dynamic oligopsony admits a characterization in the form of S-DVI:

Theorem 7.3. *(Oligopsony as a Stochastic DVI) When the stochastic differential oligopsony as in Problem 7.2 is regular in the sense of Definition 7.2, it admits an equivalent stochastic differential variational inequality.*

Proof. As noted in the previous section, we have the following optimality conditions for Firm i's problem, starting with variational inequality:

$$\left\langle H_q^i(t, w^*, q^*, d^*; \lambda, \eta), q - q^* \right\rangle + \left\langle H_d^i(t, w^*, q^*; \lambda, \eta), d - d^* \right\rangle \leq 0$$

which is

$$\nabla_{u^i} H^i(w^*, I^*, u^{*,i}, t; \lambda, \eta)^T (u^i(t) - u^{*,i}(t)) \leq 0$$

a.e. $t \in [t_0, t_f]$, $\mathbf{P}-$a.s. , $\forall u^i \in U^i$ and the corresponding adjoint dynamics:

$$d\lambda(t) = -H_w dt + (\sigma w \frac{\partial \lambda}{\partial w}) dB_t, \frac{d\eta(t)}{dt} = -H_I$$

while the state dynamics still hold:

$$\begin{aligned} dw^*(t) &= f(q^*(t), w^*(t))dt + \sigma w^*(t) dB_t, \quad w^*(t_0) = w_0 \\ \frac{dI^*(t)}{dt} &= q^*(t) - d^*(t), \quad I(t_0) = I_0 \end{aligned}$$

Notice that Definition 7.2 contains two groups of regularity conditions, hence individual player's variational inequality become a sufficient condition for the best response problem of his/her own. In addition, we know that in order for the Nash equilibrium condition to hold, all these conditions must hold simultaneously. Meanwhile, introduce the following tuples:

$$F^i(y, u, t) = \nabla_{u^i} H^i(y^*(t), u^{*,i}(t), t;)$$

where y is defined as: $y^i(u, t) = (w(u, t), I^i(u, t), \lambda^i(u, t), \eta^i(u, t))^T$; with augmented state-adjoint dynamics:

$$g^i = \begin{pmatrix} f(w,u,t) \\ p-d \\ -\nabla_w H^i(x^*(t), u^{*,i}(t), t; p^i(t), q^i(t)) \\ -\nabla_I H^i(x^*(t), u^{*,i}(t), t; p^i(t), q^i(t)) \end{pmatrix}, \quad \pi^i = \begin{pmatrix} \sigma w \\ 0 \\ (\sigma w \frac{\partial \lambda}{\partial w}) \\ 0 \end{pmatrix};$$

such that the following dynamics holds:

$$dy^i = g^i(y^i, u, t)dt + \pi^i dB_t$$

and initial-terminal conditions

$$y^i(t_0) = y^{i,0} = \begin{pmatrix} w^0 \\ I^0 \\ \lambda^i(0) \text{ free} \\ \eta^i(0) \text{ free} \end{pmatrix}; \quad y^i(t_f) = y^{i,f} = \begin{pmatrix} w(t_f) \text{ free} \\ I(t_f) \text{ free} \\ 0 \\ p_f \end{pmatrix}$$

hence

$$\begin{aligned} y &= (y^i : i = 1, \cdots, N) \\ g &= (g^i : i = 1, \cdots, N) \\ \pi &= (\pi^i : i = 1, \cdots, N) \\ F &= (F^i : i = 1, \cdots, N) \end{aligned}$$

The above augmentation of notations leads to the following stochastic differential variational inequality: find $u^* \in U$, and the corresponding (w^*, I^*) such that

$$E\left\{\sum_i \int_{t_0}^{t_f} F[w^*, I^*, u^*, t]^T (u - u^*) dt \leq 0\right\}, \forall u \in U, \quad \mathbf{P} - a.s. \tag{7.15}$$

here

$$\begin{aligned} u &\in U \subseteq (L^2[t_0, t_f])^m, \\ y(u,t) &= \arg\left\{dy = g(y,u,t)dt + \pi(y,u,t)dB_t, y(t_0) = y^0, y(t_f) = y^f\right\}. \end{aligned}$$

□

7.4 Summary and Future Work

In this chapter we examined the stochastic dynamic monopsony and ologopsony problem with raw material price as stochastic dynamics and inventory as deterministic dynamics. We analyze the necessary condition of dynamic equilibrium. We also out-lined numerical algorithms to solve for such equilibrium based on a pde solver.

