Deep learning method applied to stochastic maximum principle for optimal control

By

RANAMBININTSOA MARKA (marka.ranambinintsoa@aims.ac.rw)
African Institute for Mathematical Sciences (AIMS), Rwanda
Supervised by: Pr.Olivier Menoukeu Pamen
University of Liverpool, Department of Mathematical Sciences, United Kingdom

June 2024

AN ESSAY PRESENTED TO AIMS RWANDA IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE AWARD OF MASTER OF SCIENCE IN MATHEMATICAL SCIENCES

AIMS African Institute for Mathematical Sciences RWANDA

DECLARATION

This work was carried out at AIMS Rwanda in partial fulfilment of the requirements for a Master of Science Degree.

I hereby declare that except where due acknowledgement is made, this work has never been presented wholly or in part for the award of a degree at AIMS Rwanda or any other University.

Student :Marka RANAMBININTSOA

Supervisor:Professor, Olivier Menoukeu Pamen

ACKNOWLEDGEMENTS

I would like to extend my heartfelt thanks to several individuals and groups who have supported me throughout the journey of completing this thesis.

First and foremost, I am deeply grateful to the staff at AIMS Rwanda for their unwavering support and care. Your dedication to ensuring a conducive environment for our studies was invaluable and greatly appreciated.

I would also like to express my sincere gratitude to my supervisor, Pr. Olivier Menoukeu Pamen, for providing me with the thesis topic and guiding me throughout this process. Your insightful feedback and encouragement were crucial in shaping this work.

A special thank you goes to my tutors, Dr. Roger Ranomenjanahary, who meticulously corrected my English and helped me communicate my ideas more effectively. Your patience and expertise were indispensable.

Lastly, I am thankful to all the students at AIMS. Sharing this journey with you has been an enriching experience, and your camaraderie and support made these years memorable and rewarding.

Thank you all for your contributions and support.

Abstract

This thesis explores the integration of classical stochastic control methods with deep learning techniques to solve stochastic optimal control problems, which are pivotal in applied mathematics and financial engineering. By merging the theoretical insights of Pontryagin's Maximum Principle, duality, and convex analysis with advanced numerical methods and deep learning algorithms, we establish a robust framework for addressing high-dimensional and complex control problems under uncertainty. We present two foundational theorems on the stochastic maximum principle and reformulate the control problem into a variational form that facilitates the application of deep learning. Three algorithms are developed to solve the reformulated problem, utilizing deep neural networks and advanced optimization techniques. Numerical results show the efficacy of these algorithms in providing accurate and efficient solutions to stochastic control problems. This work underscores the significant potential of combining classical mathematical theories with modern computational approaches to advance the field of stochastic optimal control.

Contents

1. Introduction

The stochastic optimal control problem is a significant topic in applied mathematics and financial engineering, involving the optimization of a control policy in the presence of uncertainty. The integration of classical stochastic control methods with deep learning presents a promising avenue for solving complex, high-dimensional control problems. The convergence of theoretical insights from nonlinear filtering, duality, and convex analysis with modern numerical methods and deep learning algorithms offers a robust framework for addressing the challenges inherent in stochastic optimal control. Stochastic analysis is a significant mathematical field with considerable theoretical value and practical applicability. Within stochastic analysis, stochastic optimization has emerged as a substantial and rapidly expanding counterpart to deterministic optimization. Two prominent subfields within stochastic optimization are stochastic control and optimal stopping. Stochastic control encompasses a well-known challenge known as the stochastic optimal control problem, which addresses uncertainty in decision-making to either maximize or minimize an objective function.

Two fundamental tools for tackling this problem are Pontryagin's Maximum Principle and Bellman's Dynamic Programming. Pontryagin's principle, formulated in 1956 by the Russian mathematician Lev Semenovich Pontryagin, provides necessary optimality conditions for stochastic optimal control problems. Antonio Moro, A. Bensoussan, and Sonjoy K. Mitter provided foundational work in nonlinear filtering and stochastic control, addressing complex systems where the state is partially observed. Their work discussed the application of stochastic processes and control theory to filtering problems, setting a precedent for the development of more sophisticated control methods [A. Bensoussan (1981)]. Jean-Michel Bismut's contributions to duality and convex analysis in stochastic control highlighted the mathematical underpinnings necessary for understanding optimal control problems. Bismut introduced duality techniques to simplify the control problem and utilized convex analysis to address probabilistic constraints, which were essential for developing robust control policies. These theoretical advancements provided the necessary tools for tackling complex optimization problems in stochastic environments [Bismut (2014)]. Krylov and Dong focused on the convergence rates of finite-difference approximations for Bellman equations with Lipschitz coefficients, providing a theoretical foundation for numerical methods in stochastic control [Dong (2007)]. This work was complemented by Dupuis and Kushner, who detailed numerical methods for continuous-time control problems, emphasizing the practical aspects of implementing these theoretical constructs [Kushner (2001)]. Together, these contributions underscored the importance of numerical methods in solving stochastic control problems accurately and efficiently [Dong (2007) and Kushner (2001)]. The comprehensive introductions to stochastic calculus by Fima C. Klebaner and Damien Lamberton and Bernard Lapeyre were crucial for modeling and solving stochastic differential equations (SDEs) that arised in stochastic control problems [Yong (1999)]. The Euler-Maruyama scheme, highlighted in educational resources such as Euler, was a fundamental numerical method for simulating SDEs, demonstrating practical implementation aspects. These works collectively provided the mathematical and practical tools necessary for dealing with the stochastic processes inherent in control problems [Euler]. Various numerical methods exist for solving stochastic optimal control problems, including the Markov chain approximation method and probabilistic numerical methods based on dynamic programming. However, only a few of these methods were suitable for handling high-dimensional problems [Kushner (2001)]. W.E. and J Han discussed the application of deep learning to approximate solutions for stochastic control problems, marking a significant shift from traditional methods to data-driven approaches. Their work outlined how deep reinforcement learning (DRL) can be utilized to solve high-dimensional control problems by approximating the value functions and policies directly from data. This represented a major advancement in the ability to handle complex control problems that were previously infeasible with classical methods [J Han (2016)]. Peng, Ji, and Zhang proposed algorithms for solving fully-coupled forward-backward stochastic differential equations (FBSDEs) using deep learning, tackling the curse of dimensionality that plagues classical methods. Their approach leveraged neural networks to approximate the solutions of these complex systems, demonstrating the efficacy of deep learning in handling high-dimensional problems. This innovation was crucial for advancing the field of stochastic control, where high-dimensionality was a significant challenge [Shaolin Ji and Zhang (2020)]. SMP provides a set of necessary conditions for optimality in stochastic control problems. Yong and Ma and S. Peng offered detailed analyses of forward-backward SDEs and their applications to SMP. They highlighted how these systems can be utilized to derive optimal control policies, providing a theoretical basis for merging SMP with deep learning techniques [Ma (1999) , Peng (2000)]. This integration of classical principles with modern technology was essential for advancing practical applications of stochastic control. Optimization algorithms such as the limited-memory BFGS method, discussed by Nocedal and Liu, can be integrated with deep learning models to optimize parameters effectively. This was crucial for training neural networks involved in approximating solutions to stochastic control problems [Liu (1989)]. The combination of advanced optimization techniques with deep learning enhanced the efficiency and accuracy of solving complex control problems. Bernt Oksendal and J explored applications of stochastic control in various domains, including finance and engineering. His work discussed the practical implications of using SMP and deep learning for solving real-world problems, highlighting potential areas for future research [Oksendal]. These applications demonstrated the practical utility of theoretical advancements and the potential for significant impact in various fields.

This thesis employs the stochastic maximum principle alongside deep learning methods to solve stochastic optimal control problems. In the second Chapter, we discuss some notions of stochastic calculus [Klebaner (2012), Lamberton and Lapeyre (2011)] that are used on this thesis. The Chapter three divided into three parts: first, we introduce the following stochastic optimal control problem,

$$\inf_{u(.)\in\mathcal{U}_{ad}[0,T]} \mathbb{E}\left\{\int_0^T f(t,x_t,u_t)dt + h(x_T)\right\}, \tag{1.0.1}$$

with the state process given by

$$x_t = x_0 + \int_0^t b(t,x_,u_s)ds + \int_0^t \sigma(t,x_s,t_s)dW_s,$$

where b, σ, f and h are some functions satisfying some condition that will be made precise later and $\mathcal{U}_{ad}$ is the set of admissible control that will be defined.

In the second part, we shall delve into the stochastic maximum principle, along with its conditions, as introduced in Theorem 3.2.7, which serves as the main tool for solving the optimal control problem, as elaborated in [Yong (1999)]. Finally, the control problem is reformulated into a new variational control problem, which can be expressed as

$$\inf_{\bar{p}_0,\{\bar{q}_t\}_{0\le t\le T}} \mathbb{E}\left[|-h_x(\bar{x}_T)-\bar{p}_T|^2\right], \tag{1.0.2}$$

such that

$$\bar{x}_t = x_0 + \int_0^t b(s, \bar{x}_s, \bar{u}_s)ds + \int_0^t \sigma(s, \bar{x}_s, \bar{u}_s)dW_s,$$
$$\bar{p}_t = \bar{p}_0 - \int_0^t H_x(s, \bar{x}_s, \bar{p}_s, \bar{q}_s)ds + \int_0^t \bar{q}_s dW_s,$$
$$\bar{u}_t = \arg\max_{u \in U} H(t, \bar{x}_y, u, \bar{p}_t, \bar{q}_t),$$

where the process $\{\bar{q}\}_{0\leq t\leq T}$ and the initial state $\bar{p}_0$ are regarded as control, that is to facilitate the use of deep learning methods. A significant focus of this chapter is to prove of Theorem 3.2.7 and derive the new control problem. In the fourth Chapter, the numerical algorithm and the corresponding results are discussed. To tackle the numerical solution of the new control problem (3.3.3), three algorithms suitable for different cases are proposed via deep learning techniques. In the first algorithm (Algorithm 1), a single Deep Neural Network is constructed to simulate the control $\bar{q}_t$, with time t considered as part of the inputs to the neural network. Through training such a neural network, an approximate estimation of $\bar{q}_t$ is obtained, enabling the derivation of the approximate solution $(\bar{x}_t, \bar{p}_t, \bar{q}_t, \bar{u}_t)_{0\leq t\leq T}$ for the initial problem (1.0.2). To calculate the maximum condition in equation (1.0.2), the $L-BFGS$ method proposed by Lui [Liu (1989)] is utilized to approximate the optimal control $\bar{u}$. Consequently, Algorithm 1 is only suitable for handling low-dimensional problems. The second algorithm (Algorithm 2) is designed to address a broader class of stochastic optimal control problems, as discussed by Bismut [Bismut (2021)], where all coefficients are C^1 in u, and the optimal control $\bar{u}$ remains within the control domain. The objective of this algorithm is to enhance the computational efficiency of obtaining an approximate solution for the optimal control $\bar{u}$ within the maximum condition. First, the maximum condition is transformed into another type of constraint $H_u(t, x, u, p, q) = 0$. Then, two neural networks are constructed to simulate the two controls $\{\bar{q}\}_{0\leq t\leq T}$ and $\{\bar{u}\}_{0\leq t\leq T}$, respectively. This approach significantly reduces computation time, particularly benefiting high-dimensional cases where $\bar{u}$ cannot be explicitly expressed. The third algorithm is employed when the function $\bar{H}$ defined by equation (3.2.19) is known. With this algorithm, a class of high-dimensional stochastic optimal control problems can be addressed even when the optimal control $\bar{u}$ cannot be explicitly determined, as described by Shaolin Ji and Zhang [Shaolin Ji and Zhang (2020)]. It is important to note that when $\bar{u}$ has an explicit representation, as in the case of Algorithm 1, the function $\bar{H}$ can also be explicitly defined. Therefore, Algorithm 1, with an explicit representation of $\bar{u}$, essentially becomes a special case of Algorithm 3. The numerical results obtained from all three algorithms demonstrate promising performance. When the optimal control $\bar{u}$ has an explicit representation, allowing for the explicit solution of $\bar{H}$, Algorithm 3 emerges as an intuitive and superior choice. On the other hand, even in cases where $\bar{u}$ cannot be explicitly expressed, the algorithms effectively address the stochastic optimal control problem. In such cases, Algorithm 2 or 3 present preferable alternatives for high-dimensional cases, provided the conditions outlined in Sections 4.1.2 or 4.1.3 are satisfied. Otherwise, Algorithm 1 should be chosen, although it is more suitable for low-dimensional cases. In Section 4.2, the numerical results will be showcased and the performance of the three algorithms compared. Throughout this thesis, operation within the confines of a convex set is maintained.

2. Notion of stochastic calculus

Stochastic calculus is a branch of mathematics that deals with processes that involve randomness or uncertainty. It provides a framework for analyzing and modeling systems that evolve over time in a probabilistic manner. Stochastic calculus is particularly useful in fields such as finance, and economics, where random fluctuations play a significant role. Key concepts in stochastic calculus include stochastic processes, which are mathematical models that describe the evolution of random variables over time, and stochastic differential equations (SDEs), which are equations that involve both deterministic and random components. Stochastic calculus also encompasses the study of Itô calculus, which is a mathematical theory that extends calculus to deal with stochastic processes and SDEs. All of them can found in [Lamberton and Lapeyre (2011) and Klebaner (2012)].

2.1 Space probability

2.1.1 Definition. A set Ω is called a **sample space** if Ω is the set of all possible outcomes of the probabilistic experiment.

2.1.2 Definition. Let $\mathcal{F}$ be a set whose elements are subsets of Ω. Then, $\mathcal{F}$ is a σ**-algebra** if it satisfies the following axioms:

(i) $\Omega \in \mathcal{F}$.

(ii) If $A \in \mathcal{F}$, then $A^c \in \mathcal{F}$, where A^c is the complement of A.

(iii) If $\{A_n\}_{n\in\mathbb{N}} \in \mathcal{F}$, and $A = \bigcup_{n=1}^{\infty} A_n$, then $A \in \mathcal{F}$.

If $\mathcal{F}$ and $\mathcal{A}$ are two σ-algebra and $\mathcal{A} \subset \mathcal{F}$, we say that $\mathcal{A}$ is a **sub-sigma-algebra** of $\mathcal{F}$.

2.1.3 Definition. A space $(\Omega, \mathcal{F})$ is called **measurable space** if $\mathcal{F}$ is a σ-algebra of Ω.

2.1.4 Remark. Let $(\Omega, \mathcal{F})$ be a measurable space. A subset $\mathcal{E}$ of Ω is said to be measurable if $\mathcal{E} \in \mathcal{F}$.

2.1.5 Definition. A function $f : \Omega \to \bar{\mathbb{R}}$ is called **measurable** if, for each $a \in \bar{\mathbb{R}}$, $\{x \in X \subset \Omega : f(x) > a\}$ is a measurable set.

