Two Approaches to
Non-Zero-Sum Stochastic Differential Games of Control and Stopping

Qinghua Li

Advised by Professor Ioannis Karatzas

Submitted in partial fulfillment of the
requirement for the degree
of Doctor of Philosophy
in the Graduate School of Arts and Sciences

COLUMBIA UNIVERSITY

March 9, 2011

© Qinghua Li

2011

All Rights Reserved

Abstract

This dissertation takes two approaches – martingale and backward stochastic differential equation (BSDE) – to solve non-zero-sum stochastic differential games in which all players can control and stop the reward streams of the games. Existence of equilibrium stopping rules is proved under some assumptions.

The martingale part provides an equivalent martingale characterization of Nash equilibrium strategies of the games. When using equilibrium stopping rules, Isaacs' condition is necessary and sufficient for the existence of an equilibrium control set.

The BSDE part shows that solutions to BSDEs provide value processes of the games. A multidimensional BSDE with reflecting barrier is studied in two cases for its solution: existence and uniqueness with Lipschitz growth, and existence in a Markovian system with linear growth rate.

Contents

List of Tables

Acknowledgements

I am indebted to my advisor, Professor Ioannis Karatzas, a respectable scholar, enthusiastic teacher, kindhearted gentleman, and responsible senior coworker, for admitting me into his group and showing me the way to conduct research. His suggestions and supports are the most necessary condition for this dissertation. However, it is the challenging expectations from my advisor that consist of my most exciting experience in graduate school. Being a student of Professor Karatzas packs me with pride and pressure for a future career path.

I hereby announce gratitude to my family, for their century-long tradition in being decent intellectuals. My mother deserves special thanks, for providing me with a better environment of growing up, at the price of her own career in pharmacy.

Dedicated to My Youth

Chapter 1

Bibliographical Notes

Game theory is deeply rooted in history, benefiting since ancient times, yet human beings did not seem to have mathematically brought it to any higher level until early the twentieth century. Despite of the Nobel prize for John Nash's "Nash equilibrium", among other Nobel Laureates on game theory, twentieth century achievements on non-zero-sum games were not well known to the great majority until an Oscar film called *"A Beautiful Mind"* was release by *Universal Pictures* in the year 2001. If there has to be a simplest and understandable example on non-zero-sum game and Nash equilibrium, go watching the movie. Author of this dissertation was lucky enough to have heard the following dialogue (not every word exactly recorded).

(11th March 2009, Columbia University)

Kuhn : Don't learn game theory from the movie. The blonde thing is not a Nash equilibrium!
Odifreddi : How you invented the theory, I mean, the story about the blonde, was it real?
Nash : No!!!
Odifreddi : Did you apply game theory to win Alicia?
Nash : ...Yes...

(followed by 10 min's discussion on personal life and game theory)

1.1 From zero-sum to non-zero-sum games

1.1.1 Von Neumann and zero-sum games

Von Neumann moved to Princeton University in 1928, where he popularized the study of game theory. Von Neumann and coworkers' main achievements on games were about those where one player's reward is identically the other player's lost, formally called ”zero-sum” games.

In a zero-sum game, there are two players I and II, whose generic strategies are respectively denoted as s_1 and s_2, and $R(s_1, s_2)$ a reward for Player I and cost to Player II that is subject to the players' strategies (s_1, s_2). Simultaneously, Player I tries to maximize the reward and Player II minimizes it. Since Player I does not necessarily know the strategy that Player II is employing, he would maximize his reward with strategy s_1, assuming Player II makes it the least favorable by minimizing the same quantity R with strategy s_2. The resulting reward

$$\underline{V} = \sup_{s_1} \inf_{s_2} R(s_1, s_2) \tag{1.1.1}$$

is called ”lower value” of the zero-sum game. Symmetrically, Player II would minimize his cost in Player I's most favorable choice, resulting in ”upper value”

$$\overline{V} = \inf_{s_2} \sup_{s_1} R(s_1, s_2) \tag{1.1.2}$$

of the game.

Lower value of a game is apparently no larger than the upper value. When the two values identify with each other, they are called ”the value” of the game. The zero-sum game is then said to ”have a value”. The optimal pair of strategies (s_1^*, s_2^*) that achieves the upper and the lower values is called a ”saddle point”.

Definition 1.1.1 *A saddle point* (s_1^*, s_2^*) *of a zero-sum game is a single pair of strategies that attains both* sup inf *in the lower value (1.1.1) and* inf sup *in the upper value (1.1.2). When a saddle exists,*

$$\underline{V} = \sup_{s_1} \inf_{s_2} R(s_1, s_2) = \inf_{s_2} \sup_{s_1} R(s_1, s_2) = \overline{V}. \tag{1.1.3}$$

Another definition of a saddle point (s_1^*, s_2^*) respects the stability that the optimal strategies produce. When Player II employs strategy s_2^*, the strategy s_1^* had better maximize the reward R over all possible strategies for Player I. When Player I employs strategy s_1^*, the strategy s_2^* had better minimize the cost R over all possible strategies for Player II. This way, neither Player is likely to deviate from his optimal strategy.

Definition 1.1.2 *A saddle point* (s_1^*, s_2^*) *of a zero-sum game is a pair of strategies such that*

$$R(s_1, s_2^*) \le R(s_1^*, s_2^*) \le R(s_1^*, s_2). \tag{1.1.4}$$

The two definitions of a saddle are equivalent. If the pair of strategies (s_1^*, s_2^*) attains inf sup and sup inf, then

$$R(s_1^*, s_2^*) = \inf_{s_2} \sup_{s_1} R(s_1, s_2) = \inf_{s_2} R(s_1^*, s_2) \le R(s_1^*, s_2), \text{ for any } s_2, \tag{1.1.5}$$

and

$$R(s_1^*, s_2^*) = \sup_{s_1} \inf_{s_2} R(s_1, s_2) = \sup_{s_1} R(s_1, s_2^*) \ge R(s_1, s_2^*), \text{ for any } s_1. \tag{1.1.6}$$

Suppose (s_1^*, s_2^*) satisfies inequality (1.1.4). Taking supremum over all admissible strategies s_1 of Player I, and infimum over all admissible strategies s_2 of Player II, inequality (1.1.4) becomes

$$\sup_{s_1} R(s_1, s_2^*) \le R(s_1^*, s_2^*) \le \inf_{s_2} R(s_1^*, s_2). \tag{1.1.7}$$

Since $\inf_{s_2} \sup_{s_1} R(s_1, s_2) \le \sup_{s_1} R(s_1, s_2^*)$, and $\inf_{s_2} R(s_1^*, s_2) \le \sup_{s_1} \inf_{s_2} R(s_1, s_2)$, inequality (1.1.7) produces

$$\inf_{s_2} \sup_{s_1} R(s_1, s_2^*) \le R(s_1^*, s_2^*) \le \sup_{s_1} \inf_{s_2} R(s_1^*, s_2). \tag{1.1.8}$$

But that

$$\sup_{s_1} \inf_{s_2} R(s_1^*, s_2) \le \inf_{s_2} \sup_{s_1} R(s_1, s_2^*) \tag{1.1.9}$$

always holds true, there has to be

$$\sup_{s_1} \inf_{s_2} R(s_1, s_2) = R(s_1^*, s_2^*) = \inf_{s_2} \sup_{s_1} R(s_1, s_2). \tag{1.1.10}$$

As an example from finance, signing contract on one contingent claim is a zero-sum game. The writer's profit/loss is identically the buyer's loss/profit. Take a European call option $(S_T - K)^+$ for example. At maturity T, is the stock price \$ S_T is higher than the strike price K, the contract forces writer to sell the stock worthy of \$ S_T to the buyer at price \$ K. The buyer can sell the stock on the market at price \$ S_T. In this case the different \$ $(S_T - K)$ between market price and strike price is profit for the buyer and missed profit for the writer. If stock price falls below strike price at maturity, then the buyer does not need to take any action. From the writer's point of view, such a contract should not be delivered for free. He charges the buyer a price \$ P for the option to buy low. When stock price goes above strike price, the buyer wins and the seller loses \$ $((S_T - K) - P)$. When stock price goes below strike price, the buyer loses and the seller wins \$ P. How much writer of a contract should charge the buyer is the theme question answered by theories on option pricing.

1.1.2 John Nash and non-zero-sum games

Besides zero-sum games, there exist games with multiple players where the players' rewards do not necessarily sum up to a constant. Questions like how to reach some pleasant stability for all parties concerned lead to the development of non-zero-sum games.

In John Nash's 1949 one-page Nobel Prize winning paper, he wrote:

*One may define a concept of **AN n-PERSON GAME** in which each player has a finite set of pure strategies and in which a definite set of payments to the n players corresponds to each n-tuple of pure strategies, one strategy being taken by each player.*

One such n-tuple counters another if the strategy of each player in the countering n-tuple yields the highest obtainable expectation for its player against the $n-1$ strategies of the other players in the countered n-tuple. A self-countering n-tuple is called ***AN EQUILIBRIUM POINT****.*

Translating Nash's definitions into twenty-first century plain English. A non-zero-sum game is the game in which each player chooses a strategy as his best response to other players' strategies. An equilibrium is a set of strategies, such that, when applied, no player will profit from unilaterally changing his own strategy. Equivalently, the equilibrium was a fixed point of the mapping from a given set of strategies to the set of strategies as the players' best responses to the given set.

Definition 1.1.3 *In a non-zero-sum game of N Players, each player, indexed by i, can choose a strategy s_i. Player i receives a reward $R^i(s_1, \cdots, s_N)$ related to the N Players' strategies. An equilibrium $(s_1^*, \cdots, s_N^*)$ of the non-zero-sum game is a set of strategies, such that*

$$\begin{aligned}
&R^1(s_1^*, s_2^*, \cdots, s_N^*) \geq R^1(s_1, s_2^*, \cdots, s_N^*), \textit{ for any } s_1;\\
&R^2(s_1^*, s_2^*, \cdots, s_N^*) \geq R^2(s_1^*, s_2, \cdots, s_N^*), \textit{ for any } s_2;\\
&\vdots\\
&R^1(s_1^*, s_2^*, \cdots, s_N^*) \geq R^1(s_1^*, s_2^*, \cdots, s_N), \textit{ for any } s_N.
\end{aligned} \tag{1.1.11}$$

To credit Nash's formulation of this equilibrium, the equilibrium set of strategies as in Definition 1.1.3 is conventionally called "Nash equilibrium". It is indeed an equilibrium, for when imposed to all Players, no rational Player will want to change for a different strategy.

Nash equilibrium of a non-zero-sum game generalizes the Von Neumann-Morgenstern notion of saddle point of a zero-sum game.

Consider the zero-sum game in section 1.1.1, where Player I chooses strategy s_1 to maximize his reward $R(s_1, s_2)$, and Player II chooses strategy s_2 to minimize $R(s_1, s_2)$ as cost. But his minimizing the cost R is equivalent to Player II's maximizing $-R$. We may construct a 2-person non-zero-sum game with the two players' rewards $R^1 = R$ and $R^2 = -R$. When Player II employs strategy s_2^*, s_1^* maximizes Player I's reward R^1 over all possible strategies for Player I. When Player I employs strategy s_1^*, s_2^* maximizes Player I's reward R^2 over all possible strategies for Player II. Hence the pair of strategies (s_1^*, s_2^*) is a Nash equilibrium of the non-zero-sum game, which turns out to incorporate saddle point of a zero-sum game. The reasoning is summarized in Table 1.1.2.

	Player I	Player II	optimal (s_1^*, s_2^*)
0-sum	$\max_{s_1} R(s_1, s_2)$	$\min_{s_2} R(s_1, s_2)$	"saddle": $R(s_1, s_2^*) \leq R(s_1^*, s_2^*)$, $R(s_1^*, s_2^*) \leq R(s_1^*, s_2)$
0-sum	$\max_{s_1} R(s_1, s_2)$	$\max_{s_2} -R(s_1, s_2)$	$R(s_1^*, s_2^*) \geq R(s_1, s_2^*)$, $-R(s_1^*, s_2^*) \geq -R(s_1^*, s_2)$
non-0-sum	$\max_{s_1} R^1(s_1, s_2)$	$\max_{s_2} R^2(s_1, s_2)$	"equilibrium": $R^1(s_1^*, s_2^*) \geq R^1(s_1, s_2^*)$, $R^2(s_1^*, s_2^*) \geq R^2(s_1^*, s_2)$

Table 1.1.2: a saddle of a zero-sum game as a special case of an equilibrium of a non-zero-sum game.

Through the most popular one, Nash equilibrium is not the only optimality criteria for an N-player non-zero-sum game. Other optimality criteria include "efficient" and "in the core".

Definition 1.1.4 *A set of strategies $(s_1^*, \cdots, s_N^*)$ is said to be efficient of the N-player game, if for any set of strategies $(s_1, \cdots, s_N)$, there exists some Player i, such that his rewards*

$$R^i(s_1^*, \cdots, s_N^*) \geq R^i(s_1, \cdots, s_N). \tag{1.1.12}$$

The set $(s_1^*, \cdots, s_N^*)$ is said to be in the core, if for any index subset $I \subset \{1, \cdots, N\}$, there exists some Player i, such that his rewards

$$R^i(s_1^*, \cdots, s_N^*) \geq R^i(s_1, \cdots, s_N), \tag{1.1.13}$$

where $s_j = s_j^*$, for all $j \overline{\in} I$.

Efficient strategies cannot be modified to improve every player's situation. A set of strategies is in the core, if coalition within any lot cannot improve everyone in the lot while strategies of players outside of this lot remain the same. Strategies in the core are both Nash and efficient. Nash equilibrium and efficiency do not cover each other. This dissertation will focus on Nash equilibrium.

1.1.3 Popular approaches to stochastic differential games

Stochastic differential games are a family of dynamic, continuous time versions of games incorporating randomness in both the states and the rewards. States are random, described by an adapted diffusion process whose dynamics are known. To play a game, a player receives a running reward cumulated at some rate till the end of the game, and a terminal reward granted at the end of the game. The rewards are related to both the state process, and the controls at the choice of the players, as deterministic or random functions or functionals of them. A control represents a player's action in attempt to influence his rewards. Assuming his rationality, a player should certainly act in the most profitable way to his knowledge. Since the rewards can be random, they are usually measured in expectation, or some other more advanced criteria, for example variance as a measure of risk.

Depending on different settings, a game could never end, end at a finite deterministic time, or end at a random time. When the game is terminated at a random time, the random time is usually a stopping time, meaning up to any deterministic time, a player is informative enough to tell if he is to quit the game or not. One case of interest is to quit when the state process hits some boundary. The other case is letting a player determine the time to quit the game, based on his information up-to-date about the state process, about his own rewards, and even about other players' actions. In the latter case, a player is again assumed rational, seeking the best reward possible.

In a zero-sum game of timing, one player chooses a stopping time to maximize his reward received from the other player, and the other player chooses another stopping time to minimize the first player's reward as cost to him. Such a zero-sum game of stopping is called a "Dynkin game". It is the two-player game version of the optimal stopping problem, in practice the optimal exercise of an American contingent claim. Dynkin games are connected to singular controls, in the sense that, for convex cost functions, value function of the former games are derivatives of value functions of the latter. This connection was first observed by Taksar (1985) [47], followed by Fukushima and Taksar (2002) [25] in a Markovian setting by solving free-boundary problems, and Karatzas and Wang (2001) [36] in a non-Markovian setting based on weak compactness arguments.

In both zero-sum and non-zero-sum games, the existence and even choice of optimal controls largely relies on, if not equivalent to, the achievablity of the maximum or maxima of the reward functions. One may prove such achievabilities in zero-sum games, for example, in Beneš (1970) [1]. However, existence of an optimal control set that maximizes the Hamiltonians usually enters a non-zero-sum game as an equivalent condition, called Isaacs' condition, or Nash condition, for example in Davis (1979) [12].

Due to the nature of the problem, there have been at least three major approaches to solving stochastic differential games - partial differential equations, martingale techniques, and backward stochastic differential equations. Non-zero-sum stochastic differential games have not yet fallen out of these categories.

For Markovian rewards, which are functions of the current value of an underling diffusion state process, partial differential equations become a handy tool. Over the past thirty years, Bensoussan, Frehse and Friedman built a regularity theory of PDE's to study stochastic differential games. Among their extensive works, Bensoussan and Friedman (1977) considered in [6] games of optimal stopping. The existence of optimal stopping times of such games is reduced to the study of regular solutions of quasi-variational inequalities, assuming continuous and bounded running rewards and terminal rewards; Bensoussan and Frehse (2000) in [4] solved a non-zero-sum game of optimal controls, which is terminated when the state process exits a bounded domain. Their running rewards are quadratic forms of the controls. Fleming and Soner (1993) [23] give lectures on controlled Markov diffusions.

Under some regularity conditions and with uniqueness of solution in some sense, the HJB PDEs can be numerically implemented using the finite difference method. Duffie (2006) [15] is a good manual of finite difference methods for financial computations.

The martingale approach to characterizing optimal controls dates back to 1970's. The idea is exactly the one to derive Verification Theorems and Hamilton-Jacobi-Bellman equations: the expected reward is a supermartingale, and it is a martingale if and only if the control is optimal. The martingale approach allows the rewards to be path-dependant on the state process. Among others, there was a line of early works dealing with path-dependant rewards developing from optimization, through zero-sum games, and to non-zero-sum games, by Beneš (1970) [1] and (1971) [2], Duncan and Varaiya (1971) [16], Davis and Varaiya (1973) [13] and Davis (1973) [11]. See Davis (1979) [12] for a survey on the martingale method for optimal control problems.

To accommodate path-dependent rewards in games of stopping, Snell envelopes named after J. L. Snell for his 1952 work [46], instead of stopping regions for Markovian rewards (c.f. Shiryayev (1979) [45]), are used to derive optimal stopping rules. Snell envelope is the smallest supermartingale dominating the rewards, and is a martingale if stopped at the optimal stopping time. It is optimal to stop when, for the first time, terminal reward granted for early exercise meets the best expected reward over all possible stopping times. The martingale method also facilitates the study of zero-sum and non-zero-sum games of control and stopping, and is particularly useful if the rewards depend on the path of the state process. When there are terminal rewards only, Lepeltier and Etourneau (1987) in [40] used martingale techniques to provide sufficient conditions for the existence of optimal stopping times on processes that need not be Markovian. Their general theory requires some order assumption and supermartingale assumptions on the terminal reward. Karatzas and Zamfirescu (2008) in [38] took the martingale approach to characterize, then construct saddle points for zero-sum games of control and stopping. They also characterized the value processes by the semimartingale decompositions and proved a stochastic maximum principle for continuous, bounded running reward that can be a functional of the diffusion state process.

The martingale approach is very intuitive, revealing the essence of the problems.

As a tool for stochastic control theory, backward stochastic differential equations (BSDEs for short) were first proposed by Bismut in the 1970's. Pardoux and Peng (1990) proved in [43] existence and uniqueness of the solution to a BSDE with uniformly Lipschitz growth. El Karoui, Peng and Quenez (1997) [20] is a survey on BSDEs and their financial applications. Considerable attention has been devoted to studying the association between BSDEs and stochastic differential games. Cvitanić and Karatzas (1996) proved in [10] existence and uniqueness of the solution to the equation with double reflecting barriers, and associated their BSDE to a zero-sum Dynkin game. Their work generalized El Karoui, Kapoudjian, Pardoux, Peng and Quenez (1997) [19] on one-dimensional BSDE with one reflecting boundary, which captures early stopping features as that of American options. Hamadène and Lepeltier (2000) [29] and Hamadène (2006) [30] added controls to the Dynkin game studied by Cvitanić and Karatzas (1996) [10], the tool still being BSDE with double reflecting barriers. Markovian rewards of games correspond to the equations in the Markovian framework. Hamadène studied in [27] and [28] Nash equilibrium control with forward-backward SDE. In Hamadène, Lepeltier and Peng (1997) [26], the growth rates of their forward-backward SDE are linear in the value process and the volatility process, and polynomial in the state process. Their state process is a diffusion satisfying an ”$\mathbb{L}^2$-dominance” condition. These three authors solve a non-zero-sum game without stopping, based on existence result of the multi-dimensional BSDE.

BSDE’s are after all as much of an analytical tool as probabilistic. The privilege to use heavy analysis is an advantage of the BSDE approach, for it facilitates solving the optimization problems under looser technical conditions. There have also been plenty of works on numerical solutions to BSDE’s.

Readers interested in numerical methods for stochastic differential games are referred to works by H.J. Kushner and P. Dupuis.

1.2 Martingale techniques

In the stochastic differential game Problem 2.1.1 to be formulated in Chapter 2, a representative ith Player faces the optimization problem with expected reward

$$J_t(\tau, u) := \mathbb{E}^u[R_t(\tau, u)|\mathscr{F}_t], \tag{1.2.1}$$

when all other Players’ strategies are given. To simplify notation, this is a typical question of finding a stopping rule τ^* and control u^* to maximize the one-dimensional expected reward (2.2.1) over all stopping rules $\tau \in \mathscr{S}(t, T)$ and all admissible controls $u \in \mathscr{U}$. The reward process R is defined as

$$R_t(\tau, u) := \int_t^{\tau\wedge T} h(s, X, u_s)ds + L(\tau)\mathbb{1}_{\{\tau<T\}} + \xi\mathbb{1}_{\{\tau=T\}}, \tag{1.2.2}$$

Throughout this section, notations like J, R, h, L, and ξ are one-dimensional.

Classical martingale characterization of the optimization problem views optimal stopping and optimal control separately.

For every fixed admissible control u, denote

$$Q(t, u) = \mathbb{E}^u[R_0(\tau^*, u)|\mathscr{F}_t] = \sup_{\tau\in\mathscr{S}_{t,\rho}} J_t(\tau, u) + \int_0^t h(s, X, u_s)ds. \tag{1.2.3}$$

It is optimal to stop the first time when $Q(\cdot, u)$ meets $R_0(\cdot, u)$. The stopped supermartingale $Q(\cdot \wedge \tau, u)$ is a $\mathbb{P}^u$-supermartingale with respect to $\{\mathscr{F}_t\}_{0\le t\le\rho}$, and is a martingale if and only if τ is optimal. $Q(\cdot, u)$ is called the Snell envelope of $R_0(\cdot, u)$. Optimal stopping theory using Snell envelope does not require the reward being Markovian in the state process X.

For every fixed stopping time τ, denote

$$\begin{aligned} V(t,u) &:= \sup_{u\in\mathscr{U}} J_t(\tau,u) + \int_0^t h(s,X,u_s)ds \\ &= \sup_{u\in\mathscr{U}} \mathbb{E}^u[R_t(\tau,u)|\mathscr{F}_t] + \int_0^t h(s,X,u_s)ds. \end{aligned} \tag{1.2.4}$$

$V(\cdot, u)$ is a $\mathbb{P}^u$-supermartingale with respect to $\{\mathscr{F}_t\}_{0\le t\le\rho}$, for every u, and is a martingale if and only if u is optimal.

Once obtaining the supermartingale property of $V(\cdot, u)$, with the help of Doob-Meyer decomposition of super(sub)martingales, we can decompose

$$V(\cdot,u) = V(0,u) + M(\cdot,u) - A(\cdot,u) \tag{1.2.5}$$

as summation of a $\mathbb{P}^u$-martingale M and decreasing process $-A$. A martingale representation theorem further represents $M(\cdot, u) = \int_0^t (Z_s^u)' dB_s^u$ as a stochastic integral integral with respect to to the $\mathbb{P}^u$-standard Brownian motion B^u. It turns out that $Z^u = Z$ is irrelevant of u. $A(\cdot, u)$ can be shown to satisfy

$$A(t,u) - A(t,v) = -\int_0^t (H(s,X,Z_s,u_s) - H(s,X,Z_s,v_s))ds, \; 0 \le t \le \tau, \tag{1.2.6}$$

for any controls $u, v \in \mathscr{U}$. The function or functional H is the Hamiltonian defined as

$$H(t,X,z,u) := z\sigma^{-1}(t,X)f(t,X,u) + h(t,X,u). \tag{1.2.7}$$

Derived from the martingale property of the optimal control u^*, locally maximizing the Hamiltonian H equates to globally maximizing the expected reward. The latter is made of much more random noise than the former. The existence of a control u^* that maximizes the Hamiltonian is called "Isaacs' condition". Necessity of "Isaacs' condition" for maximizing the expected reward is called the "stochastic maximum principle".

For these martingale theorems to apply, most works so far assume boundedness of the rewards as a technical assumption, though the general belief is that the boundedness assumption can be relaxed. The arguments in the survey article Davis (1979) [12] indeed proceed as well if the rewards have at most polynomial growth in the supremum of the historical path of the state process.

Readers are referred to Karatzas and Shreve (1998) [34] for Snell envelopes of optimal stopping problems, to Karatzas and Shreve (1988) [33] and Revuz and Yor (1999) [44] for Brownian motion and continuous time martingales.

1.2.1 Snell envelope

A typical optimal stopping problem looks for a stopping rule τ^* that attains supremum in

$$Y(t) := \sup_{\tau\in\mathscr{S}(t,T)} \mathbb{E}[R_\tau|\mathscr{F}_t],\ 0 \le t \le T. \tag{1.2.8}$$

The terminal time T is finite. The filtration $\{\mathscr{F}_t\}_{0\le t\le T}$ satisfies the usual condition. The process $\{R_t\}_{0\le t\le T}$ is interpreted as a player's reward at every time t. The value process Y is the best expected reward possible the player could get by choosing to stop between current time t and terminal time T. If assuming R is bounded from below and right-continuous, then Y has an RCLL modification which shall still be denoted by the same symbol. The process Y is the smallest RCLL supermartingale dominating R. To credit Snell's contribution to solving this optimal stopping problem, Y is called the Snell envelope of R. If further more assuming $\mathbb{E}[\sup_{s\le t\le T} R_s|\mathscr{F}_t] < \infty$, the optimal stopping rule is

$$\tau^* = \inf\{t \le s \le T|R_s = Y_s\}, \tag{1.2.9}$$

the first time reward process R meets value process Y from below.

See Appendix D, Karatzas and Shreve (1998) [34] for detailed expositions of Snell envelope.

1.2.2 Doob-Meyer decomposition

The sum of a martingale and predictable, increasing (decreasing) process with respect to the same filtration is a supermartingale (submartingale). Whether the reverse claim is true or not raises the question of supermartingale (submartingale) decomposition.

For discrete time martingale, the answer is simple, for the two summands have been explicitly constructed.

Theorem 1.2.1 *(Doob decomposition) Any submartingale $Y = \{Y_n, \mathscr{F}_n\}_{n=0,1,\cdots}$ can be uniquely decomposed as*

$$Y_n = M_n + A_n, \tag{1.2.10}$$

the summation of a martingale $M = \{M_n, \mathscr{F}_n\}_{n=0,1,\cdots}$ and an predictable, increasing sequence $A = \{A_n, \mathscr{F}_n\}_{n=0,1,\cdots}$.

Proof. Taking $A_0 = 0$, and $A_{n+1} = A_n - Y_n + \mathbb{E}[Y_{n+1}|\mathscr{F}_n] = \sum_{k=0}^{n} \mathbb{E}([Y_{k+1}|\mathscr{F}_k] - Y_k)$. □

In continuous time, there has not been any analogue construction of the increasing (decreasing) process. A natural resort would be approximating continuous time martingales using the discrete time result. To show convergence of the approximating sequence of discrete time monotonic processes, additional assumptions are required. A most commonly used condition is a right-continuous supermartingale (submartingale) of class $\mathscr{DL}$ or class $\mathscr{D}$.

Definition 1.2.1 *The collection of all stopping times τ bounded between 0 and a finite positive number T (respectively, infinity) is denoted as $\mathscr{S}_{0,T}$ ($\mathscr{S}_{0,\infty}$). A right-continuous process $\{Y_t, \mathscr{F}_t\}_{t\geq 0}$ is said to be of class $\mathscr{D}$, if the family $\{Y_\tau\}_{\tau\in\mathscr{S}_{0,\infty}}$ is uniformly integrable; of class $\mathscr{DL}$, if the family $\{Y_\tau\}_{\tau\in\mathscr{S}_{0,T}}$ is uniformly integrable, for every $0 \leq T \leq \infty$.*

Theorem 1.2.2 *(Doob-Meyer decomposition) Let a filtration $\{\mathscr{F}_t\}_{t\geq 0}$ be right-continuous and such that $\mathscr{F}_0$ contains all $\mathbb{P}$-negligible sets in $\mathscr{F}$. If a right-continuous submartingale $Y = \{Y_t, \mathscr{F}_t\}_{t\geq 0}$ is of class $\mathscr{DL}$, then it admits the decomposition*

$$Y_t = M_t + A_t, t \geq 0, \tag{1.2.11}$$

as the summation of a right-continuous martingale $M = \{M_t, \mathscr{F}_t\}_{t\geq 0}$ and a predictable, increasing process $A = \{A_t, \mathscr{F}_t\}_{t\geq 0}$. Under the condition of predictability of process A, the decomposition is unique. Further, if Y is of class $\mathscr{D}$, the M is a uniformly integrable martingale and A is integrable.

Without the assumption of class $\mathscr{DL}$, the decomposition is also valid, but M being only a local martingale is the price to pay.

1.2.3 Martingale representation theorems

The Ito integral of an adapted, square-integrable process with respect to Brownian motion is a local martingale. Conversely, is a (local) martingale $\{M, \{\mathscr{F}_t\}\}$ a stochastic integral of some adapted, square-integrable process with respect to a certain Brownian motion? The answer is given by the martingale representation theorems.

In 1953, J. L. Doob answered yes. If M is a d-dimensional continuous local martingale on the filtered probability space $(\Omega, \mathscr{F}, \mathbb{P})$ with filtration $\{\mathscr{F}_t\}$, then one can construct, on a possibly extended $(\tilde{\Omega}, \tilde{\mathscr{F}}, \tilde{\mathbb{P}})$ with a possibly extended filtration $\{\tilde{\mathscr{F}}_t\}$, a d-dimensional Brownian motion W, and a $d \times d$ matrix Z of measurable, adapted, square-integrable process, such that $\mathbb{P}$-a.s. M has the representation

$$M_t = \int_0^t Z_s dW_s, \tag{1.2.12}$$

as the stochastic integral of Z with respect to the Brownian motion W, which is not prefixed. The Brownian motion W is constructed according to the local martingale M. Since the the original probability space $(\Omega, \mathscr{F}, \mathbb{P})$ might not be enough to support the Brownian motion required for the representation, an extension might be necessary.

Preferrably, we would like all martingales on the same filtered probability space be stochastic integrals with respect to one single Brownian motion. This is true if the (augmented) filtration is Brownian. If M is a d-dimensional RCLL, square-integrable martingale on the filtered probability space $(\Omega, \mathscr{F}, \mathbb{P})$ with (augmented) filtration $\{\mathscr{F}_t\}$ generated by a Brownian motion B, then there exists a $d \times d$ matrix Z of measurable, adapted, square-integrable process, such that $\mathbb{P}$-a.s. M has the representation

$$M_t = \int_0^t Z_s dB_s. \tag{1.2.13}$$

The Brownian motion B is the same for all $\mathbb{P}$-martingales on $\{\mathscr{F}_t\}$. It is the given Brownian motion that generate $\{\mathscr{F}_t\}$.

In the setting of Chapter 2, we need to represent the martingale part $M(\cdot,u)$ of the value process $V(\cdot,u)$ in (1.2.5) to solve the optimization problem. $M(\cdot,u)$ is a $\mathbb{P}^u$-martingale with respect to $\{\mathscr{F}_t\}$. It had better take up the integral form

$$M_t = \int_0^t Z_s^u dB_s^u, \tag{1.2.14}$$

for some $\mathbb{P}^u$-standard Brownian motion B^u. This has been confirmed with (Theorem 3.1, Fujisaki, Kallianpur and Kunita (1972) [24]). Their result is not covered by the previous two representation theorems, because the Brownian motion B^u is the prefixed drifted $\mathbb{P}$-B.M. and standard $\mathbb{P}^u$-B.M. defined in (3.1.10), and $\{\mathscr{F}_t\}$ is not necessarily generated by B^u due to randomness in the drift coefficient.

1.3 Backward stochastic differential equations

Backward stochastic differential equations were proposed for the first time in Bismut (1973) [7] as means to solve stochastic optimal control problems. The two subjects agree in terms of seeking adapted strategies to achieve a terminal goal. In the setting of Chapter 3, the value process of a BSDE turns out to be the value process of a control problem, the proof of which require boundedness of the rewards in most previous works. Terminal reward of the control problem corresponds to terminal value of the equation, Hamiltonian corresponds to the driver, and early exercise rewards corresponds to reflecting boundaries. Once linked a backward equation, not only probabilistic tools but also analytical techniques can help. This BSDE approach offers more flexibility, though somewhat less intuitive.

We will focus on how different types of BSDE's are connected to various optimal control and stopping problems, as summarized in tables (1.1) and (1.2) below.

Table 1.1: correspondence between types of stochastic differential games and types of BSDE's.

Game	BSDE
one person's optimal control	1-dim, no reflection
zero-sum game of control	1-dim, no reflection
one person's optimal stopping	1-dim, lower reflecting boundary
Dynkin game (zero-sum game of stopping)	1-dim, double reflecting boundary
N-player non-zero-sum game of control	N-dim, no reflection
risk-sensitive control	quadratic driver

Table 1.2: correspondence between parameters of stochastic differential games and parameters of BSDE's.

Game	BSDE
number of rewards to optimize	dimension
value process	value process
Hamiltonian	driver
maximum duration of game	terminal time
terminal reward	terminal value
early exercise reward	reflecting boundary
regret from suboptimal exercise time	the increasing process
Brownian noise from state process	Brownian noise added to the equation
instantaneous volatility of value process	volatility process

1.3.1 Birth of BSDE

A control problem with expected reward

$$J_t(u,v) = \mathbb{E}^{u,v}[\int_t^T h(s,X,u_s,v_s)ds + \xi|\mathscr{F}_t] \tag{1.3.1}$$

identifies with BSDE

$$Y^{u,v}(t) = \xi + \int_t^T H(s,X,Z^{u,v}(s),(u,v)(t,X,Z^{u,v}(s))ds - \int_t^T Z^{u,v}(s)dB_s, \tag{1.3.2}$$

in the sense that the two processes $J(u,v)$ and $Y^{u,v}$ coincide. For the control problem, h is an instantaneous reward rate, and ξ is the fixed terminal reward at time T. The Hamiltonian H is defined as

$$H(t,X,z,u,v) := z\sigma^{-1}(t,X)f(t,X,u,v) + h(t,X,u,v). \tag{1.3.3}$$

By starting from a simplified version which can be solved by martingale representation, Pardoux and Peng (1990) [43] used Picard iteration to show existence of an adapted solution, and similar inequalities used in the iteration to show uniqueness of such solutions, to a backward equation of the general form

$$Y(t) = \xi + \int_t^T g(s, Y(s), Z(s))ds - \int_t^T Z(s)dB_s, \tag{1.3.4}$$

where the function g is uniformly Lipschitz in the Y and Z arguments. The two parameters g and ξ in (1.3.4) are called ”terminal value” and ”driver” of the BSDE. The solution consists of ”value process” Y and ”volatility process” Z.

Both existence and uniqueness can alternatively be proven at the same time by the contraction method as in El Karoui, Peng and Quenez (1997) [20]. They first pick two arbitrary adapted, square-integrable processes Y^0 and Z^0 in the driver g to solve the equation

$$Y^1(t) = \xi + \int_t^T g(s, Y^0(s), Z^0(s))ds - \int_t^T Z^1(s)dB_s. \tag{1.3.5}$$

for (Y^1, Z^1). As in Pardoux and Peng (1990) [43], the process Z^1 comes from representation of the martingale

$$\mathbb{E}\left[\xi + \int_0^T g(s, Y^0(s), Z^0(s))ds \middle| \mathscr{F}_t\right] = \mathbb{E}\left[\xi + \int_0^T g(s, Y^0(s), Z^0(s))ds\right] + \int_0^t Z^1(s)dB_s. \tag{1.3.6}$$

The process Y^1 is defined as

$$\begin{aligned} Y^1(t) =& \mathbb{E}\left[\xi + \int_t^T g(s, Y^0(s), Z^0(s))ds \middle| \mathscr{F}_t\right] \\ =& \mathbb{E}\left[\xi + \int_0^T g(s, Y^0(s), Z^0(s))ds \middle| \mathscr{F}_t\right] - \int_0^t g(s, Y^0(s), Z^0(s))ds. \end{aligned} \tag{1.3.7}$$

The contract method argues existence and uniqueness of solution to equation (1.3.4) by proving the mapping from (Y^0, Z^0) to (Y^1, Z^1) is a contraction, thus having a unique fixed point (Y, Z). The fixed point solves equation (1.3.4).

The contraction method is equivalent to Pardoux and Peng’s 1990 proof. Besides measurabilities and integrabilities, a crucial technical assumption of the two proofs is the driver g being Lipschitz in both value process y and volatility process z, uniformly in time t.

Under those assumptions above, and in dimension one, Comparison Theorem (section 2.2, El Karoui, Peng and Quenez (1997) [20]) states that a larger terminal value and a larger driver will produce a larger value process of a BSDE. Conversely, that a larger value process has to be produced by a larger terminal value and a larger driver is called the Converse Comparison Theorem. Briand, Coquet, Hu, Mémin and Peng (2000) [8] proved a Converse Comparison Theorem for one-dimensional BSDE with Lipschitz driver. Comparison Theorems and the converse, when holding true, determines a necessary and sufficient condition for the optimal control(s).