In the future, with numerical solution of the pde ready we can examine more qualitative and computational properties of this problem.

Appendix

HJB-PDE Solution Algorithm and Package

Now we are aware that if we were given a $V(x,t)$, the optimal feedback control could be synthesized via Algorithm 3.1. However, in Step 1 of Algorithm 3.1 we only mentioned that the HJB-PDE could be solved numerically without any further details. This appendix provides a deeper review on the topic of numerical computation of HJB-PDE. We will first review existing literature on the numerical solution of (first order) HJB-PDE with background in deterministic feedback optimal control problem. Later we will introduce the ROC-HJ package by Bokanowski et al. (http://uma.ensta-paristech.fr/soft/ROC-HJ/) and describe more details about its implementation towards our numerical experiment in Chapter 3 and Chapter 6. Note that taking the DPP perspective in stochastic optimal control problem will also lead to a second order HJB-PDE. Such HJB-PDEs are beyond the scope of this thesis, please refer to Yong and Zhou (1999) [148] for some review. Another class of problems that this appendix does not cover is on those HJB-PDEs induced by infinite horizon optimal control problems. In other words, the PDE we are considering to solve numerically is the one employed by Theorem 3.2 and Theorem 3.4. Here, it is convenient to re-write the HJB-PDE of interest:

$$\frac{\partial V(x,t)}{\partial t} + \min_u \left[\nabla_x V(x,t) \cdot f(t,x,u) + f_0(t,x,u)\right] = 0 \tag{.1}$$

for $\forall (x,t) \in \mathbb{R}^n \times (t_0, t_f)$, with the boundary condition:

$$V(t_f, x) = K(x) \tag{.2}$$

Different numerical PDE methods, including the finite difference and semi-Lagrangian schemes, could be applied towards the numerical solution of HJB-PDEs. For an overview of the existing literature, please refer to Appendix A of Bardi and Capuzzo-Dolcetta (2008) [9]. In this thesis, we will mainly employ the semi-Lagrangian scheme towards the numerical solution of (3.7) along with (3.8). Toward this end, Falcone and Ferretti (2014) [50] provides a comprehensive review on the basics of the semi-Lagrangian schemes. We shall follow their book and review the main idea of a semi-Lagrangian scheme in here with a problem of 1-spatial dimension.

For a uniform grid $\{x_1, x_2, \cdots, x_I\}, \{t_1, t_2, \cdots, t_N\}$, let $v_i^n = V(x_i, t_n)$, then a discretized HJB-PDE is the following:

$$\frac{v_i^n - v_i^{n-1}}{\Delta t} + \min_u \left\{ \frac{v^n(x_i + \Delta t f(x_i, u)) - v_i^n}{\Delta x} + f_0(x_i, t_n, u) \right\} = 0 \qquad (.3)$$

hence the numerical scheme to consider is:

$$v_i^{n+1} = \min_u \{ v^n(x_i + \Delta t f(x_i, t_n, u)) + \Delta t g(x_i, t_n, u) \} \qquad (.4)$$

$$v_i^0 = v_0(x_i) \qquad (.5)$$

here the value of $v^n(x_i + \Delta t f(x_i, t_n, u))$ is to be interpolated by a Lagrangian polynomial. In each time step, the numerical evaluation of the RHS of equation (.4) is carried out by enumerating the value to minimize with regard a pre-specified, discrete set of control values. It is easy to see that there are more issues related to this topic regarding to, for instance, the convergence of different semi-Lagrianian schemes. Still, please refer to Falcone and Ferretti (2014) [50] for rigorous proofs.

In this thesis, we employ ROC-HJ package, which is currently maintained by Bokanowski et al. (http://uma.ensta-paristech.fr/soft/ROC-HJ/). Implemented in C++, this package is capable of solving HJB-PDEs induced by finite horizon, steady-state/infinite horizon problems of first and second order with both finite difference and semi-Lagragian schemes. It is also capable of handling Hamiltonian functions of min-max type which could be derived from two-person-zero-sum differential games. (see Basar and Olsder (1999) [10] and Friedman (1971) [53]). Please refer to Bokanowski et al. (2010) [28] and Altarovici et al. (2013) [3] for more details on the analysis.