2.1.6 Definition. Let P be a function that associates a real number to each element of σ-algebra $\mathcal{F}$. Then P is **probability measure** if the following conditions are satisfied:

(i) If $A \in \mathcal{F}$, then $P(A) \geq 0$.

(ii) $P(\Omega) = 1$.

(iii) If $\{A_n\}_{n\in\mathbb{N}} \in \mathcal{F}$ with $A_j \cap A_k = \emptyset$ if $j \neq k$, then

$$P\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} P(A_n).$$

2.1.7 Definition. A **probability space** is a triple $(\Omega, \mathcal{F}, P)$ where Ω is a sample space, $\mathcal{F}$ is σ-algebra and P is a probability measure.

The σ-algebra $\mathcal{F}$ is complete with respect the probability measure P, if for every subset $A \subset B$ and $B \in \mathcal{F}$ such that $P(B) = 0$, then A also belongs to $\mathcal{F}$. Therefore, we say that the space $(\Omega, \mathcal{F}, P)$ is **complete** if the σ-algebra $\mathcal{F}$ is complete.

Let $(\Omega, \mathcal{F}, P)$ be a probability space and let I be an index set with a total order (often $\mathbb{N}, \mathbb{R}^+$, or a subset of $\mathbb{R}^+$).

For every $i \in I$, let $\mathcal{A}_i$ be a sub-sigma-algebra of $\mathcal{F}$. Then $\mathbb{F} := (\mathcal{A}_i)_{i \in I}$ is called a **filtration** if $\mathcal{A}_k \subseteq \mathcal{A}_\ell$ for all $k \leq \ell$. Note that filtrations are families of σ-algebras that are ordered in a non-decreasing manner.

2.1.8 Definition. A space $(\Omega, \mathcal{F}, \mathbb{F}, P)$ is said to be **filtered probability space** if $\mathbb{F}$ is a filtration.

2.1.9 Definition. A sequence $(X(t))_{0 \leq t \leq T}$ of random variable is **adapted** to the filtration if, for any t, $X(t)$ is $\mathcal{F}_t-$measurable.

2.2 Stochastic process

The present section deals with the stochastic process. The main topics dicussed are stochastic processes and their properties, these include Brownian motion, and Ito process any dimension. The majority of the materials presented in this section are sourced from the book authored by Lamberton and Lapeyre (2011). The notation $W(t)$ represent the Brownian motion.

2.2.1 Definition. Let I be a non empty index set and let $(\Omega, \mathcal{F}, P)$ be a probability space. A family $\{W(t), t \in I\}$ of a random variable from $(\Omega, \mathcal{F}, P)$ to $\mathbb{R}^m$ is called **stochastic process**, where

$$W(t) : \Omega \times I \to \mathbb{R}^m$$

is a random variable that moves across of time.

A particularly important example of a stochastic process is a Brownian motion .

2.2.2 Definition. A **Brownian motion** is a real-valued, continuous stochastic process $(W_t)_{t \geq 0}$ with independent and stationary increments. Mathematically, we can formulate this statement as follows:

(i) Continuity: P almost surely the map $s \mapsto W_s(w)$ continue.

(ii) Independent increments: if $s \leq t$, $W(t) - W(s)$ is independent of

$\mathcal{F}_s = \sigma(W(u), u \leq s).$

(iii) Stationary increments : if $s \leq t$, $W(t) - W(s)$ and $W(t - s) - W(0)$ have the same probability law.

2.2.3 Theorem. *If $(W(t))_{t \geq 0}$ is a Brownian motion, then $W(t) - W(0)$ is a normal random variable with mean rt and variance $\sigma^2 t$, where r and σ are constant real numbers.*

2.2.4 Remark. Let $\{W(t)\}_{t\in[0,\infty]}$ a Brownian motion satisfying the following conditions:

$$W(0) = 0, \quad E[W(t)] = 0 \quad \text{and} \quad Var[W(t)] = t.$$

Any Brownian motion that holds these properties is called a **Brownian motion standard**.

From now on, the Brownian motion is assumed to be standard if nothing else is mentioned. The distribution of W_t can be expressed as:

$$\frac{1}{\sqrt{2\pi t}} \exp\left(-\frac{x^2}{2t}\right) dx, \tag{2.2.1}$$

where dx is a Lebesgue measure on $\mathbb{R}$.

2.2.5 Definition. Let $\mathbb{F} = \{\mathcal{F}_t\}_{t\geq 0}$ is a filtration and let $(\Omega, \mathcal{F}, \mathbb{F}, P)$ be a filtered probability space. An $\mathbb{F}$-adapted $\mathbb{R}^m$- value process $X(.)$ is called an **m-dimentional** $\mathbb{F}-$Brownian motion over $[0,\infty)$ if for all $0 \leq s < t$, $X(t) - X(s)$ is independent of $\mathcal{F}_s$ and is normally distributed with mean zeros and covariance $(t-s)I$, where I is a matrix identity.

2.2.6 Definition. A real value, continuous stochastic process is an $(\mathcal{F}_t)_{t>0}-$Brownian motion if the following conditions hold:

(i) For any $t \geq 0$, $X(s)$ is $\mathcal{F}_t-$ measurable.

(ii) If $s \leq t$, $X(t) - X(s)$ is independent of the sigma-algebra $\mathcal{F}_s$.

(iii) If $s \leq t$, $X(t) - X(s)$ and $X(t-s) - X(0)$ have the same probability law.

2.2.7 Remark. The first condition in the Definition 2.2.6 shows that $\sigma(X(u), u \leq t) \subset \mathcal{F}_t$. This implies that the $\mathbb{F}-$Brownian motion is a Brownian motion with respect to its natural filtration.

2.2.8 Definition. A stochastic process $\{X(t), t \geq 0\}$ is called a **martingale** if the following conditions are satisfied:

(i) $X(t)$ is $\mathcal{F}_t$-adapted,

(ii) $\mathbb{E}(|X(t)|) < \infty$,

(iii) $\mathbb{E}(X(t)|\mathcal{F}_s) = X(s)$, for $s \leq t$, a.s.

2.2.9 Example. An one-dimensional random walk is a martingale, such that

$$P_{left} = \frac{1}{2} = P_{right}.$$

2.3 Itô process

An Itô process, named after the Japanese mathematician Kiyoshi Itô, is a stochastic process that includes a component of random noise or uncertainty. It is often used in mathematical finance and stochastic calculus to model systems where randomness plays a significant role. Our primary goal now is to establish the Itô formula.

2.3.1 Definition. An **Itô process** (or a stochastic integral) is a stochastic process on $(\Omega, \mathcal{F}, P)$ adapted to $(\mathcal{F}_t)_{t>0}$ which can be written as

$$X(t) = X(0) + \int_0^t b_s ds + \int_0^t \sigma_s dW(s), \tag{2.3.1}$$

where $W(t)$ is a Brownian motion, b_t is predictable and integrable, and σ_t is predictable W-integrable process, that is

$$\int_0^t \left(\sigma_s^2 + |b_s|\right) ds < \infty, \text{ a.s.} \tag{2.3.2}$$

In compact form, we can rewrite the Itô process in the equation (2.3.1) as:

$$dX(t) = b_t dt + \sigma_t dW(t). \tag{2.3.3}$$

2.3.2 Theorem. [*Theorem 3.4.10, Lamberton and Lapeyre (2011)*] *Let $(X_t)_{0\leq t\leq T}$ be an Itô process and f be a twice continuously differentiable function. Then*

$$f(X(t)) = f(X(0)) + \int_0^t f'(X(s))ds + \frac{1}{2}\int_0^t f''(X(s))d\langle X(s), X(s)\rangle, \tag{2.3.4}$$

where by definition

$$\langle X(t), X(t)\rangle = \int_0^t \sigma_s^2 ds$$

and

$$\int_0^t f'(X(s))dX(s) = \int_0^t f'(X(s))b_s ds + \int_0^t f'(X(s))\sigma_s dW(s).$$

Notice that if the function $(t, x) \longmapsto f(t, x)$ is twice continuously differentiable with respect to x and is once continuously differentiable with respect to t, then the Itô formula can be written as:

$$f(t, X(t)) = f(0, X(0)) + \int_0^t f'_s(s, X(s))ds + \int_0^t f'_x(s, X(s))dX(s) + \frac{1}{2}\int_0^t f''_{xx}(s, X(s))d\langle X(s), X(s)\rangle.$$

2.3.3 Multidimensional Itô formula. Let $(W(t) = (W^1(t), ..., W^p(t))_{t\geq 0}$ be a $\mathbb{R}^p$-value process adapted to $\{\mathcal{F}_t\}_{t>0}$, where all the $(W^i(t))_{t\geq 0}$ are independent standard $\{\mathcal{F}_t\}_{t>0}$-Brownian motion. This is called a standard p-dimensional $\{\mathcal{F}_t\}_{t>0}$-Brownian motion.

2.3.4 Definition. An Itô process with respect to $(W(t), \mathcal{F}_t)$, where $(W(t))_{t\geq 0}$ is a standard p-dimensional $\mathcal{F}_t$ Brownian motion, is a stochastic process $(X(t))_{0\leq t\leq T}$ of the form

$$X(t) = X(0) + \int_0^t b_s ds + \sum_{i=1}^p \int_0^t \sigma_s^i dW^i(s), \tag{2.3.5}$$

where b_t and (σ_t^i) are adapted to $\mathcal{F}_t$, and

$$\int_0^t \left(|b_s| + (\sigma_s^i)^2\right) ds < \infty.$$

By applying the Definition 2.3.4, we can reformulate the Proposition 3.4.18 in Lamberton and Lapeyre (2011) in the following way.

2.3.5 Proposition. Let $(X^1(t)), ..., (X^n(t))$ be n Itô process, where each $X^i(t)$ can be expressed as

$$X^i(t) = X^i(0) + \int_0^t b_s^i ds + \sum_{j=1}^{p} \int_0^t \sigma_s^{i,j} dW^j(s). \tag{2.3.6}$$

Then, if f is twice continuously differentiable with respect to x and is continuously differentiable with respect to t, we obtain

$$\begin{aligned} f(t, X^1(t), ..., X^n(t)) = f(0, X^0(t), ..., X^n(t)) &+ \int_0^t \frac{\partial f}{\partial s}\big(s, X^1(t), ..., X^n(t)\big)\, ds \\ &+ \sum_{i=1}^{n} \int_0^t \frac{\partial f}{\partial x_i}\big(s, X^1(t), ..., X^n(t)\big)\, dX^i(t) \\ &+ \frac{1}{2}\sum_{i,j=1}^{n} \int_0^t \frac{\partial^2 f}{\partial x_i x_j}\big(s, X^1(t), ..., X^n(t)\big)\, d\langle X^i(s), X^j(s)\rangle, \end{aligned}$$

with

- $dX^i(s) = b_s^i + \sum_{j=1}^{p} \sigma_s^{i,j} dW^j(t)$,
- $d\langle X^i(s), X^j(s)\rangle = \sum_{m=1}^{p} \sigma_s^{i,m}\sigma_s^{j,m} ds$.

2.3.6 Proposition. Let Z and $\hat{Z}$ be $\mathbb{R}^n-$valued continuous process satisfying

$$\begin{cases} dZ(t) = b(t)dt + \sigma(t)dW(t), \\ d\hat{Z}(t) = \hat{b}(t)dt + \hat{\sigma}(t)dW(t), \end{cases}$$

where $b, \sigma, \hat{b}, \hat{\sigma}$ are $\{\mathcal{F}_t\}_{t>0}$ $-$adapted measurable processes taking values in $\mathbb{R}^n$, and $\{W(t)\}_{t\geq 0}$ is one-dimensional standard Brownian motion. Then

$$d\langle Z(t), \hat{Z}(t)\rangle = \langle dZ(t), \hat{Z}(t)\rangle + \langle Z(t), d\hat{Z}(t)\rangle + \langle dZ(t), d\hat{Z}(t)\rangle, \tag{2.3.7}$$

where

$$\langle dZ(t), d\hat{Z}(t)\rangle = \langle \sigma(t)dW(t), \hat{\sigma}(t)dW(t)\rangle.$$

The equation (2.3.7) can be expressed as follows

$$\begin{aligned} \langle Z(t), \hat{Z}(t)\rangle = \langle Z(0), \hat{Z}(0)\rangle &+ \int_0^t \Big\{\langle Z(s), \hat{b}(s)\rangle + \langle b(s), \hat{Z}(s)\rangle + \langle \sigma(s), \hat{\sigma}(s)\rangle\Big\}\, ds \\ &+ \int_0^t \Big\{\langle \sigma(s), \hat{Z}(s)\rangle + \langle Z(s), \hat{\sigma}(s)\rangle\Big\}\, dW(s). \end{aligned} \tag{2.3.8}$$

2.4 Stochastic Differential Equation (SDE)

We are concerned in this section with stochastic differential equation. All the material related to stochastic differential equations discussed here can be found in [Klebaner (2012) , Lamberton and Lapeyre (2011)]. Let $\{W(t), t \geq 0\}$, be a Brownian motion process.

2.4.1 Ordinary differential equation (ODE). Denote $X(t)$ as a differential function defined for $t \geq 0$ and $b(X(t),t)$ a function. For all $t \in [0,T]$,

$$\frac{dX(t)}{dt} = b(X(t),t) \quad \text{and} \quad X(0) = X_0.$$

We say that $X(t)$ is a solution of the ODE if the function $X(t)$ satisfies the above conditions.

2.4.2 Definition. An equation of the form

$$dX(t) = b(X(t),t)dt + \sigma(X(t),t)dW(t), \tag{2.4.1}$$

where $b(x,t)$ and $\sigma(x,t)$ are given functions and $X(t)$ is an unknown-process, is called a **stochastic differential equation** (SDE).

The function $b(x,t)$ is called drift coefficient and $\sigma(x,t)$ is called diffusion coefficient. Our main aim in this section is to find the expression of the unknown-process X_t. Let us examine the characteristic of the solution of the equation (2.4.1).