An optimization problem considers one control u only and the other control v disappears in (1.3.1). An optimal control u^* is chosen among all admissible controls to maximize

$$J_t(u) = \mathbb{E}^u \left[\int_t^T h(s, X, u_s) ds + \xi \middle| \mathscr{F}_t \right]. \tag{1.3.8}$$

If the rewards h and ξ are bounded, the value process Y^u of BSDE

$$Y^u(t) = \xi + \int_t^T H(s, X, Z^u(s), u(t, X, Z^u(s)) ds - \int_t^T Z^u(s) dB_s \tag{1.3.9}$$

can be shown to equal $J_t(u)$ in (1.3.8), with Hamiltonian H defined as

$$H(t, X, z, u) := z\sigma^{-1}(t, X) f(t, X, u) + h(t, X, u). \tag{1.3.10}$$

When technical conditions are satisfied, maximizing Hamiltonian H is equivalent to maximizing value process Y^u, which equals expected reward $J(u)$. Hence a control u^* is optimal if and only if u^* maximizes $H(t, x, z, u)$ among all admissible controls. Beneš (1970) [1] proved achievablity of the Hamiltonian by a measurable u^*.

In a zero-sum game with expected rewards (1.3.1), one player chooses control u to maximize $J(u, v)$, and the other player chooses control v to minimize $J(u, v)$. A saddle point is a pair of controls (u^*, v^*) such that

$$J(u, v^*) \leq J(u^*, v^*) \leq J(u^*, v). \tag{1.3.11}$$

If existing, the saddle (u^*, v^*) attains superema and infima, and identifies sup inf and inf sup in

$$\sup_u \inf_v J(u, v) = \inf_v \sup_u J(u, v). \tag{1.3.12}$$

With bounded rewards, if a control pair (u^*, v^*) satisfy

$$H(t, x, z, u, v^*) \leq H(t, x, z, u^*, v^*) \leq H(t, x, z, u^*, v), \tag{1.3.13}$$

then (u^*, v^*) is a saddle point of the zero-sum game. Existence of controls that maximize or minimize the Hamiltonians in a way like (1.3.13) is called ”Isaacs’ condition”. Necessity of Isaacs’ condition is called the ” Stochastic Maximum Principle”. Comparison Theorem of BSDE’s is used to derive sufficiency of Isaacs’ condition, and converse Comparison Theorem for the maximum principle.

A case in optimal control that receives more special treatments is the Markovian case. In the Markovian framework, where the state process X is the solution to a forward SDE

$$X_t = X_0 + \int_0^t f(s, X_s, u_s) ds + \int_0^t \sigma(s, X_s) dB_s^u, \; 0 \leq t \leq T, \tag{1.3.14}$$

and where rewards are functions of the state process X as in

$$J_t(u, v) = \mathbb{E}^{u,v} \left[\int_t^T h(s, X_s, u_s, v_s) ds + \xi \middle| \mathscr{F}_t \right], \tag{1.3.15}$$

expected reward $J_t(u,v)$ is a function of the time t and the current value of the state process $X_t = x$.

Corresponding to Markovian setting of a control problem, there is the forward-backward system of stochastic differential equations (FBSDE)

$$\begin{cases} X^{t,x}(s) = x,\ 0 \le s \le t; \\ dX^{t,x}(s) = \sigma(s, X^{t,x}(s))' dB_s,\ t < s \le T, \end{cases} \tag{1.3.16}$$

and

$$Y^{t,x}(s) = \xi(X^{t,x}(T)) + \int_s^T g(r, X^{t,x}(r), Y^{t,x}(r), Z^{t,x}(r))dr - \int_s^T Z^{t,x}(r)dB_r,\ t \le s \le T. \tag{1.3.17}$$

As an application of Ito's formula, if a function y solves the PDE

$$\begin{aligned} &\partial_t y(t,x) + \mathscr{A} y(t,x) + g(t, x, y(t,x), \sigma'(t,x)\partial_x y(t,x)) = 0; \\ &y(T,x) = \xi(x), \end{aligned} \tag{1.3.18}$$

where $\mathscr{A}$ is the infinitesimal generator

$$\mathscr{A}_{t,x} = \sum_{i,j} \frac{1}{2}(\sigma\sigma')_{ij}(t,x)\partial^2_{x_i x_j} + \sum_i f_i(t,x)\partial_{x_i}, \tag{1.3.19}$$

then

$$(Y^{t,x}_\cdot, Z^{t,x}_\cdot) = (y(\cdot, X^{t,x}_\cdot), \sigma'(t, X^{t,x}_\cdot)\partial_x y(\cdot, X^{t,x}_\cdot)) \tag{1.3.20}$$

solves the forward-backward system (1.3.16) and (1.3.17).

PDE (1.3.18) is the renowned Feynman-Kac formula that links PDE's to probability.

1.3.2 The role of reflecting boundaries

An optimal stopping problem looks for a stopping time to maximize the expected reward

$$J_t(\tau) = \mathbb{E}\left[\int_t^\tau h(s,X)ds + L(\tau)\mathbb{1}_{\{\tau<T\}} + \xi\mathbb{1}_{\{\tau=T\}} \middle| \mathscr{F}_t \right]. \tag{1.3.21}$$

In addition to running reward cumulated at rate h, if a player sticks to the end of the game, he receives a terminal reward ξ; if he decides to quit at any earlier stopping time τ, then terminal reward ξ is replaced by an early exercise reward L related to time τ of quitting.

If the early exercise reward L is progressively measurable and continuous, and as time is up if $L(T-)$ is not above terminal reward ξ, the solution to the BSDE

$$\begin{cases} Y(t) = \xi + \int_t^T h(s)ds - \int_t^T Z(s)dB_s + K(T) - K(t); \\ Y(t) \ge L(t),\ 0 \le t \le T,\ \int_0^T (Y(t) - L(t))dK(t) = 0 \end{cases} \tag{1.3.22}$$

provides the value process of the optimal stopping problem (1.3.21) and its optimal stopping rule. That is,

$$Y(t) = \sup_{\tau \in \mathscr{S}_t} \mathbb{E}\left[\int_t^\tau h(s)ds + L(\tau)\mathbb{1}_{\{\tau<T\}} + \xi \mathbb{1}_{\{\tau=T\}} \middle| \mathscr{F}_t \right]. \tag{1.3.23}$$

Since a player can always quit immediately at time t and leave with an early exercise reward $L(t)$, the maximum reward he could get never falls below L. The optimal stopping rule can be shown as

$$\tau^* = \inf\{t \leq s \leq T : Y(s) \leq L(t)\} \wedge T, \tag{1.3.24}$$

the first time when early exercise reward meets the best reward possible from below. Intuitively, the process K is interpreted as the cumulative profit missed for sticking to the game after the optimal time to quit, hence being increasing in time t. When playing the game before the optimal stopping time when Y meets L, there is no regret, so K is flat. If the player is asleep at the optimal stopping time, he suffers from earning less profit than could be, so K increase accordingly. Seeing from the equation (1.3.22), whenever the value process Y is about to drop below L, the increasing process K kicks Y up with a minimal strength.

The process L in (1.3.22) is called a ”reflecting boundary”, ”reflecting barrier”, or simply ”obstacle”. Since the value process in the optimization problem can never be smaller than the early exercise reward L, L is referred to as a ”lower reflecting boundary”. A reflecting boundary is an additional term in BSDE’s to accommodate an early exercise privilege in optimization problems. A BSDE with a reflecting boundary or reflecting boundaries is said to be reflected.

A general form of equation (1.3.22), the reflected BSDE

$$\begin{cases} Y(t) = \xi + \displaystyle\int_t^T g(s, Y(s), Z(s))ds - \int_t^T Z(s)dB_s + K(T) - K(t); \\ Y(t) \geq L(t),\ 0 \leq t \leq T,\ \displaystyle\int_0^T (Y(t) - L(t))dK(t) = 0. \end{cases} \tag{1.3.25}$$

has been solved by El Karoui, Kapoudjian, Pardoux, Peng and Quenez (1997) [19], in dimension one. With Lipschitz driver g, the solution Y to the equation (1.3.25) is connected to the optimal stopping problem as

$$Y(t) = \sup_{\tau \in \mathscr{S}_t} \mathbb{E}\left[\int_t^\tau g(s, Y(s), Z(s))ds + L(\tau)\mathbb{1}_{\{\tau<T\}} + \xi \mathbb{1}_{\{\tau=T\}} \middle| \mathscr{F}_t \right]. \tag{1.3.26}$$

The optimal stopping rule τ^* is the first hitting time of the lower reflecting boundary.

$$\tau^* = \inf\{t \leq s \leq T : Y(s) \leq L(t)\} \wedge T. \tag{1.3.27}$$

El Karoui, Kapoudjian, Pardoux, Peng and Quenez (1997) [19] demonstrated existence of solution to equation (1.3.25) with two methods - contraction and penalization.

As for equation (1.3.4) without reflection, the contraction method views solution (Y, Z) to equation(1.3.25) as a fixed point of the contraction mapping from two arbitrary adapted, square-integrable processes Y^0 and Z^0 to (Y^1, Z^1) defined via

$$\begin{cases} Y^1(t) = \xi + \int_t^T g(s, Y^0(s), Z^0(s))ds - \int_t^T Z^1(s)dB_s + K^1(T) - K^1(t); \\ Y^1(t) \geq L(t), 0 \leq t \leq T, \int_0^T (Y^1(t) - L(t))dK^1(t) = 0. \end{cases} \tag{1.3.28}$$

With the help of the theory on optimal stopping reviewed in section 1.2.1, the conditional expectation

$$\mathbb{E}\left[\int_0^{\tau^*} g(s, Y^0(s), Z^0(s))ds + L(\tau^*)\mathbb{1}_{\{\tau^*<T\}} + \xi\mathbb{1}_{\{\tau^*=T\}} \middle| \mathscr{F}_t\right] \tag{1.3.29}$$

with optimal stopping time τ^* from (1.3.27) is a supermartingale, hence admitting the Doob-Meyer decomposition of continuous time martingales

$$\begin{aligned} &\mathbb{E}\left[\int_0^{\tau^*} g(s, Y^0(s), Z^0(s))ds + L(\tau^*)\mathbb{1}_{\{\tau^*<T\}} + \xi\mathbb{1}_{\{\tau^*=T\}} \middle| \mathscr{F}_t\right] \\ =&\mathbb{E}\left[\int_0^{\tau^*} g(s, Y^0(s), Z^0(s))ds + L(\tau^*)\mathbb{1}_{\{\tau^*<T\}} + \xi\mathbb{1}_{\{\tau^*=T\}}\right] + \int_0^t Z^1(s)dB_s - K^1(t). \end{aligned} \tag{1.3.30}$$

The term K^1 is the increasing process from the decomposition, and Z^1 comes from representation of the martingale part. Define a process Y^1 by

$$Y^1(t) = \mathbb{E}\left[\int_t^{\tau^*} g(s, Y^0(s), Z^0(s))ds + L(\tau^*)\mathbb{1}_{\{\tau^*<T\}} + \xi\mathbb{1}_{\{\tau^*=T\}} \middle| \mathscr{F}_t\right]. \tag{1.3.31}$$

The triple (Y^1, Z^1, K^1) satisfies (1.3.28).

The penalization method views solution (Y, Z) to equation(1.3.25) as strong limit of solutions $\{(Y^n, Z^n)\}_{n=1}^{\infty}$ to the penalized equations

$$Y^n(t) = \xi + \int_t^T g(s, Y^n(s), Z^n(s))ds - \int_t^T Z^n(s)dB_s + n\int_t^T (Y^n(s) - L^n(s))^- ds. \tag{1.3.32}$$

BSDE (1.3.32) is the non-reflected one solved by Pardoux and Peng (1990) [43]. Proof of convergence mainly relies on Comparison Theorem to guarantee that the sequence $\{Y^n\}$ is increasing hence having a pointwise limit. Lipschitz condition on the driver g is also required for uniform $\mathbb{L}^2$ boundedness of $\{Y^n\}$.

In dimension one, El Karoui, Kapoudjian, Pardoux, Peng and Quenez (1997) [19] is able to prove the Comparison Theorem for the reflected equation.

With a reflecting boundary, Feynman-Kac formula (1.3.18) for the forward-backward Markovian system (1.3.16) and (1.3.17) is modified to be a variational inequality. If a function y solves the variational inequality

$$\begin{aligned} &\max\{L(t,x)-y(t,x), \partial_t y(t,x)+\mathscr{A}y(t,x)+g(t,x,y(t,x),\sigma'(t,x)\partial_x y(t,x))\}=0;\\ &y(T,x)=\xi(x), \end{aligned} \tag{1.3.33}$$

where $\mathscr{A}$ is the infinitesimal generator in (1.3.19), then $(Y^{t,x}_\cdot, Z^{t,x}_\cdot)$ as in (1.3.20) satisfies the system of forward equation (1.3.16) and backward equation

$$\begin{aligned} Y^{t,x}(s)=&\xi(X^{t,x}(T))+\int_s^T g(r,X^{t,x}(r),Y^{t,x}(r),Z^{t,x}(r))dr-\int_s^T Z^{t,x}(r)dB_r\\ &+K^{t,x}(T)-K^{t,x}(s),\ t\leq s\leq T. \end{aligned} \tag{1.3.34}$$

Rigorous discussion of the variational inequality can be found in section 8, El Karoui, Kapoudjian, Pardoux, Peng and Quenez (1997) [19].

A Dynkin game is a zero-sum game of stopping, initiated by Dynkin and Yushkevich (1968) [17]. Consider a Dynkin game with payoff

$$R_t(\tau,\rho)=\int_t^{\tau\wedge\rho} h(s,X)ds+L(\tau)\mathbb{1}_{\{\tau<T,\tau\leq\rho\}}+U(\rho)\mathbb{1}_{\{\rho\leq\tau\}}+\xi\mathbb{1}_{\{\tau\wedge\rho=T\}}. \tag{1.3.35}$$

Player I chooses stopping time τ at which he quits the game. Player II chooses stopping time ρ. Soon as either player quits, the game is ended. The payoff $R(\tau,\rho)$ is the amount that Player II pays Player I at the end of the game. If Player I whistles to end the game at time τ before Player II does, he receives amount $L(\tau)+\xi$ from Player II. If Player II quits the game first, he pays Player I amount $U(\rho)+\xi$. The random quantity $R(\tau,\rho)$ is reward for Player I and cost to Player II, which should therefor be maximized by Player I and minimized by Player II. To average over all scenarios, optimize instead the expected payoff

$$J_t(\tau,\rho)=\mathbb{E}[\int_t^{\tau\wedge\rho} h(s,X)ds+L(\tau)\mathbb{1}_{\{\tau<T,\tau\leq\rho\}}+U(\rho)\mathbb{1}_{\{\rho\leq\tau\}}+\xi\mathbb{1}_{\{\tau\wedge\rho=T\}}|\mathscr{F}_t]. \tag{1.3.36}$$

Saddle point of this Dynkin game is a pair of stopping times (τ^*,ρ^*), such that

$$J(\tau,\rho^*)\leq J(\tau^*,\rho^*)\leq J(\tau^*,\rho). \tag{1.3.37}$$

The saddle (τ^*,ρ^*) attains superema and infima, and identifies lower value sup inf and upper value inf sup in

$$V:=\underline{V}=\sup_\tau\inf_\rho J(\tau,\rho)=\inf_\rho\sup_\tau J(\tau,\rho)=\overline{V}. \tag{1.3.38}$$

In case Player I chooses to stop immediately at current time t, he receives payoff $L(t)$ from Player II. In case Player II chooses to stop immediately at current time t, he pays payoff $U(t)$ to Player I. When existing, value V of the game as Player I's maximum reward and Player II's minimum cost is always above L and below U. For this Dynkin game, it suffices to consider only early exercise rewards $L \leq U$.

When early exercise rewards L and U are continuous, the value process Y of the doublely reflected BSDE

$$\begin{cases} Y(t) = \xi + \int_t^T h(s,X)ds - \int_t^T Z(s)dB_s + K^+(T) - K^+(t) - (K^-(T) - K^-(t)); \\ L(t) \leq Y(t) \leq U(t), 0 \leq t \leq T, \ \int_0^T (Y(t) - L(t))dK(t) = \int_0^T (U(t) - Y(t))dK(t) = 0 \end{cases} \tag{1.3.39}$$

provides the value process V of Dynkin game with expected payoff (1.3.36). The increasing process K^+ is the minimal force that maintains value process Y above lower reflecting boundary L. K^+ is an additional term for early exercise privilege at time τ by Player I, to maximize his reward. For early exercise privilege at time ρ by Player II to minimize his cost, a minimal cumulative force K^-, which is an increasing process, pushes value process Y downwards whenever it hits upper reflecting boundary U from below.

The connection between Dynkin games and doublely reflected BSDE's was explored in Cvitanic and Karatzas (1996) [10]. They proved existence and uniqueness of solution to the equation

$$\begin{cases} Y(t) = \xi + \int_t^T g(s,Y(s),Z(s))ds - \int_t^T Z(s)dB_s + K^+(T) - K^+(t) -^- (T) - K^-(t)); \\ L(t) \leq Y(t) \leq U(t), 0 \leq t \leq T, \ \int_0^T (Y(t) - L(t))dK(t) = \int_0^T (U(t) - Y(t))dK(t) = 0. \end{cases} \tag{1.3.40}$$

with Lipschitz driver g. The authors demonstrate uniqueness of the solution with both contraction and penalization methods.

1.3.3 Growth rates beyond Lipschitz

Risk-sensitive controls were initiated by Whittle, Bensoussan and coworkers, among others. Receiving a controlled random reward R, a risk-sensitive player takes not only the expectation but also the variance of his reward into consideration. El Karoui and Hamadène (2003) in [18] link risk-sensitive control problems to BSDE's with an additional term quadratic in the volatility process.

Consider a general risk preference coefficient θ. For the Player with reward process $R_t(u)$ controlled by u, the quantity

$$\mathbb{E}^u[R_t(u)|\mathscr{F}_t] + \frac{\theta}{2} Var^u[R_t(u)|\mathscr{F}_t] \tag{1.3.41}$$

is about equal to

$$\frac{1}{\theta}\ln \mathbb{E}^u[\exp\{\theta R_t(u)\}|\mathscr{F}_t], \tag{1.3.42}$$

when absolute value $|\theta|$ is small. If $\theta > 0$ (< 0), a larger variance contributes to a larger (small) expected reward, hence a higher risk is more (less) preferable to the Player. The Player is called risk-prone if $\theta > 0$, and risk-averse if $\theta < 0$. If $\theta = 0$, the variance term disappear from the expected reward, then the Player is said to be risk-neutral. So, instead of maximizing the expected reward, our risk-sensitive Player maximizes his expected exponential reward

$$J_t(u) = \mathbb{E}^u[\exp\{\theta R_t(u)\}|\mathscr{F}_t]. \tag{1.3.43}$$

Let $R_t(u)$ take a generic form $\int_t^T h(s, X, u_s)ds + \xi$, where X is the underlying state process. Let H_θ be Hamiltonian as

$$H_\theta(t, x, z, u) := z\sigma^{-1}(t, x)f(t, x, u) + \theta h(t, x, u). \tag{1.3.44}$$

Solution (Y^u, Z^u) to the quadratic BSDE

$$Y^u(t) = \theta\xi + \int_t^T (H_\theta(s, x, Z^u(s), u(s, x, Z^u(s)) + \frac{1}{2}Z^u(s)^2)ds - \int_t^T Z^u(s)dB_s \tag{1.3.45}$$

is connected to the risk-sensitive control problem by the identity

$$e^{Y^u} = J_t(u). \tag{1.3.46}$$

We notice from expression (1.3.43) that θ is equivalent to a rescaling multiplier of the reward R_t, it suffices to to solve BSDE (1.3.45) for the case $\theta = 1$.

More generally, if the value processes (y, z) solve

$$y(t) = e^\xi + \int_t^T y(s)(g(s, \log y(s), z(s)/y(s))ds - \int_t^T z(s)dB_s, \tag{1.3.47}$$

then by Itô's formula, (Y, Z) defined via

$$\begin{cases} Y(t) = \log y(t); \\ Z(t) = z(t)/y(t) \end{cases} \tag{1.3.48}$$

solve BSDE

$$Y(t) = \theta\xi + \int_t^T (g(s, Y(s), Z(s)) + \frac{1}{2}Z(s)^2)ds - \int_t^T Z(s)dB_s. \tag{1.3.49}$$

Equation (1.3.47), thus equation (1.3.49), has a solution when the driver g and the terminal value ξ are bounded. Existence of solution to (1.3.47) is due to Pardoux and Peng (1990) [43]. Since the transformation between Y and y in (1.3.48) is monotonic, when Comparison Theorem is needed for equations with quadratic growth, one can compare solutions to equations of the form (1.3.47), then conclude by transforming back to solutions to (1.3.49). Again, comparison can be applied assuming that g and ξ are bounded.

Above is a brief illustration of connections between risk-sensitive controls and quadratic BSDE's. A zero-sum game corresponds to a one-dimensional BSDE, and a non-zero-sum game a multidimensional equation. Rigorous formulation and technical treatments to the risk-sensitive control using quadratic BSDEs can be found in El Karoui and Hamadène (2003) [18].

Kobylansky (2000) [39] considers one-dimensional BSDE's whose drives have quadratic growth rate, not necessarily a quadratic term, in the volatility process. Her basic idea was the exponential transformation (1.3.48), which requires some condition like bounded parameters. Up to an exponential change, she approximated a quadratic driver with a monotonic sequence of Lipschitz drivers. Solution to the quadratic BSDE turns out to be limit of a monotonic sequence of solutions. It was Comparison Theorem that guarantees monotonicity of solutions to the sequence of approximating equations.

Even for controls indifferent to risk, since the driver of a BSDE corresponds to the Hamiltonian of a control problem, more general growth rates of the driver allows for growth rates of the game rewards.

1.3.4 Difference in several dimensions

It would be tempting to extend all results on one dimensional BSDE's to multi-dimensions, for example Comparisons, reflections, and higher growth rates, one reason being the correspondence between multidimensional BSDE's and non-zero-sum games.

Consider an N-Plyer non-zero-sum stochastic differential game of control. Each player, indexed by i, chooses a control u_i. Player i receives a reward $R^i(u_1, \cdots, u_N)$ related to all the N Players' controls. The Players' rewards have the form

$$R^i_t(u_1, u_2, \cdots, u_N) = \int_t^T h_i(s, X, u_1, u_2, \cdots, u_N)ds + \xi_i|\mathscr{F}_t],\ 0 \le t \le T,\ i = 1, \cdots, N. \tag{1.3.50}$$

For Player i, he receives a cumulative reward at rate h_i and terminal reward ξ_i. Every Player i aims at optimizing his expected reward J^i, defined as

$$\begin{aligned} J^1(u_1, u_2, \cdots, u_N) &= \mathbb{E}^{u_1, u_2, \cdots, u_N}[R^1(u_1, u_2, \cdots, u_N)|\mathscr{F}_t]; \\ J^2(u_1, u_2, \cdots, u_N) &= \mathbb{E}^{u_1, u_2, \cdots, u_N}[R^2(u_1, u_2, \cdots, u_N)|\mathscr{F}_t]; \\ &\vdots \\ J^N(u_1, u_2, \cdots, u_N) &= \mathbb{E}^{u_1, u_2, \cdots, u_N}[R^N(u_1, u_2, \cdots, u_N)|\mathscr{F}_t]. \end{aligned} \tag{1.3.51}$$

Define Hamiltonian $H = (H_1, H_2, \cdots, H_N)$ as

$$\begin{aligned} H_i(t, X, z_i, u_1, u_2, \cdots, u_N) :=& z_i \sigma^{-1}(t, X) f(t, X, u_1, u_2, \cdots, u_N) \\ &+ h_i(t, X, u_1, u_2, \cdots, u_N),\ i = 1, \cdots, N. \end{aligned} \tag{1.3.52}$$

If the N-dimensional processes

$$Y^{u_1,u_2,\cdots,u_N} = (Y_1^{u_1,u_2,\cdots,u_N}, \cdots, Y_N^{u_1,u_2,\cdots,u_N}), \tag{1.3.53}$$

and

$$Z^{u_1,u_2,\cdots,u_N} = (Z_1^{u_1,u_2,\cdots,u_N}, \cdots, Z_N^{u_1,u_2,\cdots,u_N}) \tag{1.3.54}$$

solve the N-dimensional BSDE

$$\begin{cases} Y_1^{u_1,u_2,\cdots,u_N}(t) = \xi_1 + \int_t^T H_1(s, X, Z_1^{u_1,u_2,\cdots,u_N}(s), (u_1, u_2, \cdots, u_N)(t, X, Z^{u_1,u_2,\cdots,u_N}(s))ds \\ \qquad\qquad - \int_t^T Z^{u_1,u_2,\cdots,u_N}(s)_1 dB_s; \\ Y_2^{u_1,u_2,\cdots,u_N}(t) = \xi_2 + \int_t^T H_2(s, X, Z_2^{u_1,u_2,\cdots,u_N}(s), (u_1, u_2, \cdots, u_N)(t, X, Z^{u_1,u_2,\cdots,u_N}(s))ds \\ \qquad\qquad - \int_t^T Z^{u_1,u_2,\cdots,u_N}(s)_2 dB_s; \\ \qquad\vdots \\ Y_N^{u_1,u_2,\cdots,u_N}(t) = \xi_N + \int_t^T H_N(s, X, Z_N^{u_1,u_2,\cdots,u_N}(s), (u_1, u_2, \cdots, u_N)(t, X, Z^{u_1,u_2,\cdots,u_N}(s))ds \\ \qquad\qquad - \int_t^T Z^{u_1,u_2,\cdots,u_N}(s)_N dB_s, \end{cases} \tag{1.3.55}$$

then the value process $Y^{u_1,u_2,\cdots,u_N}$ of the BSDE provides the value process $J(u_1, u_2, \cdots, u_N)$ of the non-zero-sum game.

A multi-dimensional BSDE of the general form

$$\begin{cases} Y_1(t) = \xi_1 + \int_t^T g_1(s, X, Y(s), Z(s))ds - \int_t^T Z(s)_1 dB_s; \\ Y_2(t) = \xi_2 + \int_t^T g_2(s, X, Y(s), Z(s))ds - \int_t^T Z(s)_2 dB_s; \\ \qquad\vdots \\ Y_N(t) = \xi_N + \int_t^T g_N(s, X, Y(s), Z(s))ds - \int_t^T Z(s)_N dB_s \end{cases} \tag{1.3.56}$$

is thus of interest. The case of Lipschitz driver $g = (g_1, \cdots, g_N)$ has been covered in Pardoux and Peng (1990) [43]. One might ask for extending to drivers of higher growth rate, like Kobylanski's 2000 work [39] in dimension one. We recall that Kobylanski concluded convergence of the approximating sequence by showing its monotonicity via Comparison Theorem. But in several dimensions, Lipschitz growth is far from enough for the Comparison Theorem to hold. An equivalent condition to apply the Comparison Theorem is provided by Hu and Peng (2006) [31].

Hamadène, Lepeltier and Peng (1997) [26] worked on the Markovian case, assuming that the driver $g(t,x,y,z)$ is a continuous function, with growth rate polynomial in x, and linear in y and z. They also approximated the driver with Lipschitz drivers, first deriving a weakly convergent subsequence of solutions to the approximating equations by weak compactness, then arguing that the weak limit is in fact strong under an "$\mathbb{L}^2$-dominance" assumption. Their $\mathbb{L}^2$-dominance" assumption is not necessary and can be removed.

In order to modify a non-zero-sum game with rewards (1.3.50) to incorporate optimal stopping features, reflections have to be added to the N-dimensional (1.3.56). When a reflected BSDE had only one dimension, El Karoui, Kapoudjian, Pardoux, Peng and Quenez (1997) [19] provided two methods - contraction and penalization. The penalization method does not help with solving multi-dimensional equations, again due to the lack of Comparison Theorem. The contraction method requires at most Lipschitz growth of the drivers. In Chapter 3, we shall explore the connections between non-zero-sum games of control with optimal stopping features and multi-dimensional BSDEs with reflection. Existence and uniqueness of solutions to such equations will be shown for Lipschitz drivers. In the Markovian framework, we shall prove existence of solutions to the equations with growth rates linear in the value process and in the volatility process, and polynomial in the historical maximum of state process.

Chapter 2

Martingale Interpretation

Chapter 3 of the dissertation is adapted from Karatzas and Li (2009) [32]. In that piece of work, we solved a non-zero-sum game of control and stopping, by identifying value process of the game with solution to a multi-dimensional reflected backward stochastic differential equations (BSDE). There, we prove existence of equilibrium strategy, assuming Isaacs' condition. The main tools are analytical tricks to prove existence of solution to the BSDE, and Comparison Theorem to prove optimality of controls from Isaacs' condition. The privilege to use heavy analysis is an advantage of the BSDE approach, for it is easier to solve the optimization problems under looser technical conditions. But our concern is, that too much heavy analysis in our BSDE chapter might disguise intuitions. To remind ourselves of the nature of the problem we solved, this chapter presents an equivalent martingale characterization of Nash equilibrium point of the non-zero-sum game in question. Starting off from this martingale characterization, stochastic maximum principle becomes a handy proposition.

Without controls, the non-zero-sum game of stopping was solved by Bensoussan and Friedman in as early as 1977, using variational inequalities. Without stopping, a martingale approach to the non-zero-sum game of control can be found in Davis (1979) [12], whose treatment will help us prove sufficiency of Isaacs' condition for the existence of equilibrium controls. This chapter is partly also a follow-up of Karatzas and Zamfirescu (2008) [38], which gave martingale characterization of saddle point of a zero-sum game where one player controls and the other stops. For the existence of a saddle point, the lower value and upper value of the game have to equate each other. Karatzas and Zamfirescu (2008) [38] argued the coincidence of several stopping rules. For the existence of an equilibrium, we no longer need to balance between sup inf and inf sup, whereas the difficulty switches to maximizing more that one expected rewards with the same set of strategies. We will take the way Karatzas and Sudderth (2006) [35] passes from a game where each player's reward terminated by himself to a game of interactive stopping. But to accommodate path-dependant rewards, Snell envelopes named after Snell (1952) [46], instead of stopping regions for Markovian rewards (c.f. Shiryayev (1979) [45]), was used to derive optimal stopping rules.

2.1 Mathematical layout

The rigorous model starts with a d-dimensional Brownian motion $B(\cdot)$ with respect to its generated filtration $\{\mathscr{F}_t\}_{0\le t\le T}$ on the canonical probability space $(\Omega, \mathscr{F}, \mathbb{P})$, in which $\Omega = C^d[0,T]$ is the set of all continuous d-dimensional function on a finite deterministic time horizon $[0,T]$, $\mathscr{F} = \mathscr{B}\left(C^d[0,T]\right)$ is the Borel sigma algebra, and $\mathbb{P}$ is the Wiener measure.

For every $t \in [0,T]$, define a mapping $\phi_t : C[0,T] \to [0,T]$ by $\phi_t(y)(s) = y(s\wedge t)$, which truncates the function $y \in C[0,T]$. For any $y^0 \in C[0,T]$, the pre-image $\phi_t^{-1}(y^0)$ collects all functions in $C[0,T]$ which are identical to y^0 up to time t. A stopping rule is a mapping $\tau : C[0,T] \to [0,T]$, such that

$$\{y \in C[0,T] : \tau(y) \le t\} \in \phi_t^{-1}\left(\mathscr{B}\left(C[0,T]\right)\right). \tag{2.1.1}$$

The set of all stopping rules ranging between t_1 and t_2 is denoted by $\mathscr{S}(t_1,t_2)$.

The state process $X(\cdot)$ solves the stochastic functional equation

$$X(t) = X(0) + \int_0^t \sigma(s,X)dB_s,\ 0 \le t \le T, \tag{2.1.2}$$

where the volatility matrix $\sigma : [0,T]\times\Omega \to \mathbb{R}^d\times\mathbb{R}^d$, $(t,\omega)\mapsto\sigma(t,\omega)$, is a predictable process.

Assumption 2.1.1 *(1) The volatility matrix $\sigma(t,\omega)$ is nonsingular for every $(t,\omega) \in [0,T]\times\Omega$;*
(2) there exists a positive constant A such that

$$|\sigma_{ij}(t,\omega) - \sigma_{ij}(t,\bar{\omega})| \le A \sup_{0\le s\le t} |\omega(s) - \bar{\omega}(s)|, \tag{2.1.3}$$

for all $1 \le i,j \le d$, for all $t \in [0,T]$, $\omega,\bar{\omega} \in \Omega$.

Under Assumption 2.1.1 (2), for every initial value $X(0) \in \mathbb{R}^d$, there exists a pathwise unique strong solution to equation (3.1.2) (Theorem 14.6, Elliott (1982) [21]).

The control vector $\underline{u} = (u^1,\cdots,u^N)$ take values in some given separable product metric spaces $\mathbb{A} = (\mathbb{A}_1,\times,\mathbb{A}_N)$. We shall assume that $\mathbb{A}_1,\times,\mathbb{A}_N$ are countable unions of nonempty, compact subsets, and are endowed with the σ-algebras $\mathscr{A}_1,\times,\mathscr{A}_N$ of their respective Borel subsets. In this chapter, we use the set $\underline{\mathscr{U}} = \mathscr{U}_1\times\cdots\times\mathscr{U}_N$ of closed loop control vectors in the form of $\underline{u}_t = \underline{\mu}(t,\omega)$ that is an N-dimensional non-anticipative functional of the state process $X(\cdot)$, for $0 \le t \le T$, where $\underline{\mu} = (\mu^1\times\cdots\times\mu^N) : [0,T]\times\Omega \to \mathbb{A}$ is a deterministic measurable mapping.

We consider the predictable mapping

$$\begin{aligned} f : [0,T]\times\Omega\times\mathbb{A} &\to \mathbb{R}^d, \\ (t,\omega,\underline{\mu}(t,\omega)) &\mapsto f(t,\omega,\underline{\mu}(t,\omega)), \end{aligned} \tag{2.1.4}$$

satisfying:

Assumption 2.1.1 *(continued)*
(3) There exists a positive constant A such that

$$\left|\sigma^{-1}(t,\omega) f(t,\omega,\underline{\mu}(t,\omega))\right| \leq A, \tag{2.1.5}$$

for all $0 \leq t \leq T$, $\omega \in \Omega$, *and all the* $\mathbb{A}$*-valued representative elements* $\underline{\mu}(t,\omega)$ *of the control space* $\underline{\mathscr{U}}$.

For generic control vectors $\underline{u}_t = \underline{\mu}(t,\omega)$, define $\mathbb{P}^{\underline{u}}$, a probability measure equivalent to $\mathbb{P}$, via the Radon-Nykodim derivative

$$\left.\frac{d\mathbb{P}^{\underline{u}}}{d\mathbb{P}}\right|\mathscr{F}_t = \exp\left\{\int_0^t \sigma^{-1}(s,X) f(s,X,\underline{u}_s) dB_s - \frac{1}{2}\int_0^t |\sigma^{-1}(s,X) f(s,X,\underline{u}_s)|^2 ds\right\}. \tag{2.1.6}$$

Then by Girsanov Theorem,

$$B_t^{\underline{u}} := B_t - \int_0^t \sigma^{-1}(s,X) f(s,X,\underline{u}_s) ds,\ 0 \leq t \leq T, \tag{2.1.7}$$

is a $\mathbb{P}^{\underline{u}}$-Brownian Motion with respect to the filtration $\{\mathscr{F}_t\}_{0\leq t\leq T}$. In the probability space $(\Omega,\mathscr{F},\mathbb{P})$ and with respect to the filtration $\{\mathscr{F}_t\}_{0\leq t\leq T}$, the pair $(X, B^{\underline{u}})$ is a weak solution to the forward stochastic functional equation

$$X_t = X_0 + \int_0^t f(s,X,\underline{u}_s) ds + \int_0^t \sigma(s,X) dB_s^{\underline{u}},\ 0 \leq t \leq T. \tag{2.1.8}$$

In the three subsequent sections of this chapter, we shall study, in a sequel, optimization problems with the following rewards.

Problem 2.1.1 *($N = 1$, one player's optimization)*

$$R_t(\tau, u) := \int_t^{\tau\wedge\rho} h(s,X,u_s) ds + L(\tau)\mathbb{1}_{\{\tau<\rho\}} + \eta\mathbb{1}_{\{\tau=\rho\}}. \tag{2.1.9}$$

In Problem 2.1.1, $\underline{u}_t = u_t = \mu(t,\omega)$ is a control in $\underline{\mathscr{U}} = \mathscr{U}_1 =: \mathscr{U}$, ρ is a stopping rule in $\mathscr{S}$, and τ is a stopping time in $\mathscr{S}(t,\rho)$ for $t \leq \rho$. Both the control u and the stopping rule τ is at the player's choice. The cumulative reward rate $h : [0,T]\times\Omega\times\mathbb{A} \to \mathbb{R}$, $(t,\omega,\mu(t,\omega)) \mapsto h(t,\omega,\mu(t,\omega))$, is a predictable process in t, non-anticipative functional in $X(\cdot)$, and measurable function in $\mu(t,\omega)$. The early exercise reward $L : [0,T]\times\Omega \to \mathbb{R}$, $(t,\omega) \mapsto L(t,\omega) =: L(t)$, is a $\{\mathscr{F}_t\}_{0\leq t\leq T}$-adapted process. The terminal reward η is a real-valued $\mathscr{F}_\rho$-measurable random variable. The rewards h, L and η are a.e. bounded for all $\omega \in \Omega$, $0 \leq t \leq T$, and all admissible controls $u_t = \mu(t,\omega)$.

