PEOPLE'S DEMOCRATIC REPUBLIC OF ALGERIA
MINISTRY OF HIGHER EDUCATION AND SCIENTIFIC RESEARCH
University Mohamed Khider of Biskra
Faculty of Exact Sciences and Natural and Life Sciences

THESIS

Presented for the degree of

Doctor Sciences in Mathematics

Option: Probability

Presented by

Nour El Houda ABADA

Titre

On the variational principle for a class of stochastic control for systems governed by stochastic differential equations of mean-field type with applications

Members of the jury:

Djabrane Yahia , Pr., University of Biskra	President
Mokhtar Hafayed, Pr., University of Biskra	Supervisor
Imad Eddine Lakhdari, MCA, University of Biskra	Co-Supervisor
Abdelmoumen Tiaiba, Pr., University of M'sila,	Examiner
Youcef Djenaihi, MCA, University of Sétif	Examiner
Khalil Saadi, Pr., University of M'sila	Examiner

2022

PEOPLE'S DEMOCRATIC REPUBLIC OF ALGERIA

MINISTRY OF HIGHER EDUCATION AND SCIENTIFIC RESEARCH

University Mohamed Khider of Biskra

Faculty of Exact Sciences and Natural and Life Sciences

Thesis

Presented for the degree of

Doctor in Mathematics

Option : Probability

Presented by

Nour El Houda Abada

Title

On the variational principle for a class of stochastic control for systems governed by stochastic differential equations of mean-field type with applications

---

Members of the jury :

Djabrane Yahia *Prof, University of Biskra,* **President**

Mokhtar Hafayed, *Prof, University of Biskra* **Supervisor**

Imad Edine Lakhdari, *MCA, University of Biskra* **Co-Supervisor**

Abdelmoumen Tiaiba, *Prof, University of M'sila* **Examiner**

Youcef Djenaihi, *MCA, University of Sétif,* **Examiner**

Khalil Saadi, *Prof. University of M'sila* **Examiner**

2022

Dedicace

I dedicate this work to my $\mathcal{M}$other and my $\mathcal{F}$ather,

To my sister, my brother and my $\mathcal{H}$usband.

To my beloved children and my nephews.

Perseverance overcomes everything

Acknowledgments

First of all and foremost, praises and thanks to Allah, for his help and blessings throughout the research work to complete my thesis successfully.

I would like to express my sincere thanks to my advisor Prof. Mokhtar Hafayed for his insights, guidence, support. Many thanks not only because this work would have been not possible without his help, but above all because in these years he taught me with passion and patience the art of being a mathematician.

My sincere thanks to Professors Pr.Yahia Djabrane , Pr. Abdelmoumen Tiaiba, Dr Youcef Djenaihi and Pr. Khalil Saadi and Dr. Imad Eddine Lakhdari, because they agreed to spend their times for reading and evaluating my thesis.

I am very thankful to all my colleagues of the Mathematics Department and the Laboratory of Mathematical Analysis, Probability and Optimizations (LMAPO) in University Mohamed Khider of Biskra, Algeria.

Symbols and Acronyms

1. $\mathbb{N}$: Set of Natural numbers.
2. $\mathbb{R}$: Set of Real numbers.
3. $\mathbb{R}_+$: Set of Non-negative real numbers.
4. **a.e**. almost everywhere
5. **a.s.** almost surely
6. **càdlàg** continu à droite, limite à gauche
7. **càglàd** continu à gauche, limite à droite
8. **cf**. compare (abbreviation of Latin confer)
9. **e.g**. for example (abbreviation of Latin exempli gratia)
10. **i.e,**. that is (abbreviation of Latin id est)
11. **HJB** *The Hamilton-Jacobi-Bellman equation*
12. **SDE** : Stochastic differential equations.
13. **BSDE :** Backward stochastic differential equation.
14. **FBSDEs** : Forward-backward stochastic differential equations.
15. **FBSDEJs** : Forward-Backward stochastic differential equations with jumps.
16. **PDE** : Partial differential equation.
17. **ODE :** Ordinary differential equation.
18. $\dfrac{\partial f}{\partial x}, f_x$: The derivatives with respect to x.
19. $\mathbb{P}\otimes dt$: The product measure of $\mathbb{P}$ with the Lebesgue measure dt on $[0,T]$.
20. $\boldsymbol{E}(\cdot)$, $\boldsymbol{E}(\cdot \mid G)$ Expectation ; conditional expectation
21. $\sigma(A)$: $\sigma-$algebra generated by A.
22. I_A : Indicator function of the set A.