2.4.3 Definition. A process $X(t)$ is called a **strong solution** of the SDE, if $X(t)$ is adapted to the filtration $\{\mathcal{F}_t\}$ generated by the Brownian motion $\{W(t)\}$, for any $t > 0$, the integral $\int_0^t b(X(s),s)ds$ and $\int_0^t \sigma(X(s),s)dW(s)$ exist, with the second one being an Itô integral, and

$$X(t) = X(0) + \int_0^t b(X(s),s)ds + \int_0^t \sigma(X(s),s)dW(s). \tag{2.4.2}$$

2.4.4 Existence and uniqueness of strong solutions. Let $X(t)$ be a process satisfying

$$dX(t) = b(X(t),t)dt + \sigma(X(t),t)dW(t). \tag{2.4.3}$$

2.4.5 Definition. A function f is a locally Lipchitz in x and uniformly in t if, for every T and N, there is a constant K depending in T and N, such that for all $|x|,|y| \leq N$ and all $t \in [0,T]$,

$$|f(x,t) - f(y,t)| \leq K|x-y|.$$

2.4.6 Theorem. [$Theorem$ 5.4, Klebaner (2012)] *Suppose*

1- *The coefficients b and σ are locally Lipchitz in x and uniformly in t.*

2- *The coefficients b and σ satisfy the linear growth condition*

$$|b(x,t)| + |\sigma(x,t)| \leq K(1+|x|).$$

3- *$X(0)$ is independent of $(W(t), 0 \leq t \leq T)$, and $\mathbb{E}(X^2(0)) < \infty$.*

Then, the stochastic differential equation given by equation (2.4.3) has a unique strong solution X(t). Moreover, if $X(t)$ has continuous paths, then

$$\mathbb{E}\left(sup_{0 \leq t \leq T} X^2(t)\right) < C(1 + \mathbb{E}(X^2(0))),$$

where C depends on K and T.

2.4.7 Theorem. [**Yamada-Watanabe**] *Suppose that $b(x)$ satisfies the Lipschitz condition and $\sigma(x)$ satisfies a Hölder condition of order $\alpha \geq 1/2$, that is, there is a constant K such that*

$$|\sigma(x) - \sigma(y)| < K|x - y|^{\alpha}.$$

If a strong solution exists, then the solution is unique.

2.4.8 Definition. A process $(\hat{X}(t), t \geq 0)$ is called a **weak solution** of the SDE (2.4.1) if there exists a probability space with a filtration $(\hat{\Omega}, \hat{\mathcal{F}}, \hat{\mathbb{F}}, \hat{\mathbb{P}})$, a Brownian motion $(\hat{W}(t), t \geq 0)$, and a process $(\hat{X}(t), t \geq 0)$ adapted to that filtration, such that $\hat{X}(0)$ has a given distribution. For all t, the integral below are well defined and $\hat{X}(t)$ satisfies:

$$\hat{X}(t) = \hat{X}(0) + \int_0^t b(\hat{X}(s), s)ds + \int_0^t \sigma(\hat{X}(s), s)d\hat{W}(s). \tag{2.4.4}$$

From the Definition 2.4.8, if $X(t)$ is a strong solution of an SDE, it is indeed a weak solution. The uniqueness of strong solution implies the uniqueness of weak solution. Notice that it is possible to have a weak solution of an SDE, but the strong solution of that SDE may not exist.

2.4.9 Theorem. [$Theorem$ *5.11, Klebaner (2012)*] *If $\sigma(x,t)$ is positive and continuous, and for any $T > 0$ there is K_T, such that for all $x \in \mathbb{R}$,*

$$|b(x,t) + \sigma(x,t)| \leq K_T(1 + |x|),$$

then there exists a unique weak solution to SDE given by the equation (2.4.1) starting at any point $x \in \mathbb{R}$ at any times $s \geq 0$.

3. Stochastic optimal control.

A stochastic optimal control problem is a type of optimization problem where the system dynamics are influenced by random or stochastic factors. In a stochastic optimal control problem, both the state dynamics and possibly the cost or objective function involve uncertainty or randomness. The goal is to find a control policy that minimizes or maximizes some measure of performance over time, taking into account the stochastic nature of the system. The stochastic maximum principle extends the classical maximum principle to stochastic optimal control problems, providing necessary and sufficient conditions for optimality. It states that optimal control strategies are associated with adjoint processes satisfying a system of differential equations known as the adjoint equations, along with a Hamiltonian function. These equations capture how changes in the objective function affect the optimal control strategy, and they are coupled with the system dynamics. This section will begin by investigating the optimal control problem, elucidating its key concepts and relevance. After that, we will delve into the statement of the stochastic maximum principle, a fundamental theorem in stochastic optimal control theory. Finally, we will proceed to formulate the problem, laying the groundwork for detailed analysis and understanding.

3.1 Stochastic optimal control problem.

Let $T > 0$ and $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ be a filtered probability space, where $W : [0, T] \times \Omega \to \mathbb{R}^d$ is a $d-$dimensional standard $\mathbb{F}-$ Brownian motion on $(\Omega, \mathcal{F}, \mathbb{P})$, $\mathbb{F} = \{\mathcal{F}\}_{0 \leq t \leq T}$ is the natural filtration generated by the Brownian motion W. Suppose that $(\Omega, \mathcal{F}, \mathbb{P})$ is complete, $\mathcal{F}_0$ is contains all the $\mathbb{P}-$null sets in $\mathcal{F}$. All definitions presented in this section can be found in A. Bensoussan (1981) .

3.1.1 Definition. A measurable map $u(.) : [0, T] \to U \subset \mathbb{R}^k$ is called a **control**.

3.1.2 Definition. A controlled stochastic differential equation is defined by the following expression:

$$\begin{cases} dx_t = b(t, x_t, u_t)dt + \sigma(t, x_t, u_t)dW_t \\ x(0) = x_0, \end{cases} \tag{3.1.1}$$

where b and σ are given function.

3.1.3 Definition. A **state trajectory** corresponding to $u(.)$ is $x(.)$, which is the solution of the equation (3.1.1) in the Definition 3.1.2.

3.1.4 Definition. A control $u(.)$ is called **admissible control**, and $(x(.), u(.))$ is called an **admissible pair** if the following conditions hold:

$$\begin{aligned} &(i) \quad u : [0, T] \times \Omega \to U \text{ such that } u \in L^2_{\mathcal{F}}(0, T, \mathbb{R}^k) \\ &(ii) \quad x(.) \quad \text{is the unique solution of the equation 3.1.1 under } u(.), \end{aligned}$$

where

$$L^2_{\mathcal{F}}(0, T, \mathbb{R}^k) := \left\{ x : [0, T] \times \Omega \to \mathbb{R}^k | x \text{ is } \mathbb{F}-\text{adapted and } \mathbb{E}\left[\int_0^T |x_t|^2 dt\right] < \infty \right\}.$$

We denote by $\mathcal{U}_{ad}\,[0, T]$ the set of all admissible controls.

The expression defining a cost functional, which evaluates the performance of the control, is given by:

$$J(u(.)) = \mathbb{E}\left\{\int_0^T f(t, x_t, u_t)dt + h(x_T)\right\}, \tag{3.1.2}$$

where $f : [0, T] \times \mathbb{R}^n \times U \to \mathbb{R}$ is integrable and $h : \mathbb{R}^n \to \mathbb{R}$ is measurable.

Our **stochastic optimal control problem** is formulated as the minimization of the equation (3.1.2) over $\mathcal{U}_{ad}[0, T]$. More precisely, the goal is to find $u^*(.) \in \mathcal{U}_{ad}[0, T]$ (if it exist), such that

$$J(u^*(.)) = \inf_{u(.)\in\mathcal{U}_{ad}[0,T]} \mathbb{E}\left\{\int_0^T f(t, x_t, u_t)dt + h(x_T)\right\}. \tag{3.1.3}$$

3.1.5 Definition. Any $u^*(.) \in \mathcal{U}_{ad}[0, T]$ satisfying (3.1.3) is called an **optimal control**. The corresponding state process $x^*(.)$ is an optimal state process, and the pair $(x^*(.), u^*(.))$ constitutes an optimal pair.

3.2 Stochastic maximum principle.

The present section focuses on Pontryagin's Maximum Principle, as detailed in Yong (1999). Before delving deeper into Pontryagin's Maximum Principle, let us establish the following assumptions:

Assumption 1 : (i) The maps b, σ, f and h are measurable, and there exist a constant $L > 0$ and a modulus of continuity $\bar{\omega} : [0, \infty) \to [0, \infty)$ such that for $\phi(t, x, u) = b(t, x, u)$, $\sigma(t, x, u)$, $f(t, x, u)$, $h(x)$, we have

$$\begin{cases} |\phi(t, x, u) - \phi(t, \hat{x}, \hat{u})| \leq L|x - \hat{x}| + \bar{\omega}\|u - \hat{u}\|, \\ \qquad\qquad \text{for all } t \in [0, T], x, \hat{x} \in \mathbb{R}^n, \quad u, \hat{u} \in U, \\ |\phi(t, 0, u)| \leq L, \quad \text{for all } t \in [0, T], u \in U. \end{cases} \tag{3.2.1}$$

(ii) The maps b, σ, f and h are C^2 with respect to x. Moreover, there exist a constant $L > 0$ and modulus of continuity $\bar{\omega} : [0, \infty) \to [0, \infty)$ such that for $\phi(t, x, u) = b(t, x, u)$, $\sigma(t, x, u)$, $f(t, x, u)$, $h(x)$, we have

$$\begin{cases} |\phi_x(t, x, u) - \phi_x(t, \hat{x}, \hat{u})| \leq L|x - \hat{x}| + \bar{\omega}\|u - \hat{u}\|, \\ |\phi_{xx}(t, x, u) - \phi_{xx}(t, \hat{x}, \hat{u})| \leq \bar{\omega}\left(|x - \hat{x}| + \|u - \hat{u}\|\right), \\ \qquad \text{for all } t \in [0, T], x, \hat{x} \in \mathbb{R}^n, \quad u, \hat{u} \in U. \end{cases} \tag{3.2.2}$$

Assumption 2: The control domain U is a convex body in $\mathbb{R}^k$. The maps b, σ and f are locally Lipschitz in u, and their derivatives with respect to x are continuous in (x, u).

The following is the well-known Pontryagin's maximum principle, which gives a set of first order necessary conditions for optimal pairs.

3.2.1 Theorem. $[Theorem$ *2.1, Yong (1999)* $]$ *Suppose Assumption 1 and Assumption 2 holds. Let* $(x^*(.), u^*(.))$ *be an optimal pair of the problem 3.1.3. Then there exist a function* $p^*(.) : [0, T] \to \mathbb{R}^n$ *and* $q^*(.) : [0, T] \to \mathbb{R}^{n\times d}$ *satisfying the following:*

$$\begin{cases} dp_t^* = -\left\{b_x(t, x_t^*, u_t^*)^T p_t^* + \sum_{j=1}^d \sigma_x^j(t, x_t^*, u_t^*)^T q_{jt}^* - f_x(t, x_t^*, u_t^*)\right\} dt + q_t^* dW_t, \\ p_T^* = -h_x(x_T^*), \quad t \in [0, T], \end{cases} \tag{3.2.3}$$

and

$$H(t, x_t^*, u_t^*, p_t^*, q_t^*) = \max_{u \in U} H(t, x_t^*, u, p_t^*, q_t^*), \tag{3.2.4}$$

where

$$H(t, x_t^*, u_t^*, p_t^*, q_t^*) = \langle p_t^*, b(t, x_t, u_t) \rangle + tr\left[q_t^{*T} \sigma(t, x_t^*, u_t^*)\right] - f(t, x_t^*, u_t^*), \tag{3.2.5}$$

with $(t, x_t^*, u_t^*, p_t^*, q_t^*) \in [0, T] \times \mathbb{R}^n \times U \times \mathbb{R}^n \times \mathbb{R}^{n \times d}$.

Within Theorem 3.2.1, $p^*(.)$ and $q^*(.)$ are called adjoint variables. The equation (3.2.3) is known as the adjoint equation corresponding to the pair $(x^*(.), u^*(.))$. The equation (3.2.4) represents the maximum condition, while the equation (3.2.5) is denoted as the Hamiltonian.

Proof. We first prove the existence of $p^*(.)$ and $q^*(.)$. Let us consider the following stochastic differential equation (SDE):

$$\begin{cases} dp_t^* = -\left\{b_x(t, x_t^*, u_t^*)^T p_t^* + \sum_{j=1}^d \sigma_x^j(t, x_t^*, u_t^*)^T q_{jt}^* - f_x(t, x_t^*, u_t^*)\right\} dt + q_t^* dW_t, \\ p_T^* = -h_x(x_T^*), \quad t \in [0, T], \end{cases} \tag{3.2.6}$$

where b_x, f_x, h_x and σ_x satisfy the Assumption 1. Then, according to the Theorem 2.4.6, there exists a unique adapted solution $(p^*(.), q^*(.))$ to the SDE (3.2.6).

Next, we prove the existence of the maximum of the Hamiltonian function.

Let $u_t^\epsilon = (1-\epsilon)u_t^* + \epsilon u_t$ for $u_t \in U$ and $\epsilon \in (0, 1)$. Note that $u_t^\epsilon \in U$ since U is convex. Define $x^\epsilon(.)$ as the state process corresponding to the control $u^\epsilon(.)$.

Consider the Taylor expansions:

$$\begin{aligned} b(t, x_t^\epsilon, u_t^\epsilon) &= b(t, x_t^*, u_t^*) + b_x(t, x_t^*, u_t^*)(x_t^\epsilon - x_t^*) + \epsilon b_u(t, x_t^*, u_t^*)(u_t - u_t^*) + o(|x_t^\epsilon - x_t^*| + \epsilon|u_t - u_t^*|), \\ \sigma(t, x_t^\epsilon, u_t^\epsilon) &= \sigma(t, x_t^*, u_t^*) + \sigma_x(t, x_t^*, u_t^*)(x_t^\epsilon - x_t^*) + \epsilon \sigma_u(t, x_t^*, u_t^*)(u_t - u_t^*) + o(|x_t^\epsilon - x_t^*| + \epsilon|u_t - u_t^*|). \end{aligned}$$

Let $y_t^\epsilon = x_t^\epsilon - x_t^*$. Then,

$$u_t^\epsilon - u_t^* = \epsilon(u_t - u_t^*),$$

and we can define the SDE for y_t^ϵ as:

$$\begin{aligned} dy_t^\epsilon = &\left[b_x(t, x_t^*, u_t^*) y_t^\epsilon + \epsilon b_u(t, x_t^*, u_t^*)(u_t - u_t^*)\right] dt \\ &+ \left[\sigma_x(t, x_t^*, u_t^*) y_t^\epsilon + \epsilon \sigma_u(t, x_t^*, u_t^*)(u_t - u_t^*)\right] dW_t, \end{aligned}$$

with initial condition $y_0^\epsilon = 0$.