Problem 2.1.2 *($N = 2$, two-player game)*

$$\begin{aligned} R_t^1(\tau,\rho,u,v) &:= \int_t^{\tau\wedge\rho} h_1(s,X,u_s,v_s) ds + L_1(\tau)\mathbb{1}_{\{\tau<\rho\}} + U_1(\rho)\mathbb{1}_{\{\rho\leq\tau<T\}} + \xi_1\mathbb{1}_{\{\tau\wedge\rho=T\}}; \\ R_t^2(\tau,\rho,u,v) &:= \int_t^{\tau\wedge\rho} h_2(s,X,u_s,v_s) ds + L_2(\rho)\mathbb{1}_{\{\rho<\tau\}} + U_2(\tau)\mathbb{1}_{\{\tau\leq\rho<T\}} + \xi_2\mathbb{1}_{\{\tau\wedge\rho=T\}}. \end{aligned} \tag{2.1.10}$$

In Problem 2.1.2, $(u_t, v_t) = (\mu(t,\omega), \upsilon(t,\omega))$ is a pair of controls in $\underline{\mathscr{U}} = \mathscr{U}_1 \times \mathscr{U}_2 =: \mathscr{U} \times \mathscr{V}$, and τ and ρ are stopping rules in $\mathscr{S}(t,T)$. The control u and the stopping rule τ is at the choice of Player I. The control v and the stopping rule ρ is at the choice of Player II. Player I receives reward $R^1_t(\tau,\rho,u,v)$, and Player II receives reward $R^2_t(\tau,\rho,u,v)$. The cumulative reward rates h_1 and $h_2 : [0,T]\times\Omega\times\mathbb{A}_1\times\mathbb{A}_2 \to \mathbb{R}$, $(t,\omega,\mu(t,\omega),\upsilon(t,\omega)) \mapsto h_i(t,\omega,\mu(t,\omega),\upsilon(t,\omega))$, $i=1,2$, are predictable processes in t, non-anticipative functionals in $X(\cdot)$, and measurable functions in $\mu(t,\omega)$ and $\upsilon(t,\omega)$. The early exercise rewards $L = (L_1, L_2)' : [0,T]\times\Omega \to \mathbb{R}^2$, $(t,\omega) \mapsto L(t,\omega) =: L(t)$, and $U = (U_1, U_2)' : [0,T]\times\Omega \to \mathbb{R}^2$, $(t,\omega) \mapsto U(t,\omega) =: U(t)$ are both $\{\mathscr{F}_t\}_{0\le t\le T}$-adapted processes. The terminal reward $\xi = (\xi_1,\xi_2)'$ is a pair of real-valued $\mathscr{F}_T$-measurable random variables. The rewards $h = (h_1,h_2)'$, L, U and ξ are a.e. bounded for all $\omega\in\Omega$, $0\le t\le T$, and all admissible controls $u_t = \mu(t,\omega)$ and $v_t = \upsilon(t,\omega)$. Here and throughout this chapter the notation M' means transpose of some matrix M.

Problem 2.1.3 *(N-player game) For $i = 1,\cdots,N$, the ith Player's reward process is*

$$R^i_t(\underline{\tau},\underline{u}) := \int_t^{\tau_{min}} h_i(s,X,\underline{u}_s)ds + L_i(\tau_i)\mathbb{1}_{\{D_i\}} + U_i(\tau_{(i)})\mathbb{1}_{\{D_i^c\setminus E\}} + \xi_i\mathbb{1}_{\{E\}}, \tag{2.1.11}$$

where the events E and $D_1,\cdots,D_N$ are defines as

$$E = \{\tau_j = T, \textit{for all } j = 1,\cdots,N\}, \tag{2.1.12}$$

and

$$D_i = \{\tau_i < \textit{ all of } \tau_1,\cdots,\tau_{i-1},\tau_{i+1},\cdots,\tau_N\}, \tag{2.1.13}$$

and the stopping rules

$$\tau_{min} = \min\{\tau_1,\cdots,\tau_N\}, \tag{2.1.14}$$

and

$$\tau_{(i)} = \min\{\tau_1,\cdots,\tau_{i-1},\tau_{i+1},\cdots,\tau_N\}. \tag{2.1.15}$$

In Problem 2.1.3, $\underline{u}_t = (u^1_t,\cdots,u^N_t) = \underline{\mu}_t = (\mu^1(t,\omega),\cdots,\mu^N(t,\omega))$ is a control vector in $\underline{\mathscr{U}} = \mathscr{U}_1\times\cdots\times\mathscr{U}_N$, and $\underline{\tau} = (\tau_1,\cdots,\tau_N)$ is a vector of N stopping rules in $\mathscr{S}(t,T)$. For $i = 1,\cdots,N$, the control u^i and the stopping rule τ_i is at the choice of the ith player, who receives reward $R^i_t(\underline{\tau},\underline{u})$. The cumulative reward rate and $h = (h_1,\cdots,h_N) : [0,T]\times\Omega\times\mathbb{A} \to \mathbb{R}^N$, $(t,\omega,\underline{\mu}(t,\omega)) \mapsto h(t,\omega,\underline{\mu}(t,\omega))$, is an N-dimensional predictable process in t, non-anticipative functional in $X(\cdot)$, and measurable function in $\underline{\mu}(t,\omega)$. The early exercise rewards $L = (L_1,\cdots,L_N)' : [0,T]\times\Omega \to \mathbb{R}^N$, $(t,\omega) \mapsto L(t,\omega) =: L(t)$, and $U = (U_1,\cdots,U_N)' : [0,T]\times\Omega \to \mathbb{R}^N$, $(t,\omega) \mapsto U(t,\omega) =: U(t)$ are both $\{\mathscr{F}_t\}_{0\le t\le T}$-adapted processes. The terminal reward $\xi = (\xi_1,\cdots,\xi_N)'$ is a vector of N real-valued $\mathscr{F}_T$-measurable random variables. The rewards h, L, U and ξ are a.e. bounded for all $\omega\in\Omega$, $0\le t\le T$, and all admissible controls $\underline{u}_t = \underline{\mu}(t,\omega)$.

2.2 A representative player's optimization

In this section, we will focus on solving a representative player's optimization Problem 2.1.1 with expected reward

$$J_t(\tau,u) := \mathbb{E}^u[R_t(\tau,u)|\mathscr{F}_t], \tag{2.2.1}$$

where the reward $R_t(\tau, u)$ is defined as in (2.1.9). This is a question of discretionary stopping, finding a stopping rule τ^* and control u^* to maximize the expected reward (2.2.1), over all stopping rules τ in $\mathscr{S}(t,\rho)$ and all controls u in $\mathscr{U}$. It is the very optimization problem that a generic ith player in an N-player game faces, when all the other players' strategies are given.

The following notations will facilitate expositions in this section.

Notation 2.2.1 *(1) When a strategy (τ^*, u^*) maximizes (2.2.1), it attains suprema in*

$$Y(t) := \sup_{\tau\in\mathscr{S}(t,\rho)} \sup_{u\in\mathscr{U}} J_t(\tau, u) = J_t(\tau^*, u^*). \tag{2.2.2}$$

The process Y is called the value process of the optimal control and stopping problem with expected reward (2.2.1).

(2) For a generic admissible control u, define

$$\begin{aligned} V(t,u) :=& Y(t) + \int_0^t h(s, X, u_s)ds \\ =& \sup_{\tau\in\mathscr{S}(t,\rho)} \sup_{u\in\mathscr{U}} J_t(\tau, u) + \int_0^t h(s, X, u_s)ds \\ =& \sup_{\tau\in\mathscr{S}(t,\rho)} \sup_{u\in\mathscr{U}} \mathbb{E}^u[R_t(\tau, u)|\mathscr{F}_t] + \int_0^t h(s, X, u_s)ds. \end{aligned} \tag{2.2.3}$$

(3) Since the stopping rules are defined on every path $\omega \in \Omega$, the choice of an optimal stopping rule is irrelevant of the control applied. Define

$$Y(t,u) := \sup_{\tau\in\mathscr{S}(t,\rho)} J_t(\tau, u) \geq J_t(t,u) = L(t)\mathbb{1}_{\{t<\rho\}} + \eta\mathbb{1}_{\{t=\rho\}}. \tag{2.2.4}$$

We remember that $J_t(t,u)$ refers to the conditional expectation $\mathbb{E}^u[L(t)\mathbb{1}_{\{t\leq\rho\}} + \eta\mathbb{1}_{\{t=\rho\}}|\mathscr{F}_t]$. Because ρ is an $\{\mathscr{F}_t\}$-stopping rule, the events $\{t \leq \rho\}$ and $\{t = \rho\}$ are $\mathscr{F}_t$ measurable. On $\{t \leq \rho\}$, $J_t(t,u) = L(t)$ because L is progressively measurable. On $\{t = \rho\}$, $J_t(t,u) = \mathbb{E}^u[\eta|\mathscr{F}_\rho] = \eta$, because η is $\mathscr{F}_\rho$-measurable. This is why the last identity in (2.2.4) holds true.

(4) Also for a generic u, define

$$\begin{aligned} Q(t,u) :=& Y(t,u) + \int_0^t h(s, X, u_s)ds \\ =& \sup_{\tau\in\mathscr{S}_{t,\rho}} J_t(\tau, u) + \int_0^t h(s, X, u_s)ds \\ =& \sup_{\tau\in\mathscr{S}_{t,\rho}} \mathbb{E}^u[R_0(\tau, u)|\mathscr{F}_t]. \end{aligned} \tag{2.2.5}$$

(5) The Hamiltonian is defined as

$$H(t, \omega, z, u_t) = H(t, \omega, z, \mu(t, \omega)) := z\sigma^{-1}(t, \omega) f(t, \omega, \mu(t, \omega)) + h(t, \omega, \mu(t, \omega)), \tag{2.2.6}$$

for $0 \leq t \leq T$, $\omega \in \Omega$, z in $\mathbb{R}^d$, and all admissible controls $u_t = \mu(t, \omega)$.

2.2.1 optimal stopping

In (2.2.4), $Y(\cdot,u)$ is said to be the value process of the optimal stopping problem. The process $Q(\cdot,u)$ defined by equaitons (2.2.5) is the Snell envelope of $R_0(\cdot,u)$. It is the smallest RCLL supermartingale dominating $R_0(\cdot,u)$. A proof of results in Lemma 2.2.1 can be found in (Appendix D, Karatzas and Shreve (1998) [34]). The proofs in their book proceeded with a finite deterministic terminal time, but also good for a bounded $\{\mathscr{F}_t\}_t$-stopping time as the terminal time. To pass to the bounded random terminal time ρ which is an $\{\mathscr{F}_t\}$-stopping time, it suffices to multiply the reward with an indicator $\mathbb{1}_{\{t\leq\rho\}}$. See the remark at the end of Elliott (1976) [22].

Let $\tau^* = \tau^*_t(u,\rho)$ be an optimal stopping rule (stopping time) that attains supremum in (2.2.5), i.e.,

$$\mathbb{E}^u[R_0(\tau^*,u)|\mathscr{F}_t] = Q(t,u) = \sup_{\tau\in\mathscr{S}_{t,\rho}} J_t(\tau,u) + \int_0^t h(s,X,u_s)ds. \qquad (2.2.7)$$

The following lemma provides an equivalent characterization of τ^*.

Lemma 2.2.1 *The optimality of τ^* is equivalent to both of the following conditions altogether:*
(1)

$$Q(\tau^*,u) = R_0(\tau^*,u), \qquad (2.2.8)$$

or equivalently,

$$Y(\tau^*,u) = L(\tau^*)\mathbb{1}_{\{\tau^*<\rho\}} + \eta\mathbb{1}_{\{\tau^*=\rho\}}; \qquad (2.2.9)$$

(2) The stopped supermartingale $Q(\cdot\wedge\tau^*,u)$ is a $\mathbb{P}^u$-martingale with respect to $\{\mathscr{F}_t\}_{0\leq t\leq\rho}$.

Besides, τ^* has an explicit expression as the first hitting time

$$\begin{aligned}\tau^* &= \inf\{t\leq s<\rho|Q(s,u)=R_0(s,u)\}\wedge\rho \\ &= \inf\{t\leq s<\rho|Y(s,u)=L(s)\}\wedge\rho.\end{aligned} \qquad (2.2.10)$$

The optimal time from now on to stop the reward stream is the first time when the value process $Y(\cdot,u)$ drops down to the early exercise reward $L(\cdot)$. If the two processes never meet, then wait until the end to take a terminal reward η at time ρ.

2.2.2 optimal control and stopping

Classical theory on optimal stopping has helped us identify a stopping rule that maximizes the expected reward $J(\tau,u)$ over all stopping rules in $\mathscr{S}(t,\rho)$.

If there is a $u^*\in\mathscr{U}$ such that for the optimal stopping rule τ^*,

$$J_t(\tau^*,u^*) \geq J_t(\tau^*,u), \text{ a.s. on } [0,T]\times\Omega, \forall u\in\mathscr{U}, \qquad (2.2.11)$$

then since, from subsection 2.2.1,

$$J_t(\tau^*,u) \geq J_t(\tau,u), \forall u\in\mathscr{U}, \tau\in\mathscr{S}(t,\rho), \qquad (2.2.12)$$

the pair of strategies (τ^*, u^*) satisfies

$$J_t(\tau^*, u^*) \geq J_t(\tau, u), \forall u \in \mathscr{U}, \tau \in \mathscr{S}(t, \rho). \tag{2.2.13}$$

That (τ^*, u^*) satisfies inequality (2.2.13) is equivalent to its maximizing (2.2.1) and attaining suprema in (2.2.2).

The rest of this subsection will look for such a u^* satisfying inequality (2.2.11). To simplify notations, in proofs of this subsection, $L(\cdot)$ is redefined as

$$L(t) := L(t)\mathbb{1}_{\{t<\rho\}} + \eta\mathbb{1}_{\{t=\rho\}}, 0 \leq t \leq \rho. \tag{2.2.14}$$

Let $\mathscr{U}_t$ denote the quotient space where controls in $\mathscr{U}$ identical on $[t, T]$ are equivalent. To be rigorous, for any $u, v \in \mathscr{U}$,

$$u \sim v, \text{ if and only if } u_s = v_s, \text{ a.s. on } (s, \omega) \in [t, T] \times \Omega; \tag{2.2.15}$$

$$\mathscr{U}_t = \mathscr{U} / \sim. \tag{2.2.16}$$

Lemma 2.2.2 *(Karatzas and Zamfirescu (2008) [38])*
Suppose $0 \leq \tau_1 \leq \tau_2 \leq T$. τ_1 and τ_2 are both $\{\mathscr{F}_t\}$-stopping times. $u_s = v_s$ on $s \in [\tau_1, \tau_2]$, then for any bounded $\mathscr{F}_{\tau_2}$-measurable random variable Θ,

$$\mathbb{E}^u[\Theta|\mathscr{F}_{\tau_1}] = \mathbb{E}^v[\Theta|\mathscr{F}_{\tau_1}]. \tag{2.2.17}$$

Lemma 2.2.2 suggests,

$$\sup_{u\in\mathscr{U}} J_t(\cdot, u) = \sup_{u\in\mathscr{U}_t} J_t(\cdot, u). \tag{2.2.18}$$

To maximize $J_t(\cdot, u)$ over $u \in \mathscr{U}$, it suffices to consider the values of controls on $[t, T]$.

Lemma 2.2.3 *For any $t \in [0, T]$, and any $\tau \in \mathscr{S}(t, \rho)$, the set of random variables $\{J_t(\tau, u)\}_{u\in\mathscr{U}_t}$ is a family directed upwards, i.e., $\forall u^1, u^2 \in \mathscr{U}_t$, there exists a $u^0 \in \mathscr{U}_t$, such that*

$$J_t(\tau, u^0) = J_t(\tau, u^1) \vee J_t(\tau, u^2). \tag{2.2.19}$$

Hence there exists a sequence of controls $u^n(\tau) \in \mathscr{U}_t$, such that

$$\lim_{n\to\infty} \uparrow J_t(\tau, u^n(\tau)) = \sup_{u\in\mathscr{U}} J_t(\tau, u). \tag{2.2.20}$$

Proof. Define an $\mathscr{F}_t$-measurable set

$$A := \{\omega \in \Omega | J_t(\tau, u^1) \geq J_t(\tau, u^2)\}. \tag{2.2.21}$$

Let $u^0 = u^1\mathbb{1}_{\{A\}} + u^2\mathbb{1}_{\{A^c\}} \in \mathscr{U}_t$. Then

$$J_t(\tau, u^0) = \mathbb{E}^{u^0}[R_t(\tau, u^0)|\mathscr{F}_t] = \begin{cases} \mathbb{E}^{u^1}[R_t(\tau, u^1)|\mathscr{F}_t], \text{ on } A \\ \mathbb{E}^{u^2}[R_t(\tau, u^2)|\mathscr{F}_t], \text{ on } A^c \end{cases} = J_t(\tau, u^1) \vee J_t(\tau, u^2). \tag{2.2.22}$$

By the proposition on page 121 of Neveu (1975) [41], there exists a sequence of controls in $\mathscr{U}_t$, approximating the supremum from below. By Lemma 2.2.2, supremum of $J_t(\tau, u)$ over $\mathscr{U}_t$ is the same as supremum over $\mathscr{U}$. □

Theorem 2.2.1 *A strategy* (τ^*, u^*) *is optimal in the sense of (2.2.13), if and only if the following three conditions hold.*
(1) $Y(\tau^*) = L(\tau^*)\mathbb{1}_{\{\tau^*<\rho\}} + \eta\mathbb{1}_{\{\tau^*=\rho\}}$*;*
(2) $V(\cdot \wedge \tau^*, u^*)$ *is a* $\mathbb{P}^{u^*}$*-martingale;*
(3) For every $u \in \mathscr{U}$*,* $V(\cdot \wedge \tau^*, u)$ *is a* $\mathbb{P}^u$*-supermartingale.*

Proof. "if"

For any $\tau \le \tau^* \in \mathscr{S}_{t,\rho}$ and any $u \in \mathscr{U}$, $L(\tau) \le Y(\tau, u) = Y(\tau)$. By (2.2.2), (2.2.3), and (2.2.13),

$$R_t(\tau, u)+\int_0^t h(s, X, u_s)ds = L(\tau)+\int_0^\tau h(s, X, u_s)ds \le Y(\tau)+\int_0^\tau h(s, X, u_s)ds = V(\tau, u). \tag{2.2.23}$$

From condition *(1)*, equality holds in (2.2.23) with the choice of $\tau = \tau^*$, giving

$$R_t(\tau^*, u) + \int_0^t h(s, X, u_s)ds = V(\tau^*, u). \tag{2.2.24}$$

Then,

$$Y(t) + \int_0^t h(s, X, u_s)ds = V(t, u) \ge \mathbb{E}^u[V(\tau, u)|\mathscr{F}_t] \ge \mathbb{E}^u[R_t(\tau, u)|\mathscr{F}_t] + \int_0^t h(s, X, u_s)ds. \tag{2.2.25}$$

In (2.2.25), the identity comes from (2.2.3) the definition of V, first inequality from supermartingale property *(3)*, and second inequality from (2.2.23). From martingale property *(2)*, and identity (2.2.24), both inequalities become equalities if $u = u^*$ and $\tau = \tau^*$. Hence for any $u \in \mathscr{U}_t$, any $\tau \in \mathscr{S}(t, \rho)$

$$\mathbb{E}^u[R_t(\tau \wedge \tau^*, u)|\mathscr{F}_t] \le Y(t), \tag{2.2.26}$$

where equality attained by $u = u^*$ and $\tau = \tau^*$.

"only if"

Condition *(1)* comes from Lemma 2.2.1. $\tau^* = \tau^*(\rho, u^*)$ has the form of (2.2.10). For any $u \in \mathscr{U}$, Lemma 2.2.10 states that $Y(\tau^*, u) = L(\tau^*)\mathbb{1}_{\{\tau^*<\rho\}} + \eta\mathbb{1}_{\{\tau^*=\rho\}}$. Condition (1) is true, because $Y(\tau^*) = \sup_{u\in\mathscr{U}} Y(\tau^*, u)$.

To see the supermartingale property *(3)*, take $0 \le s \le t \le \tau^* \le T$, and an arbitrary

$u \in \mathscr{U}_t$.

$$
\begin{aligned}
&\mathbb{E}^u\left[J_t(\tau^*, u) + \int_s^t h(r, X, u_r)dr \middle| \mathscr{F}_s\right] \\
=&\mathbb{E}^u\left[\mathbb{E}^u\left[\int_t^{\tau^*\wedge\rho} h(r, X, u_r)dr + L(\tau^*)\middle|\mathscr{F}_t\right] + \int_s^t h(r, X, u_r)dr\middle|\mathscr{F}_s\right] \\
=&\mathbb{E}^u\left[\int_s^{\tau^*\wedge\rho} h(r, X, u_r)dr + L(\tau^*)\middle|\mathscr{F}_s\right] \\
=&\sup_{\tau\in\mathscr{S}_{s,\rho}} \mathbb{E}^u\left[\int_s^{\tau\wedge\rho} h(r, X, u_r)dr + L(\tau)\middle|\mathscr{F}_s\right] \\
=&J_s(\tau^*, u).
\end{aligned}
\tag{2.2.27}
$$

Since u^* is optimal,

$$
J_s(\tau^*, u) \le J_s(\tau^*, u^*). \tag{2.2.28}
$$

By Lemma 2.2.3, there exists a sequence of controls $\{u^n\}_n \in \mathscr{U}_t$, such that

$$
\lim_{n\to\infty} \uparrow J_t(\tau^*, u^n) = J_t(\tau^*, u^*). \tag{2.2.29}
$$

For every u^n, from (2.2.27) and (2.2.28),

$$
\mathbb{E}^u\left[J_t(\tau^*, u^n) + \int_s^t h(r, X, u_r)dr\middle|\mathscr{F}_s\right] \le J_s(\tau^*, u^*). \tag{2.2.30}
$$

Let $n \to \infty$ in (2.2.30). Bounded Convergence Theorem gives

$$
\mathbb{E}^u\left[J_t(\tau^*, u^*) + \int_s^t h(r, X, u_r)dr\middle|\mathscr{F}_s\right] \le J_s(\tau^*, u^*). \tag{2.2.31}
$$

Adding $\int_0^s h(r, X, u_r)dr$ to both sides of (2.2.31), and by definition of V in (2.2.3),

$$
\mathbb{E}^u[V(t, u)|\mathscr{F}_s] \le V(s, u). \tag{2.2.32}
$$

From supermartingale property *(3)*, $V(\cdot \wedge \tau^*, u^*)$ is a $\mathbb{P}^{u^*}$-supermartingale. In order that it is a $\mathbb{P}^{u^*}$-martingale, it suffices to show

$$
\mathbb{E}^{u^*}[V(\tau^*, u^*)] = V(0, u^*). \tag{2.2.33}
$$

The strategy (τ^*, u^*) is optimal, so for any $0 \le t \le \tau^*$,

$$
Y(t) = \mathbb{E}^{u^*}\left[\int_t^{\tau^*\wedge\rho} h(s, X, u_s^*)ds + L(\tau^*)\middle|\mathscr{F}_t\right]. \tag{2.2.34}
$$

It follows that

$$
\begin{aligned}
Y(0) =&\mathbb{E}^{u^*}\left[\int_0^{\tau^*\wedge\rho} h(s, X, u_s^*)ds + L(\tau^*)\right] \\
=&\mathbb{E}^{u^*}\left[\mathbb{E}^{u^*}\left[\int_t^{\tau^*\wedge\rho} h(s, X, u_s^*)ds + L(\tau^*)\middle|\mathscr{F}_t\right] + \int_0^t h(s, X, u_s^*)ds\right] \\
=&\mathbb{E}^{u^*}\left[Y(t) + \int_0^t h(s, X, u_s^*)ds\right] \\
=&\mathbb{E}^{u^*}[V(t, u^*)].
\end{aligned}
\tag{2.2.35}
$$

The last equation in (2.2.35) comes from the definition of V in (2.2.3). Remember that t can be chosen arbitrary over $[0, \tau^*]$. Equating the first and last terms in (2.2.35) gives (2.2.33). This proves *(2)*. □

Definition 2.2.1 *(Thrifty) A control u is called thrifty, if and only if $\{V(t \wedge \tau^*, u)\}_{0 \le t \le \rho}$ is a $\mathbb{P}^u$-martingale, where τ^* is defined in (2.2.10).*

This definition is drawn from a dynamic programming definition of thrifty strategies on page 48, Dubins and Savage (1965) [14].

Proposition 2.2.1 *With the choice of optimal stopping rule τ^* from (2.2.10), a strategy $u \in \mathscr{U}$ is optimal in the sense of (2.2.13), if and only if it is thrifty.*

Proof. This is a proposition from Theorem 2.2.1. □

Theorem 2.2.2 *Let τ^* as defined in (2.2.10), and $\{V(t,u)\}_{t \in [0,\tau^*]}$ defined as in (2.2.3). Then the following statements hold true*

(1) $\{V(t,u)\}_{t \in [0,\tau^*]}$ admits the Doob-Meyer Decomposition

$$V(t,u) = Y(0) - A(t,u) + M(t,u),\ 0 \le t \le \tau^*(u) \wedge \tau^*(v). \tag{2.2.36}$$

$Y(0) = V(0,u)$, for all $u \in \mathscr{U}$.
(2) $A(0,u) = 0$. $A(\cdot,u)$ is an increasing, integrable process, satisfying

$$A(t,u) - A(t,v) = -\int_0^t (H(s,X,Z_s,u_s) - H(s,X,Z_s,v_s))ds,\ 0 \le t \le \tau^*. \tag{2.2.37}$$

(3) $M(\cdot,u)$ is a right-continuous, uniformly integrable $\mathbb{P}^u$-martingale. Further more, $M(\cdot,u)$ is represented as a stochastic integral

$$M(t,u) = \int_0^t Z_s dB_s^u, \tag{2.2.38}$$

where Z is a predictable, square-integrable process irrelevant of u.

Proof. By Theorem 2.2.1, $\{V(t,u)\}_{t \in [0,\tau^*]}$ is a $\mathbb{P}^u$-supermartingale. Boundedness of the rewards guarantees that it is of class $\mathscr{D}$. It then admits the Doob-Meyer Decomposition (cf. page 24-25, Karatzas and Shreve (1988) [33])

$$V(t,u) = V(0,u) - A(t,u) + M(t,u),\ 0 \le t \le \tau^*. \tag{2.2.39}$$

By definitions of V and Y, (2.2.2) and (2.2.3), $V(0,u) = Y(0)$, for all $u \in \mathscr{U}$. The $\mathbb{P}^u$-martingale $M(\cdot,u)$ has the representation (Theorem 3.1, Fujisaki, Kallianpur, and Kunita, 1972)

$$M(t,u) = \int_0^t Z_s^u dB_s^u, \tag{2.2.40}$$

where Z^u is a predictable, square-integrable process. It remains to show Z^u is irrelevant of u. By definition of B^u in (2.1.7),

$$\begin{aligned} V(t,u) =& Y(0) - A(t,u) + \int_0^t Z_s^u dB_s^u \\ =& Y(0) - A(t,u) - \int_0^t Z_s^u \sigma^{-1}(s,X) f(s,X,u_s) ds + \int_0^t Z_s^u dB_s,\ 0 \le t \le \tau^*. \end{aligned} \tag{2.2.41}$$

Take arbitrary $u, v \in \mathscr{U}$. From (2.2.3) and replacing u by v in (2.2.41),

$$\begin{aligned} V(t,u) =& V(t,v) + \int_0^t (h(s,X,u_s) - h(s,X,v_s)) ds \\ =& Y(0) - A(t,v) - \int_0^t Z_s^v \sigma^{-1}(s,X) f(s,X,v_s) ds \\ &+ \int_0^t (h(s,X,u_s) - h(s,X,v_s)) ds + \int_0^t Z_s^v dB_s. \end{aligned} \tag{2.2.42}$$

Identifying martingale terms in (2.2.41) and in (2.2.42), because of uniqueness of martingale representation, we conclude

$$Z_\cdot^u = Z_\cdot^v =: Z_\cdot. \tag{2.2.43}$$

Identifying finite variation terms in (2.2.41) and in (2.2.42),

$$\begin{aligned} & A(t,u) - A(t,v) \\ =& -\int_0^t ((h(s,X,u_s) + Z_s \sigma^{-1}(s,X) f(s,X,u_s)) - (h(s,X,v_s) + Z_s \sigma^{-1}(s,X) f(s,X,v_s))) ds \\ =& -\int_0^t (H(s,X,Z_s,u_s) - H(s,X,Z_s,v_s)) ds,\ 0 \le t \le \tau^*(u) \wedge \tau^*(v). \end{aligned} \tag{2.2.44}$$

□

Proposition 2.2.2 *(Stochastic Maximum Principle)*
If (τ^*, u^*) is optimal, then for all $u \in \mathscr{U}$, and for all $0 \le t \le \tau^*(u^*) \wedge \tau^*(u)$,

$$H(t,X,Z_t,u_t^*) \ge H(t,X,Z_t,u_t). \tag{2.2.45}$$

Proof. This is a direct consequence of Theorem 2.2.1 and 2.2.2. The optimality of (τ^*, u^*) implies that $V(\cdot, u^*)$ is a martingale up to time τ^*, hence $A(\cdot, u^*) = 0$. Also $Y(\cdot, u^*) \ge Y(\cdot, u)$, hence $\tau^*(u^*) \ge \tau^*(u)$, for all $u \in \mathscr{U}$. By (2.2.46),

$$A(t,u) = \int_0^t (H(s,X,Z_s,u_s^*) - H(s,X,Z_s,u_s)) ds,\ 0 \le t \le \tau^*(u^*) \wedge \tau^*(u). \tag{2.2.46}$$

That $A(\cdot, u)$ being increasing forces $H(\cdot, X, Z_\cdot, u_\cdot^*) - H(\cdot, X, Z_\cdot, u_\cdot)$ to be nonnegative. □

Theorem 2.2.3 *Let τ^* be the optimal stopping rule defined as in (2.2.10). If a control $u^*_t = \mu^*(t,\omega)$ in $\mathscr{U}$ maximizes the Hamiltonian in the way of* ***Isaacs' condition***

$$H(t,\omega,z,\mu^*(t,\omega)) \geq H(t,\omega,z,\mu(t,\omega)), \tag{2.2.47}$$

for all $0 \leq t \leq \rho$, $\omega \in \Omega$, $z \in \mathbb{R}^d$, and $u_t = \mu(t,\omega)$ in $\mathscr{U}$, then u^* is optimal in the sense that

$$J_t(\tau^*, u^*) \geq J_t(\tau^*, u), \text{ for all } 0 \leq t \leq \tau^*, \text{ and } u \in \mathscr{U}. \tag{2.2.48}$$

Proof. This proof follows the treatment in (section 4, Davis (1979) [12]).
For $t \leq s \leq \tau^*$, define, for arbitrary $u \in \mathscr{U}$,

$$I_s(u) := \mathbb{E}^{u^*}\left[L(\tau^*) + \int_t^{\tau^*} h(r,X,u^*_r)dr \middle| \mathscr{F}_s\right] - \int_t^s (h(r,X,u^*_r) - h(r,X,u_r))dr. \tag{2.2.49}$$

By (2.2.1), (2.2.10), and (2.2.11), $I_t(u) = J_t(\tau^*, u^*)$, and $\mathbb{E}^u[I_{\tau^*}(u)|\mathscr{F}_t] = J_t(\tau^*, u)$. But $I_\cdot(u)$ can be represented as

$$I_s(u) = I_t(u) - \int_t^s (h(r,X,u^*_r) - h(r,X,u_r))dr + \int_t^s Z^*_r dB^{u^*}_r, \tag{2.2.50}$$

for some predictable, $\mathbb{P}^{u^*}$-square-integrable process Z^*. Remember the definitions of the Brownian motion B^u in (2.1.7) and the Hamiltonian H in (2.2.6), then

$$I_s(u) = I_t(u) - \int_t^s (H(r,X,Z^*_r,u^*_r) - H(r,X,Z^*_r,u_r))dr + \int_t^s Z^*_r dB^u_r. \tag{2.2.51}$$

Isaacs' condition (2.2.47) suggests $I_\cdot(u)$ being a $\mathbb{P}^u$-local supermartingale. Via standard localization arguments,

$$J_t(\tau^*, u^*) = I_t(u) \geq \mathbb{E}^u[I_{\tau^*}(u)|\mathscr{F}_t] = J_t(\tau^*, u). \tag{2.2.52}$$

□

2.3 The two-player games

In this section, we shall study the two-player game Problem (2.1.2) as a simplest illustration of the N-Player game Problem (2.1.3). Then, to move forward to the N-player game, it is only a matter of fancier notations.

The two players in Problem Problem (2.1.2), respectively, maximize their expected reward processes

$$\begin{aligned} J^1_t(\tau,\rho,u,v) &:= \mathbb{E}[R^1_t(\tau,\rho,u,v)|\mathscr{F}_t]; \\ J^2_t(\tau,\rho,u,v) &:= \mathbb{E}[R^2_t(\tau,\rho,u,v)|\mathscr{F}_t]. \end{aligned} \tag{2.3.1}$$

Definition 2.3.1 *(Equilibrium strategies)*
Let τ^* and ρ^* be stopping rules in $\mathscr{S}(t,T)$, and (u^*, v^*) controls in $\mathscr{U} \times \mathscr{V}$. The strategy $(\tau^*, \rho^*, u^*, v^*)$ is called an equilibrium point of Problem (2.1.2), if for all stopping rules τ and ρ in $\mathscr{S}(t,T)$, and all controls (u,v) in $\mathscr{U} \times \mathscr{V}$,

$$\begin{aligned} J_t^1(\tau^*, \rho^*, u^*, v^*) &\geq J_t^1(\tau, \rho^*, u, v^*); \\ J_t^2(\tau^*, \rho^*, u^*, v^*) &\geq J_t^2(\tau^*, \rho, u^*, v). \end{aligned} \tag{2.3.2}$$

Given the strategy (ρ^*, v^*) of Player II, Player I's strategy (τ^*, u^*) maximizes his expected reward over all stopping rules $\tau \in \mathscr{S}(t,T)$ and all controls $u \in \mathscr{U}$. Given the strategy (τ^*, u^*) of Player I, Player II's strategy (ρ^*, v^*) maximizes his expected reward over all stopping rules $\rho \in \mathscr{S}(t,T)$ and all controls $v \in \mathscr{V}$. Each Player faces the control problem with discretionary stopping, the one solved in section 2.2.

The following notation will facilitate expositions in this section.

Notation 2.3.1 *(1)*

$$\begin{aligned} Y_1(t,u) := Y_1(t,u;\rho,v) &:= \sup_{\tau \in \mathscr{S}(t,T)} J_t^1(\tau,\rho,u,v) \geq J_t^1(t,\rho,u,v); \\ Y_2(t,v) := Y_2(t,v;\tau,u) &:= \sup_{\rho \in \mathscr{S}(t,T)} J_t^2(\tau,\rho,u,v) \geq J_t^2(\tau,t,u,v). \end{aligned} \tag{2.3.3}$$

(2)

$$\begin{aligned} Y_1(t;\rho,v) &:= \sup_{\tau \in \mathscr{S}(t,T)} \sup_{u \in \mathscr{U}} J_t^1(\tau,\rho,u,v); \\ Y_2(t;\tau,u) &:= \sup_{\rho \in \mathscr{S}(t,T)} \sup_{v \in \mathscr{V}} J_t^2(\tau,\rho,u,v). \end{aligned} \tag{2.3.4}$$

(3)

$$\begin{aligned} Q_1(t,u) := Q_1(t,u;\rho,v) :=& Y_1(t,u) + \int_0^t h_1(s,X,u_s,v_s)ds \\ =& \sup_{\tau \in \mathscr{S}(t,T)} J_t^1(\tau,\rho,u,v) + \int_0^t h_1(s,X,u_s,v_s)ds = \sup_{\tau \in \mathscr{S}(t,T)} \mathbb{E}[R_0^1(\tau,\rho,u,v)|\mathscr{F}_t]; \\ Q_2(t,v) := Q_2(t,v;\tau,u) :=& Y_2(t,v) + \int_0^t h_2(s,X,u_s,v_s)ds \\ =& \sup_{\rho \in \mathscr{S}(t,T)} J_t^2(\tau,\rho,u,v) + \int_0^t h_2(s,X,u_s,v_s)ds = \sup_{\rho \in \mathscr{S}(t,T)} \mathbb{E}[R_0^2(\tau,\rho,u,v)|\mathscr{F}_t]. \end{aligned} \tag{2.3.5}$$

(4)

$$\begin{aligned} V_1(t;\rho,u,v) :=& Y_1(t;\rho,v) + \int_0^t h_1(s,X,u_s,v_s)ds; \\ V_2(t;\tau,u,v) :=& Y_2(t;\tau,u) + \int_0^t h_2(s,X,u_s,v_s)ds. \end{aligned} \tag{2.3.6}$$

(5) The Hamiltonians are defined as

$$\begin{aligned} H_1(t,\omega,z_1,u_t,v_t) &= H_1(t,\omega,z_1,\mu(t,\omega),\upsilon(t,\omega)) \\ :=& z_1\sigma^{-1}(t,\omega)f(t,\omega,\mu(t,\omega),\upsilon(t,\omega)) + h_1(t,\omega,\mu(t,\omega),\upsilon(t,\omega)); \\ H_2(t,\omega,z_2,u_t,v_t) &= H_2(t,\omega,z_2,\mu(t,\omega),\upsilon(t,\omega)) \\ :=& z_2\sigma^{-1}(t,\omega)f(t,\omega,\mu(t,\omega),\upsilon(t,\omega)) + h_2(t,\omega,\mu(t,\omega),\upsilon(t,\omega)), \end{aligned} \tag{2.3.7}$$

for $0 \le t \le T$, $\omega \in \Omega$, z_1 *and* z_2 *in* $\mathbb{R}^d$, *and all admissible controls* $u_t = \mu(t,\omega)$ *and* $v_t = \upsilon(t,\omega)$.

2.3.1 game of stopping

Let us first fix a generic pair of controls u and v for the two Players respectively. Player I chooses stopping rule $\tau \in \mathscr{S}(t,T)$, and Player II chooses stopping rule $\rho \in \mathscr{S}(t,T)$. Given a stopping rule ρ^0 of Player II, Player I seeks to maximize his expected reward $J_t^1(\tau^1,\rho^0,u,v)$ with τ^1. Given a stopping rule τ^0 of Player I, Player II seeks to maximize his expected reward $J_t^2(\tau^0,\rho^1,u,v)$ with ρ^1.