Overall, the package consist of (i) a pre-complied library of the numerical methods mentioned above; (ii) a set of header files to describe different parts of the optimal control problem: running cost, terminal cost, right hand side of state dynamics, etc., to be specified by the user (iii) a set source files mainly for input-output controls, also to be specified by the user. Together, such an user-customized library is compiled using GNU Compiler Collection (GCC) (https://gcc.gnu.org) with the aid of a third party tools chain named CMake (https://cmake.org/). Please refer to the User Manual of ROC-HJ package for more details with trouble shooting.

In the case of problems with high spatial dimensions, or when a very fine mesh of state-time-control space are needed for the numerical solution of HJB-PDE, parallel algorithms based on domain decomposition could be introduced. In the ROC-HJ package, such option for parallel computation is also available through the OpenMP (http://www.openmp.org) API. In this thesis, after compiling and making the ROC-HJ package with GCC, options for multiple threads become automatically available.

In Chapter 3 and 6 of this thesis, ROC-HJ package is implemented for the solution of HJB-PDE under the context of: (i) optimal control formulations of SO-DTA problem and (ii) revenue management competition, both with feedback information structure. The library is compiled, built for numerical experiment by gcc-6.2.0 on a computer running Mac OS X operating system with Intel (R) Core-i5 (TM) CPU and 8GB of RAM. We take the numerical solution of the value function $V(x, t)$ as output of Step 1 of Algorithm 3.1 and proceed with the followings steps in Matlab for the synthesis of feedback optimal control.

Bibliography

[1] Mustafa Abdulaal and Larry J LeBlanc. Methods for combining modal split and equilibrium assignment models. *Transportation Science*, 13(4):292–314, 1979.

[2] Elodie Adida and Georgia Perakis. Dynamic pricing and inventory control: Uncertainty and competition. *Operations Research*, 58(2):289–302, 2010.

[3] Albert Altarovici, Olivier Bokanowski, and Hasnaa Zidani. A general hamilton-jacobi framework for non-linear state-constrained control problems. *ESAIM: Control, Optimisation and Calculus of Variations*, 19(2):337–357, 2013.

[4] Jan Arnold. *Commodity procurement with operational and financial instruments*. Springer, 2010.

[5] J-P Aubin and Arrigo Cellina. *Differential inclusions: set-valued maps and viability theory*, volume 264. Springer Science & Business Media, 1984.

[6] Jean-Pierre Aubin. *Mathematical methods of game and economic theory*. Courier Corporation, 2007.

[7] Yossi Aviv. A time-series framework for supply-chain inventory management. *Operations Research*, 51(2):210–227, 2003.

[8] Martin Neil Baily. Dynamic monopsony and structural change. *The American economic review*, 65(3):338–349, 1975.

[9] Martino Bardi and Italo Capuzzo-Dolcetta. *Optimal control and viscosity solutions of Hamilton-Jacobi-Bellman equations*. Springer Science & Business Media, 2008.

[10] Tamer Basar and Geert Jan Olsder. *Dynamic noncooperative game theory*, volume 23. Siam, 1999.

[11] Mokhtar S Bazaraa, Hanif D Sherali, and Chitharanjan M Shetty. *Nonlinear programming: theory and algorithms*. John Wiley & Sons, 2013.

[12] Moshe E Ben-Akiva, Song Gao, Zheng Wei, and Yang Wen. A dynamic traffic assignment model for highly congested urban networks. *Transportation research part C: emerging technologies*, 24:62–82, 2012.

[13] A Bensoussan. Lectures on stochastic control. In *Nonlinear Filtering and Stochastic Control*, pages 1–62. Springer, 1983.

[14] Alain Bensoussan, Shaokuan Chen, and Suresh P Sethi. The maximum principle for global solutions of stochastic stackelberg differential games. *SIAM Journal on Control and Optimization*, 53(4):1956–1981, 2015.

[15] Fred Espen Benth, Valery A Kholodnyi, and Peter Laurence. *Quantitative energy finance*. Springer, 2014.

[16] Oded Berman and David Perry. An eoq model with state-dependent demand rate. *European Journal of Operational Research*, 171(1):255–272, 2006.

[17] Fernando Bernstein and Awi Federgruen. Dynamic inventory and pricing models for competing retailers. *Naval Research Logistics (NRL)*, 51(2):258–274, 2004.

[18] Steven T Berry. Estimating discrete-choice models of product differentiation. *The RAND Journal of Economics*, pages 242–262, 1994.