Next, the difference between the cost functional is:

$$J(u^\epsilon(.)) - J(u^*(.)) = \mathbb{E}\left[\int_0^T \left(f(t, x_t^\epsilon, u_t^\epsilon) - f(t, x_t^*, u_t^*)\right) dt + h(x_T^\epsilon) - h(x_T^*)\right]. \tag{3.2.7}$$

Using Taylor expansions for h and f,

$$\begin{aligned} h(x_T^\epsilon) &= h(x_T^*) + h_x(x_T^*)(x_T^\epsilon - x_T^*) + o(|x_T^\epsilon - x_T^*|), \\ f(t, x_t^\epsilon, u_t^\epsilon) &= f(t, x_t^*, u_t^*) + f_x(t, x_t^*, u_t^*)(x_t^\epsilon - x_t^*) + \epsilon f_u(t, x_t^*, u_t^*)(u_t - u_t^*) + o(|x_t^\epsilon - x_t^*| + \epsilon|u_t - u_t^*|). \end{aligned}$$

Thus,

$$J(u^\epsilon(.)) - J(u^*(.)) = \mathbb{E}\left[h_x(x_T^*)y_T^\epsilon + \int_0^T (f_x(t,x_t^*,u_t^*)y_t^\epsilon + \epsilon f_u(t,x_t^*,u_t^*)(u_t - u_t^*))\,dt\right] + o(\epsilon).$$

Using Itô's lemma that we have seen in equation (2.3.7) for p_t^* and y_t^ϵ, then

$$d\langle p_t^*, y_t^\epsilon\rangle = \langle dp_t^*, y_t^\epsilon\rangle + \langle p_t^*, dy_t^\epsilon\rangle + \langle dp_t^*, dy_t^\epsilon\rangle. \tag{3.2.8}$$

Substituting the expressions for dp_t^* and dy_t^ϵ for the equation (3.2.8), we obtain

$$\begin{aligned} d\langle p_t^*, y_t^\epsilon\rangle = &\left\langle -\left(b_x(t,x_t^*,u_t^*)^T p_t^* + \sum_{j=1}^d \sigma_x^j(t,x_t^*,u_t^*)^T q_{jt}^* - f_x(t,x_t^*,u_t^*)\right)dt + q_t^* dW_t, y_t^\epsilon\right\rangle \\ &+ \langle p_t^*, (b_x(t,x_t^*,u_t^*)y_t^\epsilon + \epsilon b_u(t,x_t^*,u_t^*)(u_t - u_t^*))\,dt\rangle \\ &+ \langle p_t^*, (\sigma_x(t,x_t^*,u_t^*)y_t^\epsilon + \epsilon\sigma_u(t,x_t^*,u_t^*)(u_t - u_t^*))\,dW_t\rangle \\ &+ \langle q_t^*, (\sigma_x(t,x_t^*,u_t^*)y_t^\epsilon + \epsilon\sigma_u(t,x_t^*,u_t^*)(u_t - u_t^*))\rangle\,dt. \end{aligned}$$

Computing expectations and integrating from 0 to T, we get

$$\begin{aligned} \mathbb{E}\left[\langle p_T^*, y_T^\epsilon\rangle\right] = &\,\mathbb{E}\left[\int_0^T \langle p_t^*, \epsilon b_u(t,x_t^*,u_t^*)(u_t - u_t^*)\rangle\,dt\right] + \mathbb{E}\left[\int_0^T \langle q_t^*, \epsilon\sigma_u(t,x_t^*,u_t^*)(u_t - u_t^*)\rangle\,dt\right] \\ &+ \mathbb{E}\left[\int_0^T \{\langle p_t^*, b_x(t,x_t^*,u_t^*)y_t^\epsilon\rangle + \langle q_t^*, \sigma_x(t,x_t^*,u_t^*)y_t^\epsilon\rangle\}\,dt\right] \\ &+ \mathbb{E}\left[\int_0^T \left\langle -f_x(t,x_t^*,u_t^*) + \left(b_x(t,x_t^*,u_t^*)^T p_t^* + \sum_{j=1}^d \sigma_x^j(t,x_t^*,u_t^*)^T q_{jt}^*\right), y_t^\epsilon\right\rangle dt\right] \\ &+ \mathbb{E}\left[\langle p_O^*, y_0^*\rangle\right]. \end{aligned}$$

Substituting $-h_x(x_T^*) = p_T^*$, $y_0^* = 0$ and rearranging terms, we obtain

$$\begin{aligned} \mathbb{E}\left[\langle -h_x(x_T^*), y_T^\epsilon\rangle\right] = &\,\epsilon\mathbb{E}\left[\int_0^T \langle p_t^*, b_u(t,x_t^*,u_t^*)(u_t - u_t^*)\rangle\,dt\right] + \epsilon\mathbb{E}\left[\int_0^T \langle q_t^*, \sigma_u(t,x_t^*,u_t^*)(u_t - u_t^*)\rangle\,dt\right] \\ &+ \mathbb{E}\left[\int_0^T \left\langle -f_x(t,x_t^*,u_t^*) + b_x(t,x_t^*,u_t^*)^T p_t^* + \sum_{j=1}^d \sigma_x^j(t,x_t^*,u_t^*)^T q_{jt}^*, y_t^\epsilon\right\rangle dt\right]. \end{aligned}$$

Therefore,

$$\begin{aligned} J(u^\epsilon(.)) - J(u^*(.)) = &-\epsilon\mathbb{E}\left[\int_0^T \langle p_t^*, b_u(t,x_t^*,u_t^*)(u_t - u_t^*)\rangle\,dt\right] - \epsilon\mathbb{E}\left[\int_0^T \langle q_t^*, \sigma_u(t,x_t^*,u_t^*)(u_t - u_t^*)\rangle\,dt\right] \\ &+ \epsilon\mathbb{E}\left[\int_0^T (f_u(t,x_t^*,u_t^*)(u_t - u_t^*))\,dt\right] + o(\epsilon). \end{aligned}$$

Using the formality of the function H, we obtain:

$$J(u^\epsilon(.)) - J(u^*(.)) = -\epsilon\mathbb{E}\left[\int_0^T (H(t,x_t^*,u_t,p_t^*,q_t^*) - H(t,x_t^*,u_t^*,p_t^*,q_t^*))\,dt\right] + o(\epsilon).$$

By optimality of u^*, we have $J(u^\epsilon(.)) - J(u^*(.)) \geq 0$. Then,

$$\epsilon \mathbb{E}\left[\int_0^T \left(H(t, x_t^*, u_t, p_t^*, q_t^*) - H(t, x_t^*, u_t^*, p_t^*, q_t^*)\right) dt\right] \leq o(\epsilon).$$

Thus,

$$H(t, x_t^*, u_t^*, p_t^*, q_t^*) = \max_{u \in U} H(t, x_t^*, u, p_t^*, q_t^*).$$

This completes the proof of Theorem 3.2.1. □

3.2.2 Definition. Any pair of $(p^*(.), q^*(.))$ satisfying (3.2.3) is called an **adapted solution** of (3.2.3).

3.2.3 Proposition. $[Lemma$ 2.3, Yong (1999)$]$ Suppose $\phi : G \subset \mathbb{R}^n \to \mathbb{R}$ an function that satisfies the Assumption 2. For any $x \in G$, if ϕ attains a local minimum or maximum at x, then $0 \in \partial\phi$, where

$$\partial\phi(x) = \left\{\xi \in \mathbb{R}^n \middle| \langle \xi, y\rangle \leq \lim_{t \to 0, z \in G, z \to x} \frac{\phi(z + ty) - \phi(z)}{t}\right\}.$$

3.2.4 Proposition. $[Lemma$ 2.4, Yong (1999)$]$ Let ϕ be a concave or convex function on $\mathbb{R}^n \times U$, where $U \subset \mathbb{R}^k$ is a convex body. Assume that $\phi(x, u)$ is differentiable in x and $\phi_x(x, u)$ is continue in (x, u). Then

$$\{(\phi_x(x^*, u^*), r)\, | r \in \partial_u \phi(x^*, u^*)\} \subseteq \partial_{x,u}(x^*, u^*),$$

for all $(x^*, u^*) \in \mathbb{R}^n \times U$.

3.2.5 Definition. If $(x^*(.), u^*(.))$ is an optimal pair and $(p^*(.), q^*(.))$ is an adapted solution of (3.2.3), then $(x^*(.), u^*(.), p^*(.), q^*(.))$ is called an **optimal 4-tuple** (admissible 4-tuple).

3.2.6 Proposition. $[Lemma$ 5.1, Yong (1999)$]$ Let the control U be a convex body in $\mathbb{R}^n$. Let $(x^*(.), u^*(.), p^*(.), u^*(.))$ be an admissible $4-$tuple, and let H be the corresponding $H-$function. Then, for any $t \in [0, T]$,

$$\partial_u H(t, x^*(t), u^*(t), p^*(t), q^*(t)) = \partial_u H(t, x^*(t), u^*(t)).$$

According to Theorem 5.2 in Yong (1999), the following are sufficient conditions for the stochastic maximum principle.

3.2.7 Theorem. *Let Assumption 1 and Assumption 2 hold. Let $(x^*(.), u^*(.), p^*(.), q^*(.))$ be an admissible 4-tuple. Suppose that $h(.)$ is convex, and $H(t, ., ., p_t^*, q_t^*)$ defined by*

$$H(t, x, u, p, q) \equiv \langle p, b(t, x, u)\rangle + tr\left[q^T \sigma(t, x, u)\right] - f(t, x, u), \tag{3.2.9}$$

with $(t, x, u, p, q) \in [0, T] \times \mathbb{R}^n \times U \times \mathbb{R}^n \times \mathbb{R}^{n \times d}$, is concave for all $t \in [0, T]$ almost surely and

$$H(t, x_t^*, u_t^*, p_t^*, q_t^*) = \max_{u \in U} H(t, x_t^*, u, p_t^*, q_t^*), \tag{3.2.10}$$

holds. Then, $(x^*(.), u^*(.))$ is an optimal pair of (3.1.3).

Proof. Since Assumption 1 and Assumption 2 hold, and $(x^*(.), u^*(.), p^*(.), q^*(.))$ is an admissible $4-$tuple, then the Theorem 2.4.9 tells us that $x^*(.)$ is unique. According to the Theorem 3.2.1 the function $H(t, ., ., p_t^*, q_t^*)$ and the adjoint variable $p^*(.)$ and $q^*(.)$ are well defined. By applying Proposition 3.2.6 and Proposition 3.2.3, and combining them with equation (3.2.10), we have

$$0 \in \partial_u H(t, x^*(t), u^*(t)) = \partial_u H(t, x^*(t), u^*(t), p^*(t), q^*(t)).$$

Using the Proposition 3.2.4, we conclude that

$$(H_x(t, x_t^*, u_t^*, p_t^*, q_t^*), 0) \in \partial_{x,u} H(t, x_t^*, u_t^*, p_t^*, q_t^*).$$

Since $H(t, ., ., p_t^*, q_t^*)$ is concave, then we obtain

$$H(t, x_t, u_t, p_t^*, q_t^*) - H(t, x_t^*, u_t^*, p_t^*, q_t^*) \leq \langle H_x(t, x_t^*, u_t^*, p_t^*, q_t^*), x_t - x_t^* \rangle.$$

Therefore, we get

$$\int_0^T \{H(t, x_t, u_t, p_t^*, q_t^*) - H(t, x_t^*, u_t^*, p_t^*, q_t^*)\} \, dt \leq \int_0^T \langle H_x(t, x_t^*, u_t^*, p_t^*, q_t^*), x_t - x_t^* \rangle dt, \tag{3.2.11}$$

for any admissible pair $(x^*(.), u^*(.))$.

Let us define $\mathcal{E}(t) = x(t) - x^*(t)$, which satisfies the following equation

$$\begin{cases} d\mathcal{E}_t = \{b_x(t, x_t^*, u_t^*)\mathcal{E}_t + \alpha_t\} \, dt + \sum_{j=1}^m \left[\sigma_x^j(t, x_t^*, u_t^*)\mathcal{E}_t + \beta_t^j\right] dW_t^j \\ \mathcal{E}(0) = 0, \end{cases} \tag{3.2.12}$$

where

$$\begin{cases} \alpha_t = -b_x(t, x_t^*, u_t^*)\mathcal{E}_t + b(t, x_t, u_t) - b(t, x_t^*, u_t^*), \\ \beta_t^j = -\sigma_x^j(t, x_t^*, u_t^*)\mathcal{E}_t + \sigma^j(t, x_t, u_t) - \sigma^j(t, x_t^*, u_t^*), \quad 1 \leq j \leq m. \end{cases} \tag{3.2.13}$$

Let $u(.)$ and $u^*(.)$ belong U. Then, we have

$$\begin{aligned} J(u^*(.)) - J(u(.)) &= \mathbb{E}\left\{\int_0^T f(t, x_t^*, u_t^*)dt + h(x_t^*)\right\} - \mathbb{E}\left\{\int_0^T f(t, x_t, u_t)dt + h(x_t)\right\} \\ &= \mathbb{E}\left\{\int_0^T \{f(t, x_t^*, u_t^*) - f(t, x_t, u_t)\} \, dt\right\} + \mathbb{E}\left[h(x_T^*) - h(x_T)\right]. \end{aligned} \tag{3.2.14}$$

Denote that

$$\begin{aligned} I_1 &= \mathbb{E}\left\{\int_0^T \{f(t, x_t^*, u_t^*) - f(t, x_t, u_t)\} \, dt\right\}, \\ I_2 &= \mathbb{E}\left[h(x_T^*) - h(x_T)\right]. \end{aligned}$$

Since h is a convex function, then we have

$$\langle h_x(x_T^*), \mathcal{E}_T \rangle \leq h(x_T) - h(x_T^*).$$

This implies that

$$\mathbb{E}\left[\langle h_x(x_T^*), \mathcal{E}_T \rangle\right] \leq I_2. \tag{3.2.15}$$

Utilizing the duality relationship between the equation (3.2.12) and the equation (3.2.3), we have

$$\mathbb{E}\left[\langle h_x(x_T^*), \mathcal{E}_T\rangle\right] = -\mathbb{E}\left[\langle p_T^*, \mathcal{E}_T\rangle\right]$$

By using the equation (2.3.8) from Proposition 2.3.6 to compute the Itô formula for $\langle p_T^*, \mathcal{E}_T\rangle$, we obtain

$$\begin{aligned}\langle p_T^*, \mathcal{E}_T\rangle = &\int_0^t \left\{\langle p_t^*, b_x(t, x_t^*, u_t^*)\mathcal{E}_t + \alpha_t\rangle + \langle q_t^*, \sum_{j=1}^m \left[\sigma_x^j(t, x_t^*, u_t^*)\mathcal{E}_t + \beta_t^j\right]\rangle\right\} dt \\ &+ \int_0^t \left\{\langle -b_x(t, x_t^*, u_t^*)^T p_t^* - \sum_{j=1}^d \sigma_x^j(t, x_t^*, u_t^*)^T q_{jt}^* + f_x(t, x_t^*, u_t^*), \mathcal{E}_t\rangle\right\} dt \\ &+ \int_0^t \left\{\langle q_t^*, \mathcal{E}_t\rangle + \langle p_t^*, \sum_{j=1}^m \left[\sigma_x^j(t, x_t^*, u_t^*)\mathcal{E}_t + \beta_t^j\right]\rangle\right\} dW_t + \langle p^*(0), \mathcal{E}(0)\rangle.\end{aligned}$$