Definition 2.3.2 *(Equilibrium stopping rules)*
Let $\tau^*,\rho^* \in \mathscr{S}(t,T)$, $u \in \mathscr{U}$, *and* $v \in \mathscr{V}$. *The pair of stopping rules* (τ^*,ρ^*) *is called an equilibrium stopping rule for the game of stopping with rewards (2.1.10), if for all* $\tau,\rho \in \mathscr{S}(t,T)$,

$$\begin{aligned} J_t^1(\tau^*,\rho^*,u,v) &\ge J_t^1(\tau,\rho^*,u,v); \\ J_t^2(\tau^*,\rho^*,u,v) &\ge J_t^2(\tau^*,\rho,u,v). \end{aligned} \tag{2.3.8}$$

Lemma 2.3.1 *That* (τ^*,ρ^*) *is a a pair of equilibrium stopping rules is equivalent to both of the following two conditions altogether.*
(1)

$$Y_1(\tau^*,u;\rho^*,v) = L_1(\tau^*)\mathbb{1}_{\{\tau^*<\rho^*\}} + U_1(\rho^*)\mathbb{1}_{\{\rho^*\le\tau^*<T\}} + \xi_1\mathbb{1}_{\{\tau^*\wedge\rho^*=T\}}, \tag{2.3.9}$$

and

$$Y_2(\rho^*,v;\tau^*,u) = L_2(\rho^*)\mathbb{1}_{\{\rho^*<\tau^*\}} + U_2(\tau^*)\mathbb{1}_{\{\tau^*\le\rho^*<T\}} + \xi_2\mathbb{1}_{\{\tau^*\wedge\rho^*=T\}}; \tag{2.3.10}$$

(2) The stopped supermartingales $Q_1(\cdot\wedge\tau^*,u;\rho^*,v)$ *and* $Q_2(\cdot\wedge\rho^*,v;\tau^*,u)$ *are* $\mathbb{P}^{u,v}$-*martingales.*

Besides, suppose in addition $L_1 \le U_1$, *and* $L_2 \le U_2$, *a.s., then if their exists a pair of stopping rules* (τ^*,ρ^*) *satisfying the equations*

$$\begin{aligned} \tau^* &= \inf\{t \le s < \rho | Y_1(s,u;\rho^*,v) = L_1(s)\} \wedge \rho^*; \\ \rho^* &= \inf\{t \le s < \rho | Y_2(s,v;\tau^*,u) = L_2(s)\} \wedge \tau^*, \end{aligned} \tag{2.3.11}$$

on first hitting times, then (τ^*,ρ^*) *are equilibrium.*

Proof. Definition 2.3.2 is equivalent to saying, that when Player II uses stopping rule ρ^*, Player I' stopping rule $\tau = \tau^*$ attains supremum in

$$Y_1(t,u;\rho^*,v) = \sup_{\tau\in\mathscr{S}(t,T)} J_t^1(\tau,\rho^*,u,v), \tag{2.3.12}$$

and when Player I uses stopping rule τ^*, Player II's stopping rule $\rho = \rho^*$ attains supremum in

$$Y_2(t,v;\tau^*,u) = \sup_{\rho\in\mathscr{S}(t,T)} J_t^2(\tau^*,\rho,u,v). \tag{2.3.13}$$

Each Player solves the optimal stopping problem in subsection 2.2.1. Applying Lemma 2.2.1 to the two Players respectively proves Lemma 2.3.1. □

Remark. The pair of equilibrium stopping rules (τ^*,ρ^*) defined in Definition 2.3.2 always exists. The equations (2.3.9) and (2.3.10) always have solutions. Let $t \in [0,T]$ be the current time, then $\tau^* = \rho^* = t$ is a trivial equilibrium that satisfies inequalities (2.3.8), and that solves the system (2.3.9) and (2.3.10). "It does not hurt if no one plays the game." Non-trivial equilibrium stopping rules are usually the ones of interest.

Theorem 2.3.1 *(non-existence of an optimal stopping rule)*
Suppose $L_1 \le U_1 + \epsilon$, and $L_2 \le U_2 + \epsilon$, a.s. for some real number $\epsilon > 0$. Under Assumption A 2.1, equilibrium stopping rules do not exist.

Proof. If (τ^*,ρ^*) were equilibrium, then τ^* would attain supremum in (2.3.12), and ρ^* would attain supremum in (2.3.13). There would have to be $\tau^* < \rho^*$, a.s., and $\rho^* < \tau^*$, a.s., which is impossible. □

2.3.2 game of control and stopping

For each of the two Players, when the other Player's strategy is also equilibrium, his equilibrium strategies solves the control problem with discretionary stopping in subsection 2.2.2. Claims in this subsection can be verified by applying Theorems 2.2.1, 2.2.2, and 2.2.3, and Propositions 2.2.1 and 2.2.2, to each of the two Players.

Theorem 2.3.2 *The set of stopping rules and controls (τ^*,ρ^*,u^*,v^*) is an equilibrium point of Problem 2.1.2, if and only if the following three conditions hold.*
(1)

$$Y_1(\tau^*;\rho^*,v^*) = L_1(\tau^*)\mathbb{1}_{\{\tau^*<\rho^*\}} + U_1(\rho^*)\mathbb{1}_{\{\rho^*\le\tau^*<T\}} + \xi_1\mathbb{1}_{\{\tau^*\wedge\rho^*=T\}}, \tag{2.3.14}$$

and

$$Y_2(\rho^*;\tau^*,u^*) = L_2(\rho^*)\mathbb{1}_{\{\rho^*<\tau^*\}} + U_2(\tau^*)\mathbb{1}_{\{\tau^*\le\rho^*<T\}} + \xi_2\mathbb{1}_{\{\tau^*\wedge\rho^*=T\}}; \tag{2.3.15}$$

(2) The two processes $V_1(\cdot\wedge\tau^*;\rho^*,u^*,v^*)$ and $V_2(\cdot\wedge\rho^*;\tau^*,u^*,v^*)$ are $\mathbb{P}^{u^*,v^*}$-martingales;
(3) For every $u \in \mathscr{U}$, $V_1(\cdot\wedge\tau^*;\rho^*,u,v^*)$ is a $\mathbb{P}^{u,v^*}$-supermartingale. For every $v \in \mathscr{V}$, $V_2(\cdot\wedge\rho^*;\tau^*,u^*,v)$ is a $\mathbb{P}^{u^*,v}$-supermartingale.

Definition 2.3.3 *(Thrifty) Suppose (τ^*,ρ^*) are a pair of stopping rules satisfying (2.3.11). A pair of controls (u,v) is called thrifty, if and only if $V_1(\cdot\wedge\tau^*;\rho^*,u,v)$ and $V_2(\cdot\wedge\rho^*;\tau^*,u,v)$ are $\mathbb{P}^{u,v}$-martingales.*

Proposition 2.3.1 *With the choice of equilibrium stopping rules* (τ^*, ρ^*) *satisfying (2.3.9) and (2.3.10), a pair of controls* $(u, v) \in \mathscr{U} \times \mathscr{V}$ *is equilibrium in the sense of Definition 3.1.71, if and only if it is thrifty.*

Theorem 2.3.3 *Suppose* $(\tau^*, \rho^*, u^*, v^*)$ *is a set of equilibrium strategies.* $\{V_1(t;\rho,u,v)\}_{t\in[0,\tau^*]}$ *and* $\{V_2(t;\tau,u,v)\}_{t\in[0,\rho^*]}$ *admit the Doob-Meyer Decompositions*

$$\begin{aligned} V_1(t;\rho,u,v) =& Y_1(0;\rho,v) - A_1(t;\rho,u,v) + M_1(t;\rho,u,v),\ 0 \le t \le \tau^*;\\ V_2(t;\tau,u,v) =& Y_2(0;\tau,v) - A_2(t;\tau,u,v) + M_2(t;\tau,u,v),\ 0 \le t \le \rho^*. \end{aligned} \tag{2.3.16}$$

$Y_1(0;\rho,v) = V_1(0;\rho,u,v)$, $Y_2(0;\tau,u) = V_2(0;\tau,u,v)$, *for all* $u \in \mathscr{U}$, $v \in \mathscr{V}$. $A_1(0;\rho,u,v) = A_2(0;\tau,u,v) = 0$. $A_1(\cdot;\rho,u,v))$ *and* $A_2(\cdot;\tau,u,v)$ *are increasing, integrable processes, satisfying*

$$\begin{aligned} A_1(t;\tau,u^1,v) - A_1(t;\tau,u^2,v) &= -\int_0^t (H_1(s,X,Z_1(s),u_s^1,v_s) - H_1(s,X,Z_1(s),u_s^2,v_s))ds,\\ &\quad 0 \le t \le \tau^*;\\ A_2(t;\rho,u,v^1) - A_2(t;\rho,u,v^2) &= -\int_0^t (H_2(s,X,Z_2(s),u_s,v_s^1) - H(s,X,Z_2(s),u_s,v_s^2))ds,\\ &\quad 0 \le t \le \rho^*. \end{aligned} \tag{2.3.17}$$

The processes $M_1(\cdot;\rho,u,v)$ *and* $M_2(\cdot;\tau,u,v)$ *are right-continuous, uniformly integrable* $\mathbb{P}^{u,v^*}$*-martingale and* $\mathbb{P}^{u^*,v}$*-martingale, respectively. Further more,* $M_1(\cdot;\rho,u,v)$ *and* $M_2(\cdot;\tau,u,v)$ *are represented as stochastic integrals*

$$\begin{aligned} M_1(t;\rho,u,v) &= \int_0^t Z_1^v(s)dB_s^{u,v};\\ M_2(t;\tau,u,v) &= \int_0^t Z_2^u(s)dB_s^{u,v}, \end{aligned} \tag{2.3.18}$$

where Z_1^v *and* Z_2^u *are predictable, square-integrable processes.* Z_1^v *is the same process for all* u*, and* Z_2^v *is the same process for all* v*.*

Proposition 2.3.2 *(Stochastic Maximum Principle)*
If $(\tau^*, \rho^*, u^*, v^*)$ *is an equilibrium point of Problem 2.1.2, then*

$$\begin{aligned} H_1(t,X,Z_1(t),u_t^*,v_t^*) \ge& H_1(t,X,Z_1(t),u_t,v_t^*),\ \text{for all } u \in \mathscr{U},\ 0 \le t \le \tau^*;\\ H_2(t,X,Z_2(t),u_t^*,v_t^*) \ge& H_2(t,X,Z_2(t),u_t^*,v_t),\ \text{for all } v \in \mathscr{V},\ 0 \le t \le \rho^*. \end{aligned} \tag{2.3.19}$$

Theorem 2.3.4 *(sufficiency of Isaacs' condition)*
Let τ^* *and* ρ^* *in* $\mathscr{S}(t,T)$ *be stopping rules satisfying (2.3.8). If the controls* $u_t^* = \mu^*(t,\omega)$ *in* $\mathscr{U}$ *and* $v_t^* = \upsilon^*(t,\omega)$ *in* $\mathscr{V}$ *satisfy* ***Isaacs' condition***

$$\begin{aligned} H_1(t,\omega,z_1,\mu^*(t,\omega),\upsilon^*(t,\omega)) &\ge H_1(t,\omega,z_1,\mu(t,\omega),\upsilon^*(t,\omega));\\ H_2(t,\omega,z_2,\mu^*(t,\omega),\upsilon^*(t,\omega)) &\ge H_2(t,\omega,z_2,\mu^*(t,\omega),\upsilon(t,\omega)), \end{aligned} \tag{2.3.20}$$

for all $0 \le t \le T$, ω *in* Ω, z_1 *and* z_2 *in* $\mathbb{R}^d$, $u_t = \mu(t,\omega)$ *in* $\mathscr{U}$ *and* $v_t = \upsilon(t,\omega)$ *in* $\mathscr{V}$, *then* u^*, v^* *are optimal in the sense that*

$$\begin{aligned} &J_t^1(\tau^*,\rho^*,u^*,v^*) \ge J_t^1(\tau^*,\rho^*,u,v^*), \text{ for all } u \in \mathscr{U}, 0 \le t \le \tau^*(u^*,v^*) \wedge \tau^*(u,v^*);\\ &J_t^2(\tau^*,\rho^*,u^*,v^*) \ge J_t^2(\tau^*,\rho^*,u^*,v), \text{ for all } v \in \mathscr{V}, 0 \le t \le \rho^*(u^*,v^*) \wedge \rho^*(u^*,v). \end{aligned} \tag{2.3.21}$$

If a pair of stopping rules satisfies the two equivalent conditions in Lemma 2.3.1, for all controls $u \in \mathscr{U}$ and $v \in \mathscr{V}$, and if the controls u^* and v^* satisfy Isaacs' condition (2.4.17), then combing (2.3.8) and (2.3.21) suggests that the strategy $(\tau^*, \rho^*, u^*, v^*)$ is an equilibrium point as in Definition 2.3.1.

2.4 The N-player games

When all the other $N-1$ players' strategies are given, a player faces the optimization problem that we have solved in section 2.2. This section will extend the two-player game Problem 2.1.2 studied in section 2.3 to the N-player version Problem 2.1.3.

Definition 2.4.1 *(Equilibrium strategies)*
Let $\underline{\tau}^* = (\tau_1^*, \cdots, \tau_N^*)$ *be a vector of stopping rules in* $\mathscr{S}(t,T)$*, and control vector* $\underline{u}^* = (u_1^*, \cdots, u_N^*)$ *in* $\underline{\mathscr{U}}$. *The strategy* $(\underline{\tau}^*, \underline{u}^*)$ *is called an equilibrium point of the* N*-Player stochastic differential game of control and stopping,*

$$J_t^i(\underline{\tau}^*,\underline{u}^*) \ge J_t^i((\tau_1^*,\cdots,\tau_{i-1}^*,\tau_i,\tau_{i+1}^*,\cdots,\tau_N^*),(u^{1,*},\cdots,u^{i-1,*},u^i,u^{i+1,*},\cdots,u^{N,*})), \tag{2.4.1}$$

for all stopping rules τ_i *in* $\mathscr{S}(t,T)$ *and all controls* u_i *in* $\mathscr{U}_i$*, for each player i,* $i = 1,\cdots,N$.

The characterization of the equilibrium point will use the following notations defined for all $i = 1,\cdots,N$.

Notation 2.4.1 *(1)*

$$Y_i(t,\underline{u}) := Y_i(t,\underline{u};\underline{\tau}) := \sup_{\tau_i \in \mathscr{S}(t,T)} J_t^i(\underline{\tau},\underline{u}); \tag{2.4.2}$$

(2)

$$Y_i(t;\underline{\tau},\underline{u}) := \sup_{\tau_i \in \mathscr{S}(t,T)} \sup_{u_i \in \mathscr{U}_i} J_t^i(\underline{\tau},\underline{u}); \tag{2.4.3}$$

(3)

$$\begin{aligned} Q_i(t,\underline{u}) :=& Q_i(t,\underline{u};\underline{\tau}) := Y_i(t,\underline{u}) + \int_0^t h_i(s,X,\underline{u}_s)ds\\ &= \sup_{\tau_i \in \mathscr{S}(t,T)} J_t^i(\underline{\tau},\underline{u}) + \int_0^t h_i(s,X,\underline{u}_s)ds = \sup_{\tau_i \in \mathscr{S}(t,T)} \mathbb{E}^{\underline{u}}[R_0^i(\underline{\tau},\underline{u})|\mathscr{F}_t]; \end{aligned} \tag{2.4.4}$$

(4)

$$V_i(t;\underline{\tau},\underline{u}) := Y_i(t;\underline{\tau},\underline{u}) + \int_0^t h_i(s,X,\underline{u}_s)ds; \tag{2.4.5}$$

(5) The Hamiltonians are defined as

$$H_i(t,\omega,z_i,\underline{u}_t) = H_i(t,\omega,z_i,\underline{\mu}(t,\omega)) := z_i\sigma^{-1}(t,\omega)f(t,\omega,\underline{\mu}(t,\omega)) + h_i(t,\omega,\underline{\mu}(t,\omega)), \tag{2.4.6}$$

for $0 \le t \le T$*,* $\omega \in \Omega$*,* z_i *in* $\mathbb{R}^d$*, and all admissible controls* $\underline{u}_t = \underline{\mu}(t,\omega)$*.*

2.4.1 game of stopping

We first fix an arbitrary control vector $\underline{u} = (u_1, \cdots, u_N)$ for the N-Players. The purpose of this subsection is to find a set of equilibrium stopping rules $\underline{\tau}^* = (\tau_1^*, \cdots, \tau_N^*)$ in the sense that

$$J_t^i(\underline{\tau}^*,\underline{u}) \ge J_t^i((\tau_1^*,\cdots,\tau_{i-1}^*,\tau_i,\tau_{i+1}^*,\cdots,\tau_N^*),\underline{u}), \text{ for all } \tau_i \in \mathscr{S}(t,T), \tag{2.4.7}$$

for all $i = 1, \cdots, N$. This is an N**-player game of stopping**. Equivalent conditions for the existence of equilibrium stopping rules with be provided for a generic vector $\underline{u}$ of controls.

Definition 2.4.2 *(Equilibrium stopping rules)*
For a generic control vector $\underline{u} = (u_1, \cdots, u_N)$ *for the* N*-Players. The set of stopping rules* $\underline{\tau}^* = (\tau_1^*, \cdots, \tau_N^*)$ *is said to be equilibrium for the* N*-player game of stopping, if*

$$J_t^i(\underline{\tau}^*,\underline{u}) \ge J_t^i((\tau_1^*,\cdots,\tau_{i-1}^*,\tau_i,\tau_{i+1}^*,\cdots,\tau_N^*),\underline{u}), \text{ for all } \tau_i \text{ in } \mathscr{S}(t,T), \tag{2.4.8}$$

for all $i = 1, \cdots, N$*.*

Lemma 2.4.1 *That* $\underline{\tau}^*$ *is a vector of equilibrium stopping rules is equivalent to both of the following conditions altogether, for all* $i = 1, \cdots, N$*,*

(1)

$$\begin{aligned} Y_i(\tau_i^*;\underline{\tau}^*,\underline{u}^*) =& R^i_{\tau^*_{min}}(\underline{\tau}^*,\underline{u}^*) \\ =& L_i(\tau_i^*)\mathbb{1}_{\{\tau_i^*<\tau_{(i)}^*\}} + U_i(\tau_{(i)}^*)\mathbb{1}_{\{\tau_{(i)}^*\le\tau_i^*<T\}} + \xi_i\mathbb{1}_{\{\tau^*_{min}=T\}}; \end{aligned} \tag{2.4.9}$$

(2) The stopped supermartingale $Q_i(\cdot \wedge \tau_i^*, \underline{u}; \underline{\tau}^*)$ *is a* $\mathbb{P}^{\underline{u}}$*–martingale.*

Besides, suppose in addition $L_i \le U_i$*, a.s., for all* $i = 1, \cdots, N$*, then if their exists a pair of stopping rules* (τ^*, ρ^*) *satisfying the equations*

$$\begin{cases} \tau_1^* = \inf\{t \le s < \rho | Y_1(\tau_1^*;\underline{\tau}^*,\underline{u}^*) = L_1(s)\} \wedge \tau_{(1)}^*; \\ \tau_2^* = \inf\{t \le s < \rho | Y_2(\tau_2^*;\underline{\tau}^*,\underline{u}^*) = L_2(s)\} \wedge \tau_{(2)}^*; \\ \quad\vdots \\ \tau_N^* = \inf\{t \le s < \rho | Y_N(\tau_N^*;\underline{\tau}^*,\underline{u}^*) = L_N(s)\} \wedge \tau_{(N)}^*, \end{cases} \tag{2.4.10}$$

on first hitting times, then $\underline{\tau}^*$ *is an equilibrium stopping rule.*

Proof. By Lemma 2.2.1. □

Theorem 2.4.1 *(non-existence of an optimal stopping rule)*
Suppose $L_i \leq U_i + \epsilon$, a.s., for all $i = 1, \cdots, N$, for some real number $\epsilon > 0$. Under Assumption A 2.1, equilibrium stopping rules do not exist.

Proof. If $\underline{\tau}^*$ were equilibrium, there would have to be $\tau_i^* < \tau_{(i)}^*$, a.s., for all $i = 1, \cdots, N$, which is impossible. □

2.4.2 game of control and stopping

Suppose $(\underline{\tau}^*, \underline{u}^*)$ is an equilibrium point of the N-player game of controls and stopping. Given all the other $N-1$ Players' stopping rules $(\tau_1^*, \cdots, \tau_{i-1}^*, \tau_{i+1}^*, \cdots, \tau_N^*)$ and controls $(u^{1,*}, \cdots, u^{i-1,*}, u^i, u^{i+1,*}, \cdots, u^{N,*})$, the strategy $(\tau_i, u^i) = (\tau_i^*, u^{i,*})$ maximizes Player i's expected reward

$$J_t^i((\tau_1^*, \cdots, \tau_{i-1}^*, \tau_i, \tau_{i+1}^*, \cdots, \tau_N^*), (u^{1,*}, \cdots, u^{i-1,*}, u^i, u^{i+1,*}, \cdots, u^{N,*})). \tag{2.4.11}$$

Player i faces a maximization problem solved in section 2.2. As consequences of Theorems 2.2.1, 2.2.2, and 2.2.3, and Propositions 2.2.1 and 2.2.2 , we have the following results for the N-player game.

Theorem 2.4.2 *The strategy $(\underline{\tau}^*, \underline{u}^*)$ is an equilibrium point of the N-player game of controls and stopping, if and only if the following three conditions hold for all $i = 1, \cdots, N$.*

(1)

$$\begin{aligned} Y_i(\tau_i^*; \underline{\tau}^*, \underline{u}^*) =& R^i_{\tau^*_{min}}(\underline{\tau}^*, \underline{u}^*) \\ =& L_i(\tau_i^*)\mathbb{1}_{\{\tau_i^* < \tau_{(i)}^*\}} + U_i(\tau_{(i)}^*)\mathbb{1}_{\{\tau_{(i)}^* \leq \tau_i^* < T\}} + \xi_i \mathbb{1}_{\{\tau^*_{min} = T\}}; \end{aligned} \tag{2.4.12}$$

(2) $V_i(\cdot \wedge \tau_i^*; \underline{\tau}^*, \underline{u}^*)$ is a $\mathbb{P}^{\underline{u}^*}$-martingale;
(3) For every $u^i \in \mathscr{U}_i$, the process $V_i(\cdot \wedge \tau_i^*; \underline{\tau}^*, (u^{1,*}, \cdots, u^{i-1,*}, u^i, u^{i+1,*}, \cdots, u^{N,*}))$ is a $\mathbb{P}^{(u^{1,*}, \cdots, u^{i-1,*}, u^i, u^{i+1,*}, \cdots, u^{N,*})}$-supermartingale.

Definition 2.4.3 *(Thrifty) Suppose $\underline{\tau}^*$ are equilibrium stopping rules. A vector $\underline{u}$ of controls is called thrifty, if and only if $V_i(\cdot \wedge \tau_i^*; \underline{\tau}^*, \underline{u})$ is a $\mathbb{P}^{\underline{u}}$-martingale for all $i = 1, \cdots, N$.*

Proposition 2.4.1 *With the choice of equilibrium stopping rules $\underline{\tau}^*$, a vector $\underline{u} \in \underline{\mathscr{U}}$ controls is equilibrium in the sense of Definition 2.4.1, if and only if it is thrifty.*

Theorem 2.4.3 *Suppose $(\underline{\tau}^*, \underline{u}^*)$ are equilibrium strategies, then the following statements are true for all $i = 1, \cdots, N$.*

(1) $V_i(\cdot \wedge \tau_i^*; \underline{\tau}^*, (u^{1,*}, \cdots, u^{i-1,*}, u^i, u^{i+1,*}, \cdots, u^{N,*}))$ *admits the Doob-Meyer Decomposition*

$$\begin{aligned} &V_i(t \wedge \tau_i^*; \underline{\tau}^*, (u^{1,*}, \cdots, u^{i-1,*}, u^i, u^{i+1,*}, \cdots, u^{N,*})) \\ =&Y_i(0; \underline{\tau}^*, (u^{1,*}, \cdots, u^{i-1,*}, u^i, u^{i+1,*}, \cdots, u^{N,*})) - A_i(t; u^i) + M_i(t; u^i),\ 0 \le t \le \tau_i^*. \end{aligned} \tag{2.4.13}$$

(2) For all $u^i \in \mathscr{U}_i$, $A_i(0; u^i) = 0$. $A_i(\cdot; u^i)$ *is an increasing, integrable process, satisfying*

$$\begin{aligned} &A_i(t; u^i) - A_i(t; v^i) \\ = &-\int_0^t (H_i(s, X, Z_i(s), (u^{1,*}, \cdots, u^{i-1,*}, u^i, u^{i+1,*}, \cdots, u^{N,*})_s) \\ &- H_i(s, X, Z_i(s), (u^{1,*}, \cdots, u^{i-1,*}, v^i, u^{i+1,*}, \cdots, u^{N,*})_s))ds,\ 0 \le t \le \tau_i^*. \end{aligned} \tag{2.4.14}$$

(3) For all $u_i \in \mathscr{U}_i$, $M_i(0; u^i) = 0$. $M_i(\cdot; u^i)$ *is a right-continuous, uniformly integrable* $\mathbb{P}^{(u^{1,*},\cdots,u^{i-1,*},u^i,u^{i+1,*},\cdots,u^{N,*})}$*-martingale.* $M_i(\cdot; u_i)$ *is represented as the stochastic integral*

$$M_i(t; u_i) = \int_0^t Z_i^{(u^{1,*},\cdots,u^{i-1,*},u^i,u^{i+1,*},\cdots,u^{N,*})}(s)dB_s^{(u^{1,*},\cdots,u^{i-1,*},u^i,u^{i+1,*},\cdots,u^{N,*})}, \tag{2.4.15}$$

where $Z_i^{(u^{1,*},\cdots,u^{i-1,*},u^i,u^{i+1,*},\cdots,u^{N,*})}$ *is a predictable, square-integrable process identical for all* $u^i \in \mathscr{U}_i$.

Proposition 2.4.2 *(Stochastic Maximum Principle)*
If $(\underline{\tau}^*, \underline{u}^*)$ *is an equilibrium point of the* N*-player game of controls and stopping, then, for all* $i = 1, \cdots, N$,

$$H_i(t, X, Z_i(t), \underline{u}_t^*) \ge H_i(t, X, Z_i(t), (u_1^*, \cdots, u_{i-1}^*, u_i, u_{i+1}^*, \cdots, u_N^*)_t), \tag{2.4.16}$$

for all $u^i \in \mathscr{U}_i$, $0 \le t \le \tau^*$.

Theorem 2.4.4 *(Sufficiency of Isaacs' condition)*
Let $\underline{\tau}$ *be a vector of equilibrium stopping rules. If a control vector* $\underline{u}^* = \underline{\mu}(t, \omega)$ *in* $\underline{\mathscr{U}}$ *satisfy* ***Isaacs' condition***

$$H_i(t, \omega, z_i, \underline{\mu}^*(t, \omega)) \ge H_i(t, \omega, z_i, (\mu^{1,*}, \cdots, \mu^{i-1,*}, \mu^i, \mu^{i+1,*}, \cdots, u^{N,*})(t, \omega)), \tag{2.4.17}$$

for all $0 \le t \le T$, $\omega \in \Omega$, $z_i \in \mathbb{R}^d$, $u_t^i = \mu^i(t, \omega)$ *in* $\mathscr{U}_i$, *for all* $i = 1, \cdots, N$, *then* $\underline{u}^*$ *is equilibrium in the sense that*

$$J_t^i(\underline{\tau}^*, \underline{u}^*) \ge J_t^i(\underline{\tau}^*, (u^{1,*}, \cdots, u^{i-1,*}, u^i, u^{i+1,*}, \cdots, u^{N,*})),\ \text{for all } u^i \in \mathscr{U}_i, \tag{2.4.18}$$

for all $i = 1, \cdots, N$. *Combining (2.4.7) and (2.4.18), the set of strategies* $(\underline{\tau}^*, \underline{u}^*)$ *is an equilibrium point by Definition 2.4.1.*

Chapter 3

BSDE Approach

This chapter considers non-zero-sum games with features of both stochastic control and optimal stopping, for a process of diffusion type, via the backward SDE approach. Running rewards, terminal rewards and early exercise rewards are all included. The running rewards can be functionals of the diffusion state process. Since the Nash equilibrium of an N-player non-zero-sum game is technically not more difficult than a two-player non-zero-sum game, only notationally more tedious, the number of players is assumed to be two, for concreteness.

Section 3.1 solves two games of control and stopping. The controls enter the drift of the underlying state process.

In the first game of section 3.1, each player controls and stops, and his stopping time terminates his own reward stream only. The value processes of both players are part of the solution to a multi-dimension BSDE with reflecting barrier. The instantaneous volatilities of the two players' value processes are explicitly expressed in the solution. Existence of the solution to general forms of the multi-dimensional BSDE with reflecting barrier will be proven in section 3.2 and section 3.3. Then, in the Markovian framework, the instantaneous volatilities can enter the controls as arguments, in which case the game is said to observe volatilities in addition to the other two arguments, namely time and the state-process.

In the second game of section 3.1, there are interactions of stopping. The time for each player to quit the game is the earliest of his own stopping time and the stopping time of the other player. Using the original definition of equilibrium introduced by Nash in 1949, the second game will be reduced to first solving games of the first type, then proving convergence of an iterated sequence of stopping times. The argument for convergence is monotonicity, hinted at Karatzas and Sudderth (2006) [35]: earlier stopping implies smaller value processes, and smaller value processes imply even earlier stopping. This technicality, reluctantly, assumes one pair of the terminal rewards is increasing. Due to the restriction of the comparison theorem to dimension one, convergence of the iteration will be proven for closed loop controls and Markovian controls only, without observing the volatilities.

Section 3.2 proves existence and uniqueness of the solution to a multi-dimensional BSDE with reflecting barrier, a general form of the one that accompanies Game 3.1.1. Section 3.3 discusses extension of the existence of solutions to equations of ultra-Lipschitz growth.

In our Game 3.1.1 where each player terminates his own reward, one may argue the optimality of stopping times via the semimartingale decomposition of the value processes. The BSDE approach here proposes a multi-dimensional BSDE whose value processes in the solution provide the value processes of the non-zero-sum games. News both good and bad is that general existence result of solutions to multi-dimensional BSDE with reflecting barrier still remains a widely open question. As is proven in Hu and Peng (2006) [31], in several dimensions, the comparison theorem is very restricted, so the penalization method which solves the one-dimensional counterpart problem does not help. Without Lipschitz growth condition, convergence arguments of the usual Picard type iteration cannot proceed, either. In a Markovian framework, this paper proves the Markovian structure of solutions to multidimensional reflected BSDEs with Lipschitz growth, and uses this Markovian structure as a starting point to extend existence result to equations with growth rates linear in the value and volatility processes, and polynomial in state process.

3.1 Two games of control and stopping

In the non-zero sum games of control and stopping to be discussed in this chapter, each player receives a reward. Based on their up-to-date information, the two players I and II, respectively, first choose their controls u and v, then the times τ and ρ to stop their own reward streams. The controls u and v are two processes that enter the dynamics of the underlying state process for the rewards. The optimality criterion for our non-zero-sum games is that of a Nash equilibrium, in which each player's expected reward is maximized when the other player maximizes his. In taking conditional expectations of the rewards, the change-of-measure setup to be formulated fixes one single Brownian filtration and one single state process for all controls u and v. Hence when optimizing the expected rewards over the control sets, there is no need to keep in mind the filtration or the state process.

Let us set up the rigorous model. We start with a d-dimensional Brownian motion $B(\cdot)$ with respect to its generated filtration $\{\mathscr{F}_t\}_{0\leq t\leq T}$ on the canonical probability space $(\Omega, \mathscr{F}, \mathbb{P})$, in which $\Omega = C^d[0,T]$ is the set of all continuous d-dimensional function on a finite deterministic time horizon $[0,T]$, $\mathscr{F} = \mathscr{B}\left(C^d[0,T]\right)$ is the Borel sigma algebra, and $\mathbb{P}$ is the Wiener measure.

For every $t \in [0,T]$, define a mapping $\phi_t : C[0,T] \to [0,T]$ by $\phi_t(y)(s) = y(s \wedge t)$, which truncates the function $y \in C[0,T]$. For any $y^0 \in C[0,T]$, the pre-image $\phi_t^{-1}(y^0)$ collects all functions in $C[0,T]$ which are identical to y^0 up to time t. A stopping rule is a mapping $\tau : C[0,T] \to [0,T]$, such that

$$\{y \in C[0,T] : \tau(y) \le t\} \in \phi_t^{-1}(\mathscr{B}(C[0,T])). \tag{3.1.1}$$

The set of all stopping rules ranging between t_1 and t_2 is denoted by $\mathscr{S}(t_1, t_2)$.

In the **path-dependent** case, the state process $X(\cdot)$ solves the stochastic functional equation

$$X(t) = X(0) + \int_0^t \sigma(s, X) dB_s,\ 0 \le t \le T, \tag{3.1.2}$$

where the volatility matrix $\sigma : [0,T] \times \Omega \to \mathbb{R}^d \times \mathbb{R}^d$, $(t,\omega) \mapsto \sigma(t,\omega)$, is a predictable process. In particular in the **Markovian** case, the volatility matrix $\sigma : [0,T] \times \mathbb{R}^d \to \mathbb{R}^d \times \mathbb{R}^d$, $(t, \omega(t)) \mapsto \sigma(t, \omega(t))$, is a deterministic mapping, then the state process equation (3.1.2) becomes the stochastic differential equation

$$X(t) = X(0) + \int_0^t \sigma(s, X(s)) dB_s,\ 0 \le t \le T. \tag{3.1.3}$$

The Markovian case is indeed a special case of path-dependence. Since it will receive some extra attention later at the end of subsection 3.1.2, we describe the Markovian framework separately from the more general path-dependent case.

Assumption 3.1.1 *(1) The volatility matrix $\sigma(t,\omega)$ is nonsingular for every $(t,\omega) \in [0,T] \times \Omega$;*
(2) there exists a positive constant A such that

$$|\sigma_{ij}(t,\omega) - \sigma_{ij}(t,\bar{\omega})| \le A \sup_{0 \le s \le t} |\omega(s) - \bar{\omega}(s)|, \tag{3.1.4}$$

for all $1 \le i, j \le d$, for all $t \in [0,T]$, $\omega, \bar{\omega} \in \Omega$.

Under Assumption 3.1.1 (2), for every initial value $X(0) \in \mathbb{R}^d$, there exists a pathwise unique strong solution to equation (3.1.2) (Theorem 14.6, Elliott (1982) [21]).

The controls u and v take values in some given separable metric spaces $\mathbb{A}_1$ and $\mathbb{A}_2$, respectively. We shall assume that $\mathbb{A}_1$ and $\mathbb{A}_2$ are countable unions of nonempty, compact subsets, and are endowed with the σ-algebras $\mathscr{A}_1$ and $\mathscr{A}_2$ of their respective Borel subsets. The controls u and v are said
(i) to be **open loop**, if $u_t = \mu(t,\omega)$ and $v_t = \upsilon(t,\omega)$ are $\{\mathscr{F}_t\}_{0\le t\le T}$-adapted processes on $[0,T]$, where $\mu : [0,T] \times \Omega \to \mathbb{A}_1$ and $\upsilon : [0,T] \times \Omega \to \mathbb{A}_2$ are non-anticipative measurable mappings;
(ii) to be **closed loop**, if $u_t = \mu(t, X)$ and $v_t = \upsilon(t, X)$ are non-anticipative functionals of the state process $X(\cdot)$, for $0 \le t \le T$, where $\mu : [0,T] \times \Omega \to \mathbb{A}_1$ and $\upsilon : [0,T] \times \Omega \to \mathbb{A}_2$ are deterministic measurable mappings;
(iii) to be **Markovian**, if $u_t = \mu(t, X(t))$ and $v_t = \upsilon(t, X(t))$, for $0 \le t \le T$, where $\mu : [0,T] \times \mathbb{R}^d \to \mathbb{A}_1$ and $\upsilon : [0,T] \times \mathbb{R}^d \to \mathbb{A}_2$ are deterministic measurable functions.

In the path-dependent case, the set $\mathscr{U} \times \mathscr{V}$ of admissible controls are taken as all the closed loop controls. The techniques that we shall use to solve for the optimal closed loop controls also apply to the open loop controls, so the extension of the results from closed loop to open loop is only a matter of more complicated notations. The discussion will be restricted within the class of closed loop controls for clarity of the exposition. In the Markovian case, the set $\mathscr{U} \times \mathscr{V}$ of admissible controls are taken as all the Markovian controls. Markovian controls are a subset of closed loop controls.

We consider the predictable mapping

$$\begin{aligned} f : [0,T] \times \Omega \times \mathbb{A}_1 \times \mathbb{A}_2 &\to \mathbb{R}^d, \\ (t, \omega, \mu(t,\omega), \upsilon(t,\omega)) &\mapsto f(t, \omega, \mu(t,\omega), \upsilon(t,\omega)), \end{aligned} \tag{3.1.5}$$

in the path-dependent case, and the deterministic measurable mapping

$$\begin{aligned} f : [0,T] \times \Omega \times \mathbb{A}_1 \times \mathbb{A}_2 &\to \mathbb{R}^d, \\ (t, \omega, \mu(t,\omega(t)), \upsilon(t,\omega(t))) &\mapsto f(t, \omega(t), \mu(t,\omega(t)), \upsilon(t,\omega(t))), \end{aligned} \tag{3.1.6}$$

in the Markovian case, satisfying:

Assumption 3.1.1 *(continued)*
(3) There exists a positive constant A such that

$$\left|\sigma^{-1}(t,\omega) f(t, \omega, \mu(t,\omega), \upsilon(t,\omega))\right| \le A, \tag{3.1.7}$$

for all $0 \le t \le T$, $\omega \in \Omega$, *and all the* $\mathbb{A}_1 \times \mathbb{A}_2$*-valued representative elements* $(\mu(t,\omega), \upsilon(t,\omega))$ *of the control spaces* $\mathscr{U} \times \mathscr{V}$.