[19] Dimitri P Bertsekas, Dimitri P Bertsekas, Dimitri P Bertsekas, and Dimitri P Bertsekas. *Dynamic programming and optimal control*, volume 1. Athena Scientific Belmont, MA, 1995.

[20] Dimitris Bertsimas and Georgia Perakis. Dynamic pricing: A learning approach. In *Mathematical and computational models for congestion charging*, pages 45–79. Springer, 2006.

[21] Venkataraman Bhaskar, Alan Manning, and Ted To. Oligopsony and monopsonistic competition in labor markets. *The Journal of Economic Perspectives*, 16(2):155–174, 2002.

[22] Venkataraman Bhaskar and Ted To. Oligopsony and the distribution of wages. *European Economic Review*, 47(2):371–399, 2003.

[23] C Bishop. Pattern recognition and machine learning (information science and statistics), 1st edn. 2006. corr. 2nd printing edn, 2007.

[24] Jean-Michel Bismut. Conjugate convex functions in optimal stochastic control. *Journal of Mathematical Analysis and Applications*, 44(2):384–404, 1973.

[25] Jean-Michel Bismut. Growth and optimal intertemporal allocation of risks. *Journal of Economic Theory*, 10(2):239–257, 1975.

[26] Gabriel Bitran and René Caldentey. An overview of pricing models for revenue management. *Manufacturing & Service Operations Management*, 5(3):203–229, 2003.

[27] Tudor Bodea and Mark Ferguson. *Segmentation, revenue management and pricing analytics*. Routledge, 2014.

[28] Olivier Bokanowski, Nicolas Forcadel, and Hasnaa Zidani. Reachability and minimal times for state constrained nonlinear problems without any controllability assumption. *SIAM Journal on Control and Optimization*, 48(7):4292–4316, 2010.

[29] Michael J Brennan. The supply of storage. *The American Economic Review*, 48(1):50–72, 1958.

[30] Alberto Bressan. *Lecture Notes on Functional Analysis*. American Mathematical Society, 2012.

[31] Alberto Bressan and Ke Han. Existence of optima and equilibria for traffic flow on networks. *Networks and Heterogeneous Media*, 8(3):627–648, 2013.

[32] Alberto Bressan and Benedetto Piccoli. *Introduction to the mathematical theory of control*, volume 2. American institute of mathematical sciences Springfield, 2007.

[33] Luigi Brighi and Reinhard John. Characterizations of pseudomonotone maps and economic equilibrium. *Journal of Statistics and Management Systems*, 5(1-3):253–273, 2002.

[34] George W Brown. Iterative solution of games by fictitious play. *Activity analysis of production and allocation*, 13(1):374–376, 1951.

[35] Arthur E Bryson and Yu-Chi Ho. Applied optimal control: optimization, estimation, and control [m]. *Washington, DC: Hemisphere*, 1975.

[36] Patrick Chan and Ronnie Sircar. Bertrand and cournot mean field games. *Applied Mathematics & Optimization*, 71(3):533–569, 2015.

[37] Kalyan Chatterjee and William Samuelson. Bargaining under incomplete information. *Operations Research*, 31(5):835–851, 1983.

[38] Fangruo Chen, Awi Federgruen, and Yu-Sheng Zheng. Near-optimal pricing and replenishment strategies for a retail/distribution system. *Operations Research*, 49(6):839–853, 2001.

[39] Zhigang Chen and Rebecca Lent. Supply analysis in an oligopsony model. *American Journal of Agricultural Economics*, 74(4):973–979, 1992.

[40] Roberto Cominetti, José R Correa, and Nicolás E Stier-Moses. The impact of oligopolistic competition in networks. *Operations Research*, 57(6):1421–1437, 2009.

[41] Engelbert Dockner. *Differential games in economics and management science*. Cambridge University Press, 2000.

[42] Xiaodong Du, L Yu Cindy, and Dermot J Hayes. Speculation and volatility spillover in the crude oil and agricultural commodity markets: A bayesian analysis. *Energy Economics*, 33(3):497–503, 2011.

[43] Darrell Duffie. *Dynamic Asset Pricing Theory*. Princeton University Press, 2010.

[44] Wedad Elmaghraby and Pınar Keskinocak. Dynamic pricing in the presence of inventory considerations: Research overview, current practices, and future directions. *Management Science*, 49(10):1287–1309, 2003.

[45] Philip Elmer-DeWitt. How apple became a monopsonist, July 2011. http://fortune.com/2011/07/05/how-apple-became-a-monopsonist/.