23. $\mathcal{F}^{Y}$: The filtration generated by the process Y.

24. $W(\cdot), B(\cdot)$: Brownian motions

25. $\mathcal{F}_t^B$ the natural filtration generated by the brownian motion $B(\cdot)$,

26. $\mathrm{F}_1 \vee \mathrm{F}_2$ denotes the σ-field generated by $\mathrm{F}_1 \cup \mathrm{F}_2$.

27. $(\Omega, \mathcal{F}, \mathbb{P})$ probability space

28. $\{\mathcal{F}_t\}_{t\geq 0}$: filtration

29. $\left(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\in[0,T]}, \mathbb{P}\right)$ filtered probability space.

30. $\mathbb{L}^p(\mathcal{F})$: the set of $\mathbb{R}^n$-valued $\mathcal{F}-$measurable random variables X such that
$$\boldsymbol{E}(\,|X|^p) < \infty.$$

31. $\mathbb{L}^p_{\mathcal{G}}(\Omega, \mathbb{R}^n\ )$: the set of $\mathbb{R}^n$-valued $\mathcal{G}-$measurable random variables X such that
$$\boldsymbol{E}(\,|X|^p) < \infty.$$

32. $\mathbb{L}^p_{\mathcal{F}}([0,T]\,, \mathbb{R}^n)$: the set of all $(\mathcal{F}_t)_{t\geq 0}$-adapted $\mathbb{R}^n$-valued processes X such that
$$\boldsymbol{E}\int_0^T |X(t)|^p\, dt < \infty.$$

33. $\mathbb{L}^\infty_{\mathcal{F}}([0,T]\,, \mathbb{R}^n)$: the set of all $(\mathcal{F}_t)_{t\geq 0}$-adapted $\mathbb{R}^n$-valued processes X essentially bounded processes.

34. $(u(\cdot), \xi(\cdot))$: continous-singular control.

35. $\partial_\mu g$: the derivatives with respect to measur μ.

36. $\mathcal{D}_\zeta g(\mu_0)$: the *Fréchet-derivative* of g at μ_0 in the direction ξ.

Résumé

Cette thèse de doctorat s'inscrit dans le cadre de la théorie de contrôle optimal stochastique. Le thème central est l'optimisation stochastique afin d'établir des conditions nécessaires d'un contrôle optimal sous forme du principe du maximum stochastique de type de Pontryagin.

D'une part, et plus précisement, nous étudions des problèmes de contrôle stochastique optimal singulier partiellement observés de type mean-field (*McKean-Vlasov*) général avec des corrélations entre le système et l'observation $Y(\cdot)$. Dans ce travail, la variable de contrôle $(u(\cdot), \xi(\cdot))$ a deux composantes, la première $u(\cdot)$ est absolument continue et la seconde $\xi(\cdot)$ est une variation bornée, non décroissante continue à droite avec limit à gauche (*càdlàg*).

Le système stochastique étudié est gouverné par une équation différentielle stochastique contrôlée de type Itô où les coefficients de la dynamique dépendent du processus d'état ainsi que de sa loi de probabilité $\mathbb{P}_{x^{u,\xi}(t)}$ et de la variable de contrôle continue $u(\cdot)$, définit par :

$$\begin{cases} \mathrm{d}x^{u,\xi}(t) = & f(t, x^{u,\xi}(t), \mathbb{P}_{x^{u,\xi}(t)}, u(t))\mathrm{d}t + \sigma(t, x^{u,\xi}(t), \mathbb{P}_{x^{u,\xi}(t)}, u(t))\mathrm{d}W(t) \\ & \quad +g(t, x^{u,\xi}(t), \mathbb{P}_{x^{v,\xi}(t)}, v(t))\mathrm{d}\widetilde{W}(t) + G(t)\mathrm{d}\xi(t), \\ x^{u,\xi}(0) = & x_0, \quad t \in [0, T]. \end{cases}$$