For generic controls $u_t = \mu(t,\omega)$ and $v_t = \upsilon(t,\omega)$, define $\mathbb{P}^{u,v}$, a probability measure equivalent to $\mathbb{P}$, via the Radon-Nikodym derivative

$$\left.\frac{d\mathbb{P}^{u,v}}{d\mathbb{P}}\right|\mathscr{F}_t = \exp\left\{\int_0^t \sigma^{-1}(s,X) f(s,X,u_s,v_s) dB_s - \frac{1}{2}\int_0^t |\sigma^{-1}(s,X) f(s,X,u_s,v_s)|^2 ds\right\}. \tag{3.1.8}$$

Then, by the Girsanov Theorem,

$$B_t^{u,v} := B_t - \int_0^t \sigma^{-1}(s,X) f(s,X,u_s,v_s) ds,\ 0 \le t \le T \tag{3.1.9}$$

is a $\mathbb{P}^{u,v}$-Brownian Motion on $[0,T]$ with respect to the filtration $\{\mathscr{F}_t\}_{0\le t\le T}$. In the Markovian case, equation (3.1.9) can be written as

$$B_t^{u,v} = B_t - \int_0^t \sigma^{-1}(s,X(s)) f(s,X(s),\mu(s,X(s)),\upsilon(s,X(s))) ds,\ 0 \le t \le T. \tag{3.1.10}$$

In the probability space $(\Omega, \mathscr{F}, \mathbb{P})$ and with respect to the filtration $\{\mathscr{F}_t\}_{0\le t\le T}$, the pair $(X, B^{u,v})$ is a weak solution to the forward stochastic functional equation

$$X(t) = X(0) + \int_0^t f(s,X,u_s,v_s) ds + \int_0^t \sigma(s,X) dB_s^{u,v},\ 0 \le t \le T, \tag{3.1.11}$$

in the path-dependent case, and a weak solution to the forward stochastic differential equation

$$X(t) = X(0) + \int_0^t f(s, X(s), \mu(s, X(s)), \upsilon(s, X(s)))ds + \int_0^t \sigma(s, X(s))dB_s^{u,v}, \; 0 \le t \le T, \tag{3.1.12}$$

in the Markovian case.

When playing the game, the two players choose first their admissible controls u in $\mathscr{U}$ and v in $\mathscr{V}$, then for any given $t \in [0, T]$, they chose τ_t and ρ_t from $\mathscr{S}(t, T)$, times for them to quit the game. The pair of control and stopping rule (u, τ) is up to player I and the pair (v, ρ) is up to player II. For starting the game at time t, applying controls u and v, and quitting the game at τ_t and ρ_t respectively, the players receive rewards $R_t^1(\tau_t, \rho_t, u, v)$ and $R_t^2(\tau_t, \rho_t, u, v)$. To average over uncertainty, their respective reward processes are measured by the conditional $\mathbb{P}^{u,v}$-expectations

$$\mathbb{E}^{u,v}[R_t^1(\tau_t, \rho_t, u, v)|\mathscr{F}_t] \text{ and } \mathbb{E}^{u,v}[R_t^2(\tau_t, \rho_t, u, v)|\mathscr{F}_t]. \tag{3.1.13}$$

In the non-zero-sum games, the two players seek first admissible control strategies u^* in $\mathscr{U}$ and v^* in $\mathscr{V}$, and then stopping rules τ_t^* and ρ_t^* from $\mathscr{S}(t, T)$, to maximize their expected rewards, in the sense that

$$\begin{aligned} &\mathbb{E}^{u^*,v^*}[R_t^1(\tau_t^*, \rho_t^*, u^*, v^*)|\mathscr{F}_t] \ge \mathbb{E}^{u,v^*}[R_t^1(\tau_t, \rho_t^*, u, v^*)|\mathscr{F}_t], \; \forall \tau_t \in \mathscr{S}(t, T), \; \forall u \in \mathscr{U}; \\ &\mathbb{E}^{u^*,v^*}[R_t^2(\tau_t^*, \rho_t^*, u^*, v^*)|\mathscr{F}_t] \ge \mathbb{E}^{u^*,v}[R_t^1(\tau_t^*, \rho_t, u^*, v)|\mathscr{F}_t], \; \forall \rho_t \in \mathscr{S}(t, T), \; \forall v \in \mathscr{V}. \end{aligned} \tag{3.1.14}$$

The interpretation is as follows: when player II employs strategy (ρ_t^*, v^*), the strategy (τ_t^*, u^*) maximizes the expected reward of player I over all possible strategies on $\mathscr{S}(t, T) \times \mathscr{U}$; and vice versa, when player I employs strategy (τ_t^*, u^*), the strategy (ρ_t^*, v^*) is optimal for player II over all possible strategies on $\mathscr{S}(t, T) \times \mathscr{V}$. The set of controls and stopping rules $(\tau^*, \rho^*, u^*, v^*)$ is called the equilibrium point, or **Nash equilibrium**, of the game. For notational simplicity, denote

$$V^i(t) := \mathbb{E}^{u^*,v^*}[R_t^i(\tau_t^*, \rho_t^*, u^*, v^*)|\mathscr{F}_t], \tag{3.1.15}$$

the value process of the game for each player $i = 1, 2$.

In subsections 3.1.1-3.1.2 and subsection 3.1.3, we shall consider two games, which differ in the forms of the rewards R^1 and R^2.

Game 3.1.1

$$\begin{aligned} R_t^1(\tau_t, \rho_t, u, v) = R_t^1(\tau_t, u, v) &:= \int_t^{\tau_t} h_1(s, X, u_s, v_s)ds + L_1(\tau_t)\mathbb{1}_{\{\tau_t<T\}} + \xi_1\mathbb{1}_{\{\tau_t=T\}}; \\ R_t^2(\tau_t, \rho_t, u, v) = R_t^2(\rho_t, u, v) &:= \int_t^{\rho_t} h_2(s, X, u_s, v_s)ds + L_2(\rho_t)\mathbb{1}_{\{\rho_t<T\}} + \xi_2\mathbb{1}_{\{\rho_t=T\}}. \end{aligned} \tag{3.1.16}$$

Game 3.1.2

$$\begin{aligned} & R_t^1(\tau_t, \rho_t, u, v) \\ := & \int_t^{\tau_t \wedge \rho_t} h_1(s, X, u_s, v_s) ds + L_1(\tau_t) \mathbb{1}_{\{\tau_t < \rho_t\}} + U_1(\rho_t) \mathbb{1}_{\{\rho_t \le \tau_t < T\}} + \xi_1 \mathbb{1}_{\{\tau_t \wedge \rho_t = T\}}; \\ & R_t^2(\tau_t, \rho_t, u, v) \\ := & \int_t^{\tau_t \wedge \rho_t} h_2(s, X, u_s, v_s) ds + L_2(\rho_t) \mathbb{1}_{\{\rho_t < \tau_t\}} + U_2(\tau_t) \mathbb{1}_{\{\tau_t \le \rho_t < T\}} + \xi_2 \mathbb{1}_{\{\tau_t \wedge \rho_t = T\}}. \end{aligned} \tag{3.1.17}$$

Rewards from both games are summations of cumulative rewards at rates $h = (h_1, h_2)'$, early exercise rewards $L = (L_1, L_2)'$ and $U = (U_1, U_2)'$, and terminal rewards $\xi = (\xi_1, \xi_2)'$. Here and throughout this chapter the notation M' means transpose of some matrix M. The cumulative reward rates h_1 and $h_2 : [0, T] \times \Omega \times \mathbb{A}_1 \times \mathbb{A}_2 \to \mathbb{R}$, $(t, X, \mu(t, \omega), \upsilon(t, \omega)) \mapsto h_i(t, X, \mu(t, \omega), \upsilon(t, \omega))$, $i = 1, 2$, are predictable processes in t, non-anticipative functionals in $X(\cdot)$, and measurable functions in $\mu(t, \omega)$ and $\upsilon(t, \omega)$. The early exercise rewards $L : [0, T] \times \Omega \to \mathbb{R}^2$, $(t, \omega) \mapsto L(t, \omega) =: L(t)$, and $U : [0, T] \times \Omega \to \mathbb{R}^2$, $(t, \omega) \mapsto U(t, \omega) =: U(t)$ are both $\{\mathscr{F}_t\}_{0 \le t \le T}$-adapted processes. The terminal reward $\xi = (\xi_1, \xi_2)'$ is a pair of real-valued $\mathscr{F}_T$-measurable random variables. In the Markovian case, the rewards take the form $h(t, X, u_t, v_t) = \underline{h}(t, X(t), \mu(t, X(t)), \upsilon(t, X(t)))$, $L(t) = \underline{L}(t, X(t))$, $U(t) = \underline{U}(t, X(t))$, and $\xi = \underline{\xi}(X(T))$, for all $0 \le t \le T$ and some deterministic functions $\underline{L} : [0, T] \times \mathbb{R}^d \to \mathbb{R}$, $\underline{U} : [0, T] \times \mathbb{R}^d \to \mathbb{R}$, and $\underline{\xi} : \mathbb{R}^d \to \mathbb{R}^2$.

Assumption 3.1.2 *(1) The early exercise reward processes L and U are continuous, progressively measurable. In Game 3.1.1, assume $L(T) \le \xi$ holds a.e. on Ω. In Game 3.1.2, assume $L(t, \omega) \le U(t, \omega) \le \xi(\omega)$, a.e. $(t, \omega) \in [0, T] \times \Omega$, and also assume, for $i = 1, 2$, that the reward processes $U_i(\cdot)$, whose terminal values are defined as $U_i(T) = \xi_i$, are increasing processes.*
(2) There exist some constants $p \ge 1$ and $C_{rwd} > 0$, such that

$$|h(t, \omega, \mu(t, \omega), \upsilon(t, \omega))| + |L(t, \omega)| + |U(t, \omega)| + |\xi(\omega)| \le C_{rwd} \left(1 + \sup_{0 \le s \le t} |\omega(s)|^{2p}\right), \tag{3.1.18}$$

a.e. for all $\omega \in \Omega$, $0 \le t \le T$, and all admissible controls $u_t = \mu(t, \omega)$ and $v_t = \upsilon(t, \omega)$.

From the rewards and the coefficients of the state process, we define the Hamiltonians associated with our games as

$$\begin{aligned} & H_1(t, \omega, z_1, u_t, v_t) = H_1(t, \omega, z_1, \mu(t, \omega), \upsilon(t, \omega)) \\ := & z_1 \sigma^{-1}(t, \omega) f(t, \omega, \mu(t, \omega), \upsilon(t, \omega)) + h_1(t, \omega, \mu(t, \omega), \upsilon(t, \omega)); \\ & H_2(t, \omega, z_2, u_t, v_t) = H_2(t, \omega, z_2, \mu(t, \omega), \upsilon(t, \omega)) \\ := & z_2 \sigma^{-1}(t, \omega) f(t, \omega, \mu(t, \omega), \upsilon(t, \omega)) + h_2(t, \omega, \mu(t, \omega), \upsilon(t, \omega)), \end{aligned} \tag{3.1.19}$$

for $0 \le t \le T$, $\omega \in \Omega$, z_1 and z_2 in $\mathbb{R}^d$, and all admissible controls $u_t = \mu(t, \omega)$ and $v_t = \upsilon(t, \omega)$. From Assumption 3.1.1 (3), the Hamiltonians are Lipschitz functions in z_1 and z_2, uniformly over all $0 \le t \le T$, $\omega \in \Omega$, and all admissible controls $u_t = \mu(t, \omega)$ and $v_t = \upsilon(t, \omega)$.

Assumption 3.1.3 *(Isaacs' condition) There exist admissible controls $u_t^* = \mu^*(t,\omega)$ in $\mathscr{U}$ and $v_t^* = \upsilon^*(t,\omega)$ in $\mathscr{V}$, such that*

$$\begin{aligned} &H_1(t,\omega,z_1,\mu^*(t,\omega),\upsilon^*(t,\omega)) \geq H_1(t,\omega,z_1,\mu(t,\omega),\upsilon^*(t,\omega));\\ &H_2(t,\omega,z_2,\mu^*(t,\omega),\upsilon^*(t,\omega)) \geq H_2(t,\omega,z_2,\mu^*(t,\omega),\upsilon(t,\omega)), \end{aligned} \tag{3.1.20}$$

for all $0 \leq t \leq T$, $\omega \in \Omega$, $(z_1,z_2) \in \mathbb{R}^{2\times d}$, and all admissible controls $u_t = \mu(t,\omega)$ and $v_t = \upsilon(t,\omega)$.

The Isaacs' conditions on the Hamiltonians are "local" optimality conditions, formulated in terms of every point (t,z_1,z_2) in Euclidean space and every path ω in the function space Ω. Theorems 3.1.1 and 3.1.2 take the local conditions on the Hamiltonians and transform them into "global" optimization statements involving each higher-dimensional objects, such as stopping times, stochastic processes, etc., cumulated in the Euclidean space and averaged over the probability space. This implication is endowed by the continuous-time setting, contrasted to some discrete-time optimization problems where local maximization need not lead to global maximization.

When linking value processes of the games to the solutions to BSDEs, we shall discuss the solutions in the following spaces $\mathbb{M}^2(m;0,T)$ and $\mathbb{L}^2(m\times d;0,T)$ of processes, defined as

$$\begin{aligned} &\mathbb{M}^k(m;t,T)\\ &:= \left\{ m\text{-dimensional predictable RCLL process } \phi(\cdot) \text{ s.t. } \mathbb{E}\left[\sup_{[t,T]} \phi_s^2\right] \leq \infty \right\}, \end{aligned} \tag{3.1.21}$$

and

$$\begin{aligned} &\mathbb{L}^k(m\times d;t,T)\\ &:= \left\{ m\times d\text{-dimensional predictable RCLL process } \phi(\cdot) \text{ s.t. } \mathbb{E}\left[\int_t^T \phi_s^2 dt\right] \leq \infty \right\}, \end{aligned} \tag{3.1.22}$$

for $k = 1,2$, and $0 \leq t \leq T$.

3.1.1 Each player's reward terminated by himself

This subsection studies Game 3.1.1 where a player's time to quit is determined by his own decision. We shall demonstrate that the solution to a two-dimensional BSDE with reflecting barrier provides to the two players' value processes. The optimal stopping rules will be derived from reflecting conditions of the BSDE. The optimal controls come from Isaacs' condition, Assumption 3.1.3 on the Hamiltonians, which plays here the role of the driver of the corresponding BSDE.

The solution to the following system of BSDEs

$$\begin{cases} Y_1^{u,v}(t) = \xi_1 + \int_t^T H_1(s, X, Z_1^{u,v}(s), u_s, v_s)ds - \int_t^T Z_1^{u,v}(s)dB_s + K_1^{u,v}(T) - K_1^{u,v}(t), \\ \qquad Y_1^{u,v}(t) \geq L_1(t), 0 \leq t \leq T; \ \int_0^T (Y_1^{u,v}(t) - L_1(t))dK_1^{u,v}(t) = 0; \\ Y_2^{u,v}(t) = \xi_2 + \int_t^T H_2(s, X, Z_2^{u,v}(s), u_s, v_s)ds - \int_t^T Z_2^{u,v}(s)dB_s + K_2^{u,v}(T) - K_2^{u,v}(t), \\ \qquad Y_2^{u,v}(t) \geq L_2(t), 0 \leq t \leq T; \ \int_0^T (Y_2^{u,v}(t) - L_2(t))dK_2^{u,v}(t) = 0, \end{cases} \tag{3.1.23}$$

provides the players' value processes in Game 3.1.1, with the proper choice of controls $u = u^*$ and $v = v^*$ mandated by Isaacs' condition. From now on, a BSDE with reflecting barrier in the form of (3.1.23) will be denoted as $(T, \xi, H(u, v), L)$ for short. The solution to this BSDE is a triplet of processes $(Y^{u,v}, Z^{u,v}, K^{u,v})$, satisfying $Y^{u,v}(\cdot) \in \mathbb{M}^2(2; 0, T)$, $Z^{u,v}(\cdot) \in \mathbb{L}^2(2 \times d; 0, T)$, and $K^{u,v}(\cdot) = (K_1^{u,v}(\cdot), K_2^{u,v}(\cdot))'$ a pair of continuous increasing processes in $\mathbb{M}^2(2; 0, T)$.

We focus on the game aspect in this section, making use of results like existence of the solution to the BSDE, one-dimensional comparison theorem and continuous dependence theorems to be proven in section 3.2 and section 3.3. The proofs of claims will not rely on developments in this section.

Theorem 3.1.1 *Let $(Y^{u,v}, Z^{u,v}, K^{u,v})$ solve BSDE (3.1.23) with parameters $(T, \xi, H(u, v), L)$. Define the stopping rules*

$$\tau_t^*(y; r) := \inf\{s \in [t, r] : y(s) \leq L_1(s)\} \wedge r, \tag{3.1.24}$$

and

$$\rho_t^*(y; r) = \inf\{s \in [t, r] : y(s) \leq L_2(s)\} \wedge r, \tag{3.1.25}$$

for $y \in C[0, T]$ and $r \in [t, T]$. Let the stopping times $\tau_t(u, v) := \tau_t^* \left(Y_1^{u,v}(\cdot); T\right)$ and $\rho_t(u, v) := \rho_t^* \left(Y_2^{u,v}(\cdot); T\right)$, and the controls $u^* \in \mathscr{U}$ and $v^* \in \mathscr{V}$ satisfy Isaacs' condition Assumption 3.1.3. The quadruplet $(\tau(u^*, v^*), \rho(u^*, v^*), u^*, v^*)$ is a Nash equilibrium for Game 3.1.1. Furthermore, $V_i(t) = Y_i^{u^*,v^*}(t)$, $i = 1, 2$.

Proof. Let $(Y^{u,v}, Z^{u,v}, K^{u,v})$ solve BSDE (3.1.23) with parameters $(T, \xi, H(u, v), L)$. Since $Z^{u,v}$ is square-integrable with respect to the $\mathbb{P}$ measure, not necessarily square-integrable with respect to the $\mathbb{P}^{u,v}$ measure, the processes $\int_t^{\cdot} Z_1^{u,v}(s)dB_s^{u,v}$ and $\int_t^{\cdot} Z_2^{u,v}(s)dB_s^{u,v}$ are local $\mathbb{P}$-martingales, not necessarily $\mathbb{P}^{u,v}$-martingales. For every $n = 1, 2, \cdots$, let

$$T_1^n := \inf\left\{s \in [t, T] : |Z_1^{u,v}(s)| \geq n\right\} \wedge T \tag{3.1.26}$$

be the localizing sequences of stopping times. The localized processes $\int_t^{\cdot \wedge T_1^n} Z_1^{u,v}(s)dB_s^{u,v}$ and $\int_t^{\cdot \wedge T_1^n} Z_2^{u,v}(s)dB_s^{u,v}$ are $\mathbb{P}^{u,v}$-martingales on $[0, T]$. As $n \to \infty$, $T_1^n \to T$, hence

$$\int_t^{\cdot \wedge T_1^n} Z_1^{u,v}(s)dB_s^{u,v} \to \int_t^{\cdot} Z_1^{u,v}(s)dB_s^{u,v}, \tag{3.1.27}$$

and

$$\int_t^{\cdot\wedge T_1^n} Z_2^{u,v}(s)dB_s^{u,v} \to \int_t^{\cdot} Z_2^{u,v}(s)dB_s^{u,v}, \tag{3.1.28}$$

almost everywhere. Taking a stopping rule $\tau_t \in \mathscr{S}(t,T)$, and integrating $dY_1^{u,v}$ from t to $\tau_t \wedge T_1^n$,

$$\begin{aligned} Y_1^{u,v}(t) =& Y_1^{u,v}(\tau_t \wedge T_1^n) + \int_t^{\tau_t\wedge T_1^n} H_1(s,X,Z_1^{u,v}(s),u_s,v_s)ds \\ &- \int_t^{\tau_t\wedge T_1^n} Z_1^{u,v}(s)dB_s + K_1^{u,v}(\tau_t\wedge T_1^n) - K_1^{u,v}(t) \\ =& Y_1^{u,v}(\tau_t \wedge T_1^n) + \int_t^{\tau_t\wedge T_1^n} h_1(s,X,u_s,v_s)ds \\ &- \int_t^{\tau_t\wedge T_1^n} Z_1^{u,v}(s)dB_s^{u,v} + K_1^{u,v}(\tau_t\wedge T_1^n) - K_1^{u,v}(t) \end{aligned} \tag{3.1.29}$$

Taking conditional expectation $\mathbb{E}^{u,v}[\cdot|\mathscr{F}_t]$, since $Y_1^{u,v}(\cdot) \geq L_1(\cdot)$, $Y_1^{u,v}(T) = \xi_1$, and $K_1(\cdot)$ is an increasing process,

$$\begin{aligned} Y_1^{u,v}(t) =& \mathbb{E}^{u,v}\left[Y_1^{u,v}(\tau_t\wedge T_1^n) + \int_t^{\tau_t\wedge T_1^n} h_1(s,X,u_s,v_s)ds + K_1^{u,v}(\tau_t\wedge T_1^n) - K_1^{u,v}(t) \middle| \mathscr{F}_t \right] \\ \geq& \mathbb{E}^{u,v}\left[L_1(\tau_t\wedge T_1^n)\mathbb{1}_{\{\tau_t\wedge T_1^n<T\}} + \xi_1\mathbb{1}_{\{\tau_t\wedge T_1^n=T\}} + \int_t^{\tau_t\wedge T_1^n} h_1(s,X,u_s,v_s)ds \middle| \mathscr{F}_t \right]. \end{aligned} \tag{3.1.30}$$

According to the reflecting condition in BSDE (3.1.23), $K_1^{u,v}(\tau_t(u,v)\wedge T_1^n) = K_1^{u,v}(t)$, because $K_1^{u,v}(\cdot)$ is flat on $\{(\omega,t) \in (\Omega\times[0,T]) : Y_1^{u,v}(t) \neq L_1(t)\}$. On $\{\tau_t(u,v) < T\}$, $Y_1^{u,v}(\tau_t(u,v)) = L_1(\tau_t(u,v))$; on $\{\tau_t(u,v) = T\}$, $Y_1^{u,v}(\tau_t(u,v)) = \xi_1$. Then,

$$\begin{aligned} & Y_1^{u,v}(t) \\ =& \mathbb{E}^{u,v}\left[Y_1^{u,v}(\tau_t(u,v)\wedge T_1^n) + \int_t^{\tau_t(u,v)\wedge T_1^n} h_1(s,X,u_s,v_s)ds \middle| \mathscr{F}_t \right] \\ =& \mathbb{E}^{u,v}\left[Y_1^{u,v}(T_1^n)\mathbb{1}_{\{T_1^n<\tau_t(u,v)\}} \middle| \mathscr{F}_t \right] \\ &+ \mathbb{E}^{u,v}\left[L_1(\tau_t(u,v))\mathbb{1}_{\{\tau_t(u,v)\leq T_1^n,\tau_t(u,v)<T\}} + \xi_1\mathbb{1}_{\{\tau_t(u,v)\leq T_1^n,\tau_t(u,v)=T\}} + \int_t^{\tau_t(u,v)\wedge T_1^n} h_1(s,X,u_s,v_s)ds \middle| \mathscr{F}_t \right]. \end{aligned} \tag{3.1.31}$$

From Assumption 3.1.3 (2), both rewards inside the last conditional expectations in (3.1.30) and (3.1.31) are bounded by

$$(1+T)C_{\text{rwd}}\left(1 + \sup_{0\leq s\leq T}|X(s)|^{2p}\right). \tag{3.1.32}$$

But since $(X, B^{u,v})$ is a weak solution to the stochastic functional equation (3.1.11), there exists (cf. page 306 of Karatzas and Shreve (1988) [33]) a constant C such that

$$\mathbb{E}^{u,v}\left[\sup_{0\leq s\leq T}|X(s)|^{2p}\right] \leq C\left(1 + |X(0)|^{2p}\right) < \infty. \tag{3.1.33}$$

We then apply the dominated convergence theorem to the last conditional expectations in (3.1.30) and (3.1.31), to get

$$\begin{aligned}
&\lim_{n\to\infty} \mathbb{E}^{u,v}\left[L_1(\tau_t \wedge T_1^n)\mathbb{1}_{\{\tau_t\wedge T_1^n<T\}} + \xi_1\mathbb{1}_{\{\tau_t\wedge T_1^n=T\}} + \int_t^{\tau_t\wedge T_1^n} h_1(s,X,u_s,v_s)ds \middle| \mathscr{F}_t \right] \\
=&\mathbb{E}^{u,v}\left[L_1(\tau_t)\mathbb{1}_{\{\tau_t<T\}} + \xi_1\mathbb{1}_{\{\tau_t=T\}} + \int_t^{\tau_t} h_1(s,X,u_s,v_s)ds \middle| \mathscr{F}_t \right],
\end{aligned} \tag{3.1.34}$$

and

$$\begin{aligned}
&\lim_{n\to\infty} \mathbb{E}^{u,v}\left[L_1(\tau_t(u,v))\mathbb{1}_{\{\tau_t(u,v)\le T_1^n,\tau_t(u,v)<T\}} + \xi_1\mathbb{1}_{\{\tau_t(u,v)\le T_1^n,\tau_t(u,v)=T\}} + \int_t^{\tau_t(u,v)\wedge T_1^n} h_1(s,X,u_s,v_s)ds \middle| \mathscr{F}_t \right] \\
=&\mathbb{E}^{u,v}\left[L_1(\tau_t(u,v))\mathbb{1}_{\{\tau_t(u,v)<T\}} + \xi_1\mathbb{1}_{\{\tau_t(u,v)=T\}} + \int_t^{\tau_t(u,v)} h_1(s,X,u_s,v_s)ds \middle| \mathscr{F}_t \right].
\end{aligned} \tag{3.1.35}$$

For the fixed $t \in [0,T]$, denote

$$\theta(s,u_s,v_s) := \sigma^{-1}(s,X)f(s,X,u_s,v_s),\ t \le s \le T. \tag{3.1.36}$$

Then, from the change of measure (3.1.8) and the Bayes rule,

$$\begin{aligned}
&\mathbb{E}^{u,v}\left[Y_1^{u,v}(T_1^n)\mathbb{1}_{\{T_1^n<\tau_t(u,v)\}} \middle| \mathscr{F}_t \right] \\
=&\mathbb{E}\left[\exp\left\{ \int_t^{T_1^n} \theta(s,u_s,v_s)dB_s - \frac{1}{2}\int_t^{T_1^n} |\theta(s,u_s,v_s)|^2 ds \right\} Y_1^{u,v}(T_1^n)\mathbb{1}_{\{T_1^n<\tau_t(u,v)\}} \middle| \mathscr{F}_t \right].
\end{aligned} \tag{3.1.37}$$

Both random variables inside the expectations in (3.1.37) converge to zero a.e., as n tends to infinity. Furthermore,

$$\begin{aligned}
&\exp\left\{ \int_t^{T_1^n} \theta(s,u_s,v_s)dB_s - \frac{1}{2}\int_t^{T_1^n} |\theta(s,u_s,v_s)|^2 ds \right\} |Y_1^{u,v}(T_1^n)|\mathbb{1}_{\{T_1^n<\tau_t(u,v)\}} \\
\le& \sup_{t\le s\le T} \exp\left\{ \int_t^{s} \theta(r,u_r,v_r)dB_r - \frac{1}{2}\int_t^{s} |\theta(r,u_r,v_r)|^2 dr \right\} |Y_1^{u,v}(s)|,
\end{aligned} \tag{3.1.38}$$

and

$$\begin{aligned}
&\mathbb{E}\left[\sup_{t\le s\le T} \exp\left\{ \int_t^{s} \theta(r,u_r,v_r)dB_r - \frac{1}{2}\int_t^{s} |\theta(r,u_r,v_r)|^2 dr \right\} |Y_1^{u,v}(s)| \middle| \mathscr{F}_t \right] \\
\le&\mathbb{E}\left[\sup_{t\le s\le T} \exp\left\{ \int_t^{s} 2\theta(r,u_r,v_r)dB_r - \frac{1}{2}\int_t^{s} 2|\theta(r,u_r,v_r)|^2 dr \right\} \middle| \mathscr{F}_t \right]^{1/2} \mathbb{E}\left[\sup_{t\le s\le T} \left(Y_1^{u,v}(s)\right)^2 \middle| \mathscr{F}_t \right]^{1/2}.
\end{aligned} \tag{3.1.39}$$

By the dominated convergence theorem, in order that (3.1.37) converge to zero, it suffices that the right hand side of (3.1.39) be finite. From the definition of the solutions to reflected BSDEs, as in section 3.2 and section 3.3, we know that

$$\mathbb{E}\left[\sup_{t\le s\le T} \left(Y_1^{u,v}(s)\right)^2 \right] < \infty \tag{3.1.40}$$

holds, so it remain to show that

$$\mathbb{E}\left[\sup_{t\le s\le T}\exp\left\{\int_t^s 2\theta(r,u_r,v_r)dB_r-\frac{1}{2}\int_t^s 2|\theta(r,u_r,v_r)|^2dr\right\}\right]<\infty. \tag{3.1.41}$$

Because $|\theta(s,u_s,v_s)|$ is bounded by the constant A, from Assumption 3.1.1 (3) and identity (3.1.36), we know that the process

$$\exp\left\{\int_t^\cdot 2\theta(s,u_s,v_s)dB_s-\frac{1}{2}\int_t^\cdot 2|\theta(s,u_s,v_s)|^2ds\right\} \tag{3.1.42}$$

is a.e. bounded by the constant e^{A^2T} times the exponential $\mathbb{P}$-martingale

$$Q(\cdot):=\exp\left\{\int_t^\cdot 2\theta(s,u_s,v_s)dB_s-\frac{1}{2}\int_t^\cdot 4|\theta(s,X,u_s,v_s)|^2ds\right\} \tag{3.1.43}$$

on $[0,T]$ with quadratic variation process

$$\langle Q\rangle(\cdot)=4\int_t^\cdot\left(Q^2(s)\int_t^s|\theta(r,u_r,v_r)|^2dr\right)ds. \tag{3.1.44}$$

But

$$\begin{aligned}&Q^2(\cdot)\int_t^\cdot|\theta(s,u_s,v_s)|^2ds\\ \le&A^2Te^{4A^2T}\exp\left\{\int_t^\cdot 4\theta(s,u_s,v_s)dB_s-\frac{1}{2}\int_t^\cdot 16|\theta(s,u_s,v_s)|^2ds\right\}.\end{aligned} \tag{3.1.45}$$

By the Burkholder-Davis-Gundy inequalities and inequality (3.1.45), there exists a constant C, such that $\mathbb{E}\left[\sup_{t\le s\le T}Q(s)\right]$ is dominated by

$$\begin{aligned}&2CAT^{1/2}e^{2A^2T}\mathbb{E}\left[\left(\int_t^T\exp\left\{\int_t^s 4\theta(r,u_r,v_r)dB_r-\frac{1}{2}\int_t^s 16|\theta(r,u_r,v_r)|^2dr\right\}ds\right)^{1/2}\right]\\ \le&2CAT^{1/2}e^{2A^2T}\left(\int_t^T\mathbb{E}\left[\exp\left\{\int_t^s 4\theta(r,u_r,v_r)dB_r-\frac{1}{2}\int_t^s 16|\theta(r,u_r,v_r)|^2dr\right\}\right]ds\right)^{1/2}\\ =&2CAT^{1/2}e^{2A^2T}(T-t)^{1/2}.\end{aligned}$$

This proves (3.1.41).
We may now state that

$$\begin{aligned}&\mathbb{E}^{u,v}\left[Y_1^{u,v}(T_1^n)\mathbb{1}_{\{T_1^n<\tau_t(u,v)\}}\middle|\mathscr{F}_t\right]\\ =&\mathbb{E}\left[\exp\left\{\int_t^{T_1^n}\theta(s,u_s,v_s)dB_s-\frac{1}{2}\int_t^{T_1^n}|\theta(s,u_s,v_s)|^2ds\right\}Y_1^{u,v}(T_1^n)\mathbb{1}_{\{T_1^n<\tau_t(u,v)\}}\middle|\mathscr{F}_t\right]\\ \to&0,\text{ as } n\to 0.\end{aligned} \tag{3.1.46}$$

The expressions (3.1.30), (3.1.31), (3.1.34), (3.1.35) and (3.1.46) together lead to

$$Y_1^{u,v}(t) \geq \mathbb{E}^{u,v}\left[L_1(\tau_t)\mathbb{1}_{\{\tau_t<T\}} + \xi_1\mathbb{1}_{\{\tau_t=T\}} + \int_t^{\tau_t} h_1(s,X,u_s,v_s)ds \middle| \mathscr{F}_t \right], \tag{3.1.47}$$

and

$$Y_1^{u,v}(t) = \mathbb{E}^{u,v}\left[L_1(\tau_t(u,v))\mathbb{1}_{\{\tau_t(u,v)<T\}} + \xi_1\mathbb{1}_{\{\tau_t(u,v)=T\}} + \int_t^{\tau_t(u,v)} h_1(s,X,u_s,v_s)ds \middle| \mathscr{F}_t\right], \tag{3.1.48}$$

which mean that

$$Y_1^{u,v}(t) = \mathbb{E}^{u,v}[R_t^1(\tau_t(u,v),\rho_t,u,v)|\mathscr{F}_t] \geq \mathbb{E}^{u,v}[R_t^1(\tau_t,\rho_t,u,v)|\mathscr{F}_t], \tag{3.1.49}$$

for all $\rho_t \in \mathscr{S}(t,T)$ and all $\tau_t \in \mathscr{S}(t,T)$.

To derive optimality of the controls (u^*,v^*) from Isaacs' condition Assumption 3.1.3, applying the comparison theorem (Theorem 3.2.2 and 3.3.3) to the first component of BSDE (3.1.23) gives $Y_1^{u^*,v^*}(\cdot) \geq Y_1^{u,v^*}(\cdot)$ a.e. on $[0,T]\times\Omega$. From the identity in (3.1.49),

$$\begin{aligned} &\mathbb{E}^{u,v^*}[R_t^1(\tau_t(u^*,v^*),\rho_t(u^*,v^*),u^*,v^*)|\mathscr{F}_t] = Y_1^{u^*,v^*}(t) \\ \geq &Y_1^{u,v^*}(t) = \mathbb{E}^{u,v^*}[R_t^1(\tau_t(u,v^*),\rho_t(u,v^*),u,v^*)|\mathscr{F}_t]. \end{aligned} \tag{3.1.50}$$

As a conjunction of (3.1.49) and (3.1.50), for all $\tau_t \in \mathscr{S}(t,T)$,

$$\begin{aligned} &\mathbb{E}^{u^*,v^*}[R_t^1(\tau_t(u^*,v^*),\rho_t(u^*,v^*),u^*,v^*)|\mathscr{F}_t] \\ \geq&\mathbb{E}^{u,v^*}[R_t^1(\tau_t(u,v^*),\rho_t(u,v^*),u,v^*)|\mathscr{F}_t] \\ \geq&\mathbb{E}^{u,v^*}[R_t^1(\tau_t,\rho_t(u,v^*),u,v^*)|\mathscr{F}_t]. \end{aligned} \tag{3.1.51}$$

The above arguments proceed with arbitrary stopping times $\rho_t \in \mathscr{S}(t,T)$, because player II's stopping time ρ_t does not enter player I's reward.

By symmetry between the two players,

$$Y_2^{u^*,v^*} = \mathbb{E}^{u^*,v^*}[R_t^2(\tau_t(u^*,v^*),\rho_t(u^*,v^*),u^*,v^*)|\mathscr{F}_t], \tag{3.1.52}$$

and

$$\mathbb{E}^{u^*,v^*}[R_t^2(\tau_t(u^*,v^*),\rho_t(u^*,v^*),u^*,v^*)|\mathscr{F}_t] \geq \mathbb{E}^{u^*,v}[R_t^2(\tau_t(u^*,v^*),\rho_t,u^*,v)|\mathscr{F}_t]. \tag{3.1.53}$$

Combining (3.1.50), (3.1.51), (3.1.52) and (3.1.53) implies, that the quadruplet (τ^*,ρ^*,u^*,v^*) is a Nash equilibrium and their value processes $V(\cdot)$ are identified with the solution to a BSDE with reflecting barrier with parameters $(T,\xi,H(u^*,v^*),L)$. The optimal controls (u^*,v^*) are chosen according to Isaacs' condition Assumption 3.1.3. Both players stop respectively according to the pair of rules (τ_t^*,ρ_t^*), as soon as their expected rewards hit the early stopping rewards $L_1(\cdot)$ and $L_2(\cdot)$ for the first time. □

Remark 3.1.1 *The absence of $L_i(\cdot)$ from the reward is equivalent to that the ith player never stops until time T, $i=1,2$. The corresponding BSDE for his optimal reward exhibits no reflecting barrier.*

Remark 3.1.2 *If the deterministic time T is replaced by a bounded $\{\mathscr{F}_t\}_{0\leq t\leq T}$-stopping time, it technically does not make any difference to results in this subsection.*

3.1.2 Controls observing volatility

This subsection discusses whether the inclusion of instantaneous volatilities of the value processes into the controls will expand the admissible control sets.

For the rewards considered in this chapter, when using control u and v, the $\mathbb{P}^{u,v}$-conditional expected rewards are $\mathbb{P}^{u,v}$-Brownian semimartingales with respect to the filtration $\{\mathscr{F}_t\}_{0\le t\le T}$, having the decompositions

$$\begin{aligned}\mathbb{E}^{u,v}[R^1_t(\tau,\rho,u,v)|\mathscr{F}_t] =& A^{u,v}_1(t) + M^{u,v}_1(t) = A^{u,v}_1(t) + \int_0^t Z^{u,v}_1(s)dB^{u,v}_1(s);\\ \mathbb{E}^{u,v}[R^2_t(\tau,\rho,u,v)|\mathscr{F}_t] =& A^{u,v}_2(t) + M^{u,v}_2(t) = A^{u,v}_2(t) + \int_0^t Z^{u,v}_2(s)dB^{u,v}_2(s).\end{aligned} \tag{3.1.54}$$

The processes $A^1(\cdot)$ and $A^2(\cdot)$ have finite variation. The processes $M^1(\cdot)$ and $M^2(\cdot)$ are $\mathbb{P}^{u,v}$-local martingales with respect to $\{\mathscr{F}_t\}_{0\le t\le T}$. The predictable, square-integrable processes $Z^{u,v}_1(\cdot)$ and $Z^{u,v}_2(\cdot)$ from martingale representation are called instantaneous volatility processes, the very integrand processes of the stochastic integrals in the BSDE (3.1.23). Because they naturally show up in the BSDEs solved by value process of the game, we may include the instantaneous volatilities $Z^{u,v}_1(\cdot)$ and $Z^{u,v}_2(\cdot)$ as arguments of the controls u and v, in the hope of making more informed decisions. Going one step further, in the case of risk-sensitive controls initiated by Whittle, Bensoussan and coworkers, among others, for example Bensoussan, Frehse and Nagai (1998) [5], the players are sensitive not only to the expectations, but also to the variances of their rewards. El Karoui and Hamadène (2003) identified in [18] risk-sensitive controls to BSDEs with quadratic growth in $Z^{u,v}_1(\cdot)$ and $Z^{u,v}_2(\cdot)$, which made the problem very tractable. Their value processes are different from the risk-indifferent case only up to an exponential transformation. Is it better to emphasize sensitivity to volatilities by including them as arguments of the controls?