[46] Jacob Engwerda. *LQ dynamic optimization and differential games*. John Wiley & Sons, 2005.

[47] Lawrence C Evans. *Partial Differential Equations: Second Edition*. American Mathematical Society, 2010.

[48] Francisco Facchinei, Andreas Fischer, and Veronica Piccialli. On generalized nash games and variational inequalities. *Operations Research Letters*, 35(2):159–164, 2007.

[49] Francisco Facchinei and Jong-Shi Pang. *Finite-dimensional variational inequalities and complementarity problems*. Springer Science & Business Media, 2007.

[50] Maurizio Falcone and Roberto Ferretti. *Semi-Lagrangian approximation schemes for linear and Hamilton-Jacobi equations*, volume 133. SIAM, 2014.

[51] Anthony V Fiacco and Garth P McCormick. *Nonlinear programming: sequential unconstrained minimization techniques*, volume 4. Siam, 1990.

[52] Abdulai Fofana and Shabbar Jaffry. Measuring oligopsony power of uk salmon retailers. *Marine Resource Economics*, pages 485–506, 2008.

[53] Avner Friedman. *Differential games*. Courier Corporation, 2013.

[54] Terry L Friesz. *Dynamic optimization and differential games*, volume 135. Springer Science & Business Media, 2010.

[55] Terry L Friesz, David Bernstein, Tony E Smith, Roger L Tobin, and BW Wie. A variational inequality formulation of the dynamic network user equilibrium problem. *Operations Research*, 41(1):179–191, 1993.

[56] Terry L Friesz, Taeil Kim, Changhyun Kwon, and Matthew A Rigdon. Approximate network loading and dual-time-scale dynamic user equilibrium. *Transportation Research Part B: Methodological*, 45(1):176–207, 2011.

[57] Terry L Friesz, Changhyun Kwon, Tae Il Kim, Lifan Fan, and Tao Yao. Competitive robust dynamic pricing in continuous time with fixed inventories. *arXiv preprint arXiv:1208.4374*, 2012.

[58] Terry L Friesz, Javier Luque, Roger L Tobin, and Byung-Wook Wie. Dynamic network traffic assignment considered as a continuous time optimal control problem. *Operations Research*, 37(6):893–901, 1989.

[59] Terry L Friesz, Matthew A Rigdon, and Reetabrata Mookherjee. Differential variational inequalities and shipper dynamic oligopolistic network competition. *Transportation Research Part B: Methodological*, 40(6):480–503, 2006.

[60] Drew Fudenberg and Jean Tirole. Game theory. *Cambridge, Massachusetts*, 393, 1991.

[61] Guillermo Gallego and Ming Hu. Dynamic pricing of perishable assets under competition. *Management Science*, 60(5):1241–1259, 2014.

[62] Mauro Garavello, Ke Han, and Benedetto Piccoli. *Models for Vehicular Traffic on Networks*, volume 1. American institute of mathematical sciences Springfield, 2016.

[63] Nathan Gartner and Chronis Stamatiadis. Integration of dynamic traffic assignment with real-time traffic adaptive control system. *Transportation Research Record: Journal of the Transportation Research Board*, (1644):150–156, 1998.

[64] Rajna Gibson and Eduardo S Schwartz. Stochastic convenience yield and the pricing of oil contingent claims. *The Journal of Finance*, 45(3):959–976, 1990.

[65] Kenneth Gilbert. An arima supply chain model. *Management Science*, 51(2):305–310, 2005.

[66] Igor Vladimirovich Girsanov. *Lectures on mathematical theory of extremum problems*, volume 67. Springer Science & Business Media, 1972.

[67] Nicolas Hadjisavvas, Sándor Komlósi, and Siegfried S Schaible. *Handbook of generalized convexity and generalized monotonicity*, volume 76. Springer Science & Business Media, 2006.

[68] Patrick T Harker. Generalized nash games and quasi-variational inequalities. *European journal of Operational research*, 54(1):81–94, 1991.

[69] Chris Harris, Sam Howison, and Ronnie Sircar. Games with exhaustible resources. *SIAM Journal on Applied Mathematics*, 70(7):2556–2581, 2010.

[70] Richard F Hartl, Suresh P Sethi, and Raymond G Vickson. A survey of the maximum principles for optimal control problems with state constraints. *SIAM review*, 37(2):181–218, 1995.