Nous supposons que le processus d'état $x^{u,\xi}(t)$ ne peut pas être observé directement, mais les contrôleurs peuvent observer un processus de bruit associé $Y(\cdot)$, régit par l'équation suivante :

$$\begin{cases} dY(t) & = h(t, x^{u,\xi}(t), u(t))dt + \mathrm{d}\widetilde{W}(t) \\ Y(0) & = 0, \end{cases}$$

où $\widetilde{W}(t)$ est un processus stochastique dépendant du contrôle $u(\cdot)$, et $Y(\cdot)$ le processus

d'observation. On definit $\mathcal{F}_t^Y-$martingale $\rho^u(t)$ qui est une solution de l'equation suivante :

$$\begin{cases} \mathrm{d}\rho^u(t) = \rho^u(t)h\left(t, x^u(t), u(t)\right)\mathrm{d}Y(t), \\ \rho^u(0) = 1. \end{cases}$$

D'aprés *le théorème de dérivation de Radon-Nikodym*, cette martingale a permis de définir une nouvelle probabilité notée $\mathbb{P}^u$, qui dépend de $u\left(\cdot\right)$ et donnée par :

$$\left.\frac{\mathrm{d}\mathbb{P}^u}{\mathrm{d}\mathbb{P}}\right|_{\mathcal{F}_t^Y} = \rho^u(t).$$

La fonctionnelle de coût $J(u(\cdot), \xi(\cdot))$ peut s'écrire sous forme

$$\begin{aligned} J(u(\cdot), \xi(\cdot)) = \boldsymbol{E}\Bigg[\int_0^T \rho^u(t)l(t, x^{u,\xi}(t), \mathbb{P}_{x^{u,\xi}(t)}, u(t))\mathrm{d}t + \rho^u(T)\psi(x^{u,\xi}(T), \mathbb{P}_{x^{u,\xi}(T)}) \\ + \int_{[0,T]} \rho^u(t)M(t)\mathrm{d}\xi(t)\Bigg]. \end{aligned}$$

Par l'utilisation des techniques variationnelles convexes classiques, nous établissons un ensemble de conditions nécessaires de contrôle singulier optimal sous la forme du principe du maximum. Notre résultat principal est prouvé en appliquant le *théorème de Girsanov* et les dérivées par rapport à une mesure (ou la loi de probabilité) au sense de P. Lions.

D'autre part, nous établissons des conditions nécessaires du second-ordre pour un contrôle stochastique mixed continu-singulier $\left(u\left(\cdot\right), \xi\left(\cdot\right)\right)$, où le système est gouverné par des systèmes différentiels stochastiques contrôlés non linéaires. Le principe du maximum ponctuel du second-ordre en termes de martingale par rapport à la variable de temps est prouvé. Le domaine de contrôle est supposé convexe. Notons que dans ce travail que les termes de dérivée et les termes de diffusion des systèmes dépendent de la variable de contrôle continue $u\left(\cdot\right)$. Notre résultat est prouvé en utilisant des techniques variationnelles sous certaines conditions de convexité.

Cette thèse s'articule autour de trois chapitres :

Le premier chapitre est essentiellement un rappel. Nous présentons quelques concepts et résultats qui nous permettrons d'aborder notre travail ; tels que les processus stochastiques, les filtrations, l'espérance conditionnelle, les martingales, les formules d'Itô, les différentes méthodes de résolution d'un problème de contrôle optimal stochastiques (principe du maximum stochastique [1] et le principe de la programmation dynamique[2]), ainsi que les différentes classes de contrôle stochastique, ... etc.