Among the set of closed loop controls, including instantaneous volatilities into the controls means finding all deterministic measurable functionals $\mu : [0,T]\times\Omega\times\mathbb{R}^d\times\mathbb{R}^d\to\mathbb{A}_1$ and $\upsilon : [0,T]\times\Omega\times\mathbb{R}^d\times\mathbb{R}^d\to\mathbb{A}_2$, such that when applying the controls $u_t = \mu(t,X,Z_1(t),Z_2(t))$ and $v_t = \upsilon(t,X,Z_1(t),Z_2(t))$, for some $\{\mathscr{F}_t\}_{0\le t\le T}$-adapted processes $Z_1(\cdot)$ and $Z_2(\cdot)$, the resulted instantaneous volatilities $Z^{u,v}_1(\cdot)$ and $Z^{u,v}_2(\cdot)$ in the semimartingale decomposition (3.1.54) coincide with arguments $Z_1(\cdot)$ and $Z_2(\cdot)$ of μ and υ.

Including instantaneous volatilities into Markovian controls means the same as what is described in the previous paragraph, except that $\mu : [0,T]\times\mathbb{R}^d\times\mathbb{R}^d\times\mathbb{R}^d\to\mathbb{A}_1$ and $\upsilon : [0,T]\times\mathbb{R}^d\times\mathbb{R}^d\times\mathbb{R}^d\to\mathbb{A}_2$ are deterministic measurable functions, and that $u_t = \mu(t,X(t),Z_1(t),Z_2(t))$ and $v_t = \upsilon(t,X(t),Z_1(t),Z_2(t))$. This is the case about which we are going to have more to say.

The Hamiltonians in this case become

$$
\begin{aligned}
&H_1(t,\omega(t),z_1,(\mu,\upsilon)(t,\omega(t),z_1,z_2))\\
=&z_1\sigma^{-1}(t,\omega(t))f(t,\omega(t),(\mu,\upsilon)(t,\omega(t),z_1,z_2))+h_1(t,\omega(t),(\mu,\upsilon)(t,\omega(t),z_1,z_2));\\
&H_2(t,\omega(t),z_2,(\mu,\upsilon)(t,\omega(t),z_1,z_2))\\
=&z_2\sigma^{-1}(t,\omega(t))f(t,\omega(t),(\mu,\upsilon)(t,X(t),z_1,z_2))+h_2(t,\omega(t),(\mu,\upsilon)(t,\omega(t),z_1,z_2)),
\end{aligned}
\tag{3.1.55}
$$

for $0\le t\le T$, $\omega\in\Omega$, z_1 and z_2 in $\mathbb{R}^d$, and $\mathbb{A}_1\times\mathbb{A}_2$-valued measurable functions (μ,υ). From Assumption 3.1.1 (3) and Assumption 3.1.2 (2), the Hamiltonians are liner in z_1 and z_2, and polynomial in $\sup_{0\le s\le t}|\omega(s)|$. To be more specific, we have

$$
|H_i(t,\omega(t),z_1,z_2,(\mu,\upsilon)(t,\omega(t),z_1,z_2))|\le A|z_i|+C_{\text{rwd}}\left(1+\sup_{0\le s\le t}|\omega(s)|^{2p}\right),
\tag{3.1.56}
$$

for $i=1,2$, all $0\le t\le T$, $\omega\in\Omega$, z_1 and z_2 in $\mathbb{R}^d$, and $\mathbb{A}_1\times\mathbb{A}_2$-valued measurable functions (μ,υ). The growth rates of the Hamiltonians (3.1.55) satisfy Assumption 3.3.1 (2) for the driver of the BSDE (3.3.2). With all other assumption on the coefficients also satisfied, by Theorem 3.3.2, there exists a solution $(Y^{\mu,\upsilon},Z^{\mu,\upsilon},K^{\mu,\upsilon})$ to the following equation

$$
\begin{cases}
Y_1^{\mu,\upsilon}(t)=\xi_1+\displaystyle\int_t^T H_1(s,X(s),Z_1^{\mu,\upsilon}(s),(\mu,\upsilon)(s,X(s),Z_1^{\mu,\upsilon}(s),Z_2^{\mu,\upsilon}(s)))ds\\
\qquad\qquad -\displaystyle\int_t^T Z_1^{\mu,\upsilon}(s)dB_s+K_1^{\mu,\upsilon}(T)-K_1^{\mu,\upsilon}(t),\\
\qquad\qquad Y_1^{\mu,\upsilon}(t)\ge L_1(t),\ 0\le t\le T;\ \displaystyle\int_0^T(Y_1^{\mu,\upsilon}(t)-L_1(t))dK_1^{\mu,\upsilon}(t)=0;\\
Y_2^{\mu,\upsilon}(t)=\xi_2+\displaystyle\int_t^T H_2(s,X(s),Z_2^{\mu,\upsilon}(s),(\mu,\upsilon)(s,X(s),Z_1^{\mu,\upsilon}(s),Z_2^{\mu,\upsilon}(s)))ds\\
\qquad\qquad -\displaystyle\int_t^T Z_2^{\mu,\upsilon}(s)dB_s+K_2^{\mu,\upsilon}(T)-K_2^{\mu,\upsilon}(t),\\
\qquad\qquad Y_2^{\mu,\upsilon}(t)\ge L_2(t),\ 0\le t\le T;\ \displaystyle\int_0^T(Y_2^{\mu,\upsilon}(t)-L_2(t))dK_2^{\mu,\upsilon}(t)=0.
\end{cases}
\tag{3.1.57}
$$

Assumption 3.1.4 *(Isaacs' condition) There exist deterministic functions* $\mu^*:[0,T]\times\mathbb{R}^d\times\mathbb{R}^d\times\mathbb{R}^d\to\mathbb{A}_1$ *and* $\upsilon^*:[0,T]\times\mathbb{R}^d\times\mathbb{R}^d\times\mathbb{R}^d\to\mathbb{A}_2$*, such that*

$$
\begin{aligned}
H_1(t,x,z_1,(\mu^*,\upsilon^*)(t,x,z_1,z_2))&\ge\sup_{\bar z_1,\bar z_2\in\mathbb{R}^d}H_1(t,x,z_1,(\mu,\upsilon^*)(t,x,\bar z_1,\bar z_2));\\
H_2(t,x,z_2,(\mu^*,\upsilon^*)(t,x,z_1,z_2))&\ge\sup_{\bar z_1,\bar z_2\in\mathbb{R}^d}H_2(t,x,z_2,(\mu^*,\upsilon)(t,x,\bar z_1,\bar z_2)),
\end{aligned}
\tag{3.1.58}
$$

for all $0\le t\le T$*,* x*,* z_1 *and* z_2 *in* $\mathbb{R}^d$*, and all* $\mu:[0,T]\times\mathbb{R}^d\times\mathbb{R}^d\times\mathbb{R}^d\to\mathbb{A}_1$ *and* $\upsilon:[0,T]\times\mathbb{R}^d\times\mathbb{R}^d\times\mathbb{R}^d\to\mathbb{A}_2$.

Associated with coefficients f and σ of the state process $X(\cdot)$ and with the rewards h, $L(\cdot)$ and ξ, the admissible set $\mathscr{U} \times \mathscr{V} = \{(u, v)\}$ of Markovian controls that observe volatilities are defined as the collection of all

$$(u_t, v_t) = (\mu, \upsilon)(t, X(t), Z_1^{\mu,\upsilon}(t), Z_2^{\mu,\upsilon}(t)), \tag{3.1.59}$$

for measurable functions $\mu : [0, T]\times\mathbb{R}^d\times\mathbb{R}^d\times\mathbb{R}^d \to \mathbb{A}_1$ and $\upsilon : [0, T]\times\mathbb{R}^d\times\mathbb{R}^d\times\mathbb{R}^d \to \mathbb{A}_2$. In particular,

$$(u_t^*, v_t^*) = (\mu^*, \upsilon^*)(t, X(t), Z_1^{\mu^*,\upsilon^*}(t), Z_2^{\mu^*,\upsilon^*}(t)), \tag{3.1.60}$$

$$(u_t, v_t^*) = (\mu, \upsilon^*)(t, X(t), Z_1^{\mu,\upsilon^*}(t), Z_2^{\mu,\upsilon^*}(t)), \tag{3.1.61}$$

and

$$(u_t^*, v_t) = (\mu^*, \upsilon)(t, X(t), Z_1^{\mu^*,\upsilon}(t), Z_2^{\mu^*,\upsilon}(t)). \tag{3.1.62}$$

Assumption 3.1.4 implies Isaacs' condition, Assumption 3.1.3. Then we reach the same statements as in Theorem 3.1.1, the only difference being $(Y^{u,v}, Z^{u,v}, K^{u,v})$ replaced by $(Y^{\mu,\upsilon}, Z^{\mu,\upsilon}, K^{\mu,\upsilon})$, and BSDE (3.1.23) replaced by BSDE (3.1.57).

In fact, by Theorem 3.3.1, there exist deterministic measurable mappings $\beta_1^{\mu,\upsilon}$ and $\beta_2^{\mu,\upsilon} : [0, T] \times \mathbb{R}^d \to \mathbb{R}^d$, such that $Z_1^{\mu,\upsilon}(t) = \beta_1^{\mu,\upsilon}(t, X(t))$, and $Z_2^{\mu,\upsilon}(t) = \beta_2^{\mu,\upsilon}(t, X(t))$, for all $0 \le t \le T$. Hence (3.1.59) becomes

$$(u_t, v_t) = (\mu, \upsilon)(t, X(t), \beta_1^{\mu,\upsilon}(t, X(t)), \beta_2^{\mu,\upsilon}(t, X(t))), \tag{3.1.63}$$

a pair of Markovian controls.

3.1.3 Rewards terminated by either player

In this subsection, Game 3.1.2 is studied. One player's time to quit the game is determined by the conjunction of both players' stopping rules. As soon as one player stops, the Game 3.1.2 is terminated. When quitting the game, player I receives reward

$$R_0^1(\tau, \rho, u, v) = \int_0^{\tau\wedge\rho} h_1(s, u, v)ds + \begin{cases} L_1(\tau), \text{ if player 1 stops first;} \\ U_1(\rho), \text{ if player 2 stops first;} \\ \xi_1, \text{ if neither stops before time } T; \end{cases} \tag{3.1.64}$$

whereas player II receives reward

$$R_0^2(\tau, \rho, u, v) = \int_0^{\tau\wedge\rho} h_2(s, u, v)ds + \begin{cases} U_2(\tau), \text{ if player 1 stops first;} \\ L_2(\rho), \text{ if player 2 stops first;} \\ \xi_2, \text{ if neither stops before time } T. \end{cases} \tag{3.1.65}$$

Optimal controls for Game 3.1.2 will again be the pair (u^*, v^*) from Isaacs' condition, Assumption 3.1.3. The interaction of stopping rules seems complicated. Let us temporarily ignore the controls and focus on reducing the game of stopping to a tractable formulation.

For any fixed stopping rules τ_t^0 and ρ_t^0 in $\mathscr{S}(t,T)$, let player I choose stopping rule $\tau_t = \tau_t^1$ from $\mathscr{S}(t,T)$ and player II choose $\rho_t = \rho_t^1$ from $\mathscr{S}(t,T)$ to maximize their respective rewards

$$\begin{aligned}
&R_t^1(\tau_t, \rho_t^0, u, v)\\
:=&\int_t^{\tau_t\wedge\rho_t^0} h_1(s,X,u_s,v_s)ds + L_1(\tau_t)\mathbb{1}_{\{\tau_t<\rho_t^0\}} + U_1(\rho_t^0)\mathbb{1}_{\{\rho_t^0\le\tau_t<T\}} + \xi_1\mathbb{1}_{\{\tau_t\wedge\rho_t^0=T\}};\\
&R_t^2(\tau_t^0, \rho_t, u, v)\\
:=&\int_t^{\tau_t^0\wedge\rho_t} h_2(s,X,u_s,v_s)ds + L_2(\rho_t)\mathbb{1}_{\{\rho_t<\tau_t^0\}} + U_2(\tau_t^0)\mathbb{1}_{\{\tau_t^0\le\rho_t<T\}} + \xi_2\mathbb{1}_{\{\tau_t^0\wedge\rho_t=T\}},
\end{aligned} \tag{3.1.66}$$

in conditional $\mathbb{P}^{u,v}$-expectations. With a little abuse of the notation $U_1(\cdot)$ and $U_2(\cdot)$ as in Assumption 3.1.2 (1), rewrite

$$\begin{aligned}
U_1(\rho_t^0)\mathbb{1}_{\{\rho_t^0\le\tau_t<T\}} + \xi_1\mathbb{1}_{\{\tau_t\wedge\rho_t^0=T\}} &= U_1(\rho_t^0)\mathbb{1}_{\{\tau_t\ge\rho_t^0\}};\\
U_2(\tau_t^0)\mathbb{1}_{\{\tau_t^0\le\rho_t<T\}} + \xi_2\mathbb{1}_{\{\tau_t^0\wedge\rho_t=T\}} &= U_2(\tau_t^0)\mathbb{1}_{\{\rho_t\ge\tau_t^0\}}.
\end{aligned} \tag{3.1.67}$$

But suggested by (3.1.66), on $\{\tau_t \ge \rho_t^0\}$, player I's running reward is cut off at time ρ_t^0, and terminal reward remains $U_1(\rho_t^0)$ anyway, so he will not profit from sticking to the game after time ρ_t^0. Symmetrically, player II will not profit from stopping after τ_t^0. Because of the indifference to late stopping, maximizing expected rewards (3.1.66) is equivalent to choosing $\tau_t = \tau_t^1$ from $\mathscr{S}(t,\rho_t^0)$ and $\rho_t = \rho_t^1$ from $\mathscr{S}(t,\tau_t^0)$ to maximize the conditional $\mathbb{P}^{u,v}$-expectations of

$$\begin{aligned}
&\int_t^{\tau_t\wedge\rho_t^0} h_1(s,X,u_s,v_s)ds + L_1(\tau_t)\mathbb{1}_{\{\tau_t<\rho_t^0\}} + U_1(\rho_t^0)\mathbb{1}_{\{\tau_t=\rho_t^0\}};\\
&\int_t^{\tau_t^0\wedge\rho_t} h_2(s,X,u_s,v_s)ds + L_2(\rho_t)\mathbb{1}_{\{\rho_t<\tau_t^0\}} + U_2(\tau_t^0)\mathbb{1}_{\{\rho_t=\tau_t^0\}}.
\end{aligned} \tag{3.1.68}$$

In the spirit of Nash's 1949 original definition of equilibrium, the equilibrium stopping rules (τ_t^*, ρ_t^*) of Game 3.1.2 is a fixed point of the mapping

$$\begin{aligned}
\Gamma :&\mathscr{S}(t,T)\times\mathscr{S}(t,T) \to \mathscr{S}(t,T)\times\mathscr{S}(t,T),\\
&(\tau_t^0,\rho_t^0) \mapsto (\tau_t^1,\rho_t^1).
\end{aligned} \tag{3.1.69}$$

To show existence of equilibrium stopping rules, it suffices to prove a.e. convergence of iteration via Γ, starting from a certain initial stopping rule.

This reduction will solve Game 3.1.2 by approximating it with a sequence of much simpler optimization problems. The optimization is in a simplified form of Game 3.1.1, hence it can be associated with a BSDE with reflecting barrier. The admissible set $\mathscr{U}\times\mathscr{V}$ of controls are still closed loop. At every step of the iteration, there is no interaction in either controls or stopping. Without interaction, the resulting two-dimensional BSDE for the players consists in fact of two separate one-dimensional equations. Hence the comparison theorem for one-dimensional equations applies to the derivation of the pair of equilibrium controls (u^*, v^*) from Assumption 3.1.3 at every step of the iteration. So (u^*, v^*) should also be equilibrium in the limit. The first time when the value process hits the lower reflecting boundary is the optimal time to stop.

Lemma 3.1.1 *Let the players' rewards be as in (3.1.68), the value process $V(\cdot)$ as in (3.1.15), and (u^*, v^*) as in Isaacs' condition, Assumption 3.1.3. The triplet $(Y^{u,v}, Z^{u,v}, K^{u,v})$ satisfies $Y^{u,v}(\cdot) \in \mathbb{M}^2(2; 0, T)$, $Z^{u,v}(\cdot) \in \mathbb{L}^2(2\times d; 0, T)$, and $K^{u,v}(\cdot)$ continuous increasing in $\mathbb{M}^2(2; 0, T)$ solves*

$$\begin{cases} Y_1^{u,v}(t) = U_1(\rho_t^0) + \int_t^{\rho_t^0} H_1(s, X, Z_1^{u,v}(s), u_s, v_s)ds - \int_t^{\rho_t^0} Z_1^{u,v}(s)dB_s \\ \qquad + K_1^{u,v}(\rho_t^0) - K_1^{u,v}(t),\ 0 \le t \le \rho_t^0; \\ \qquad Y_1^{u,v}(t) \ge L_1(t),\ t \in [0, \rho_t^0];\ \int_0^{\rho_t^0} (Y_1^{u,v}(t) - L_1(t))dK_1^{u,v}(t) = 0; \\ Y_2^{u,v}(t) = U_2(\tau_t^0) + \int_t^{\tau_t^0} H_2(s, X, Z_2^{u,v}(s), u_s, v_s)ds - \int_t^{\tau_t^0} Z_2^{u,v}(s)dB_s \\ \qquad + K_2^{u,v}(\rho_t^0) - K_2^{u,v}(t),\ 0 \le t \le \tau_t^0; \\ \qquad Y_2^{u,v}(t) \ge L_2(t),\ t \in [0, \tau_t^0];\ \int_0^{\tau_t^0} (Y_2^{u,v}(t) - L_2(t))dK_2^{u,v}(t) = 0. \end{cases} \tag{3.1.70}$$

For player I, choose the stopping time $\tau_t^1 := \tau_t^*\left(Y_1^{u,v}(\cdot); \rho_t^0\right)$, and for player II, choose the stopping time $\rho_t^1 := \rho_t^*\left(Y_2^{u,v}(\cdot); \tau_t^0\right)$, where the stopping rules τ^* and ρ^* are defined in (3.1.24) and (3.1.25). The quadruplet $(\tau^1, \rho^1, u^*, v^*)$ is optimal in the sense that

$$\begin{aligned} &\mathbb{E}^{u^*,v^*}[R_t^1(\tau_t^1(u^*, v^*), \rho_t^0, u^*, v^*)|\mathscr{F}_t] \ge \mathbb{E}^{u,v^*}[R_t^1(\tau_t, \rho_t^0, u, v^*)|\mathscr{F}_t],\ \forall \tau_t \in \mathscr{S}(t, \rho_t^0),\ \forall u \in \mathscr{U}; \\ &\mathbb{E}^{u^*,v^*}[R_t^2(\tau_t^0, \rho_t^1(u^*, v^*), u^*, v^*)|\mathscr{F}_t] \ge \mathbb{E}^{u^*,v}[R_t^2(\tau_t^0, \rho_t, u^*, v)|\mathscr{F}_t],\ \forall \rho_t \in \mathscr{S}(t, \tau_t^0),\ \forall v \in \mathscr{V}. \end{aligned} \tag{3.1.71}$$

Furthermore, $V_i(t) = Y_i^{u^*,v^*}(t)$, $0 \le t \le T$, $i = 1, 2$.

Proof. Apply Theorem 3.1.1 to each individual player. □

The following arguments proceed for $h_i \ge 0$. If in general $h_i \ge -c$ bounded from below, then the arguments should be tailored by shifting upwards the rewards and value processes.

Now we start an iteration via Γ, defined by (3.1.69), with $\tau_t^0 = \rho_t^0 = T$, and $Y_1^0(\cdot) = Y_2^0(\cdot) = +\infty$. Put the controls $(u, v) = (u^*, v^*)$. As in Lemma 3.1.1, τ^1 and ρ^1 are the two players' optimal stopping rules when their respective terminal times are ρ^0 and τ^0. In the language of the fixed point formulation, $(\tau_t^1, \rho_t^1) = \Gamma(\tau_t^0, \rho_t^0)$. Apparently, $\tau_t^1 \vee \rho_t^1 \le \tau_t^0 \wedge \rho_t^0 = T$. Assume

$$\tau_t^n \vee \rho_t^n \le \tau_t^{n-1} \wedge \rho_t^{n-1} \tag{3.1.72}$$

for n. Denote τ_t^{n+1} and ρ_t^{n+1} as their stopping rules that attain the superema in

$$Y_1^{n+1}(t) = \sup_{\tau_t \in \mathscr{S}(t,\rho_t^n)} \mathbb{E}^{u^*,v^*}\left[\int_t^{\tau_t \wedge \rho_t^n} h_1(s,X,u_s^*,v_s^*)ds + L_1(\tau_t)\mathbb{1}_{\{\tau_t<\rho_t^n\}} + U_1(\rho_t^n)\mathbb{1}_{\{\tau_t=\rho_t^n\}} \middle| \mathscr{F}_t\right] \tag{3.1.73}$$

and

$$Y_2^{n+1}(t) = \sup_{\rho_t \in \mathscr{S}(t,\tau_t^n)} \mathbb{E}^{u^*,v^*}\left[\int_t^{\tau_t^n \wedge \rho_t} h_2(s,X,u_s^*,v_s^*)ds + L_2(\rho_t)\mathbb{1}_{\{\rho_t<\tau_t^n\}} + U_2(\tau_t^n)\mathbb{1}_{\{\rho_t=\tau_t^n\}} \middle| \mathscr{F}_t\right], \tag{3.1.74}$$

given the two players' respective terminal times are ρ^n and τ^n, then $(\tau_t^{n+1},\rho_t^{n+1}) = \Gamma(\tau_t^n,\rho_t^n)$ in the fixed point language. There exists a pair of stopping rules (τ^{n+1},ρ^{n+1}) that attains the suprema in (3.1.73) and (3.1.74), by replacing the notations (τ^0,ρ^0) with (τ^n,ρ^n) and (τ^1,ρ^1) with (τ^{n+1},ρ^{n+1}) in Lemma 3.1.1. According to Lemma 3.1.1, together with (3.1.9) and (3.1.19), for $n = 1,2,\cdots$, the processes $Y^n(\cdot) \in \mathbb{M}^2(2;0,T)$, $Z^n(\cdot) \in \mathbb{L}^2(2\times d;0,T)$, and $K^n(\cdot)$ continuous increasing in $\mathbb{M}^2(2;0,T)$ satisfy

$$\begin{cases} dY_1^n(s) = -H_1(s,X,Z_1^n(s),u_s^*,v_s^*)ds + Z_1^n(s)dB_s - dK_1^n(s) \\ \qquad = -h_1(s,X,u_s^*,v_s^*)ds + Z_1^n(s)dB_s^{u^*,v^*} - dK_1^n(s),\ t \le s \le \rho_t^{n-1}; \\ Y_1^n(\rho_t^{n-1}) = U_1(\rho_t^{n-1}); \\ dY_2^n(s) = -H_2(s,X,Z_2^n(s),u_s^*,v_s^*)ds + Z_2^n(s)dB_s - dK_2^n(s) \\ \qquad = -h_2(s,X,u_s^*,v_s^*)ds + Z_2^n(s)dB_s^{u^*,v^*} - dK_2^n(s),\ t \le s \le \tau_t^{n-1}; \\ Y_2^n(\tau_t^{n-1}) = U_2(\tau_t^{n-1}). \end{cases} \tag{3.1.75}$$

Integrating dY_1^n from ρ_t^n to ρ_t^{n-1}, and dY_2^n from τ_t^n to τ_t^{n-1}, then taking $\mathbb{P}^{u^*,v^*}-$expectations, and conditioning on $\mathscr{F}_{\rho_t^n}$ and $\mathscr{F}_{\tau_t^n}$, respectively, we obtain

$$\begin{aligned} Y_1^n(\rho_t^n) =& \mathbb{E}^{u^*,v^*}\Bigg[U_1(\rho_t^{n-1}) + \int_{\rho_t^n}^{\rho_t^{n-1}} h_1(s,X,u_s^*,v_s^*)ds - \int_{\rho_t^n}^{\rho_t^{n-1}} Z_1^n(s)dB_s^{u^*,v^*} \\ &+ K_1^n(\rho^{n-1}) - K_1^n(\rho_t^n)\Bigg|\mathscr{F}_{\rho_t^n}\Bigg] \\ \ge& \mathbb{E}^{u^*,v^*}[U_1(\rho_t^{n-1})|\mathscr{F}_{\rho_t^n}] \ge \mathbb{E}^{u^*,v^*}[U_1(\rho_t^n)|\mathscr{F}_{\rho_t^n}] = U_1(\rho_t^n); \\ Y_2^n(\tau_t^n) =& \mathbb{E}^{u^*,v^*}\Bigg[U_2(\tau_t^{n-1}) + \int_{\tau_t^n}^{\tau_t^{n-1}} h_2(s,X,u_s^*,v_s^*)ds - \int_{\tau_t^n}^{\tau_t^{n-1}} Z_2^n(s)dB_s^{u^*,v^*} \\ &+ K_2^n(\tau_t^{n-1}) - K_2^n(\tau_t^n)\Bigg|\mathscr{F}_{\tau_t^n}\Bigg] \\ \ge& \mathbb{E}^{u^*,v^*}[U_2(\tau_t^{n-1})|\mathscr{F}_{\tau_t^n}] \ge \mathbb{E}^{u^*,v^*}[U_2(\tau_t^n)|\mathscr{F}_{\tau_t^n}] = U_2(\tau_t^n). \end{aligned} \tag{3.1.76}$$

The first pair of inequalities in the above two entries come from the nonnegativity assumptions of h_1 and h_2, and the fact that $K_1(\cdot),K_2(\cdot)$ are increasing processes, once more with the help of the same localization technique in the proof of Theorem 3.1.1. The second pair of inequalities come from the induction assumption $\tau_t^n \vee \rho_t^n \le \tau_t^{n-1} \wedge \rho_t^{n-1}$, and the monotonicity assumption of $U(\cdot)$ in Assumption 3.1.2 (1). One can get rid of conditional expectations as in the final pair of equalities, because by Assumption 3.1.2 (1) the process $U(\cdot)$ is progressively measurable with respect to the filtration $\{\mathscr{F}_t\}_{0\le t\le T}$. But $U_1(\rho_t^n) = Y_1^{n+1}(\rho_t^n)$, and $U_2(\tau_t^n) = Y_2^{n+1}(\tau_t^n)$, hence $Y_1^{n+1}(\rho_t^n) \le Y_1^n(\rho_t^n)$, and $Y_2^{n+1}(\tau_t^n) \le Y_2^n(\tau_t^n)$. By the comparison theorem (Theorem 3.2.2 and Theorem 3.3.3) in dimension one, $Y_1^{n+1}(s) \le Y_1^n(s)$, for all $t \le s \le \rho_t^n$, and $Y_2^{n+1}(s) \le Y_2^n(s)$, for all $t \le s \le \tau_t^n$. From Lemma 3.1.1, for all $n = 1, 2, \cdots$, the optimal stopping times $\tau_t^{n+1} := \tau_t^*\left(Y_1^{n+1}(\cdot);\rho_t^n\right) \le \rho_t^n$, and $\rho_t^{n+1} := \rho_t^*\left(Y_2^{n+1}(\cdot);\tau_t^n\right) \le \tau_t^n$, where the stopping rules τ^* and ρ^* are defined in (3.1.24) and (3.1.25). Then $Y^{n+1}(\cdot) \le Y^n(\cdot)$ implies $\tau_t^{n+1} \le \tau_t^n$, and $\rho_t^{n+1} \le \rho_t^n$. Finally, we have finished the $(n+1)$th step of mathematical induction by concluding

$$\tau_t^{n+1} \vee \rho_t^{n+1} \le \tau_t^n \wedge \rho_t^n. \tag{3.1.77}$$

The sequences $\{Y^n(\cdot)\}_n$, $\{\tau_t^n\}_n$ and $\{\rho_t^n\}_n$ from the induction are all decreasing, thus have pointwise limits $Y^*(\cdot)$, τ_t^* and ρ_t^*.

By analogy with the argument used to prove Theorem 3.1.1, we have

$$\begin{aligned}
Y_1^n(t) \ge &\int_t^{\tau_t\wedge\rho_t^{n-1}} h_1(s,X,u_s,v_s^*)ds + L_1(\tau_t)\mathbb{1}_{\{\tau_t<\rho_t^n\}} + U_1(\rho_t^{n-1})\mathbb{1}_{\{\tau_t=\rho_t^{n-1}\}}\\
&+ \int_t^{\tau_t\wedge\rho_t^{n-1}} Z_1^n(s)dB_s^{u,v^*};\\
Y_2^n(t) \ge &\int_t^{\tau_t^{n-1}\wedge\rho_t} h_2(s,X,u_s^*,v_s)ds + L_2(\rho_t)\mathbb{1}_{\{\rho_t<\tau_t^{n-1}\}} + U_2(\tau_t^{n-1})\mathbb{1}_{\{\rho_t=\tau_t^{n-1}\}}\\
&+ \int_t^{\rho_t\wedge\tau_t^{n-1}} Z_2^n(s)dB_s^{u^*,v}.
\end{aligned} \tag{3.1.78}$$

The inequalities in (3.1.78) become equalities, if $\tau_t = \tau_t^n$, $u = u^*$ and $\rho_t = \rho_t^n$, $v = v^*$. First taking corresponding conditional expectations of (3.1.78) with respect to $\mathscr{F}_t$, the stochastic integrals vanish still by the localization technique as in proof of Theorem 3.1.1. Then letting $n \to \infty$, and using the equivalence between maximizing (3.1.66) and maximizing (3.1.68), we arrive at

$$\begin{aligned}
Y_1^*(t) &= \mathbb{E}^{u^*,v^*}[R_t^1(\tau_t^*,\rho_t^*,u^*,v^*)|\mathscr{F}_t] \ge \mathbb{E}^{u,v^*}[R_t^1(\tau_t,\rho^*,u,v^*)|\mathscr{F}_t],\ \forall\tau_t\in\mathscr{S}(t,T),\ \in\mathscr{U};\\
Y_2^*(t) &= \mathbb{E}^{u^*,v^*}[R_t^2(\tau_t^*,\rho_t^*,u^*,v^*)|\mathscr{F}_t] \ge \mathbb{E}^{u^*,v}[R_t^1(\tau_t^*,\rho_t,u^*,v)|\mathscr{F}_t],\ \forall\rho_t\in\mathscr{S}(t,T),\ \in\mathscr{V},
\end{aligned} \tag{3.1.79}$$

with rewards R^1 and R^2 as in (3.1.66).

The inductive procedure produces a Nash equilibrium (τ^*,ρ^*,u^*,v^*) for Game 3.1.2. The equilibrium controls (u^*,v^*) come from Isaacs' condition, Assumption 3.1.3. The equilibrium stopping rules (τ^*,ρ^*) are the limits of the iterative sequence of optimal stopping rules, thus provide a fixed point of the mapping Γ defined in (3.1.69).

Theorem 3.1.2 *Under Assumptions 3.1.1, 3.1.2 and 3.1.3, if h is bounded from below, then the limit (τ^*,ρ^*,u^*,v^*) from the iteration is an equilibrium point of Game 3.1.2.*

Remark 3.1.3 *Game 3.1.2 always has a trivial equilibrium* $(t, t, 0, 0)$. *The iterative procedure in this section can be numerically implemented to determine if the limiting equilibrium point* $(\tau^*, \rho^*, u^*, v^*)$ *is trivial or not.*

3.2 A multidimensional reflected BSDE with Lipschitz growth

Starting from this section, we solve multidimensional BSDEs with reflecting barriers, the type of BSDEs associated with Game 3.1.1, and provide two useful properties of the equations, the comparison theorem in dimension one and the theorem about continuous dependence of the solution on the terminal values. The discussions on the BSDEs are postponed until here, only to finish the game part first. Proofs of results to be stated from now on in this paper do not depend on any earlier arguments.

This section assumes the following the following Lipschitz growth condition and integrability conditions on the parameters of the equations.

Assumption 3.2.1 *(1) The driver g is a mapping* $g : [0, T] \times \mathbb{R}^m \times^{m\times d} \to \mathbb{R}^m$, $(t, y, z) \mapsto g(t, y, z)$. *For every fixed* $y \in \mathbb{R}^m$ *and* $z \in \mathbb{R}^{m\times d}$, *the process* $\{g(t, y, z)\}_{0\le t\le T}$ *is* $\{\mathscr{F}_t\}_{0\le t\le T}$*-predictable. For all* $t \in [0, T]$, $g(t, y, z)$ *is uniformly Lipschitz in y and z, i.e. there exists a constant* $b > 0$, *such that*

$$|g(t, y, z) - g(t, \overline{y}, \overline{z})| \le b(||y - \overline{y}|| + ||z - \overline{z}||), \tag{3.2.1}$$

for all $t \in [0, T]$, $y \in \mathbb{R}^m$ *and* $z \in \mathbb{R}^{m\times d}$. *Furthermore,*

$$\mathbb{E}\left[\int_0^T g(t, 0, 0)^2 dt\right] < \infty. \tag{3.2.2}$$

(2) The random variable ξ *is* $\mathscr{F}_T$*-measurable and square-integrable. The lower reflecting boundary L is continuous, progressively measurable, and satisfies*

$$\mathbb{E}\left[\sup_{[0,T]} L^+(t)^2\right] < \infty. \tag{3.2.3}$$

Also, $L(T) \le \xi$, *a.e. on* Ω.

Under Assumption 3.2.1, this section proves existence and uniqueness of solution (Y, Z, K) to the following BSDE

$$\begin{cases} Y(t) = \xi + \int_t^T g(s, Y(s), Z(s))ds - \int_t^T Z(s)dB_s + K(T) - K(t); \\ Y(t) \ge L(t),\ 0 \le t \le T,\ \int_0^T (Y(t) - L(t))dK(t) = 0, \end{cases} \tag{3.2.4}$$

in the spaces

$$\begin{aligned}
&Y(\cdot) = (Y_1(\cdot), \cdots, Y_m(\cdot))' \in \mathbb{M}^2(m; 0, T) \\
=&\left\{m\text{-dimensional RCLL predictable process } \phi(\cdot) \text{ s.t. } \mathbb{E}\left[\sup_{[0,T]} \phi_t^2\right] \le \infty\right\}; \\
&Z(\cdot) = (Z_1(\cdot), \cdots, Z_m(\cdot))' \in \mathbb{L}^2(m \times d; 0, T) \\
=&\left\{m \times d\text{-dimensional RCLL predictable process } \phi(\cdot) \text{ s.t. } \mathbb{E}\left[\int_0^T \phi_t^2 dt\right] \le \infty\right\}; \\
&K(\cdot) = (K_1(\cdot), \cdots, K_m(\cdot))'\text{: continuous, increasing process in } \mathbb{M}^2(m; 0, T),
\end{aligned} \tag{3.2.5}$$

where the positive integer m is the dimension of the equation. The backward equation and the reflecting condition in (3.2.4) should be interpreted component-wise. It means that, for every $i = 1, \cdots, m$, in the ith dimension,

$$\begin{cases} Y_i(t) = \xi_i + \int_t^T g_i(s, Y(s), Z(s))ds - \int_t^T Z_i(s)dB_s + K_i(T) - K_i(t); \\ Y_i(t) \ge L_i(t), 0 \le t \le T, \ \int_0^T (Y_i(t) - L_i(t))dK_i(t) = 0. \end{cases} \tag{3.2.6}$$

The process $Y_i(\cdot)$ is motivated by the Brownian noise $B(\cdot)$ as the "fuel", whose amount is determined by a "control" $Z_i(\cdot)$. The driver g_i leads $Y_i(\cdot)$ towards the "final destination" ξ_i. Whenever the ith component $Y_i(\cdot)$ drops to the lower reflecting boundary $L_i(\cdot)$, it receives a force $K_i(\cdot)$ that kicks it upwards. When $Y_i(\cdot)$ stays above level $L_i(\cdot)$, the force $K_i(\cdot)$ does not apply. The process $K_i(\cdot)$ stands for the minimum cumulative exogenous energy required to keep $Y_i(\cdot)$ above level $L_i(\cdot)$. The m equations compose a system of m "vehicles" whose "drivers" track each other. For notational simplicity, the vector form (3.2.4) is used as a shorthand.