[71] Ying Hu. Stochastic maimum principle. In John Baillieul and Tariq Samad, editors, *Encyclopedia of Systems and Control*. Springer-Verlag, London, 2014.

[72] John C Hull. *Options, futures, and other derivatives.* Pearson Education India, 2006.

[73] Aleksandr Davidovich Ioffe and Vladimir Mikhaïlovitch Tikhomirov. Theory of extremal problems. 1979.

[74] Rufus Isaacs. *Differential games: a mathematical theory with applications to warfare and pursuit, control and optimization*. Courier Corporation, 1999.

[75] Pushkin Kachroo, MV Aerde, and Kumar Krishen. Feedback control design for intelligent transportation systems. 1998.

[76] Pushkin Kachroo and Kaan Özbay. Feedback control solutions to network level user-equilibrium real-time dynamic traffic assignment problems. *Networks and Spatial Economics*, 5(3):243–260, 2005.

[77] Pushkin Kachroo and Kaan MA Özbay. *Feedback control theory for dynamic traffic assignment*. Springer Science & Business Media, 2012.

[78] Veijo Kaitala. Equilibria in a stochastic resource management game under imperfect information. *European Journal of Operational Research*, 71(3):439–453, 1993.

[79] Ioannis Karatzas and Steven Shreve. *Brownian motion and stochastic calculus*, volume 113. Springer Science & Business Media, 2012.

[80] Larry S Karp and Alex F McCalla. Dynamic games and international trade: an application to the world corn market. *American Journal of Agricultural Economics*, 65(4):641–650, 1983.

[81] Ani L Katchova, Ian M Sheldon, and Mario J Miranda. A dynamic model of oligopoly and oligopsony in the us potato-processing industry. *Agribusiness*, 21(3):409–428, 2005.

[82] Masoumeh Kazemi Zanjani, Mustapha Nourelfath, and Daoud Ait-Kadi. A multi-stage stochastic programming approach for production planning with uncertainty in the quality of raw materials and demand. *International Journal of Production Research*, 48(16):4701–4723, 2010.

[83] T Tony Ke, Zuo-Jun Max Shen, and J Miguel Villas-Boas. Search for information on multiple products. *Management Science*, 2016.

[84] Elias Koutsoupias and Christos Papadimitriou. Worst-case equilibria. In *Stacs*, volume 99, pages 404–413. Springer, 1999.

[85] Harold W Kuhn, John C Harsanyi, Reinhard Selten, Jorgen W Weibull, and Eric van Damme. The work of john nash in game theory. *Economic Sciences, 1991-1995*, 3:160, 1997.

[86] Ankur A Kulkarni and Uday V Shanbhag. New insights on generalized nash games with shared constraints: Constrained and variational equilibria. pages 151–156, 2009.

[87] Vidyadhar Kulkarni and Keqi Yan. Production-inventory systems in stochastic environment and stochastic lead times. *Queueing Systems*, 70(3):207–231, 2012.

[88] Jean-Michel Lasry and Pierre-Louis Lions. Mean field games. *Japanese Journal of Mathematics*, 2(1):229–260, 2007.

[89] Andrew Ledvina and Ronnie Sircar. Dynamic bertrand oligopoly. *Applied Mathematics & Optimization*, 63(1):11–44, 2011.

[90] J_L_ Lions and Guido Stampacchia. Variational inequalities. *Communications on pure and applied mathematics*, 20(3):493–519, 1967.

[91] Michael Ludkovski and Ronnie Sircar. Exploration and exhaustibility in dynamic cournot games. *European Journal of Applied Mathematics*, 23(03):343–372, 2012.

[92] Michael Ludkovski and Xuwei Yang. Dynamic cournot models for production of exhaustible commodities under stochastic demand. In *Commodities, Energy and Environmental Finance*, pages 371–396. Springer, 2015.

[93] David G Luenberger and Yinyu Ye. *Linear and nonlinear programming*, volume 2. Springer, 1984.

[94] Jin Ma, Philip Protter, and Jiongmin Yong. Solving forward-backward stochastic differential equations explicitlyâĂŤa four step scheme. *Probability Theory and Related Fields*, 98(3):339–359, 1994.

[95] Constantinos Maglaras and Joern Meissner. Dynamic pricing strategies for multiproduct revenue management problems. *Manufacturing & Service Operations Management*, 8(2):136–148, 2006.