After completing this computation and using $\mathcal{E}(0) = 0$, then we have

$$\begin{aligned}\langle p_T^*, \mathcal{E}_T\rangle = &\int_0^T \left\{\langle f_x(t, x_t^*, u_t^*), \mathcal{E}_t\rangle + \langle p_t^*, \alpha_t\rangle + \sum_{j=1}^m \langle q^{*j}, \beta_t^j\rangle\right\} dt \\ &+ \int_0^T \left\{\langle q_t^*, \mathcal{E}_t\rangle + \langle p_t^*, \sum_{j=1}^m \left[\sigma_x^j(t, x_t^*, u_t^*)\mathcal{E}_t + \beta_t^j\right]\rangle\right\} dW_t.\end{aligned}$$

Since $W_t \sim \mathcal{N}(0,t)$, then we have $\mathbb{E}\left[W(T)\right] = \mathbb{E}\left[W(0)\right] = 0$. Therefore, we have the following equality

$$\mathbb{E}\left[\langle h_x(x_T^*), \mathcal{E}_T\rangle\right] = -\mathbb{E}\int_0^T \left\{\langle f_x(t, x_t^*, u_t^*), \mathcal{E}_t\rangle + \langle p_t^*, \alpha_t\rangle + \sum_{j=1}^m \langle q^{*j}, \beta_t^j\rangle\right\} dt.$$

The expression of f_x from the equation (3.2.9), can be written as

$$f_x(t, x^*, u^*) = -H_x(t, x^*, u^*, p, q) + \langle p, b_x(t, x^*, u^*)\rangle + tr\left[q^T \sigma_x(t, x^*, u^*)\right]. \tag{3.2.16}$$

From the expression of f_x in the equation (3.2.16), we obtain

$$\begin{aligned}\mathbb{E}\left[\langle h_x(x_T^*), \mathcal{E}_T\rangle\right] =&\mathbb{E}\int_0^T \langle H_x(t, x_t^*, u_t^*, p_t, q_t), \mathcal{E}_t\rangle dt - \mathbb{E}\int_0^T \{\langle p_t, b(t, x_t, u_t) - b(t, x_t^*, u_t^*)\rangle \\ &- \sum_{j=1}^m \langle q_t^j, \sigma^j(t, x_t, u_t) - \sigma^j(t, x_t^*, u_t^*)\rangle\} dt.\end{aligned}$$

By using the equation (3.2.11), then we get

$$\begin{aligned}\mathbb{E}\left[\langle h_x(x_T^*), \mathcal{E}_T\rangle\right] \geq &\mathbb{E}\int_0^T \{H(t, x_t, u_t, p_t, q_t) - H(t, x_t^*, u_t^*, p_t, q_t)\} dt \\ &- \mathbb{E}\int_0^T \left\{\langle p_t, b(t, x_t, u_t) - b(t, x_t^*, u_t^*)\rangle + \sum_0^T \langle q_t^j, \sigma^j(t, x_t, u_t) - \sigma^j(t, x_t^*, u_t^*)\rangle\right\} dt.\end{aligned}$$

Then, we obtain

$$\mathbb{E}\left[\langle h_x(x_T^*), \mathcal{E}_T\rangle\right] \geq -\mathbb{E}\int_0^T \{f(t, x_t, u_t) - f(t, x_t^*, u_t^*)\}\, dt = -I_1. \tag{3.2.17}$$

By combining the equations (3.2.17), (3.2.15), and (3.2.14), we obtain

$$J(u^*(.)) - J(u(.)) \leq 0, \quad \text{which implie that} \quad J(u^*(.)) \leq J(u(.)).$$

Finally, since $u(.) \in U$ is arbitrary, then we have

$$J(u^*(.)) = \inf_{u\in U} J(u(.)) = \inf_{u\in U} \mathbb{E}\left\{\int_0^T f(t, x_t, u_t)dt + h(x_t)\right\}.$$

□

Now, let $(x^*(.), u^*(.))$ be a given optimal pair. We introduce the adjoint BSDE as (3.2.3), where $p^*(.)$ and $q^*(.)$ two adapted processes that need to solved.

3.2.8 Remark. The partial differentials of the Hamiltonian H satisfy $b(t, x, u) = H_p(t, x, u, p, q)$ and $\sigma(t, x, u, p, q) = H_q(t, x, p, q)$, where

$$\frac{\partial H}{\partial_p} = H_p, \quad \frac{\partial H}{\partial_q} = H_q.$$

By applying Remark 3.2.8 and combining it with the equations (3.1.1), (3.2.3), and (3.2.10), we obtain the following:

$$\begin{cases} dx_t^* = b(t, x_t^*, u_t^*)dt + \sigma(t, x_t^*, u_t^*)dW_t, \\ dp_t^* = -H_x(t, x_t^*, u_t^*, p_t^*, q_t^*)dt + q_t^* dW_t, \\ x_0^* = x_0 \quad , \quad p_T^* = -h_x(x_T^*), \\ H(t, x_t^*, u_t^*, p_t^*, q_t^*) = \max_{u\in U} H(t, x_t^*, u, p_t^*, q_t^*), \end{cases} \tag{3.2.18}$$

where $(x^*(.), u^*(.), p^*(.), q^*(.))$ is a $4-$tuple which is a solution of the equation (3.2.18).

3.2.9 Definition. The equation (3.2.18) is called a **stochastic Hamiltonian system**, also referred to as a Forward-Backward Stochastic Differential Equation (FBSDE).

3.2.10 Remark. According to Theorem 3.2.7, Remark 3.2.8 and the equation (3.2.18), there exists a function $\bar{H} : [0, T] \times \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^{n\times d} \to \mathbb{R}$ such that

$$\bar{H}(t, x, p, q) = \max_{u\in U} H(t, x, u, p, q), \tag{3.2.19}$$

which is independent of the control u. The equation (3.2.19) is said to be a **maximum principle**.

Throughout this essay, we focus on studying problems with convex control domains that correspond to the first-order adjoint equations.

3.3 Problem formulation

As we aim to apply the stochastic maximum principle with a deep learning method, it is necessary to reformulate problem (3.1.3) as a new problem based on the stochastic maximum principle and its corresponding stochastic Hamiltonian system. Suppose that there exists a solution $(x^*(.), u^*(.), p^*(.), q^*(.))$ for the FBSDE which is the equation (3.2.18).

It is known, the FBSDE can be regarded as a stochastic optimal control problem, as described in [Ma (1999)]. Based on this idea, we have the following state equation with a maximum condition, which is equivalent to (3.2.18):

$$\begin{cases} d\bar{x}_t = b(t, \bar{x}_t, \bar{u}_t)dt + \sigma(t, \bar{x}_t, \bar{u}_t)dW_t, \\ d\bar{p}_t = -H_x(t, \bar{x}_t, \bar{u}_t, \bar{p}_t, \bar{q}_t)dt + \bar{q}_t dW_t, \\ \bar{x}_0 = x_0 \quad , \quad \bar{p}_0 = \bar{p}_0, \\ H(t, \bar{x}_t, \bar{u}_t, \bar{p}_t, \bar{q}_t) = \max_{u \in U} H(t, \bar{x}_t, u, \bar{p}_t, \bar{q}_t), \end{cases} \tag{3.3.1}$$

where $(\bar{p}_0, \bar{q})$ is the pair of control value in $\mathbb{R}^n \times \mathbb{R}^n$.

We assume that there exists a function $\hat{u}(t, x, p, q)$ such that $\bar{u}_t$ can be represented as

$$\bar{u}_t = \hat{u}(t, \bar{x}_t, \bar{p}_t, \bar{q}_t) = \arg\max_{u \in U} H(t, \bar{x}_y, u, \bar{p}_t, \bar{q}_t). \tag{3.3.2}$$

Now, we present the new variational problem, serving as a reformulation of control problem (3.1.3). Thus, its formulation is as follows:

$$\inf_{\bar{p}_0, \{\bar{q}_t\}_{0 \le t \le T}} \mathbb{E}\left[| - h_x(\bar{x}_T) - \bar{p}_T|^2 \right], \tag{3.3.3}$$

such that

$$\begin{aligned} \bar{x}_t &= x_0 + \int_0^t b(s, \bar{x}_s, \bar{u}_s)ds + \int_0^t \sigma(s, \bar{x}_s, \bar{u}_s)dW_s, \\ \bar{p}_t &= \bar{p}_0 - \int_0^t H_x(s, \bar{x}_s, \bar{p}_s, \bar{q}_s)ds + \int_0^t \bar{q}_s dW_s, \\ \bar{u}_t &= \arg\max_{u \in U} H(t, \bar{x}_y, u, \bar{p}_t, \bar{q}_t). \end{aligned}$$

Suppose that $\bar{H}$ in the equation (3.2.19) is differentiable with respect to x, p and q, we have

$$\begin{cases} \bar{H}_p(t, x, p, q) = H_p(t, x, u^*, p, q) = b(t, x, u^*), \\ \bar{H}_q(t, x, p, q) = H_q(t, x, u^*, p, q) = \sigma(t, x, u^*), \\ \bar{H}_x(t, x, p, q) = H_x(t, x, u^*, p, q), \\ u^* = \arg\max_{u \in U} H(t, x, u, p, q), \end{cases} \tag{3.3.4}$$

for any $(t, x, u, p, q) \in [0, T] \times \mathbb{R}^n \times U \times \mathbb{R}^n \times \mathbb{R}^{n \times d}$. Then the FBSDE (3.2.18) can be rewritten as

$$\begin{cases} dx_t^* = \bar{H}_p(t, x_t^*, p_t^*, q_t^*)dt + \bar{H}_q(t, x_t^*, p_t^*, q_t^*)dW_t, \\ dp_t^* = -\bar{H}_x(t, x_t^*, p_t^*, q_t^*)dt + q_t^* dW_t, \\ x_0^* = x_0, \quad p_T^* = -h_x(x_T^*), \end{cases} \tag{3.3.5}$$

which is a FBSDE without constraint. Therefore the problem (3.3.3) is equivalent to

$$\inf_{\bar{p}_0,\{\bar{q}_t\}_{0\le t\le T}} \mathbb{E}\left[|-h_x(\bar{x}_T)-\bar{p}_T|^2\right] \tag{3.3.6}$$

such that

$$\begin{aligned}
\bar{x}_t &= x_0 + \int_0^t \bar{H}_p(s,\bar{x}_s,\bar{p}_s,\bar{q}_s)ds + \int_0^t \bar{H}_q(s,\bar{x}_s,\bar{p}_s,\bar{q}_s)dW_s,\\
\bar{p}_t &= \bar{p}_0 - \int_0^t \bar{H}_x(s,\bar{x}_s,\bar{p}_s,\bar{q}_s)ds + \int_0^t \bar{q}_s dW_s.
\end{aligned}$$

3.3.1 Proposition. Let h be a function continuous with respect to x. The FBSDE equation (3.3.5) is solvable if and only if one can find an $p^*(.)\in\mathbb{R}$ and $u^*(.)\in\mathcal{U}[0,T]$ such that the equation (3.3.1) admits a strong solution $(x^*(.),p^*(.))$ satisfying $p^*(T)=-h(x^*(T))$.

3.3.2 Definition. The cost functional of stochastic optimal control problem (3.3.6) be expressed as

$$V(0,x_0) = \inf_{\bar{p}_0,\{\bar{q}_t\}_{0\le t\le T}} \mathbb{E}\left[|-h_x(\bar{x}_T)-\bar{p}_T|^2\right].$$

3.3.3 Theorem. [*Proposition 1.1, Ma (1999)*] *Assume $h_x(x)$ is continuous with respect to x. Suppose also that the map $\bar{H}$ defined by (3.2.19) is differentiable in x,p,q, and its derivatives are continuous in x,p,q, such that for $\phi=\bar{H}_x,\bar{H}_p,\bar{H}_q$, and there exists constant $C>0$,*

$$\begin{aligned}
|\phi(t,x,p,q)-\phi(t,x',p',q')| &\le C\left(|x-x'|+|p-p'|+|q-q'|\right),\\
|\phi(t,0,0,0)|,|\bar{H}_q(t,x,p,0)| &\le C,
\end{aligned}$$

for all $x',x,p,p'\in\mathbb{R}^n, q,q'\in\mathbb{R}^{n\times d}$. Then the FBSDE (3.3.5) is solvable over $[0,T]$ if and only if $V(0,x_0)=0$.

Proof. Firstly, suppose that the FBSD (3.3.5) is solvable over $[0,T]$. Then Proposition 3.3.1 tell us that there exits a solution $(x^*(.),p^*(.))$ satisfying $p^*(T)=-h_x(x^*(T))$. Then, we obtain

$$\begin{aligned}
-h_x(x^*(T))-p^*(T)=0 &\iff |-h_x(x^*(T))-p^*(T)|^2=0\\
&\iff \mathbb{E}\left[|-h_x(x^*(T))-p^*(T)|^2\right]=0\\
&\iff \inf_{\bar{p}_0,\{\bar{q}_t\}_{0\le t\le T}} \mathbb{E}\left[|-h_x(x^*(T))-p^*(T)|^2\right]=0.
\end{aligned}$$

Therefore the last equation above implies that $V(0,x_0)=0$.