Lemma 3.2.1 *For any processes* $(Y^0(\cdot), Z^0(\cdot)) \in \mathbb{L}^2(m; 0, T) \times \mathbb{L}^2(m \times d; 0, T)$, *there exist unique* $(Y^1(\cdot), Z^1(\cdot)) \in \mathbb{M}^2(m; 0, T) \times \mathbb{L}^2(m \times d; 0, T)$, *and* $K^1(\cdot) \in \mathbb{M}^2(m; 0, T)$, *such that*

$$\begin{cases} dY^1(t) = -g(t, Y^0(t), Z^0(t))dt + Z^1(t)dB_t - dK^1(t), 0 \le t \le T; \\ Y^1(T) = \xi; \\ Y^1(t) \ge L(t), 0 \le t \le T, \ \int_0^T (Y^1(t) - L(t))dK^1(t) = 0. \end{cases} \tag{3.2.7}$$

Proof. For any $i = 1, \cdots, m$, in the ith dimension, by Corollary 3.7 of El Karoui, Kapoudjian, Pardoux, Peng and Quenez (1997) [19], there exists a unique solution $(Y_i^1(\cdot), Z_i^1(\cdot)) \in \mathbb{M}^2(1; 0, T) \times \mathbb{L}^2(d; 0, T)$, and a continuous, increasing process $K_i^1(\cdot) \in$

$\mathbb{M}^2(1;0,T)$, to the one-dimensional reflected BSDE

$$\begin{cases} dY_i^1(t) = -g_i(t, Y^0(t), Z^0(t))dt + Z_i^1(t)dB_t - dK_i^1(t),\ 0 \le t \le T; \\ Y_i^1(T) = \xi_i; \\ Y_i^1(t) \ge L_i(t),\ 0 \le t \le T,\ \displaystyle\int_0^T (Y_i^1(t) - L_i(t))dK_i^1(t) = 0. \end{cases} \tag{3.2.8}$$

The processes $Y^1(\cdot) := (Y_1^1(\cdot), \cdots, Y_m^1(\cdot))'$, $Z^1(\cdot) := (Z_1^1(\cdot), \cdots, Z_m^1(\cdot))'$, and $K^1(\cdot) := (K_1^1(\cdot), \cdots, K_m^1(\cdot))'$ form the desired triplet. □

To prove existence and uniqueness of the solution to the multi-dimensional BSDE (3.2.4) with reflecting barrier, it suffices to show that the mapping

$$\begin{aligned} \Lambda :&\mathbb{L}^2(m;0,T) \times \mathbb{L}^2(m \times d;0,T) \to \mathbb{L}^2(m;0,T) \times \mathbb{L}^2(m \times d;0,T), \\ &(Y^0, Z^0) \mapsto (Y^1, Z^1), \end{aligned} \tag{3.2.9}$$

is a contraction.

Theorem 3.2.1 *The mapping Λ is a contraction from $\mathbb{L}^2(m;0,T) \times \mathbb{L}^2(m \times d;0,T)$ to $\mathbb{L}^2(m;0,T) \times \mathbb{L}^2(m \times d;0,T)$.*

Proof. For a progressively measurable process $\phi(\cdot)$, the norm $||\phi||_2 := \sqrt{\mathbb{E}\Big[\int_0^T \phi_t^2 dt\Big]}$ is equivalent to the norm $||\phi||_{2,\beta} := \sqrt{\mathbb{E}\Big[\int_0^T e^{\beta t}\phi_t^2 dt\Big]}$. We prove the contraction statement under the norm $\|\cdot\|_{2,\beta}$. Suppose $(Y^0(\cdot), Z^0(\cdot))$ and $(\overline{Y}^0(\cdot), \overline{Z}^0(\cdot))$ are both in $\mathbb{M}^2(m;0,T) \times \mathbb{L}^2(m \times d;0,T)$. Denote $(Y^1(\cdot), Z^1(\cdot)) = \Lambda(Y^0(\cdot), Z^0(\cdot))$ and $(\overline{Y}^1(\cdot), \overline{Z}^1(\cdot)) = \Lambda(\overline{Y}^0(\cdot), \overline{Z}^0(\cdot))$. Applying Itô's rule to $e^{\beta t}(Y^1(t) - \overline{Y}^1(t))^2$, and integrating the derivative from t to T,

$$\begin{aligned} &e^{\beta t}(Y^1(t) - \overline{Y}^1(t))^2 + \beta \int_t^T e^{\beta s}(Y^1(s) - \overline{Y}^1(s))^2 ds + \int_t^T e^{\beta s}(Z^1(s) - \overline{Z}^1(s))^2 ds \\ =&2\int_t^T e^{\beta s}(Y^1(s) - \overline{Y}^1(s))(g(s, Y^0(s), Z^0(s)) - g(s, \overline{Y}^0(s), \overline{Z}^0(s)))ds \\ &+ 2\int_t^T e^{\beta s}(Y^1(s) - \overline{Y}^1(s))(Z^1(s) - \overline{Z}^1(s))ds + 2\int_t^T e^{\beta s}(Y^1(s) - \overline{Y}^1(s))(dK^1(s) - d\overline{K}^1(s)) \\ &+ 2\int_t^T e^{\beta s}(Y^1(s) - \overline{Y}^1(s))(Z^1(s) - \overline{Z}^1(s))dB_s. \end{aligned} \tag{3.2.10}$$

Because g is uniformly Lipschitz,

$$|g(s, Y^0(s), Z^0(s)) - g(s, \overline{Y}^0(s), \overline{Z}^0(s))| \le b|Y^0(s) - \overline{Y}^0(s)| + b|Z^0(s) - \overline{Z}^0(s)|. \tag{3.2.11}$$

For every constant $\alpha_1 > 0$,

$$\begin{aligned} &2e^{\beta s}(Y^1(s) - \overline{Y}^1(s))(g(s, Y^0(s), Z^0(s)) - g(s, \overline{Y}^0(s), \overline{Z}^0(s))) \\ \le&\alpha_1 e^{\beta s}(Y^1(s) - \overline{Y}^1(s))^2 + \frac{2b^2}{\alpha_1}e^{\beta s}(Y^0(s) - \overline{Y}^0(s))^2 + \frac{2b^2}{\alpha_1}e^{\beta s}(Z^0(s) - \overline{Z}^0(s))^2. \end{aligned} \tag{3.2.12}$$

For every constant $\alpha_2 > 0$,

$$2\int_t^T e^{\beta s}(Y^1(s)-\bar{Y}^1(s))(Z^1(s)-\bar{Z}^1(s))ds \\ \leq \alpha_2 e^{\beta s}(Y^1(s)-\bar{Y}^1(s))^2 + \frac{1}{\alpha_2}e^{\beta s}(Z^1(s)-\bar{Z}^1(s))^2. \tag{3.2.13}$$

Since, by definition of the mapping Λ, $Y^1(\cdot) \geq L(\cdot)$, and $\bar{Y}^1(\cdot) \geq L(\cdot)$, $(Y^1(\cdot)-L(\cdot))dK^1(\cdot) = (\bar{Y}^1(\cdot)-L(\cdot))d\bar{K}^1(\cdot) \equiv 0$, and $K^1(\cdot)$ and $\bar{K}^1(\cdot)$ are increasing, we have,

$$\begin{aligned} &\int_t^T e^{\beta s}(Y^1(s)-\bar{Y}^1(s))(dK^1(s)-d\bar{K}^1(s)) \\ =&\int_t^T e^{\beta s}((Y^1(s)-L(s))-(\bar{Y}^1(s)-L(s)))(dK^1(s)-d\bar{K}^1(s)) \\ \leq& -\int_t^T e^{\beta s}((Y^1(s)-L(s))d\bar{K}^1(s) - \int_t^T e^{\beta s}(\bar{Y}^1(s)-L(s))dK^1(s) \\ \leq& 0. \end{aligned} \tag{3.2.14}$$

Combining (3.2.10), (3.2.12), (3.2.13) and (3.2.14), letting $t = 0$, and taking expectation on both sides of the inequality,

$$(\beta-\alpha_1-\alpha_2)\mathbb{E}\left[\int_t^T e^{\beta s}(Y^1(s)-\bar{Y}^1(s))^2 ds\right] + \left(1-\frac{1}{\alpha_2}\right)\mathbb{E}\left[\int_t^T e^{\beta s}(Z^1(s)-\bar{Z}^1(s))^2 ds\right] \\ \leq \frac{2b^2}{\alpha_1}\mathbb{E}\left[\int_t^T e^{\beta s}(Y^0(s)-\bar{Y}^0(s))^2 ds\right] + \frac{2b^2}{\alpha_1}\mathbb{E}\left[\int_t^T e^{\beta s}(Z^0(s)-\bar{Z}^0(s))^2 ds\right]. \tag{3.2.15}$$

Because α_1, α_2 and β are arbitrary, we may let $\alpha_1 = 8b^2$, $\alpha_2 = 2$, and $\beta = \alpha_1 + \alpha_2 + \frac{1}{2} = 8b^2 + \frac{5}{2}$, then from (3.2.15),

$$||Y^1-\bar{Y}^1||^2_{2,\beta} + ||Z^1-\bar{Z}^1||^2_{2,\beta} \leq \frac{1}{2}||Y^0-\bar{Y}^0||^2_{2,\beta} + \frac{1}{2}||Z^0-\bar{Z}^0||^2_{2,\beta}. \tag{3.2.16}$$

The mapping Λ is indeed a contraction. □

Proposition 3.2.1 *The BSDE (3.2.4) with reflecting barrier has a unique solution in* $\mathbb{M}^2(m;0,T)\times\mathbb{L}^2(m\times d;0,T)$.

Proof. The solution is the unique fixed-point, say $(Y(\cdot),Z(\cdot))$, of the contraction Λ. Since $(Y(\cdot),Z(\cdot)) \in \mathbb{L}^2(m;0,T)\times\mathbb{L}^2(m\times d;0,T)$, $(Y(\cdot),Z(\cdot)) = \Lambda(Y(\cdot),Z(\cdot))$ is also in $\mathbb{M}^2(m;0,T)\times\mathbb{L}^2(m\times d;0,T)$ by Lemma 3.2.1. □

Theorem 3.2.2 *(Comparison Theorem, El Karoui, Kapoudjian, Pardoux, Peng and Quenez (1997) [19])*
Suppose (Y,Z,K) solves (3.2.4) with parameter set (ξ,g,L), and $(\bar{Y},\bar{Z},\bar{K})$ solves (3.2.4) with parameter set $(\bar{\xi},\bar{g},\bar{L})$. Let dimension of the equations be $m = 1$. Under Assumption 3.2.1, except that the uniform Lipschitz condition only needed for either g or $\bar{g}$, if

(1) $\xi \le \bar{\xi}$, *a.e.;*
(2) $\underline{g}(t,y,z) \le \bar{g}(t,y,z)$, *a.e.* $(t,\omega) \in [0,T]\times\Omega$, $\forall (y,z) \in \mathbb{R}\times\mathbb{R}^d$*; and*
(3) $L(t) \le \bar{L}(t)$, *a.e.* $(t,\omega) \in [0,T]\times\Omega$,
then

$$Y(t) \le \bar{Y}(t),\ a.e.\ (t,\omega) \in [0,T]\times\Omega. \tag{3.2.17}$$

Theorem 3.2.3 *(Continuous Dependence Property)*
Under Assumption 3.2.1, suppose that (Y,Z,K) *solves RBSDE (3.2.4), and that* $(\bar{Y},\bar{Z},\bar{K})$ *solves*

$$\begin{cases} \bar{Y}(t) = \bar{\xi} + \int_t^T g(s,\bar{Y}(s),\bar{Z}(s))ds - \int_t^T \bar{Z}(s)dB_s + \bar{K}(T) - \bar{K}(t); \\ \bar{Y}(t) \ge L(t),\ 0 \le t \le T,\ \int_0^T (\bar{Y}(t) - L(t))d\bar{K}(t) = 0, \end{cases} \tag{3.2.18}$$

then there exists a constant number C*, such that for all* $0 \le t \le T$*,*

$$\begin{aligned} &\mathbb{E}[(Y(t)-\bar{Y}(t))^2] + \mathbb{E}\left[\int_0^T (Y(s)-\bar{Y}(s))^2 ds\right] \\ &+ \mathbb{E}\left[\int_0^T (Z(s)-\bar{Z}(s))^2 ds\right] + \mathbb{E}[(K(t)-\bar{K}(t))^2] \\ \le& C\mathbb{E}[(\xi-\bar{\xi})^2]. \end{aligned} \tag{3.2.19}$$

Proof. Applying Itô's rule to $e^{\beta t}(Y(t)-\bar{Y}(t))^2$, integrating from t to T, and then repeating the methods in proof of Theorem 3.2.1,

$$\begin{aligned} &e^{\beta t}(Y(t)-\bar{Y}(t))^2 + \beta\int_t^T e^{\beta s}(Y(s)-\bar{Y}(s))^2 ds + \int_t^T e^{\beta s}(Z(s)-\bar{Z}(s))^2 ds \\ =& e^{\beta T}(\xi-\bar{\xi})^2 + 2\int_t^T e^{\beta s}(Y(s)-\bar{Y}(s))(g(s,Y(s),Z(s)) - g(s,\bar{Y}(s),\bar{Z}(s)))ds \\ &+ 2\int_t^T e^{\beta s}(Y(s)-\bar{Y}(s))(Z(s)-\bar{Z}(s))ds + 2\int_t^T e^{\beta s}(Y(s)-\bar{Y}(s))(dK(s)-d\bar{K}(s)) \\ &+ 2\int_t^T e^{\beta s}(Y(s)-\bar{Y}(s))(Z(s)-\bar{Z}(s))dB_s \\ \le& e^{\beta T}(\xi-\bar{\xi})^2 + (\alpha_1+\alpha_2+b)\int_t^T e^{\beta s}(Y(s)-\bar{Y}(s))^2 ds \\ &+ \left(\frac{b^2}{\alpha_1}+\frac{1}{\alpha_2}\right)\int_t^T e^{\beta s}(Z(s)-\bar{Z}(s))^2 ds + 2\int_t^T e^{\beta s}(Y(s)-\bar{Y}(s))(Z(s)-\bar{Z}(s))dB_s. \end{aligned} \tag{3.2.20}$$

Rearranging the terms in (3.2.20), and taking expectations,

$$\begin{aligned}&e^{\beta t}\mathbb{E}[(Y(t)-\overline{Y}(t))^2]+(\beta-b-\alpha_1-\alpha_2)\mathbb{E}\Big[\int_t^T e^{\beta s}(Y(s)-\overline{Y}(s))^2ds\Big]\\&+\Big(1-\frac{b^2}{\alpha_1}-\frac{1}{\alpha_2}\Big)\mathbb{E}\Big[\int_t^T e^{\beta s}(Z(s)-\overline{Z}(s))^2ds\Big]\\ \le& e^{\beta T}\mathbb{E}[(\xi-\overline{\xi})^2].\end{aligned} \tag{3.2.21}$$

In (3.2.21), letting $\alpha_1 = 4b^2$, $\alpha_2 = 4$, and $\beta = b + \alpha_1 + \alpha_2 + \frac{1}{2} = 4b^2 + b + \frac{9}{2}$ gives

$$\begin{aligned}&e^{\beta t}\mathbb{E}[(Y(t)-\overline{Y}(t))^2]+\frac{1}{2}\mathbb{E}\Big[\int_t^T e^{\beta s}(Y(s)-\overline{Y}(s))^2ds\Big]+\frac{1}{2}\mathbb{E}\Big[\int_t^T e^{\beta s}(Z(s)-\overline{Z}(s))^2ds\Big]\\ \le& e^{\beta T}\mathbb{E}[(\xi-\overline{\xi})^2], \text{ for all } 0\le t\le T.\end{aligned} \tag{3.2.22}$$

Hence both

$$\mathbb{E}[(Y(t)-\overline{Y}(t))^2] \le e^{\beta T}\mathbb{E}[(\xi-\overline{\xi})^2], \text{ for all } 0\le t\le T, \tag{3.2.23}$$

and

$$\mathbb{E}\Big[\int_0^T (Y(s)-\overline{Y}(s))^2ds\Big]+\mathbb{E}\Big[\int_0^T (Z(s)-\overline{Z}(s))^2ds\Big] \le 2e^{\beta T}\mathbb{E}[(\xi-\overline{\xi})^2] \tag{3.2.24}$$

hold true.
It remains to estimate the $\mathbb{L}^2$-norm of $(K(t)-\overline{K}(t))$. Integrating dY and $d\overline{Y}$ from 0 to t gives

$$K(t) = Y(0) - Y(t) - \int_0^t g(s, Y(s), Z(s))ds + \int_0^t Z(s)dB_s, \tag{3.2.25}$$

and

$$\overline{K}(t) = \overline{Y}(0) - \overline{Y}(t) - \int_0^t g(s, \overline{Y}(s), \overline{Z}(s))ds + \int_0^t \overline{Z}(s)dB_s. \tag{3.2.26}$$

Then, there exists a constant number C_1, such that

$$\begin{aligned}&(K(t)-\overline{K}(t))^2\\ \le& C_1\Big((Y(0)-\overline{Y}(0))^2+(Y(t)-\overline{Y}(t))^2+t\int_0^t (g(s,Y(s),Z(s))-g(s,\overline{Y}(s),\overline{Z}(s)))^2ds\\&+\int_0^t (Z(s)-\overline{Z}(s))^2dB_s\Big).\end{aligned} \tag{3.2.27}$$

Taking expectation on both sides of (3.2.27), by Lipschitz condition Assumption 3.2.1 and Itô's isometry, for all $0 \leq t \leq T$,

$$\begin{aligned}
&\mathbb{E}[(K(t) - \overline{K}(t))^2] \\
\leq & C_1\Big(\mathbb{E}[(Y(0) - \overline{Y}(0))^2] + \mathbb{E}[(Y(t) - \overline{Y}(t))^2] + 2Tb^2\mathbb{E}\Big[\int_0^T (Y(t) - \overline{Y}(t))^2 dt\Big] \\
&+ (2Tb^2 + 1)\mathbb{E}\Big[\int_0^T (Z(t) - \overline{Z}(t))^2 dt\Big]\Big) \\
\leq & 4C_1(Tb^2 + 1)e^{\beta T}\mathbb{E}[(\xi - \overline{\xi})^2],
\end{aligned} \tag{3.2.28}$$

last inequality from (3.2.23) and (3.2.24). □

3.3 Markovian system with linear growth rate

This section shows existence of the solution to the multidimensional BSDE with reflecting barrier within a Markovian framework. The growth rate of the forward equation is assumed polynomial in the state process X, and linear in both the value process Y and the volatility process Z. The comparison theorem in dimension one and continuous dependence property of the value process and the volatility process on the terminal condition is also provided.

The Markovian system of forward-backward SDE's in question is the following pair of equations.

$$\begin{cases} X^{t,x}(s) = x, \ 0 \leq s \leq t; \\ dX^{t,x}(s) = f(s, X^{t,x}(s))ds + \sigma(s, X^{t,x}(s))dB_s, \ t < s \leq T. \end{cases} \tag{3.3.1}$$

$$\begin{cases} Y^{t,x}(s) = \xi(X^{t,x}(T)) + \int_s^T g(r, X^{t,x}(r), Y^{t,x}(r), Z^{t,x}(r))dr - \int_s^T Z^{t,x}(r)dB_r \\ \qquad + K^{t,x}(T) - K^{t,x}(s); \\ Y^{t,x}(s) \geq L(s, X^{t,x}(s)), \ t \leq s \leq T, \ \int_t^T (Y^{t,x}(s) - L(s, X^{t,x}(s)))dK^{t,x}(s) = 0. \end{cases} \tag{3.3.2}$$

For any $x \in \mathbb{R}^l$, the SDE (3.3.1) has a unique strong solution, under Assumption 3.3.1 (1) below (cf. page 287, Karatzas and Shreve (1988) [33]). A solution to the forward-backward system (3.3.1) and (3.3.2) is a triplet of processes $(Y^{t,x}, Z^{t,x}, K^{t,x})$ satisfying (3.3.2), where $Y^{t,x} \in \mathbb{M}^2(m; 0, T)$, $Z^{t,x} \in \mathbb{L}^2(m \times d; 0, T)$, and $K^{t,x}$ is a continuous, increasing process in $\mathbb{M}^2(m; 0, T)$. The superscript (t, x) on X, Y, Z, and K indicates the state x of the underlying process X at time t. It will be omitted for notational simplicity.

Assumption 3.3.1 *(1) In (3.3.1), the drift $f : [0, T] \times \mathbb{R}^l \to \mathbb{R}^l$, and volatility $\sigma : [0, T] \times \mathbb{R}^l \to \mathbb{R}^{l \times d}$, are deterministic, measurable mappings, locally Lipschitz in x uniformly over all $t \in [0, T]$. And for all $(t, x) \in [0, T] \times \mathbb{R}^l$, $|f(t, x)|^2 + |\sigma(t, x)|^2 \leq C(1 + |x|^2)$, for some constant C.*
(2) In (3.3.2), the driver g is a deterministic measurable mapping $g : [0, T] \times \mathbb{R}^l \times$

$\mathbb{R}^{m\times d} \to \mathbb{R}^m$, $(t,x,y,z) \mapsto g(t,x,y,z)$. And for all $(t,x,y,z) \in [0,T]\times\mathbb{R}^l\times\mathbb{R}^m\times\mathbb{R}^{m\times d}$, $|g(t,x,y,z)| \le b(1+|x|^p+|y|+|z|)$, for some positive constant b.
(3) For every fixed $(t,x) \in [0,T]\times\mathbb{R}^l$, the mapping $g(t,x,\cdot,\cdot)$ is continuous.
(4) The terminal value $\xi : \mathbb{R}^l \to \mathbb{R}^m$, $x \mapsto \xi(x)$, is a deterministic measurable mapping. The lower reflecting boundary $L : [0,T]\times\mathbb{R}^l \to \mathbb{R}^m$, $(s,x) \mapsto L(s,x)$ is deterministic measurable mapping continuous in (s,x). They satisfy $\mathbb{E}[\xi(X(T))^2] < \infty$, $\mathbb{E}\left[\sup_{[0,T]} L^+(s,X(s))^2\right] < \infty$, and $L(T,X(T)) \le \xi(X(T))$, a.e. on Ω.

Theorem 3.3.1 *Suppose that Assumption 3.3.1 holds, except the growth rate condition on g. If the driver $g(s,x,y,z)$ in the reflected BSDE (3.3.2) is Lipschitz in y and z, uniformly over all $s \in [0,T]$ and all $x \in \mathbb{R}^l$, then there exist measurable deterministic functions $\alpha : [0,T]\times\mathbb{R}^l \to \mathbb{R}^m$, and $\beta : [0,T]\times\mathbb{R}^l \to \mathbb{R}^{m\times d}$, such that for any $0 \le t \le s \le T$, $Y^{t,x}(s) = \alpha(s,X^{t,x}(s))$, and $Z^{t,x}(s) = \beta(s,X^{t,x}(s))$. The solutions to the BSDE are functions of the state process X.*

Proof. First, the one-dimensional case $m = 1$. There exist measurable, deterministic functions $a^n : [0,T]\times\mathbb{R}^l \to \mathbb{R}$, $b^n : [0,T]\times\mathbb{R}^l \to \mathbb{R}^d$, such that for any $0 \le t \le s \le T$, the solution $(Y^{(t,x),n}, Z^{(t,x),n})$ to the penalized equation

$$\begin{aligned} Y^{(t,x),n}(s) =& \xi(X^{t,x}(T)) + \int_s^T g(r,X^{t,x}(r),Y^{(t,x),n}(r),Z^{(t,x),n}(r))dr - \int_s^T Z^{(t,x),n}(r)dB_r \\ &+ n\int_s^T (Y^{(t,x),n}(r) - L(r,X^{t,x}(r)))^- dr \end{aligned} \tag{3.3.3}$$

can be expressed as $Y^{(t,x),n}(s) = a^n(s,X^{t,x}(s))$, and $Z^{(t,x),n}(s) = b^n(s,X^{t,x}(s))$; in particular, $Y^{(t,x),n}(t) = a^n(t,x)$. This is the Markovian property of solutions to one one-dimensional forward-backward SDE's with Lipschitz driver, stated as Theorem 4.1 in El Karoui, Peng and Quenez (1997) [20]. Their proof uses the Picard iteration and the Markov property of the iterated sequence of solutions, the latter being an interpretation of Theorem 6.27 on page 206 of Çinlar, Jacod, Protter and Sharpe (1980) [9]. Analyzed in section 6, El Karoui, Kapoudjian, Pardoux, Peng and Quenez (1997) [19], its solution $(Y^{(t,x),n}, Z^{(t,x),n})$ converges to some limit $(Y^{t,x}, Z^{t,x})$ in $\mathbb{M}^2(m;t,T)\times\mathbb{L}^2(m\times d;t,T)$. The penalization term $n\int_0^s (Y^{(t,x),n}(r) - L(r,X^{t,x}(r)))^- dr$ also has an $\mathbb{M}^2(m;0,T)$-limit $K^{t,x}(s)$. The triplet $(Y^{t,x}, Z^{t,x}, K^{t,x})$ solves the system (3.3.1) and (3.3.2). But the convergences are also almost everywhere on $\Omega\times[t,T]$, so

$$Y^{t,x}(s) = \lim_{n\to\infty} Y^{(t,x),n}(s) = \limsup_{n\to\infty}(a^n(s,X^{t,x}(s))) = \limsup_{n\to\infty}(a^n)(s,X^{t,x}(s)) =: a(s,X^{t,x}(s)), \tag{3.3.4}$$

and

$$Z^{t,x}(s) = \lim_{n\to\infty} Z^{(t,x),n}(s) = \limsup_{n\to\infty}(b^n(s,X^{t,x}(s))) = \limsup_{n\to\infty}(b^n)(s,X^{t,x}(s)) =: b(s,X^{t,x}(s)). \tag{3.3.5}$$

Back to a general dimension m. By Theorem 3.2.1 and Proposition 3.2.1, the sequence $(Y^{n+1}, Z^{n+1}) = \Lambda(Y^n, Z^n)$, $n = 0,1,2,\cdots$, iterated via the mapping Λ as in (3.2.1), converges to (Y,Z) a.e. on $\Omega\times[t,T]$ and in $\mathbb{M}^2(m;0,T)\times\mathbb{L}^2(m\times d;0,T)$. If one can prove $Y^1(s)$ and $Z^1(s)$ are functions of $(s,X(s))$, so is every $(Y^n(s),Z^n(s))$ by induction. Then the theorem holds, because (Y,Z) is the pointwise limit of $\{(Y^n(s),Z^n(s))\}_n$. The claim is indeed true. Starting with $Y^{(t,x),0}(s)=\alpha^0(s,X(s))$, and $Z^{(t,x),0}(s)=\beta^0(s,X(s))$, for any measurable, deterministic functions $\alpha^0:[0,T]\times\mathbb{R}^l\to\mathbb{R}^m$, and $\beta^0:[0,T]\times\mathbb{R}^l\to\mathbb{R}^{m\times d}$ satisfying $\alpha^0(\cdot,X^{t,x}(\cdot))\in\mathbb{M}^2(m;0,T)$, and $\beta^0(\cdot,X^{t,x}(\cdot))\in\mathbb{L}^2(m\times d;0,T)$. In an arbitrary ith dimension, $1\le i\le m$,

$$\begin{cases} Y_i^1(s) = \xi_i(X^{t,x}(T)) + \int_s^T g_i(r,X^{t,x}(r),\alpha^0(r,X(r)),\beta^0(r,X(r)))dr \\ \qquad - \int_s^T Z_i^1(r)dB_r + K_i^1(T) - K_i^1(s); \\ Y_i^1(s) \ge L_i(s,X^{t,x}(s)),\ t\le s\le T,\ \int_t^T (Y_i^1(s)-L_i(s,X^{t,x}(s)))dK_i^1(s)=0. \end{cases} \tag{3.3.6}$$

From the one-dimensional result, there exist measurable, deterministic functions α_i^1 : $[0,T]\times\mathbb{R}^l\to\mathbb{R}$, and $\beta_i^1:[0,T]\times\mathbb{R}^l\to\mathbb{R}^d$, such that $Y_i^{(t,x),1}(s)=\alpha_i^1(s,X^{t,x}(s))$, and $Z_i^{(t,x),1}(s)=\beta_i^1(s,X^{t,x}(s))$, for all $0\le t\le s\le T$. Let $\alpha^1=(\alpha_1^1,\cdots,\alpha_m^1)'$, and $\beta^1=(\beta_1^1,\cdots,\beta_m^1)'$, then $Y^{(t,x),1}(s)=\alpha^1(s,X^{t,x}(s))$, and $Z^{(t,x),1}(s)=\beta^1(s,X^{t,x}(s))$, for all $0\le t\le s\le T$. □

Remark 3.3.1 *To prove the above theorem, besides using the notion of "additive martingales" as in Çinlar et al (1980) [9], the two deterministic functions can also be obtained by solving a multi-dimensional variational inequality following the four-step-scheme proposed by Ma, Protter and Yong (1994) [*__?__*].*

The rest of this section will be devoted to proving existence of solutions to the reflected forward-backward system (3.3.1) and (3.3.2) under the Assumption 3.3.1. We shall construct a specific sequence of Lipschitz drivers g^n to approximate the linear-growth driver g. The corresponding sequence of solutions will turn out to converge to the system (3.3.1) and (3.3.2). We then approximate the continuous linear growth driver g by a sequence of Lipschitz functions g^n.

Let $\bar{\psi}$ be an infinitely differentiable mapping from $\mathbb{R}^m\times\mathbb{R}^{m\times d}$ to $\mathbb{R}$, such that

$$\bar{\psi}(y,z)=\begin{cases}1, & |y|^2+|z|^2\le 1;\\ 0, & |y|^2+|z|^2\ge 4,\end{cases} \tag{3.3.7}$$

and ψ a rescaling of $\bar{\psi}$ by a multiplicative constant such that

$$\int_{\mathbb{R}^m\times\mathbb{R}^{m\times d}}\psi(y,z)dydz=1. \tag{3.3.8}$$

The function ψ is a kernel conventionally used to smooth out non-differentiability, for example, by Karatzas and Ocone (1992) [**?**], or to approximate functions of higher growth rate, for example, by Hamadène, Lepeltier and Peng (1997) [26].

The approximating sequence g^n is defined as

$$g^n(t,x,y,z) = n^2\psi\left(\frac{y}{n},\frac{z}{n}\right)\int_{\mathbb{R}^m\times\mathbb{R}^{m\times d}} g(t,x,y_1,z_1)\overline{\psi}(n(y-y_1),n(z-z_1))dy_1dz_1. \quad (3.3.9)$$

According to Hamadène, Lepeltier and Peng (1997) [26], the sequence of functions g^n has the properties:
(a) g^n is Lipschitz with respect to (y,z), uniformly over all $(t,x)\in[0,T]\times\mathbb{R}^l$;
(b) $|g^n(t,x,y,z)| \le b(1+|x|^p+|y|+|z|)$, for all $(t,x,y,z)\in[0,T]\times\mathbb{R}^l\times\mathbb{R}^m\times\mathbb{R}^{m\times d}$, for some positive constant b;
(c) $|g^n(t,x,y,z)| \le b_n(1+|x|^p)$, for all $(t,x,y,z)\in[0,T]\times\mathbb{R}^l\times\mathbb{R}^m\times\mathbb{R}^{m\times d}$, for some positive constant b_n;
(d) for any $(t,x)\in[0,T]\times\mathbb{R}$, and for any compact set $S\subset\mathbb{R}^m\times\mathbb{R}^{m\times d}$,

$$\sup_{(y,z)\in S}|g^n(t,x,y,z)-g(t,x,y,z)| \to 0, \text{ as } n\to 0. \quad (3.3.10)$$

Proposition 3.3.1 *The BSDE with reflecting barrier*

$$\begin{cases} Y^n(s) = \xi(X(T)) + \int_s^T g^n(r,X(r),Y^n(r),Z^n(r))dr - \int_s^T Z^n(r)dB_r + K^n(T) - K^n(s); \\ Y^n(s) \ge L(s,X(s)),\ t\le s\le T,\ \int_t^T (Y^n(s)-L(s,X(s)))dK^n(s) = 0 \end{cases} \quad (3.3.11)$$

has a unique solution (Y^n,Z^n,K^n). Furthermore, there exist measurable, deterministic functions α^n and β^n, such that $Y^n(s)=\alpha^n(s,X(s))$, and $Z^n(s)=\beta^n(s,X(s))$, for all $0\le s\le T$.

Proof. This is a direct consequence of the uniform Lipschitz property of g^n, Proposition 3.2.1 and Theorem 3.3.1. □

Lemma 3.3.1 *Suppose (Y,Z,K) solves the BSDE (3.3.2) with reflecting barrier. Assume (2) and (4) of Assumption 3.3.1. Then there exists a positive constant C, such that*

$$\mathbb{E}\left[\sup_{0\le s\le T} Y(s)^2 + \int_t^T Z(r)^2ds + K(T)^2\right] \le C(1+|x|^{2(p\vee 1)}). \quad (3.3.12)$$

The constant C does not depend on t, but depends on m, b, T, $\mathbb{E}[\xi(X(T))^2]$ and $\mathbb{E}\left[\sup_{[0,T]} L^+(t,X(t))^2\right]$.