[96] A G Malliaris and W A Brock. *Stochastic methods in economics and finance*. Elsevier Science, 1982.

[97] Victor Martínez-de Albéniz and Kalyan Talluri. Dynamic price competition with fixed capacities. *Management Science*, 57(6):1078–1093, 2011.

[98] Andreu Mas-Colell. *The theory of general economic equilibrium: A differentiable approach*. Number 9. Cambridge University Press, 1989.

[99] Helmut Maurer. *On the minimum principle for optimal control problems with state constraints*. Rechenzentrum d. Univ., 1979.

[100] Amir H Meimand. *Differential Stackelberg Games and Their Application to Dynamic Pricing, Production Planning Network Design, and Logistics*. PhD thesis, The Pennsylvania State University, 2013.

[101] Michel Minoux and Steven Vajda. *Mathematical programming: theory and algorithms*. Wiley New York, 1986.

[102] Reetabrata Mookherjee. *A computable theory of dynamic games and its applications*. PhD thesis, The Pennsylvania State University, 2006.

[103] Anna Nagurney. *Network economics: A variational inequality approach*, volume 10. Springer Science & Business Media, 2013.

[104] Serguei Netessine and Robert A Shumsky. Revenue management games: Horizontal and vertical competition. *Management Science*, 51(5):813–831, 2005.

[105] MS Nikol?skii. Convergence of the gradient projection method in optimal control problems. *Computational Mathematics and Modeling*, 18(2):148–156, 2007.

[106] DO Norris. Nonlinear programming applied to state-constrained optimization problems. *Journal of mathematical analysis and applications*, 43(1):261–272, 1973.

[107] Afshin Oroojlooyjadid, Lawrence Snyder, and Martin Takáč. Applying deep learning to the newsvendor problem. *arXiv preprint arXiv:1607.02177*, 2016.

[108] Jong-Shi Pang and David E Stewart. Differential variational inequalities. *Mathematical Programming*, 113(2):345–424, 2008.

[109] Markos Papageorgiou. Dynamic modeling, assignment, and route guidance in traffic networks. *Transportation Research Part B: Methodological*, 24(6):471–495, 1990.

[110] Srinivas Peeta and T-H Yang. Stability issues for dynamic traffic assignment. *Automatica*, 39(1):21–34, 2003.

[111] Srinivas Peeta and Athanasios K Ziliaskopoulos. Foundations of dynamic traffic assignment: The past, the present and the future. *Networks and Spatial Economics*, 1(3-4):233–265, 2001.

[112] Shige Peng. A general stochastic maximum principle for optimal control problems. *SIAM Journal on control and optimization*, 28(4):966–979, 1990.

[113] Georgia Perakis and Guillaume Roels. The price of anarchy in supply chains: Quantifying the efficiency of price-only contracts. *Management Science*, 53(8):1249–1268, 2007.

[114] Georgia Perakis and Anshul Sood. Competitive multi-period pricing for perishable products: A robust optimization approach. *Mathematical Programming*, 107(1-2):295–335, 2006.

[115] Lev Semenovich Pontryagin. *Mathematical theory of optimal processes*. CRC Press, 1987.

[116] Martin L Puterman. *Markov decision processes: discrete stochastic dynamic programming*. John Wiley & Sons, 2014.

[117] Nicolas Querou and Mabel Tidball. Incomplete information, learning, and natural resource management. *European Journal of Operational Research*, 204(3):630–638, 2010.

[118] Bin Ran, David E Boyce, and Larry J LeBlanc. A new class of instantaneous dynamic user-optimal traffic assignment models. *Operations Research*, 41(1):192–202, 1993.

[119] Klaus Ritter. Duality for nonlinear programming in a banach space. *SIAM Journal on Applied Mathematics*, 15(2):294–302, 1967.

[120] Julia Robinson. An iterative method of solving a game. *Annals of mathematics*, pages 296–301, 1951.

[121] R Tyrrell Rockafellar. Saddle points of hamiltonian systems in convex lagrange problems having a nonzero discount rate. *Journal of Economic Theory*, 12(1):71–113, 1976.

[122] Ralph Tyrell Rockafellar. *Convex analysis*. Princeton university press, 1970.

[123] J Ben Rosen. Existence and uniqueness of equilibrium points for concave n-person games. *Econometrica: Journal of the Econometric Society*, pages 520–534, 1965.

[124] Tim Roughgarden. *Selfish routing and the price of anarchy*, volume 174. MIT press Cambridge, 2005.