Conversely, assume that $V(0,x_0)=0$, then

$$\begin{aligned}
\inf_{\bar{p}_0,\{\bar{q}_t\}_{0\le t\le T}} \mathbb{E}\left[|-h_x(x^*(T))-p^*(T)|^2\right]=0 &\iff \mathbb{E}\left[|-h_x(x^*(T))-p^*(T)|^2\right]=0,\\
&\qquad \text{for all } \bar{p}_0,\{\bar{q}_t\}_{0\le t\le T}\\
&\iff |-h_x(x^*(T))-p^*(T)|^2=0,\\
&\qquad \text{for all } \bar{p}_0,\{\bar{q}_t\}_{0\le t\le T}\\
&\iff -h_x(x^*(T))-p^*(T)=0,\\
&\qquad \text{for all } \bar{p}_0,\{\bar{q}_t\}_{0\le t\le T}.
\end{aligned}$$

Then we see $(x^*(.), p^*(.))$ satisfies $p^*(T) = -h_x(x^*(T))$. Since

$$\begin{aligned}|\phi(t,x,p,q) - \phi(t,x',p',q')| &\leq C\left(|x-x'| + |p-p'| + |q-q'|\right),\\ |\phi(t,0,0,0)|, |\bar{H}_q(t,x,p,0)| &\leq C,\end{aligned}$$

where $C > 0$, $\phi = \bar{H}_x$, $\bar{H}_p$, $\bar{H}_q$, then we can apply the Theorem 2.4.6 to conclude that $(x^*(.), p^*(.))$ is a strong solution of the equation (3.3.1). Therefore, the Proposition 3.3.1 allows us to say that the FBSDE (3.3.5) is solvable over $[0, T]$. □

3.3.4 Corollary. Suppose the assumptions in Theorem 3.3.3 and Theorem 3.2.7 hold. If there exists a solution $(\bar{x}_t, \bar{u}_t, \bar{p}_t, \bar{q}_t)_{0\leq t\leq T}$ of the equation (3.3.1) satisfying

$$\mathbb{E}\left[|-h_x(\bar{x}_t) - \bar{p}_T|^2\right] = 0, \tag{3.3.7}$$

then $(\bar{x}_t, \bar{u}_t, \bar{p}_t, \bar{q}_t)$ is a solution of the equation (3.2.18). Moreover, the cost functional (3.1.3) can be obtained by

$$J(u^*(.)) = J(\bar{u}(.)) = \mathbb{E}\left\{\int_0^T f(t, \bar{x}_t, \bar{u}_t)dt + h(\bar{x}_T)\right\}. \tag{3.3.8}$$

3.3.5 Remark. According to the Theorem 3.3.3 and the Corollary 3.3.4, we can prove that the optimal control $\bar{u}$ of (3.1.3) can be obtained when $\mathbb{E}\left[|-h_x(\bar{x}_t) - \bar{p}_T|^2\right] = 0$.

4. Numerical Algorithms and Results

In the previous chapter, we have introduced the sufficient conditions of the Stochastic Maximum Principle and reformulated the Hamiltonian system as a new variational problem. This current chapter focuses on the development of a numerical algorithm aimed at solving this variation problem through deep learning techniques. In this context, we showcase the effectiveness of neural networks in handling the inherent complexities of our problem, demonstrating their proficiency in generating solutions. Furthermore, we will introduce three distinct algorithms tailored to tackle this challenge [Shaolin Ji (7 September 2022)].

4.1 Numerical Algorithms

In this section, we present the three algorithms used to solve the new variational problem. The details of these algorithms can be found in [Shaolin Ji (7 September 2022)].

4.1.1 Algorithm 1: Numerical Algorithm with 1-NNet. Let π be a partition of the time interval, $0 = t_0 < t_1 < t_2 < \cdots < t_{N-1} < t_N = T$ of $[0, T]$. Define $\Delta t_i = t_{i+1} - t_i$ and $\Delta W_{t_i} = W_{t_i+1} - W_{t_i}$, where $W_{t_i} \sim \mathcal{N}(0, t_i)$, for $i = 0, 1, 2, ..., N-1$.

According this transformation above, the Euler scheme of the forward SDE given by the equation (3.3.1) can be written as:

$$\begin{cases} \bar{x}^{\pi}_{t_i+1} - \bar{x}^{\pi}_{t_i} = b(t_i, \bar{x}^{\pi}_{t_i}, \bar{u}^{\pi}_{t_i})\Delta t_i + \sigma(t_i, \bar{x}^{\pi}_{t_i}, \bar{u}^{\pi}_{t_i})\Delta W_{t_i}, \\ \bar{p}^{\pi}_{t_i+1} - \bar{p}^{\pi}_{t_i} = -H_x(t_i, \bar{x}^{\pi}_{t_i}, \bar{u}^{\pi}_{t_i}, \bar{p}^{\pi}_{t_i}, \bar{q}^{\pi}_{t_i}) + \bar{q}^{\pi}_{t_i}\Delta W_{t_i}, \\ \bar{x}^{\pi}_0 = x_0, \quad \bar{p}^{\pi}_0 = \bar{p}_0, \end{cases} \tag{4.1.1}$$

where the characteristic of Euler scheme can found in [Euler]. Combine the equation (4.1.13) with the maximum condition provided by equation (3.3.2) we can see that

$$\begin{cases} \bar{x}^{\pi}_{t_i+1} = \bar{x}^{\pi}_{t_i} + b(t_i, \bar{x}^{\pi}_{t_i}, \bar{u}^{\pi}_{t_i})\Delta t_i + \sigma(t_i, \bar{x}^{\pi}_{t_i}, \bar{u}^{\pi}_{t_i})\Delta W_{t_i}, \\ \bar{p}^{\pi}_{t_i+1} = \bar{p}^{\pi}_{t_i} - H_x(t_i, \bar{x}^{\pi}_{t_i}, \bar{u}^{\pi}_{t_i}, \bar{p}^{\pi}_{t_i}, \bar{q}^{\pi}_{t_i}) + \bar{q}^{\pi}_{t_i}\Delta W_{t_i}, \\ \bar{x}^{\pi}_0 = x_0, \quad \bar{p}^{\pi}_0 = \bar{p}_0, \\ \bar{u}^{\pi}_{t_i} = \arg\max_{u \in U} H(t_i, \bar{x}^{\pi}_{t_i}, u, \bar{p}^{\pi}_{t_i}, \bar{q}^{\pi}_{t_i}). \end{cases} \tag{4.1.2}$$

We have seen that $\bar{q}^{\pi}_{t_i}$ represents a new control problem in the equation (3.3.3). Let us assume $\bar{q}^{\pi}_{t_i}$ satisfy

$$\bar{q}^{\pi}_{t_i} = \phi^1(t_i, \bar{x}^{\pi}_{t_i}, \bar{u}^{\pi}_{t_i}, \bar{p}^{\pi}_{t_i}; \theta^1_{t_i}), \tag{4.1.3}$$

where ϕ^1 represent the feedback control $\bar{q}^{\pi}_{t_i}$ in term of the state, control and adjoint variable. Then, we conclude that $\bar{q}^{\pi}_{t_i}$ acts as a feedback control for the states $\bar{x}^{\pi}_{t_i}$, $\bar{p}^{\pi}_{t_i}$ and the control $\bar{u}^{\pi}_{t_i}$. It is important to note that $\bar{u}^{\pi}_{t_i}$ serves as the control for the original control problem (3.1.1), which must be determined using the maximum condition (3.3.2) within equation (3.3.1).

According to the equations (4.1.3) and (4.1.2), we can see that $\bar{q}^{\pi}_{t_i}$ and $\bar{u}^{\pi}_{t_i}$ are interdependent. Therefore, while the control $\bar{u}^{\pi}_{t_i}$ are determined by the maximization of the function Hamiltonian H, then there exist an unknown function ϕ which represent the feedback control $\bar{q}^{\pi}_{t_i}$ depend on t_i, $\bar{x}^{\pi}_{t_i}$, $\bar{p}^{\pi}_{t_i}$. In that case, we can write

$$\bar{q}^{\pi}_{t_i} = \phi(t_i, \bar{x}^{\pi}_{t_i}, \bar{p}^{\pi}_{t_i}; \theta^1_{t_i}), \tag{4.1.4}$$

where ϕ is a new unknown function.

Now, we develop a neural network (1-NNet) for simulating the feedback control $\bar{q}^{\pi}$. The network consists of five layers including one $(1+n+n)-$dimensional input layer, three $(10+n+n)-$dimensional hidden layers and a $(n \times d)-$dimensional output layer. All parameters of the network are represented as θ. We define the loss function as

$$loss = \frac{1}{M}\sum_{j=1}^{M}\left[| - h_x(\bar{x}_T^{\pi}) - \bar{p}_T^{\pi}|^2 \right], \tag{4.1.5}$$

where M is the number of sample.

For convenience, the time interval $[0, T]$ is partitioned evenly, i.e. $\Delta t_i = t_{i+1} - t_i = T/N$ for all $i = 0, 1, ..., N$. The pseudo-code for solving the stochastic optimal control problem is given as follow:

Algorithm 1 Numerical algorithm with 1-NNet

Require: The Brownian motion ΔW_{t_i}, initial parameters $(\theta^0, \bar{p}_0^{0,\pi})$, learning rate ν;
Ensure: The 4-tuple processes $(\bar{x}_{t_i}^{l,\pi}, \bar{u}_{t_i}^{l,\pi}, \bar{p}_{t_i}^{l,\pi}, \bar{q}_{t_i}^{l,\pi})$.

1: **for** $l = 0$ to maxstep **do**
2: $\quad \bar{x}_0^{l,\pi} = x_0$, $\bar{p}_0^{l,\pi} = \bar{p}_0^l$;
3: $\quad$ **for** $i = 0$ to $N-1$ **do**
4: $\quad\quad \bar{q}_{t_i}^{l,\pi} = \phi(t_i, \bar{x}_{t_i}^{l,\pi}, \bar{p}_{t_i}^{l,\pi}; \theta^l)$;
5: $\quad\quad \bar{u}_{t_i}^{l,\pi} = \arg\max_{u \in U} H(t_i, \bar{x}_{t_i}^{l,\pi}, u, \bar{p}_{t_i}^{l,\pi}, \bar{q}_{t_i}^{l,\pi})$;
6: $\quad\quad \bar{x}_{t_i+1}^{l,\pi} = \bar{x}_{t_i}^{l,\pi} + b(t_i, \bar{x}_{t_i}^{l,\pi}, \bar{u}_{t_i}^{l,\pi})\Delta t_i + \sigma(t_i, \bar{x}_{t_i}^{l,\pi}, \bar{u}_{t_i}^{l,\pi})\Delta W_{t_i}$;
7: $\quad\quad \bar{p}_{t_i+1}^{l,\pi} = \bar{p}_{t_i}^{l,\pi} - H_x(t_i, \bar{x}_{t_i}^{l,\pi}, \bar{u}_{t_i}^{l,\pi}, \bar{p}_{t_i}^{l,\pi}, \bar{q}_{t_i}^{l,\pi})\Delta t_i + \bar{q}_{t_i}^{l,\pi}\Delta W_{t_i}$;
8: $\quad$ **end for**
9: $\quad J(\bar{u}^{l,\pi}(.)) = \frac{1}{M}\sum_{j=1}^{M}\left[\frac{T}{N}\sum_{i=0}^{N-1} f(t_i, \bar{x}_{t_i}^{l,\pi}, \bar{u}_{t_i}^{l,\pi}) + h(\bar{x}_T^{l,\pi})\right]$;
10: $\quad loss = \frac{1}{M}\sum_{j=1}^{M}\left[| - h_x(\bar{x}_T^{l,\pi}) - \bar{p}_T^{l,\pi}|^2\right]$;
11: $\quad (\theta^{l+1}, \bar{p}^{l+1,\pi}) = (\theta^l, \bar{p}^{l,\pi}) - \nu\nabla loss.$
12: **end for**

According to Corollary 3.3.4, the quadruple $(\bar{x}_{t_i}^{l,\pi}, \bar{u}_{t_i}^{l,\pi}, \bar{p}_{t_i}^{l,\pi}, \bar{q}_{t_i}^{l,\pi})$ converges to the true solution $(\bar{x}_{t_i}, \bar{u}_{t_i}, \bar{p}_{t_i}, \bar{q}_{t_i})$ when the loss goes to zero.

In this algorithm, the solution of the maximum condition in line 5 is an extremum problem of multivariate functions, and in most cases, it has no analytical solution. This implies that we cannot obtain the explicit value of $\bar{u}_{t_i}^{l,\pi}$ to calculate the forward process. Which mean that we need to calculate the approximate value of $\bar{u}$. Then, $\bar{u}_i$ can be approximated through the Hamiltonian function H.

4.1.2 Algorithm 2: Numerical Algorithm with 2-NNet. In algorithm 1, we have seen the absence of an explicit solution for the maximum condition, necessitating the use of approximations. Now, we proceed to derive the approximation condition suitable for high-dimensional cases. To tackle this issue, we introduce a numerical algorithm utilizing two neural networks (2-NNet). In this case, we focus on a specific type of stochastic optimal control problem where the control domain $U = \mathbb{R}^k$, all coefficients are C^1 with respect to u, and the optimal control $\bar{u}$ lies within the boundary of the control domain. Under these conditions, the equation

(3.2.10) implies that

$$H_u(t, x_t^*, u_t^*, p_t^*, q_t^*) = 0, \quad \text{for all } u \in U, \text{ for all } t \in [0, T], \quad a.s. \tag{4.1.6}$$

Thus, the corresponding stochastic Hamiltonian system (3.2.18) can be represented as

$$\begin{cases} dx_t^* = b(t, x_t^*, u_t^*)dt + \sigma(t, x_t^*, u_t^*)dW_t, \\ dp_t^* = -H_x(t, x_t^*, u_t^*, p_t^*, q_t^*)dt + q_t^* dW_t, \\ x_0^* = x_0 \quad , \quad p_T^* = -h_x(x_T^*), \\ H_u(t, x_t^*, u_t^*, p_t^*, q_t^*) = 0, \quad \text{for all } t \in [0, T]. \end{cases} \tag{4.1.7}$$

The work of Bismut stated that the Hamiltonian system $H_u(t, x_t^*, u_t^*, p_t^*, q_t^*) = 0$ corresponds to a wide range of stochastic optimal control problems [Bismut (2021), Bismut (2014)]. According to Corollary 3.3.4 in order to solve the Hamiltonian system (4.1.7), we can it into the following new control problem

$$\inf_{\bar{p}_0, \{\bar{q}_t\}_{0\le t\le T}, \{\bar{u}_t\}_{0\le t\le T}} \mathbb{E}\left[|-h_x(\bar{x}_T) - \bar{p}_T|^2 + \lambda \int_0^T H_u(t, \bar{x}_t, \bar{u}_t, \bar{p}_t, \bar{q}_t)^2 dt \right], \tag{4.1.8}$$

such that

$$\begin{aligned} \bar{x}_t &= x_0 + \int_0^t b(s, \bar{x}_s, \bar{u}_s)ds + \int_0^t \sigma(s, \bar{x}_s, \bar{u}_s)dW_s, \\ \bar{p}_t &= \bar{p}_0 - \int_0^t H_x(s, \bar{x}_s, \bar{p}_s, \bar{q}_s)ds + \int_0^t \bar{q}_s dW_s, \end{aligned}$$

where $\bar{p}_0$, $\{\bar{q}_t\}_{0\le t\le T}$, $\{\bar{u}_t\}_{0\le t\le T}$ are the controls and λ is the hyper-parameter. For the same raison as in Corollary 3.3.4, if the cost functional (4.1.8) converge to zero, then the $4-tuple$ $(\bar{x}_t, \bar{u}_t, \bar{p}_t, \bar{q}_t)$ will converge to $(x_t^*, u_t^*, p_t^*, q_t^*)$.