Proof. First prove that, for some constant C', we have

$$\mathbb{E}\left[Y(s)^2 + \int_t^T Z(r)^2ds + K(T)^2\right] \le C'(1+|x|^{2(p\vee 1)}), \text{ for all } 0\le s\le T. \quad (3.3.13)$$

Applying Itô's rule to $Y(\cdot)^2$, and integrating from s to T, we get

$$\begin{aligned}
&Y(s)^2 + \int_s^T Z(r)^2 dr \\
=&\xi(X(T))^2 + 2\int_s^T Y(r)g(r, X(r), Y(r), Z(r))dr - 2\int_s^T Y(r)Z(r)dB(r) + 2\int_s^T L(r, X(r))dK(r).
\end{aligned} \tag{3.3.14}$$

Taking expectations of (3.3.14), and using Assumption 3.3.1 (2), we obtain

$$\begin{aligned}
&\mathbb{E}\left[Y(s)^2 + \int_s^T Z(r)^2 dr\right] \\
\leq&\mathbb{E}[\xi(X(T))^2] + 2b\mathbb{E}\left[\int_s^T |Y(r)|(1 + |X(r)|^p + |Y(r)| + |Z(r)|)dr\right] \\
&+ 2\mathbb{E}\left[\int_s^T L(r, X(r))dK(r)\right] \\
\leq&\mathbb{E}[\xi(X(T))^2] + 2\mathbb{E}\left[\int_s^T (1 + |X(r)|^{2p})dr\right] + C_1(b)\mathbb{E}\left[\int_s^T |Y(r)|^2 dr\right] \\
&+ \frac{1}{4}\mathbb{E}\left[\int_s^T |Z(r)|^2 dr\right] + 2\mathbb{E}\left[\int_s^T L(r, X(r))dK(r)\right].
\end{aligned} \tag{3.3.15}$$

For any $\alpha > 0$,

$$2\int_t^T L(s, X(s))dK(s) \leq 2\left(\sup_{[0,T]} L(s, X(s))\right)K(T) \leq \frac{1}{\alpha}K(T)^2 + \alpha \sup_{[0,T]} L^+(s, X(s))^2. \tag{3.3.16}$$

Combine (3.3.15) and (3.3.16), and apply Gronwall's Lemma to $Y(\cdot)$,

$$\begin{aligned}
&\mathbb{E}\left[Y(s)^2 + \frac{3}{4}\int_s^T Z(r)^2 dr\right] \\
\leq&C_2(b, T)\left(1 + \mathbb{E}[\xi(X(T))^2] + \mathbb{E}\left[\int_s^T |X(r)|^{2p} dr\right] + \frac{1}{\alpha}K(T)^2 + \alpha \sup_{[0,T]} L^+(s, X(s))^2\right).
\end{aligned} \tag{3.3.17}$$

If rewriting (3.3.2) from t to T, $K(\cdot)$ can be expressed in terms of $Y(\cdot)$ and $Z(\cdot)$ by

$$K(T) = Y(t) - \xi(X(T)) - \int_t^T g(s, X(s), Y(s), Z(s))ds + \int_t^T Z(s)dB_s, \tag{3.3.18}$$

and hence because of the linear growth Assumption 3.3.1 (2), we have

$$\begin{aligned}\mathbb{E}[K(T)^2] = & C_3\mathbb{E}\left[Y(t)^2 + \xi(X(T))^2 + \int_t^T g(s,X(s),Y(s),Z(s))^2 ds + \int_t^T Z(s)^2 ds\right] \\ \leq & C_4(b)\left(\mathbb{E}\left[Y(t)^2 + \xi(X(T))^2 + 1 + \int_t^T |X(s)|^{2p} ds\right] + \mathbb{E}\left[\int_t^T |Y(s)|^2 ds\right]\right. \\ & \left. + \mathbb{E}\left[\int_t^T |Z(s)|^2 ds\right]\right). \end{aligned} \tag{3.3.19}$$

Bound $\mathbb{E}[|Y(s)|^2]$ and $\mathbb{E}\left[\int_t^T |Z(s)|^2 ds\right]$ in (3.3.19) by (3.3.17),

$$\begin{aligned}\mathbb{E}[K(T)^2] \leq C_5(b,t,T)\Bigg(&\mathbb{E}\left[\xi(X(T))^2 + 1 + \int_t^T |X(s)|^{2p} ds\right] \\ &+ \frac{1}{\alpha}\mathbb{E}[K(T)^2] + \alpha\mathbb{E}\left[\sup_{[0,T]} L^+(s,X(s))^2\right]\Bigg). \end{aligned} \tag{3.3.20}$$

Let $\alpha = 4C_5(b,t,T)$, and collect $\mathbb{E}[K(T)^2]$ terms on both sides of (3.3.20),

$$\mathbb{E}[K(T)^2] \leq C_6(b,t,T)\mathbb{E}\left[\xi(X(T))^2 + 1 + \int_t^T |X(s)|^{2p} ds + \sup_{[0,T]} L^+(s,X(s))^2\right]. \tag{3.3.21}$$

Finally, (3.3.17) and (3.3.21) altogether gives

$$\begin{aligned}&\mathbb{E}\left[Y(s)^2 + \int_s^T Z(r)^2 ds + K(T)^2\right] \\ \leq & C_7(b,t,T)\left(1 + \mathbb{E}[\xi(X(T))^2] + \mathbb{E}\left[\int_t^T |X(r)|^{2p} dr\right] + \mathbb{E}\left[\sup_{[0,T]} L^+(s,X(s))^2\right]\right). \end{aligned} \tag{3.3.22}$$

From page 306 of Karatzas and Shreve (1988) [33], for $p \geq 1$,

$$\mathbb{E}\left[\sup_{[0,T]} |X^{t,x}(s)|^{2p}\right] \leq C_8(1 + |x|^{2p}). \tag{3.3.23}$$

Then the constant C' in (3.3.13) can be chosen as

$$C' = \left(\sup_{0\leq t\leq T} C_7(b,t,T)\right)\max\left\{1 + \mathbb{E}[\xi(X(T))^2] + \mathbb{E}\left[\sup_{[0,T]} L^+(s,X(s))^2\right], C_8 T\right\} < \infty. \tag{3.3.24}$$

To bound the $\mathbb{L}^2$ supremum norm of $Y(\cdot)$, taking first supremum over $s \in [0,T]$ then expectation, on both sides of (3.3.14), using Burkholder-Davis-Gundy inequality, and combining with (3.3.16),

$$\begin{aligned}
&\mathbb{E}\bigg[\sup_{[0,T]} Y(s)^2 + \int_t^T Z(r)^2 dr\bigg]\\
\leq&\mathbb{E}[\xi(X(T))^2] + 2b\mathbb{E}\bigg[\sup_{[0,T]}\int_s^T |Y(r)|(1+|X(r)|^p+|Y(r)|+|Z(r)|)dr\bigg]\\
&+C_9(m)\mathbb{E}\bigg[\sqrt{\int_t^T |Y(r)|^2\cdot|Z(r)|^2dr}\bigg] + 2\mathbb{E}\bigg[\int_s^T L(r,X(r))dK(r)\bigg]\\
\leq&\mathbb{E}[\xi(X(T))^2] + b\mathbb{E}\bigg[\int_t^T |Y(r)|^2dr\bigg] + b\mathbb{E}\bigg[\int_t^T (1+|X(r)|^p+|Y(r)|+|Z(r)|)^2dr\bigg]\\
&+C_9(m)\mathbb{E}\bigg[\sup_{[0,T]}|Y(s)|\sqrt{\int_t^T |Z(r)|^2dr}\bigg] + \mathbb{E}[K(T)^2] + \mathbb{E}\bigg[\sup_{[0,T]} L^+(s,X(s))^2\bigg]\\
\leq&\mathbb{E}[\xi(X(T))^2] + C_{10}(b)\mathbb{E}\bigg[\int_t^T (1+|X(r)|^{2p}+|Y(r)|^2+|Z(r)|^2)dr\bigg]\\
&+\frac{1}{2}\mathbb{E}\bigg[\sup_{[0,T]}|Y(s)|^2\bigg] + 2C_9(m)^2\mathbb{E}\bigg[\int_t^T |Z(r)|^2dr\bigg] + \mathbb{E}[K(T)^2] + \mathbb{E}\bigg[\sup_{[0,T]} L^+(s,X(s))^2\bigg].
\end{aligned} \tag{3.3.25}$$

Equation (3.3.25) implies that

$$\begin{aligned}
&\frac{1}{2}\mathbb{E}\bigg[\sup_{[0,T]} Y(s)^2\bigg]\\
\leq&\mathbb{E}[\xi(X(T))^2] + C_{10}(b)\mathbb{E}\bigg[\int_t^T (1+|X(r)|^{2p}+|Y(r)|^2+|Z(r)|^2)dr\bigg]\\
&+2C_9(m)^2\mathbb{E}\bigg[\int_t^T |Z(r)|^2dr\bigg] + \mathbb{E}[K(T)^2] + \mathbb{E}\bigg[\sup_{[0,T]} L^+(s,X(s))^2\bigg].
\end{aligned} \tag{3.3.26}$$

Inequalities (3.3.13), (3.3.23) and (3.3.26) conclude the lemma. □

Proposition 3.3.2 *There exists a positive constant C, such that for $0 \leq t \leq T$, $n = 1, 2, \cdots$,*

$$\alpha^n(t,x) = Y^{(t,x),n}(t) = \mathbb{E}[Y^{(t,x),n}(t)|\mathscr{F}_t] \leq C(1+|x|^{p\vee 1}). \tag{3.3.27}$$

Proposition 3.3.3 *The sequence $\{g^n(\cdot, X(\cdot), Y^n(\cdot), Z^n(\cdot))\}_n$ is uniformly bounded in the $\mathbb{L}^2(m; t, T)$-norm, and the sequence $\{K^n(\cdot)\}_n$ is uniformly bounded in the $\mathbb{M}^2(m; t, T)$-norm, both uniformly over all n. As $n \to \infty$, $g^n(\cdot, X(\cdot), Y^n(\cdot), Z^n(\cdot))$ weakly converges to some limit $G(\cdot)$ in $\mathbb{L}^2(m; t, T)$ along a subsequence, and $K^n(\cdot)$ weakly converges to some limit $K(\cdot)$ in $\mathbb{M}^2(m; t, T)$ along a subsequence, for every $s \in [t, T]$.*

Proof. It suffices to show the uniform boundedness of $\{g^n(\cdot, X(\cdot), Y^n(\cdot), Z^n(\cdot))\}_n$ in $\mathbb{L}^2(m; t, T)$ and of $\{K^n(T)\}_n$ in $\mathbb{L}^2(m)$, which is a result of the linear growth property (b) and Lemma 3.3.1. The $\mathbb{L}^2(m)$ uniform boundedness of $\{K^n(T)\}_n$ means that there exists

a constant $C < \infty$, such that $\mathbb{E}[|K^n(T)|^2] < C$. Since $K^n(\cdot)$ is required to be an increasing process starting from $K^n(t) = 0$, then for all $t \le s \le T$, $\mathbb{E}[|K^n(s)|^2] \le \mathbb{E}[|K^n(T)|^2] < C$.
□

With the help of weak convergence along a subsequence, we proceed to argue that the weak limits are also strong, thus deriving a solution to BSDE (3.3.2). For notational simplicity, the weakly convergent subsequences are still indexed by n. The passing from weak to strong convergence makes use of the Markovian structure of the system described by Theorem 3.3.1, which states that the valued process $Y^n(s)$ is a deterministic function of time s and state process $X(s)$ only.

Lemma 3.3.2 *The approximating sequence of solutions* $\{(Y^{(t,x),n}, Z^{(t,x),n})\}_n$ *is Cauchy in* $\mathbb{L}^2(m; t, T) \times \mathbb{L}^2(m \times d; t, T)$*, thus having a limit* $(Y^{t,x}, Z^{t,x})$ *in* $\mathbb{L}^2(m; t, T) \times \mathbb{L}^2(m \times d; t, T)$ *and a.e. on* $[t, T] \times \Omega$*.*

Proof. For any $t \in [0, T]$, any $x \in \mathbb{R}^l$, and any $n = 1, 2, \cdots$, $Y^{(t,x),n}(t) = \alpha^n(t, x)$ is deterministic. First prove the convergence of $\{\alpha^n(t, x)\}_n$ by showing it is Cauchy. From equation (3.3.11) comes the following inequality,

$$\begin{aligned}
&|\alpha^n(t,x) - \alpha^k(t,x)| = |Y^n(t) - Y^k(t)| \\
\le& \left|\mathbb{E}\left[\int_t^T (g^n(s, X(s), Y^n(s), Z^n(s)) - g^k(s, X(s), Y^k(s), Z^k(s)))ds\right]\right| \qquad (3.3.28)\\
&+ |\mathbb{E}[K^n(T) - K^k(T)]| + |\mathbb{E}[K^n(t) - K^k(t)]|.
\end{aligned}$$

By the weak convergence from Proposition 3.3.3, all the three summands on the right hand side of the above inequality converge to zero, as n and k both go to infinity. Denote the limit of $\alpha^n(t, x)$ as $\alpha(t, x)$, which is consequently deterministic and measurable, because $\alpha^n(\cdot, \cdot)$ is measurable. Theorem 3.3.1 states that for any $t \le s \le T$, $Y^{(t,x),n}(s) = \alpha^n(s, X^{t,x}(s))$. Because of the pointwise convergence of $\alpha^n(\cdot, \cdot)$, $Y^{(t,x),n}(s)$ converges to some $Y^{(t,x)}(s)$, a.e. $(s, \omega) \in [t, T] \times \Omega$, as $n \to \infty$. Proposition 3.3.2 states that there exists a positive constant C, such that for $0 \le t \le T$, $n = 1, 2, \cdots$,

$$|Y^{(t,x),n}(s)| = |\alpha^n(s, X_s^{t,x})| \le C(1 + |X_s^{t,x}|^{p\vee 1}), \qquad (3.3.29)$$

the last term of which is square-integrable by (3.3.23). Then it follows from the dominated convergence theorem that the convergence of $Y^{(t,x),n}(s)$ is also in $\mathbb{L}^2(m; t, T)$.

Apply Itô's rule to $(Y^{(t,x),n}(s) - Y^{(t,x),k}(s))^2$, and integrate from s to T. The reflecting conditions that leads to the inequality (3.2.14) gives

$$\begin{aligned}
&(Y^n(s) - Y^k(s))^2 + \int_s^T (Z^n(r) - Z^k(r))^2 dr \\
\le& \int_s^T (Y^n(r) - Y^k(r))(g^n(r, X(r), Y^n(r), Z^n(r)) - g^k(r, X(r), Y^k(r), Z^k(r)))dr \qquad (3.3.30)\\
&+ \int_s^T (Y^n(r) - Y^k(r))(Z^n(r) - Z^k(r))dB_r.
\end{aligned}$$

Taking expectation of (3.3.30),

$$
\begin{aligned}
&\mathbb{E}[(Y^n(s)-Y^k(s))^2]+\mathbb{E}\left[\int_s^T (Z^n(r)-Z^k(r))^2dr\right]\\
\leq&\mathbb{E}\left[\int_s^T (Y^n(r)-Y^k(r))(g^n(r,X(r),Y^n(r),Z^n(r))-g^k(r,X(r),Y^k(r),Z^k(r)))dr\right]\\
\leq&\mathbb{E}\left[\int_s^T (Y^n(r)-Y^k(r))^2dr\right]^{\frac{1}{2}}\\
&\cdot\mathbb{E}\left[\int_s^T (g^n(s,X(r),Y^n(r),Z^n(r))-g^k(r,X(r),Y^k(r),Z^k(r)))^2dr\right]^{\frac{1}{2}}.
\end{aligned}
\tag{3.3.31}
$$

In order to prove convergence of $\{Z^n(\cdot)\}_n$, it suffices to prove uniform boundedness of $\mathbb{E}\left[\int_t^T g^n(s,X(s),Y^n(s),Z^n(s))^2ds\right]$, for all n, which is part of Proposition 3.3.3. The $\mathbb{L}^2(m\times d;t,T)$-convergence of $\{Z^n(\cdot)\}_n$ implies almost sure convergence along a subsequence, also denoted as $\{Z^n(\cdot)\}_n$ to simplify notations. □

We have identified a strongly convergent subsequence of $\{(Y^n,Z^n)\}_n$, also denoted as $\{(Y^n,Z^n)\}_n$. Let's remind ourselves that (Y^n,Z^n) solves the system (3.3.1) and (3.3.11), so if the weak limit $G(\cdot)$ of $g^n(\cdot,X(\cdot),Y^n(\cdot),Z^n(\cdot))$ is also the strong limit, and if $G(\cdot)$ has the form $g(\cdot,X(\cdot),Y(\cdot),Z(\cdot))$, then the limit (Y,Z,K) indeed solves the forward-backward system (3.3.1) and (3.3.2).

Lemma 3.3.3 *As $n\to\infty$, $g^n(s,X(s),Y^n(s),Z^n(s))\to g(s,X(s),Y(s),Z(s))$, in $\mathbb{L}^2(m;t,T)$ and a.e. on $[t,T]\times\Omega$.*

Proof. The method is the same as that on page 122 of Hamadène, Lepeltier and Peng (1997) [26]. The proof is briefly repeated here for completeness.

$$
\begin{aligned}
&\mathbb{E}\left[\int_t^T |g^n(s,X(s),Y^n(s),Z^n(s))-g(s,X(s),Y(s),Z(s))|ds\right]\\
\leq&\mathbb{E}\left[\int_t^T |g^n(s,X(s),Y^n(s),Z^n(s))-g(s,X(s),Y^n(s),Z^n(s))|\mathbb{1}_{\{|Y^n(s)+Z^n(s)|\geq A\}}ds\right]\\
&+\mathbb{E}\left[\int_t^T |g^n(s,X(s),Y^n(s),Z^n(s))-g(s,X(s),Y^n(s),Z^n(s))|\mathbb{1}_{\{|Y^n(s)+Z^n(s)|\leq A\}}ds\right]\\
&+\mathbb{E}\left[\int_t^T |g(s,X(s),Y^n(s),Z^n(s))-g(s,X(s),Y(s),Z(s))|ds\right].
\end{aligned}
\tag{3.3.32}
$$

By linear growth Assumption 3.3.1 (2) for g and property (b) for g^n, and Lemma 3.3.1, both $|g^n(s,X(s),Y^n(s),Z^n(s))-g(s,X(s),Y^n(s),Z^n(s))|$ and $|g(s,X(s),Y^n(s),Z^n(s))-g(s,X(s),Y(s),Z(s))|$ are uniformly bounded in $\mathbb{L}^2(m;0,T)$ for all n. The first term on the right hand side of (3.3.32) is at most of the order $\frac{1}{A}$, thus vanishing as A goes to infinity. Recalling property (d), for fixed A, the second term vanishes as $n \to \infty$. Because of its uniform boundedness in $\mathbb{L}^2(m; t, T)$, the integrand in the third term is uniformly integrable for all n, so expectation of the integral again goes to 0 as $n \to \infty$. The a.e. convergent subsequence of $g^n(s, X(s), Y^n(s), Z^n(s))$ is also indexed by n to simplify notations. □

Proposition 3.3.4 *The $\mathbb{L}^2(m; t, T)$ convergence and the a.e. convergence of $\{Y^{(t,x),n}(s)\}_n$ to $Y^{(t,x)}(s)$ are uniform over all $s \in [t, T]$.*

Proof. To see uniform convergence of $\{Y^n\}$, applying Itô's rule to $(Y^n(s) - Y(s))^2$, integrating from s to T, taking supremum over $0 \leq s \leq T$ and then expectation, by Burkholder-Davis-Gundy inequality,

$$\begin{aligned}
&\mathbb{E}\left[\sup_{[0,T]}(Y^n(s) - Y(s))^2\right] + \mathbb{E}\left[\int_t^T (Z^n(r) - Z(r))^2 dr\right] \\
\leq &\mathbb{E}\left[\sup_{[0,T]}\int_s^T (Y^n(r) - Y(r))(g^n(r, X(r), Y^n(r), Z^n(r)) - g(r, X(r), Y(r), Z(r)))dr\right] \\
&+ \mathbb{E}\left[\left(\int_t^T (Y^n(r) - Y(r))^2 (Z^n(r) - Z(r))^2 dr\right)^{\frac{1}{2}}\right] \\
\leq &\mathbb{E}\left[\sup_{s\in[0,T]}\left(\int_s^T (Y^n(r) - Y(r))^2 dr\right)^{\frac{1}{2}}\left(\int_s^T (g^n(s, X(r), Y^n(r), Z^n(r)) - g(r, X(r), Y(r), Z(r)))^2 dr\right)^{\frac{1}{2}}\right] \\
&+ \mathbb{E}\left[\sup_{s\in[0,T]}\{|Y^n(s) - Y(s)|\}\left(\int_t^T (Z^n(r) - Z(r))^2 dr\right)^{\frac{1}{2}}\right] \\
\leq &\left(\mathbb{E}\left[\int_t^T (Y^n(r) - Y(r))^2 dr\right]\right)^{\frac{1}{2}}\left(\mathbb{E}\left[\int_t^T (g^n(s, X(r), Y^n(r), Z^n(r)) - g(r, X(r), Y(r), Z(r)))^2 dr\right]\right)^{\frac{1}{2}} \\
&+ \frac{1}{4}\mathbb{E}\left[\sup_{[0,T]}|Y^n(s) - Y(s)|^2\right] + \mathbb{E}\left[\int_t^T (Z^n(r) - Z(r))^2 dr\right].
\end{aligned} \tag{3.3.33}$$

Equation (3.3.33) implies

$$\begin{aligned}
&\frac{3}{4}\mathbb{E}\left[\sup_{[0,T]}(Y^n(s) - Y(s))^2\right] \\
\leq &\left(\mathbb{E}\left[\int_t^T (Y^n(r) - Y(r))^2 dr\right]\right)^{\frac{1}{2}}\left(\mathbb{E}\left[\int_t^T (g^n(s, X(r), Y^n(r), Z^n(r)) - g(r, X(r), Y(r), Z(r)))^2 dr\right]\right)^{\frac{1}{2}}.
\end{aligned} \tag{3.3.34}$$

By Proposition 3.3.3, by linear growth properties (b) of g^n and Assumption 3.3.1 (2) on g, and by Lemma 3.3.1, the second multiplier on the right hand side of (3.3.34) is bounded by a constant, uniformly over all n. By Lemma 3.3.2, the first multiplier on the right hand side of (3.3.34) converges to zero as $n \to \infty$. Hence

$$\lim_{n\to\infty}\mathbb{E}\left[\sup_{[0,T]}(Y^n(s) - Y(s))^2\right] = 0. \tag{3.3.35}$$

□

Proposition 3.3.5 *The process $K^n(\cdot)$ converges to some limit $K(\cdot)$ in $\mathbb{M}^1(m;t,T)$, uniformly over all $s \in [t,T]$, and a.e. on $[t,T]\times\Omega$.*

Proof. Define

$$\bar{K}(s) := Y(t) - Y(s) - \int_t^s g(r,X(r),Y(r),Z(r))dr + \int_t^s Z(r)dB_r,\ t \le s \le T, \quad (3.3.36)$$

where $Y(\cdot)$, $Z(\cdot)$ and g are the limits of $Y^n(\cdot)$, $Z^n(\cdot)$ and g^n. From (3.3.11),

$$K^n(s) = Y^n(t) - Y^n(s) - \int_t^s g^n(r,X(r),Y^n(r),Z^n(r))dr + \int_t^s Z^n(r)dB_r. \quad (3.3.37)$$

Need to show that

$$\mathbb{E}\left[\sup_{s\in[0,T]}\left|K^n(s) - \bar{K}(s)\right|\right] \to 0, \quad (3.3.38)$$

as $n \to \infty$.

For all $n = 1, 2, \cdots,$

$$\begin{aligned}&\mathbb{E}\left[\sup_{[0,T]}\left|K^n(s) - \bar{K}(s)\right|\right]\\ \le&\mathbb{E}\left[\,|Y^n(t) - Y(t)|\,\right] + \mathbb{E}\left[\sup_{[0,T]}|Y^n(s) - Y(s)|\right] + \mathbb{E}\left[\sup_{s\in[0,T]}\left|\int_t^s (Z^n(r) - Z(r))dB_r\right|\right]\\ &+ \mathbb{E}\left[\int_t^T |g^n(r,X(r),Y^n(r),Z^n(r)) - g(r,X(r),Y(r),Z(r))|\,dr\right].\end{aligned} \quad (3.3.39)$$

As $n \to \infty$, the first three summands in (3.3.39) go to zero, by Lemma 3.3.2, Proposition 3.3.4 and Lemma 3.3.3. From Burkholder-Davis-Gundy inequality, there exists a constant C universal for all n, such that

$$\mathbb{E}\left[\sup_{s\in[0,T]}\left|\int_t^s (Z^n(r) - Z(r))dB_r\right|\right] \le C\mathbb{E}\left[\left(\int_t^T |Z^n(r) - Z(r)|^2 dr\right)^{\frac{1}{2}}\right], \quad (3.3.40)$$

the right hand side of which converges to zero as $n \to \infty$, by Lemma 3.3.2.
The a.e. convergent subsequence is still denoted as $\{K^n(\cdot)\}_n$ to simplify notations. The strong limit $\bar{K}(\cdot)$ coincides with the weak limit $K(\cdot)$ in Proposition 3.3.3. □

Proposition 3.3.6 *The processes $Y(\cdot)$ and $K(\cdot)$ satisfy the reflection conditions $Y(\cdot) \ge L(\cdot, X(\cdot))$ and $\int_t^T (Y(s) - L(s,X(s)))dK(s) = 0$.*

Proof. Since (Y^n, Z^n, K^n) solves (3.3.11), $Y^n(\cdot)$ and $K(\cdot)$ satisfy the reflecting conditions $Y^n(s) \ge L(s,X(s))$, $t \le s \le T$, and $\int_t^T (Y^n(s) - L(s,X(s)))dK^n(s) = 0$. Since $Y^n(\cdot)$

converges to $Y(\cdot)$ pointwisely on $[0,T]\times\Omega$, that $Y(\cdot)\geq L(\cdot,X(\cdot))$ holds true. It remains to prove

$$\int_t^T (Y(s)-L(s,X(s)))dK(s) = \int_t^T (Y^n(s)-L(s,X(s)))dK^n(s). \tag{3.3.41}$$

To wit,

$$\begin{aligned}
&\left|\int_t^T (Y^n(s)-L(s,X(s)))dK^n(s) - \int_t^T (Y(s)-L(s,X(s)))dK(s)\right| \\
&\leq \left|\int_t^T (Y^n(s)-Y(s))dK^n(s)\right| + \left|\int_t^T (Y(s)-L(s,X(s)))d(K(s)-K^n(s))\right| \qquad (3.3.42)\\
&\leq \left|\sup_{s\in[0,T]}\{Y^n(s)-Y(s)\}K^n(T)\right| + \left|\int_t^T (Y(s)-L(s,X(s)))d(K(s)-K^n(s))\right|.
\end{aligned}$$

Let n tend to zero. By Proposition 3.3.4, the first summand in the last line of (3.3.42) converges to $|0\cdot K(T)|=0$, a.e. on Ω. Proposition 3.3.5 implies that $K^n(s)$ converges to $K(s)$ in probability, uniformly over all $s\in[t,T]$, so the measure $dK^n(s)$ weakly converges to $dK(s)$ in probability, uniformly over all $s\in[t,T]$. It follows that the second summand in the last line of (3.3.42) converges to zero, a.e. on Ω. □

We may now conclude the following existence result.

Theorem 3.3.2 *Under Assumption 3.3.1, there exists a solution (Y,Z,K) to the BSDE (3.3.2) with reflecting barrier in the Markovian framework.*

Proof. The solutions $\{(Y^n,Z^n,K^n)\}_n$ to the approximating equations (3.3.11) have limits (Y,Z,K). The triplet (Y,Z,K) is a solution to the Markovian system (3.3.1) and (3.3.2). □

Theorem 3.3.3 *(Comparison Theorem)*
Suppose $(Y^{t,x},Z^{t,x},K^{t,x})$ solves forward-backward system (3.3.1) and (3.3.2) with parameter set (ξ,g,L), and $(\overline{Y}^{t,x},\overline{Z}^{t,x},\overline{K}^{t,x})$ solves the forward-backward system (3.3.1) and (3.3.2) with parameter set $(\overline{\xi},\overline{g},\overline{L})$. Let dimension of the equations be $m=1$. Under Assumption 3.3.1 for both sets of parameters, if
(1) $\xi(x)\leq\overline{\xi}(x)$, a.e., $\forall x\in\mathbb{R}^l$;
(2) $g(s,x,y,z)\leq\overline{g}(s,x,y,z)$, for all $t\leq s\leq T$, and all $(x,y,z)\in\mathbb{R}^l\times\mathbb{R}\times\mathbb{R}^d$; and
(3) $L(s,x)\leq\overline{L}(s,x)$, for all $t\leq s\leq T$, and all $x\in\mathbb{R}^l$,
then

$$Y^{t,x}(s)\leq\overline{Y}^{t,x}(s), \textit{ for all } t\leq s\leq T. \tag{3.3.43}$$

Proof. Let $\{g^n\}_n$ and $\{\overline{g}^n\}_n$ be, respectively, the uniform Lipschitz sequences approximating g and $\overline{g}$ as in (3.3.9). According to Property (a), both g^n and $\overline{g}^n$ are Lipschitz in (y,z), for all t and x. We notice that (2) in the conditions of this theorem implies that

$$g^n(s,x,y,z)\leq\overline{g}^n(s,x,y,z), \tag{3.3.44}$$

for all $t \leq s \leq T$, and all $(x, y, z) \in \mathbb{R}^l \times \mathbb{R}^m \times \mathbb{R}^{m\times d}$, via construction (3.3.9). Let $(Y^{(t,x),n}, Z^{(t,x),n}, K^{(t,x),n})$ be solution to system (3.3.1) and (3.3.2) with parameter set (ξ, g^n, L), and $(\overline{Y}^{(t,x),n}, \overline{Z}^{(t,x),n}, \overline{K}^{(t,x),n})$ be solution to system (3.3.1) and (3.3.2) with parameter set $(\overline{\xi}, \overline{g}^n, \overline{L})$. By Theorem 3.2.2,

$$Y^{(t,x),n}(s) \leq \overline{Y}^{(t,x),n}(s),\ t \leq s \leq T. \tag{3.3.45}$$

But as $n \to \infty$, proven earlier in this section,

$$Y^{(t,x),n}(\cdot) \to Y^{t,x}(\cdot),\ \overline{Y}^{(t,x),n}(\cdot) \to \overline{Y}^{t,x}(\cdot),\ \text{a.e. on } [t,T]\times\Omega \text{ and in } \mathbb{L}^2(m;t,T), \tag{3.3.46}$$

so

$$Y^{t,x}(s) \leq \overline{Y}^{t,x}(s),\ t \leq s \leq T. \tag{3.3.47}$$

□

Theorem 3.3.4 *(Continuous Dependence Property)*
Under Assumption 3.3.1, if $(Y^{t,x}, Z^{t,x}, K^{t,x})$ solves the system (3.3.1) and (3.3.2), and $(\overline{Y}^{t,x}, \overline{Z}^{t,x}, \overline{K}^{t,x})$ solves the system (3.3.1) and

$$\begin{cases} \overline{Y}^{t,x}(s) = \overline{\xi}(X^{t,x}(T)) + \int_s^T g(r, X^{t,x}(r), \overline{Y}^{t,x}(r), \overline{Z}^{t,x}(r))dr - \int_s^T \overline{Z}^{t,x}(r)dB_r \\ \qquad + \overline{K}^{t,x}(T) - \overline{K}^{t,x}(s); \\ \overline{Y}^{t,x}(s) \geq L(s, X^{t,x}(s)),\ t \leq s \leq T,\ \int_t^T (\overline{Y}^{t,x}(s) - L(s, X^{t,x}(s)))d\overline{K}^{t,x}(s) = 0, \end{cases} \tag{3.3.48}$$

then

$$\begin{aligned} &\mathbb{E}[(Y^{t,x}(s) - \overline{Y}^{t,x}(s))^2] + \mathbb{E}\left[\int_s^T (Z^{t,x}(r) - \overline{Z}^{t,x}(r))^2 dr\right] \\ \leq &\mathbb{E}[|\xi - \overline{\xi}|^2] + C\mathbb{E}\left[\int_s^T (Y^{t,x}(r) - \overline{Y}^{t,x}(r))^2 dr\right]^{\frac{1}{2}},\ 0 \leq t \leq s \leq T. \end{aligned} \tag{3.3.49}$$

Proof. Apply Itô's rule to $(Y^{t,x} - \overline{Y}^{t,x})^2$, and integrate from s to T. Use Lemma 3.3.1 and Assumption 3.3.1 (2). □

Remark 3.3.2 *When the driver g is concerned about in Assumption 3.3.1, 3.3.1 (2) (linear growth rates in y and z, and polynomial growth rate in x) is crucial in bounding the $\mathbb{L}^2$-norms thus proving convergence of a Lipschitz approximating sequence. Continuity Assumption 3.3.1 (3) is only for convenience, because a measurable function can always be approximated by continuous functions of the same growth rate.*

Remark 3.3.3 *The results in section 3.2 and section 3.3 are valid for any arbitrary filtered probability space that can support a d-dimensional Brownian motion. In particular, in the canonical space set up at the beginning of section 3.1, we may replace Assumption 3.3.1 (1) and (2) with the more general Assumption 3.3.1 (1') and (2'), while still getting exactly the same statements in section 3.3 with tiny modifications of the proofs. Assumption 3.3.1 corresponds to Assumption 3.1.1 on the state process $X(\cdot)$ in (3.1.3). The growth rate (3.1.56) of the Hamiltonians (3.1.55) satisfies Assumption 3.3.1 (2').*

Assumption 3.3.1 *(1') In (3.3.1), the drift* $f : [0,T] \times C^l[0,\infty) \to \mathbb{R}^l$, $(t,\omega) \mapsto f(t,\omega(t))$*, and volatility* $\sigma : [0,T] \times C^l[0,\infty) \to \mathbb{R}^{l\times d}$, $(t,\omega) \mapsto \sigma(t,\omega(t))$*, are deterministic, measurable mappings such that*

$$|f(t,\omega(t)) - f(t,\bar{\omega}(t))| + |\sigma(t,\omega(t)) - \sigma(t,\bar{\omega}(t))| \le C \sup_{0\le s\le t} |\omega(s) - \bar{\omega}(s)|, \tag{3.3.50}$$

and

$$|f(t,\omega(t))|^2 + |\sigma(t,\omega(t))|^2 \le C\left(1 + \sup_{0\le s\le t} |\omega(s)|^2\right), \tag{3.3.51}$$

with some constant C *for all* $0 \le t \le T$, ω *and* $\bar{\omega}$ *in* $C^l[0,\infty)$.
(2') In (3.3.2), the driver g *is a deterministic measurable mapping* $g : [0,T]\times C^l[0,\infty)\times \mathbb{R}^m\times^{m\times d} \to \mathbb{R}^m$, $(t,\omega,y,z) \mapsto g(t,\omega(t),y,z)$*. And*

$$|g(t,\omega(t),y,z)| \le b\left(1 + \sup_{0\le s\le t} |\omega(s)|^p + |y| + |z|\right), \tag{3.3.52}$$

with some positive constant b *for all* $(t,\omega,y,z) \in [0,T] \times C^l[0,\infty) \times \mathbb{R}^m \times \mathbb{R}^{m\times d}$.

Bibliography

[1] Beneš, V. E. (1970). Existence of Optimal Strategies Based on Specified Information for a Class of Stochastic Decision Problems. *SIAM J. Control and Optimization*. Vol. 8, 179-188.

[2] Beneš, V. E. (1971). Existence of Optimal Stochastic Control Laws. *SIAM J. Control Optim.* 9, 446-472.

[3] Bensoussan, A., and A. Friedman (1977). Nonzero-Sum Stochastic Differential Games With Stopping Times and Free Boundary Problems. *Transactions of the American Mathematical Society*, Vol. 231, No. 2, 275-327.

[4] Bensoussan, A., and J. Frehse (2000). Stochastic Games for N Players. *Journal of Optimization Theory and Applications*, Vol. 105, No. 3, 543-565.

[5] Bensoussan, A., J. Frehse, and H. Nagai, 1998. Some Results on Risk-Sensitive Control with Full Observation. *J. Appl. Math. Optim.* 37 (1), 1-41.

[6] Bensoussan, A., and A. Friedman (1977). Nonzero-Sum Stochastic Differential Games With Stopping Times and Free Boundary Problems. *Transactions of the American Mathematical Society*, Vol. 231, No. 2, 275-327.

[7] Bismut, J. M. (1973). Conjugate Convex Functions in Optimal Stochastic Control. *J. Math. Anal. Appl.*. 44, 384-404.

[8] Briand, Coquet, Hu, Mémin, and Peng (2000). A Converse Comparison Theorem for BSDEs and Related Properties of g-Expectation. *Electronic Communications in Probability*. 5, 101-117.

[9] Çinlar, E., J. Jacod, P. Protter, and M. J. Sharpe (1980). Semimartingales and Markov Processes. *Z. Warscheinlichkeitstheorie verw. Gebiete*. 54, 161-219.

[10] Cvitanić, J., and I. Karatzas (1996). Backward Stochastic Differential Equations with Reflection and Dynkin Games. *The Annals of Probability*. Vol. 24, No. 4, 2024-2056.

[11] Davis, M. H. A. (1973). On the Existence of Optimal Policies in Stochastic Control. *SIAM J. Control Optim.* 11, 587-594.

[12] Davis, M. H. A. (1979). Martingale Methods in Stochastic Control. *Lecture Notes in Control and Information Sciences*. 16. Springer-Verlag, Berlin.

[13] Davis, M. H. A., and P. P. Varaiya (1973). Dynamic Programming Conditions for Partially Observable Stochastic Systems. *SIAM J. Control Optim.* 11, 226-261.

[14] Dubins, L. E., and L. J. Savage (1965). *How to Gamble If You Must: Inequalities for Stochastic Processes*. McGraw-Hill.

[15] Duffie, D. J. (2006). *Finite difference methods in financial engineering : a partial differential equation approach*. Wiley.

[16] Duncan, T. E., and P. P. Varaiya (1971). On the Solutions of a Stochastic Control System. *SIAM J. Control Optim.* 9, 354-371.

[17] Dynkin, E. B., and Yushkevich, A. A. (1968). *Theorems and Problems in Markov Processes.* Plenum Press, New York.

[18] El Karoui, N., and S. Hamadène (2003). BSDEs and Risk-Sensitive Control, Zero-Sum and Nonzero-Sum Game Problems of Stochastic Functional Differential Equations. *Stochastic Processes and their Applications* 107 145-169.

[19] El Karoui, N., C. Kapoudjian, E. Pardoux, S. Peng, and M. C. Quenez (1997). Reflected Solutions of Backward SDE'S, and Related Obstacle Problems for PDE's. *The Annals of Probability*. Vol. 25, No. 2, 702-737.

[20] El Karoui, N., S. Peng, and M. C. Quenez (1997). Backward Stochastic Differential Equations in Finance. *Mathematical Finance*. Vol. 7, No. 1, 1-71.

[21] Elliott, R. J. (1982). *Stochastic Calculus and Applications.* Springer, New York

[22] Elliott, R. J. (1976). The Existence of Value in Stochastic Differential Games. *SIAM J. Control Optim.* Vol. 14, No. 1, 85-94.

[23] Fleming, W. H., and H. M. Soner (1993). *Controlled Markov Processes and Viscosity Solutions*. Springer, New York.

[24] Fujisaki, M., G. Kallianpur, and H. Kunita (1972). Stochastic Differential Equations for the Non Linear Filtering Problem. *Osaka J. Math.* 9, 19-40

[25] Fukushima, M., and M. I. Taksar (2002). Dynkin Games via Dirichlet Forms and Singular control of One-Dimensional Diffusions. *SIAM J. Control Optimization*. Vol 41, No. 3, 682-699.

[26] Hamadène, S., J-P. Lepeltier, and S. Peng (1997). BSDEs with Continuous Coefficients and Stochastic Differential Games. *Pitman Research Notes in Mathematics Series* 364.

[27] Hamadène, S. (1998). Backward-forward SDE's and Stochastic Differential. Games. *Stochastic Processes and their Applications* 77, 1-15.

[28] Hamadène, S. (1999). Nonzero Sum Linear-Quadratic Stochastic Differential Games and Backward-forward Equations. *Stochastic Analysis and Applications*. Vol. 17, No. 1, 117-130.

[29] Hamadène, S., J-P. Lepeltier (2000). Reflected BSDE's and Mixed Game Problem *Stochastic Processes and their Applications* 85, 177-188.

[30] Hamadène, S. (2006). Mixed Zero-sum Stochastic Differential Game and American Game Options. *SIAM J. Control and Optimization*. Vol. 45, No. 2, 496-518.

[31] Hu, Ying, and Shige Peng. On the Comparison Theorem for Multidimensional BSDEs. *C. R. Acad. Sci. Paris, Ser.* I 343 (2006) 135-140.

[32] Karatzas, I., and Q. Li (2009). A BSDE Approach to Non-Zero-Sum Stochastic Differential Games of Control and Stopping. *Submitted*.

[33] Karatzas, I., and S. E. Shreve. *Brownian Motion and Stochastic Calculus*. Springer-Verlag, 1988.

[34] Karatzas, I., and S. E. Shreve (1998). *Methods of Mathematical Finance*. Springer-Verlag, 1998.

[35] Karatzas, I., and W. Sudderth (2006). Stochastic Games of Control and Stopping for a Linear Diffusion. *WSPC - Proceedings*.

[36] Karatzas, I., and H. Wang (2001). Connections Between Bounded-Variation Control and Dynkin Games. *IOS Press, Amsterdam.* 353-362.

[37] Karatzas, I., and I-M. Zamfirescu (2006). Martingale Approach to Stochastic Control with Discretionary Stopping. *Appl Math Optim* 53, 163-184.

[38] Karatzas, I., and I-M. Zamfirescu (2008). Martingale Approach to Stochastic Differential Games of Control and Stopping. *The Annals of Probability*. Vol. 36, No. 4, 1495-1527.

[39] Kobylanski, M. (2000). Backward stochastic differential equations and partial differential equations with quadratic growth. *The Annals of Probability*. Vol. 28, No. 2, 558-602.

[40] Lepeltier J.P., and E. Etourneau (1987). A Problem of Non-zero Sum Stopping Game. *Lecture notes in control and information sciences; 96*. Springer-Verlag, 1987.

[41] Neveu, J. (1975). *Discrete Parameter Martingales*. Elsevier Science, June 1975.

[42] Ocone, D. L., and A. P. Weerasinghe (2008). A Degenerate Variance Control Problem with Discretionary Stopping. *IMS Collections. Markov Processes and Related Topics: A Festchrift for Thomas G. Kurtz.* Vol. 4 (2008), 155-167.

[43] Pardoux, E., and S. Peng (1990). Adapted Solution of a Backward Stochastic Differential Equation. *Systems & Control Letters* 14 55-61.

[44] Revuz, D, and M. Yor (1999). *Continuous martingales and Brownian motion*. Springer.

[45] Shiryayev, A. N. *Optimal stopping rules.* Applications of Math. Vol. 8, Springer-Verlag, 1979.

[46] Snell, J. L. (1952). Applications of Martingale Systems Theorems. *Transactions of the American Mathematical Society*. Vol. 73, No. 2, 293-312.

[47] Taksar, M. I. (1985). Average Optimal Singular Control and a Related Stopping Problem. *Mathematics of Operations Research*. Vol. 10, No. 1, 63-81.