[125] Halsey Lawrence Royden and Patrick Fitzpatrick. *Real analysis*. Macmillan New York, 1988.

[126] Cynthia Rudin and Gah-Yi Vahn. The big data newsvendor: Practical insights from machine learning. *Available at SSRN 2559116*, 2014.

[127] Heinz Schättler. Optimal control with state space constraints. In *Encyclopedia of Optimization*, pages 968–974. Springer, 2009.

[128] Eduardo S Schwartz. The stochastic behavior of commodity prices: Implications for valuation and hedging. *The Journal of Finance*, 52(3):923–973, 1997.

[129] William E Schworm. Monopsonistic control of a common property renewable resource. *Canadian Journal of Economics*, pages 275–287, 1983.

[130] Atle Seierstad and Knut Sydsaeter. *Optimal control theory with economic applications*. Elsevier North-Holland, Inc., 1986.

[131] Suvrajeet Sen and Julia L Higle. An introductory tutorial on stochastic linear programming models. *Interfaces*, 29(2):33–61, 1999.

[132] SP Sethi and GL Thompson. Applied optimal control: Applications to management science. *Nijhoff, Boston*, 1981.

[133] Halil Mete Soner. Optimal control with state-space constraint i. *SIAM Journal on Control and Optimization*, 24(3):552–561, 1986.

[134] Halil Mete Soner. Optimal control with state-space constraint. ii. *SIAM journal on control and optimization*, 24(6):1110–1122, 1986.

[135] Wenjing Song, Ke Han, Yiou Wang, Terry L. Friesz, and Enrique del Castillo. Statistical metamodeling of dynamic network loading. *Transportation Research Part B: Methodological*, 2017.

[136] Kalyan T Talluri and Garrett J Van Ryzin. *The theory and practice of revenue management*, volume 68. springer, 2006.

[137] T Tian and JC Dunn. On the gradient projection method for optimal control problems with nonnegative l^2 inputs. *SIAM journal on control and optimization*, 32(2):516–537, 1994.

[138] Christian H Timpe and Josef Kallrath. Optimal planning in large multi-site production networks. *European Journal of Operational Research*, 126(2):422–435, 2000.

[139] Jean Tirole. *The theory of industrial organization*. MIT press, 1988.

[140] Shang Quang Tun and Sung Xung Li. Differential games of n players in stochastic systems with controlled diffusion terms. *Automation and Remote Control*, 69(5):874–890, 2008.

[141] Eliecer E Vargas and Dean F Schreiner. Modeling monopsony markets with regional cge: The oklahoma forest products industry case. *Journal of Regional Analysis and Policy*, 29(2):51–74, 1999.

[142] Xavier Vives. *Oligopoly pricing: old ideas and new tools*. MIT press, 2001.

[143] John Von Neumann and Oskar Morgenstern. *Theory of games and economic behavior*. Princeton university press, 2007.

[144] S Travis Waller and Athanasios K Ziliaskopoulos. A chance-constrained based stochastic dynamic traffic assignment model: Analysis, formulation and solution algorithms. *Transportation Research Part C: Emerging Technologies*, 14(6):418–427, 2006.

[145] Brian Wieland, Pat Mastrantonio, Sean P Willems, and Karl G Kempf. Optimizing inventory levels within intel's channel supply demand operations. *Interfaces*, 42(6):517–527, 2012.

[146] L. Williams, X. Liu, R. Mookherjee, and S. Zhong. Method and system for generating pricing recommendations, February 12 2013. US Patent 8,374,906.

[147] K Yan and V Kulkarni. Optimal inventory policies under stochastic production and demand rates. *Stochastic Models*, 24(2):173–190, 2008.

[148] Jiongmin Yong and Xun Yu Zhou. *Stochastic controls: Hamiltonian systems and HJB equations*, volume 43. Springer Science & Business Media, 1999.

[149] Jianyu Yu and Zohra Bouamra-Mechemache. Production standards, competition and vertical relationship. *European Review of Agricultural Economics*, 43(1):79–111, 2016.

Vita

Yiou Wang

Yiou Wang is a Ph.D. candidate in the Department of Industrial and Manufacturing Engineering at the Pennsylvania State University (Penn State). He received his B.S. in Physics from Peking University in Beijing, China in 2010.

Yiou is a student member of INFORMS.

He can be reached by e-mail: yiouwang2012@gmail.com.