Then, the Euler scheme is given as

$$\begin{cases} \bar{x}_{t_i+1}^\pi = \bar{x}_{t_i}^\pi + b(t_i, \bar{x}_{t_i}^\pi, \bar{u}_{t_i}^\pi)\Delta t_i + \sigma(t_i, \bar{x}_{t_i}^\pi, \bar{u}_{t_i}^\pi)\Delta W_{t_i}, \\ \bar{p}_{t_i+1}^\pi = \bar{p}_{t_i}^\pi - H_x(t_i, \bar{x}_{t_i}^\pi, \bar{u}_{t_i}^\pi, \bar{p}_{t_i}^\pi, \bar{q}_{t_i}^\pi) + \bar{q}_{t_i}^\pi \Delta W_{t_i}, \\ \bar{x}_0^\pi = x_0, \quad \bar{p}_0^\pi = \bar{p}_0, \end{cases} \tag{4.1.9}$$

and the loss function is expressed as

$$loss = \frac{1}{M}\sum_{j=1}^{M}\left[|-h_x(\bar{x}_T^\pi) - \bar{p}_T^\pi|^2 + \lambda \sum_{i=0}^{N-1} H_u(t_i, \bar{x}_{t_i}^\pi, \bar{u}_{t_i}^\pi, \bar{p}_{t_i}^\pi, \bar{q}_{t_i}^\pi)^2 \right], \tag{4.1.10}$$

where M is the number of sample, and the time-divided coefficient T/N is merged into the coefficient λ. Algorithm 2 provides the pseudo-code. In practical computations, we determine the hyper-parameter λ through trial and error.

Differing from Algorithm 1, we consider the two processes $\{\bar{u}_{t_i}^\pi, \bar{q}_{t_i}^\pi\}_{0\le i\le N-1}$ as controls. This implies that two neural networks (2-NNets) should be developed to simulate $\bar{u}_{t_i}^\pi$ and $\bar{q}_{t_i}^\pi$ respectively. Similarly to Algorithm 1, we regard $\bar{u}_{t_i}^\pi$ and $\bar{q}_{t_i}^\pi$ as feedback controls of the state $\bar{x}_{t_i}^\pi$ the time t_i. Therefore, we construct a common neural network for all time steps.

$$\bar{q}_{t_i}^\pi = \phi^1(t_i, \bar{x}_{t_i}^\pi; \theta_{t_i}^q), \tag{4.1.11}$$
$$\bar{u}_{t_i}^\pi = \phi^2(t_i, \bar{x}_{t_i}^\pi; \theta_{t_i}^u) \tag{4.1.12}$$

Algorithm 2 Numerical algorithm with 2-NNets

Require: The Brownian motion ΔW_{t_i}, initial parameters $(\theta^{q,0}, \theta^{u,0}, \bar{p}_0^{0,\pi})$, learning rate ν, and hyper-parameter λ;

Ensure: The processes $\bar{x}_{t_i}^{l,\pi}$ and $\bar{p}_T^{l,\pi}$.

for $l = 0$ to maxstep **do**

  $\bar{x}_0^{l,\pi} = x_0$, $\bar{p}_0^{l,\pi} = \bar{p}_0^l$, and $H = 0$;

  for $i = 0$ to $N-1$ **do**

    $\bar{q}_{t_i}^{l,\pi} = \phi^1(t_i, \bar{x}_{t_i}^{l,\pi}; \theta^{q,l})$;

    $\bar{u}_{t_i}^{l,\pi} = \phi^2(t_i, \bar{x}_{t_i}^{l,\pi}; \theta^{u,l})$;

    $\bar{x}_{t_i+1}^{l,\pi} = \bar{x}_{t_i}^{l,\pi} + b(t_i, \bar{x}_{t_i}^{l,\pi}, \bar{u}_{t_i}^{l,\pi})\Delta t_i + \sigma(t_i, \bar{x}_{t_i}^{l,\pi}, \bar{u}_{t_i}^{l,\pi})\Delta W_{t_i}$;

    $\bar{p}_{t_i+1}^{l,\pi} = \bar{p}_{t_i}^{l,\pi} - H_x(t_i, \bar{x}_{t_i}^{l,\pi}, \bar{u}_{t_i}^{l,\pi}, \bar{p}_{t_i}^{l,\pi}, \bar{q}_{t_i}^{l,\pi})\Delta t_i + \bar{q}_{t_i}^{l,\pi}\Delta W_{t_i}$;

    $H = H + H_u(t_i, \bar{x}_{t_i}^{l,\pi}, \bar{u}_{t_i}^{l,\pi}, \bar{p}_{t_i}^{l,\pi}, \bar{q}_{t_i}^{l,\pi})^2$

  end for

  $J(\bar{u}^{l,\pi}(.)) = \frac{1}{M}\sum_{j=1}^{M}\left[\frac{T}{N}\sum_{i=0}^{N-1} f(t_i, \bar{x}_{t_i}^{l,\pi}, \bar{u}_{t_i}^{l,\pi}) + h(\bar{x}_T^{l,\pi})\right]$;

  $loss = \frac{1}{M}\sum_{j=1}^{M}\left[| - h_x(\bar{x}_T^{l,\pi}) - \bar{p}_T^{l,\pi}|^2 + \lambda H\right]$;

  $(\theta^{q,l+1}, \theta^{u,l+1}, \bar{p}^{l+1,\pi}) = (\theta^{q,l}, \theta^{u,l}, \bar{p}^{l,\pi}) - \nu\nabla loss.$

end for

The $2-$NNets contain both one $(n+1)-$dimensional input layers and three $(n+10)-$dimensional hidden layers, the output layers are $(n \times d)-$dimensional for $\bar{q}_{t_i}^{\pi}$ and $K-$dimensional for $\bar{u}_{t_i}^{\pi}$ respectively. The loss function is given as (4.1.10).

In Algorithm 2, instead of solving the maximum condition explicitly or approximately, we only focus on the consider the control condition $H_u(t, x, u, p, q) = 0$. Thus, Algorithm 2 can effectively handle a broad range of high-dimensional problems, even when the optimal control $\bar{u}$ cannot be explicitly determined.

4.1.3 Algorithm 3: Numerical Algorithm with Explicit Expression of $\bar{\mathbf{H}}$. The Algorithm 2 mentioned above provides a method for solving a general kind of high- dimensional stochastic optimal control problems, particularly when the optimal control $\bar{u}$ has not an explicit solution. In this subsection, we introduce another algorithm for solving high-dimensional cases. We have seen that as long as the function $\bar{H}$ is known, we can solve a class of high-dimensional stochastic optimal control problems through the deep-learning method, as stated in [Shaolin Ji and Zhang (2020)].

The pseudo-code of algorithm 3 is written as:

Now, we consider the stochastic optimal control problem (3.1.3) and assume that $\bar{H}$ defined in (3.2.19) is given. Then, the corresponding Hamiltonian system with $\bar{H}$ is given as follows

$$\begin{cases} dx_t^* = \bar{H}_p(t, x_t^*, p_t^*, q_t^*)dt + \bar{H}_q(t, x_t^*, p_t^*, q_t^*)dW_t, \\ dp_t^* = -\bar{H}_x(t, x_t^*, p_t^*, q_t^*)dt + q_t^* dW_t, \\ x_0^* = x_0, \quad p_T^* = -h_x(x_T^*), \end{cases} \tag{4.1.13}$$

which is essentially a fully coupled $FBSDE$. If (4.1.13) satisfies the monotonic conditions [Peng (1999)], we can solve it in high dimensions using the deep-learning method proposed in [Shaolin Ji and Zhang (2020)] for solving fully coupled $FBSDEs$. Then the optimal state processes $(x^*(.), p^*(.), q^*(.))$ can be obtained.

Algorithm 3 Numerical algorithm with explicit of $\bar{H}$

Require: The Brownian motion ΔW_{t_i}, initial parameters $(\theta^0, \bar{p}_0^{0,\pi})$, learning rate ν;
Ensure: The triple processes $(\bar{x}_{t_i}^{l,\pi}, \bar{p}_{t_i}^{l,\pi}, \bar{q}_{t_i}^{l,\pi})$.

  for $l = 0$ to maxstep **do**
    $\bar{x}_0^{l,\pi} = x_0$, $\bar{p}_0^{l,\pi} = \bar{p}_0^l$;
3: **for** $i = 0$ to $N-1$ **do**
      $\bar{q}_{t_i}^{l,\pi} = \phi(t_i, \bar{x}_{t_i}^{l,\pi}, \bar{p}_{t_i}^{l,\pi}; \theta^l)$;
      $\bar{x}_{t_i+1}^{l,\pi} = \bar{x}_{t_i}^{l,\pi} + b(t_i, \bar{x}_{t_i}^{l,\pi}, \bar{u}_{t_i}^{l,\pi})\Delta t_i + \sigma(t_i, \bar{x}_{t_i}^{l,\pi}, \bar{u}_{t_i}^{l,\pi})\Delta W_{t_i}$;
6:    $\bar{p}_{t_i+1}^{l,\pi} = \bar{p}_{t_i}^{l,\pi} - H_x(t_i, \bar{x}_{t_i}^{l,\pi}, \bar{u}_{t_i}^{l,\pi}, \bar{p}_{t_i}^{l,\pi}, \bar{q}_{t_i}^{l,\pi})\Delta t_i + \bar{q}_{t_i}^{l,\pi}\Delta W_{t_i}$;
    end for
    $loss = \frac{1}{M}\sum_{j=1}^{M}\left[|-h_x(\bar{x}_T^{l,\pi}) - \bar{p}_T^{l,\pi}|^2\right]$;
9: $(\theta^{l+1}, \bar{p}^{l+1,\pi}) = (\theta^l, \bar{p}^{l,\pi}) - \nu\nabla loss$;
  end for
  $\bar{u}_{t_i}^{l,\pi} = \arg\max_{u\in U} H(t_i, \bar{x}_{t_i}^{l,\pi}, u, \bar{p}_{t_i}^{l,\pi}, \bar{q}_{t_i}^{l,\pi})$;
12: $J(\bar{u}^{l,\pi}(.)) = \frac{1}{M}\sum_{j=1}^{M}\left[\frac{T}{N}\sum_{i=0}^{N-1} f(t_i, \bar{x}_{t_i}^{l,\pi}, \bar{u}_{t_i}^{l,\pi}) + h(\bar{x}_T^{l,\pi})\right]$.

Since we obtained $(x^*(.), p^*(.), q^*(.))$, we get the optimal control $u^*(.)$ through the maximum condition,

$$u_t^* = \arg\max_{u\in U} H(t, x_t^*, u, p_t^*, q_t^*), \tag{4.1.14}$$

where H is given by the equation (3.2.9).

For solving the extremum problem described at line 11 in this Algorithm, we can use similar methods as those used in in Algorithm 1, such as $BFGS$, $SLSQP$. However, an important difference is that we need to calculate the extremum problem $(x^*(.), p^*(.), q^*(.))$ once throughout the entire algorithm. This is because the optimal state processes have already been obtained by solving the Hamiltonian system (4.1.13) using deep learning.

4.2 Numerical Results

For this numerical results, we present two illustrative examples to determine the optimal algorithm. Both examples are evaluated with time-points $N = 25$, a batch size of 64, and a test set sample size of 512. To ensure robustness, we compute the mean of the numerical results over 5 independent runs for each algorithm.

4.2.1 Example 1. In this example, we consider a low-dimensional case.

Let us define the control domain $U = \mathbb{R}^n$, $n = 5$. Let consider the following system control:

$$\begin{cases} dx_t = (3 - \frac{1}{3}x_t + u_t)dt + (\frac{1}{2}x_t + u_t)dW_t, \\ x(0) = 1, \end{cases} \tag{4.2.1}$$

and the cost functional is defined as

$$J(0, x_0, u(.)) = \mathbb{E}\left\{\frac{1}{2}\int_0^T \left[\langle\frac{1}{2}x_t, x_t\rangle + \langle 2u_t, u_t\rangle\right] dt + \frac{1}{2}\langle Qx_T, x_T\rangle\right\}, \tag{4.2.2}$$

where Q is a deterministic matrix taking value in $\mathbb{R}^{n\times n}$, and the Hamiltonian H can expressed as

$$H(t,x,u,p,q) = \langle p, 3-\frac{1}{3}x+u\rangle + \langle q, \frac{1}{2}x+u\rangle - \frac{1}{4}\langle x,x\rangle - \langle u,u\rangle. \tag{4.2.3}$$

For this example, the explicit solution of the control u^* is given by the following expression

$$u^* = \frac{1}{2}(p+q). \tag{4.2.4}$$

Then, we proceed to compute the function $\bar{H}$ since we possess the explicit solution for u^*.

$$\bar{H}(t,x,p,q) = \max_{u\in U} H(t,x,u,p,q). \tag{4.2.5}$$

Therefore, we obtain

$$\begin{aligned}
\bar{H}(t,x,p,q) &= \langle p, 3-\frac{1}{3}x+u^*\rangle + \langle q, \frac{1}{2}x+u^*\rangle - \frac{1}{4}\langle x,x\rangle - \langle u^*,u^*\rangle\\
&= \langle p, 3-\frac{1}{3}x\rangle + \langle p,u^*\rangle + \langle q,\frac{1}{2}x\rangle + \langle q,u^*\rangle - \frac{1}{4}\langle x,x\rangle - \langle u^*,u^*\rangle\\
&= \langle p, 3-\frac{1}{3}x\rangle + \langle q,\frac{1}{2}x\rangle + \langle p+q,u^*\rangle - \frac{1}{4}\langle x,x\rangle - \langle u^*,u^*\rangle\\
&= \langle p, 3-\frac{1}{3}x\rangle + \langle q,\frac{1}{2}x\rangle - \frac{1}{4}\langle x,x\rangle + \frac{1}{2}\langle p+q,p+q\rangle - \frac{1}{4}\langle p+q,p+q\rangle\\
&= \langle p, 3-\frac{1}{3}x\rangle + \langle q,\frac{1}{2}x\rangle - \frac{1}{4}\langle x,x\rangle + \frac{1}{4}\langle p+q,p+q\rangle.
\end{aligned}$$

The explicit solution for u^* and the corresponding Hamiltonian system is

$$\begin{cases}
dx_t^* = (3-\frac{1}{3}x_t^*+u_t^*)dt + (\frac{1}{2}x_t^*+u_t^*)dW_t,\\
dp_t^* = -(-\frac{1}{2}x_t^* - \frac{1}{3}p_t^* + \frac{1}{2}q_t^*)dt + q_t^* dW_t,\\
x_0^* = 1, \quad p_T^* = -Qx_T^*,\\
u_T^* = \frac{1}{2}(p_t^*+q_t^*).
\end{cases} \tag{4.2.6}$$

Let us define

$$\begin{aligned}
b(t,x_t,u_t) = 3-\frac{1}{3}x_t+u_t, \quad \sigma(t,x_t,u_t) = \frac{1}{2}x_t+u_t,\\
f(t,x_t,u_t) = \frac{1}{4}\langle x_t,x_t\rangle + \langle u_t,u_t\rangle, \quad h(x_T) = \frac{1}{2}\langle Qx_T,x_T\rangle,\\
H_x(t,x,u,p,q) = -\frac{1}{3}p + \frac{1}{2}q - \frac{1}{2}x, \quad H_u(t,x,u,p,q) = p+q-2u.
\end{aligned}$$

The functions b, σ, f and h are linear, exhibiting properties similar to polynomial functions. Therefore, they are continuous and differentiable, satisfying Assumption 1 and Assumption 2. Consequently, equation (4.2.6) adheres to the monotonicity condition and has a unique solution $(x^*(.), p^*(.), q^*(.))$.