Dans le deuxième chapitre, nous avons établi et prouvé les conditions nécessaires vérifiées par un contrôle optimal stochastique partiellement observé, pour un système différentiel gouverné par des équations différentielles stochastiques EDSs de type mean-field avec des corrélations entre le système et l'observation. Les coefficients de notre système dépendent du processus d'état ainsi que de sa loi de probabilité. Le domaine de contrôle stochastique est supposé convexe. La méthode utilisée est basée sur la dérivée par rapport à une mesure de probabilité. Les résultats obtenus dans le chapitre §2, sont tous nouveaux et font l'objet d'un premier article intitulé :

Nour El Houda Abada & Mokhtar Hafayed, & Shahlar Meherrem : On Partially observed optimal singular control of McKean-Vlasov stochastic systems : maximum principle approach, *Mathematical Methods in the Applied Sciences, Wiley & Jonson 2022* , Math Meth Appl Sci. 2022 ;1-21.**DOI : 10.1002/mma.8373**.

Dans le troisième chapitre, nous avons obtenu les conditions nécessaires d'optimalité de second-order sous forme d'un principe du maximum stochastique. Le système est gouverné par des équations différentielles stochastiques de la forme $t \in [0, T]$

$$\begin{cases} \mathrm{d}x^{u,\xi}(t) = f\left(t, x^{u,\xi}(t), u(t)\right)\mathrm{d}t + \sigma\left(t, x^{u,\xi}(t), u(t)\right)\mathrm{d}W(t) + G(t)\mathrm{d}\xi(t), \\ x^{u,\xi}(0) = x_0. \end{cases}$$

[1] L.S. Pontryagin, V.G. Boltanski and R.V. Gamkrelidze *(1962), The mathematical theory of optimal processes. Intersciene N.Y.*

[2] Bellman, R., Glicksberg, I., and Gross, O. *On some variational problems occurring in the theory of dynamic programming, Rend. Circ. Mat. Palermo (2), 3 (1954), 1-35.*

Le cout a minimizer est donné par

$$J\left(u(\cdot),\xi(\cdot)\right)=\boldsymbol{E}\left[h(x^{u,\xi}(T))+\int_0^T \ell(t,x^{u,\xi}(t),u(t))\mathrm{d}t+\int_{[0,T]} M(t)\mathrm{d}\xi(t)\right].$$

Dans cette partie de notre travail, la variable de contrôle $\left(u\left(\cdot\right),\xi\left(\cdot\right)\right)$ a deux composantes, la première est absolument continue et la seconde est une variation bornée, non décroissante continue à droite avec une limite à gauche (*càdlàg*). Nous établissons les conditions nécessaires du second ordre pour un problème de contrôle continu-singulier optimal. Le domaine de contrôle est nécessairement convexe. Un principe de maximum ponctuel du second-ordre en termes de martingale par rapport à la variable de temps est démontré. Des techniques variationnelles, certains théorèmes de Lebesgue sur les différenciations, les mesures et les intégrations, avec quelques estimations appropriées sont appliqués pour etablir nos résultats. Notre problème de contrôle optimal fournit également un modèle intéressant dans de nombreuses applications telles que l'économie et la finance mathématique. Nos résultats prouvés géneralisent les résultats obtenus dans l'article : "*Zhang H.& Zhang X. : Pointwise second-order necessary conditions for stochastic optimal controls, Part I : The case of convex control constraint, SIAM J. Control Optim. 53(4), 2267-2296 (2015)*", à une classe de problèmes de contrôle stochastique singulier $(u(\cdot),\xi(\cdot))$. Lorsque les conditions nécessaires d'optimalité du premier ordre sont singulières dans un certain sens, les conditions nécessaires du second ordre viendront naturellement. La nouveauté de notre travail est que sous certaines hypothèses, nous fournissons des conditions nécessaires ponctuelles du second-ordre qui sont nouvelles pour le cas du controle stochastique singulier $(u(\cdot),\xi(\cdot))$. Le principe du maximum de second-ordre établi dans cet article peut être utilisé pour choisir les contrôles candidats afin qu'ils soient optimaux à partir de la singularité de nos contrôles stochastiques. Habituellement, afin de dériver le principe du maximum de second ordre d'optimalité, il faut supposer que la condition du premier ordre dégénère dans un certain sens. Les résultats obtenus dans le chapitre §3 sont tous nouveaux et font l'objet d'un deuxième article intitulé :

Nour El Houda Abada, Mokhtar Hafayed : Stochastic pointwise second-order maximum principle for optimal continuous-singular control using variational approach, *International Journal Modelling Identification and Control,* accepté, 2022

Abstract

This thesis is concerned with stochastic singular optimal control. The central theme is the necessary conditions, in the form of the Pontryagin's stochastic maximum for optimality. Recently, the main purpose of this thesis is to derive a set of necessary conditions of optimality in the form of Pontryagin maximum principle. The control variable is a pair $(u(\cdot), \xi(\cdot))$ of measurable $\mathbb{A}_1 \times \mathbb{A}_2$-valued, $\mathbb{F}-$adapted processes, where $\mathbb{A}_1$ is a closed convex subset of $\mathbb{R}^m$ and $\mathbb{A}_2 := [0, \infty)^m$ such that $\xi(\cdot)$ is of bounded variation, nondecreasing continuous on the right with left limits.

This thesis is structured around three chapters :

The first chapter is essentially a reminder. we presents some concepts and results that allow us to prove our results, such as stochastic processes, conditional expectation, martingales, Itô formulas, different methods of solving of optimal control (maximum principle [3] and dynamical programming principle[4]) and class of stochastic control, ...etc.

Recently, in the second chapter of this thesis, we study partially observed optimal stochastic singular control problems of general mean-field with correlated noises between the system and the observation. The control variable has two components, the first being absolutely continuous and the second is a bounded variation, non decreasing continuous on the right with left limits. The dynamic system is governed by Itô-type controlled stochastic differential equation. The coefficients of the dynamic depend on the state process as well as of its probability law and the continuous control variable.

$$\begin{cases} \mathrm{d}x^{v,\xi}(t) = & f(t, x^{v,\xi}(t), \mathbb{P}_{x^{v,\xi}(t)}, v(t))\mathrm{d}t + \sigma(t, x^{v,\xi}(t), \mathbb{P}_{x^{v,\xi}(t)}, v(t))\mathrm{d}W(t) \\ & \quad + g(t, x^{v,\xi}(t), \mathbb{P}_{x^{v,\xi}(t)}, v(t))\mathrm{d}\widetilde{W}(t) + G(t)\mathrm{d}\xi(t), \\ x^{v,\xi}(0) = & x_0, \quad t \in [0, T], \end{cases}$$

[3] L.S. Pontryagin, V.G. Boltanski and R.V. Gamkrelidze *(1962), The mathematical theory of optimal processes. Intersciene N.Y.*

[4] Bellman, R., Glicksberg, I., and Gross, O. *On some variational problems occurring in the theory of dynamic programming, Rend. Circ. Mat. Palermo (2), 3 (1954), 1-35.*

where $\mathbb{P}_{x^{v,\xi}} = \mathbb{P}\circ\left(x^{v,\xi}\right)^{-1}$ denotes the law of the random variable $x^{v,\xi}(\cdot)$. We assume that the state process $x^{v,\xi}(\cdot)$ cannot be observed directly, but the controllers can observe a related noisy process $Y(\cdot)$, which is governed by the following equation :

$$\begin{cases} dY(t) & = h(t, x^{v,\xi}(t), v(t))dt + \mathrm{d}\widetilde{W}(t) \\ Y(0) & = 0, \end{cases}$$

We define the $\mathcal{F}_t^Y-$martingale $\rho^v(t)$ which is the solution of the equation

$$\begin{cases} \mathrm{d}\rho^v(t) = \rho^v(t)h\left(t, x^v(t), v(t)\right)\mathrm{d}Y(t), \\ \rho^v(0) = 1. \end{cases}$$

This martingale allowed to define a new probability, denoted by $\mathbb{P}^v$ on the space $(\Omega, \mathcal{F})$, to emphasize the fact that it depend on the control $v(\cdot)$. It is given by the Radon-Nikodym derivative :

$$\left.\frac{\mathrm{d}\mathbb{P}^v}{\mathrm{d}\mathbb{P}}\right|_{\mathcal{F}_t^Y} = \rho^v(t).$$

Hence, by Girsanov's theorem and hypothesis (C1) and (C2), $\mathbb{P}^v$ is a new probability measure of density $\rho^v(t)$. The process

$$\widetilde{W}(t) = Y(t) - \int_0^t h(s, x^{v,\xi}(s), v(s))\mathrm{d}s,$$

is a standard Brownian motion independent of $W(\cdot)$ and x_0 on the new probability space $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbb{P}^v)$.