Now, we need to introduce the Riccati equation, which arises from equation (4.2.6). Suppose that $p_t^* = -K_t x_t^*$, $q_t^* = -M_t x_t^*$ are solutions of equation (4.2.6). By combining these solutions, we derive a Riccati equation

$$\begin{cases}
\dot{K}_t - \frac{1}{2}K_t^2 - \frac{2}{3}K_t + \left(\frac{1}{2}E_n - \frac{1}{2}K_t\right)M_t + \frac{1}{2}E_n = 0,\\
\frac{1}{2}K_t^2 - \frac{1}{2}K_t + (K_t+E_n)\,M_t = 0,\\
K_T = Q,
\end{cases} \tag{4.2.7}$$

where $\dot{K}_t$ is the derivative of K_t, and E_n is the unit matrix. Now, we can get its numerical solution by using the 'RK45' method in python. Therefore, the numerical solution is used as a benchmark to compared with our algorithms. In this example, we set $Q = E_n$ as an identity matrix and T as 0.1. The numerical solution of equation (4.2.7) with RK45 is $p_0^* = -0.999327$.

Now, let us examine below the differences between the three algorithms. After computing the three algorithms using a Dell laptop with a Core $i5$ processor, we have the following results:

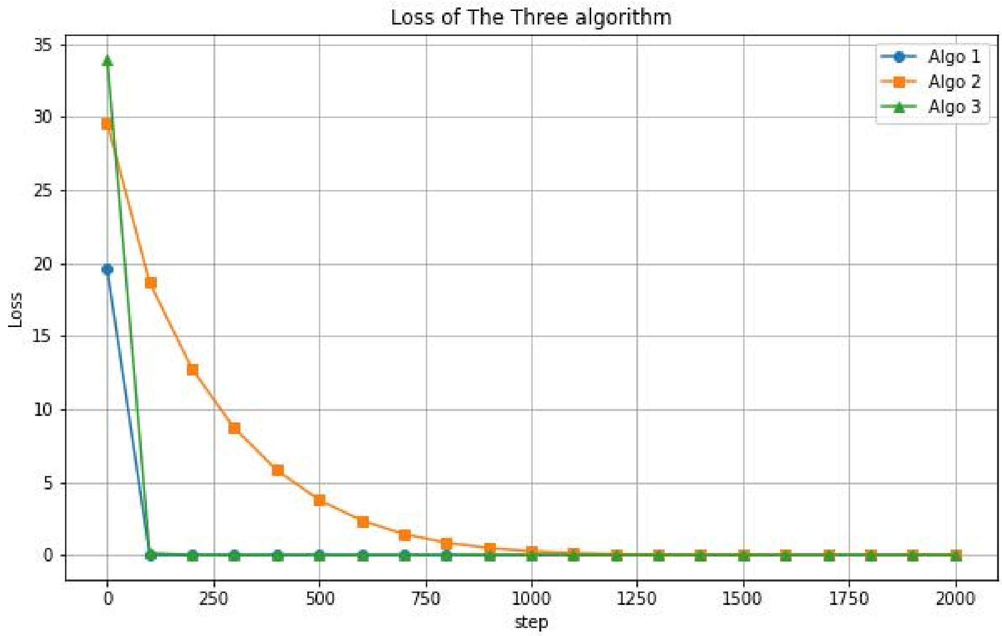

Figure 4.1: The Loss of three algorithms.

Table 4.1: Table of Cost functional after max step.

Method	**Time**	p_0	**Cost functional**	**Max Step**	**Relative error**
Algorithm 1	7000	-1.2285	3.9605	2000	0.229 33
Algorithm 2	56	-1.2179	3.9579	2000	0.218 72
Algorithm 3	40	-1.1374	3.6521	2000	0.138 17

The results depicted in Figure 4.1 illustrate the convergence of the loss functions for all three algorithms towards 0, with algorithms 1 and 3 demonstrating notably faster convergence. About the cost functional we can see that they are not big difference, but to choose of the algorithm, we should look the time and the relative error.

Now, look at the time. The algorithm 1 surpasses 7000 seconds, algorithm 3 only requires 40 seconds, and algorithm 2 takes 56 seconds. This discrepancy clearly indicates algorithm 3 as the most efficient in terms of time consumption, even the relative error is small for the algorithm 3. Consequently, the comprehensive evidence unequivocally favors algorithm 3 as the optimal performer in this specific scenario.

It is worth noting that algorithm 1 is typically suitable for low-dimensional problems, yet in this case, we have the explicit solution of u^* and $\bar{H}$, which is the special case of algorithm 3, as we've observed earlier. Thus, this example demonstrates the effectiveness of algorithm 3.

4.2.2 Example 2. This example involves a high-dimensional control $\bar{u}$ without an explicit solution. Let us consider the following stochastic control problem:

$$\begin{cases} dx_t = \sin u_t dt + x_t dW_t, \\ x(0) = 1, \end{cases}$$

with cost functional

$$J(u(.)) = \mathbb{E}\left\{\int_0^T \langle u_t, u_t\rangle dt + \frac{1}{2}\langle x_T, x_T\rangle\right\},$$

where the control domain is $U = \mathbb{R}^n$, $T = 0.1$. The Hamiltonian H is

$$H(t, x, u, p, q) = \langle p, \sin u\rangle + \langle q, x\rangle - \langle u, u\rangle,$$

which is also a multi-dimensional transcendent equation, and does not have an explicit representation of the both $\bar{u}$ and the function $\bar{H}$.

In this case, neither Algorithm 1 nor Algorithm 3 is suitable. Algorithm 1 struggles due to the high dimensionality, while Algorithm 3 is not applicable as we lack an explicit solution for $\bar{u}$. Therefore, we primarily provide the numerical results of Algorithm 2 in this example. In Algorithm 2, we only need to consider the control condition $H_u(t, x, u, p, q) = 0$. We observe that the derivative of the Hamiltonian H with respect to u is:

$$H_u(t, x, u, p, q) = p\cos u - 2u,$$

which is also a multi-dimensional transcendent equation.

Let us consider the constraint $H_u(t, x, u, p, q) = 0$. Then, the Hamiltonian system is:

$$\begin{cases} dx_t^* = \sin u_t^* dt + x_t^* dW_t, \\ dp_t^* = -q_t^* dt + q_t^* dW_t, \\ x_0^* = 1, \quad p_T^* = -h_x(x_T^*), \\ p_t^* \cos u_t^* - 2u_t^* = 0. \end{cases}$$

Let us define

$$\begin{aligned} b(t, x_t, u_t) &= \sin u_t, \quad \sigma(t, x_t, u_t) = x_t, \\ f(t, x_t, u_t) &= \langle u_t, u_t\rangle, \quad h(x_T) = \frac{1}{2}\langle x_T, x_T\rangle. \end{aligned}$$

We solely employ Algorithm 2 for our analysis and introduce another algorithm as a reference to assess the convergence of our proposed method. The algorithm by Han and E, as detailed in [J Han (2016)], directly addresses stochastic optimal control problems using deep learning. The following figure displays the cost curves for both our Algorithm 2 and the reference algorithm by Han and E. We set $n = 100$, and $\lambda = 0.1$.

The close proximity of the cost functional values between the two algorithms seen in Figure 4.2 indicates that our Algorithm 2 is highly effective in solving optimal control problems in high-dimensional spaces, even in the absence of an explicit solution for u^*.

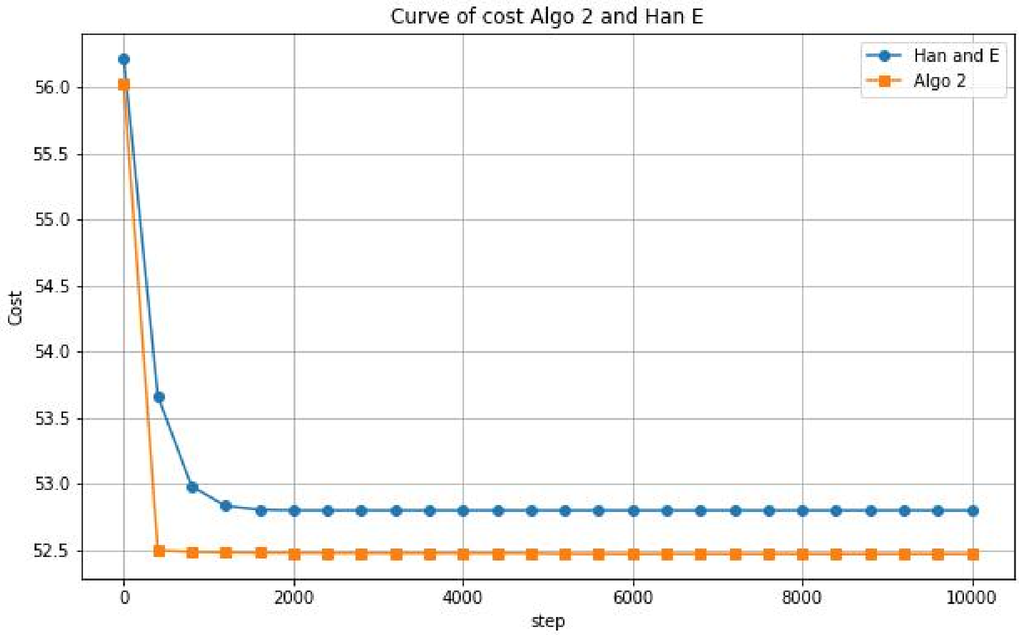

Figure 4.2: Cost functional.

Remark: The choice of algorithm can be determined based on the nature of the function H in the maximum condition and the characteristics of the stochastic optimal control problem. When $\bar{u}$ can be explicitly determined, Algorithm 3 is advantageous as it transforms the optimal stochastic control problem into solving FBSDEs. However, in cases where $\bar{u}$ lacks an explicit representation, any of the three proposed algorithms can be utilized. If the optimal control problem aligns with the conditions of Algorithm 2 or 3, then Algorithm 2 or 3 would be the preferred choices; otherwise, Algorithm 1 could be employed.

5. Conclusion

In this thesis, our primary objective has been to solve stochastic optimal control problems. We leveraged the stochastic maximum principle as our main tool for addressing these challenges. This principle has extended the classical maximum principle to stochastic optimal control problems, offering necessary conditions for optimality. To begin, we have delved into the stochastic maximum principle, exploring its conditions and implications, essential for solving optimal control problems. We then have reformulated the control problem into a new variational control problem, laying the groundwork for further analysis. Also, we have used numerical algorithms designed to solve the new control problem. We have proposed three distinct algorithms tailored to different scenarios, each leveraging deep learning techniques. The first algorithm has constructed a single deep neural network (DNN) to approximate the control, suitable for low-dimensional problems. The second algorithm has addressed a broader class of stochastic optimal control problems, enhancing computational efficiency by utilizing neural networks to simulate controls. This approach has been particularly beneficial for high-dimensional cases where the optimal control cannot be explicitly expressed. The third algorithm has been employed when certain conditions are met, allowing for the resolution of high-dimensional stochastic optimal control problems. It effectively has addressed cases where the optimal control cannot be explicitly determined. Numerical results from all three algorithms have demonstrated promising performance. When the optimal control can be explicitly represented, Algorithm 3 has emerged as an intuitive choice. However, in cases where explicit representation is not feasible, Algorithms 2 or 3 has offered preferable alternatives for high-dimensional problems. Algorithm 1 remains suitable for low-dimensional cases. Throughout this work, we have maintained a focus on coherence and consistency, ensuring robust analyses within the framework of stochastic optimal control.

References

Antonio Moro A. Bensoussan, Sonjoy K. Mitter. *Nonlinear filtering and stochastic control Proceedings*. 1981.

Jean-Michel Bismut. *An Introductory Approach to Duality in Optimal Stochastic Control*. Society for Industrial and Applied Mathematics, 2014.

Jean-Michel Bismut. *Convex analysis and probability*. Université Paris VI, 1973, 2021.

Matthew Davey. *Error-correction using Low-Density Parity-Check Codes*. Phd, University of Cambridge, 1999.

Krylov N.V. Dong, H. *The rate of convergene of finite-difference approximations for parabolic bellman equation with lipschitz coefficients in cylindrical domains.* Appl. Math. Optim. 56(1), 37, 2007.

Euler. Euler-maruyama scheme. youtube, https://www.youtube.com/watch?v=sgagUAP4IJQ, 2022.

W.E. J Han. *Deep Learning Approximation for Stochastic Control Problems, Deep Reinforcement Learning Workshop*. 2016.

Fima C Klebaner. *Introduction to stochastic Calculus with applications*. Imperial College Press, 2012.

Dupuis P.G. Kushner, H. *Numerical Methods for stochastic Control Problems in Continuous Time*. Springer, Berlin, 2001.

Damien Lamberton and Bernard Lapeyre. *Introduction to stochastic calculus applied to finance*. Chapman and Hall/CRC, 2011.

Nocedal J. Liu, D.C. On the limited memory bfgs method for large scale optimization. *Math. Program.*, 1989.

Yong Ma, J. *Forward-Backward Stochastic Differential Equations and Their Applications*. Springer Science and Business Media, 1999.

Bernt Oksendal. *Optimal stopping and stochastic control differential games for jump diffusions*.

S. Peng. Problem of eigenvalues of stochastic hamiltonian systems with boundary conditions. *Stoch. Process. Appl. 88(2)*, 2000.

Wu Z. Peng, S. *Fully coupled forward-backward stochastic differential equations and applications to optimal control*. 1999.

Ying Peng Shaolin Ji, Shige Peng and Xichuan Zhang. Three algorithms for solving high-dimensional fully-coupled fbsdes through deep learning. *School of Mathematics, Shandong University, 250100, China*, 2020.

Ying Peng Xichuan Zhang Shaolin Ji, Shige Peng. Solving stochastic optimal control problem via stochastic maximum principle with deep learning method. *Springer Science+Business Media, LLC, part of Springer Nature 2022*, 7 September 2022.

Zhou Yong, J. *Stochastic Controls-Hamiltonian System and HJB Equation*. Springer, 1999.