By using Radon-Nikodym derivative, and the martingale property of $\rho^v(t)$, the cost functional can be written as

$$J(v(\cdot),\xi(\cdot))=\boldsymbol{E}\left[\int_0^T \rho^v(t)l(t,x^{v,\xi}(t),\mathbb{P}_{x^{u,\xi}(t)},v(t))\mathrm{d}t+\rho^v(T)\psi(x^{v,\xi}(T),\mathbb{P}_{x^{v,\xi}(T)})\right.$$
$$\left.+\int_{[0,T]}\rho^v(t)M(t)\mathrm{d}\xi(t)\right].$$

In terms of a classical convex variational techniques, we establish a set of necessary contiditions of optimal singular control in the form of maximum principle. Our main result is proved by applying Girsanov's theorem and the derivatives with respect to probability law in P.L. Lions' sense. To illustrate our theoretical result, we study partially observed linear quadratic singular control problem of mean-field type. The results obtained in Chapter §2 are all new and are the subject of a first article entitled :

Nour El Houda Abada & Mokhtar Hafayed & Shahlar Meherrem : On Partially observed optimal singular control of McKean-Vlasov stochastic systems : maximum principle approach, *Mathematical Methods in the Applied Sciences, Wiley & Jonson 2022,* **DOI : 10.1002/mma.8373**.

In the third chapter, we study stochastic singular optimal control problem. We establish a set of second-order necessary conditions for optimal continuous-singular stochastic control, where the systems is governed by nonlinear controlled Itô stochastic differential systems.

$$\begin{cases} \mathrm{d}x^{u,\xi}(t)=f\left(t,x^{u,\xi}(t),u(t)\right)\mathrm{d}t+\sigma\left(t,x^{u,\xi}(t),u(t)\right)\mathrm{d}W(t)+G(t)\mathrm{d}\xi(t), \\ x^{u,\xi}(0)=x_0. \end{cases}$$

The expected cost to be minimized over the class of admissible controls has the form

$$J\left(u(\cdot),\xi(\cdot)\right)=\boldsymbol{E}\left[h(x^{u,\xi}(T))+\int_0^T \ell(t,x^{u,\xi}(t),u(t))\mathrm{d}t+\int_{[0,T]}M(t)\mathrm{d}\xi(t)\right].$$

Here the control variable is a pair $(u(\cdot),\xi(\cdot))$ of measurable $\mathbb{A}_1\times\mathbb{A}_2$-valued, $\mathbb{F}-$adapted processes, where $\mathbb{A}_1$ is a closed convex subset of $\mathbb{R}^m$ and $\mathbb{A}_2 := [0, \infty)^m$ such that $\xi(\cdot)$ is of bounded variation, nondecreasing continuous on the right with left limits. The process $x^{u,\xi}(\cdot)$ is the state variable valued in $\mathbb{R}^n$ associated to $(u(\cdot), \xi(\cdot))$. This construction allows us to define integrals of the form $\int_{[0,T]} G(t)\mathrm{d}\xi(t)$ and $\int_{[0,T]} M(t)\mathrm{d}\xi(t)$.

The control process has two components, the first being absolutely continuous and the second is a bounded variation, non decreasing continuous on the right with left limits. Pointwise second order maximum principle in terms of the martingale with respect to the time variable is proved. The control domain is assumed to be convex. In this chapter, the continuous control variable enters into both the drift and the diffusion terms of the control systems. Variational techniques, some Lebesgue theorems in differentiations, measure and integrations, with some appropriate estimates are applied to derive our results.

Our continuous-singular control problem under studied provides also an interesting models in many applications such as economics and mathematical finance. This paper extends the results obtained in "*Zhang H., Zhang X. : Pointwise second-order necessary conditions for stochastic optimal controls, Part I : The case of convex control constraint, SIAM J. Control Optim. 53(4), 2267-2296 (2015)*" to a class of continuous-singular stochastic control problems.

The main novelty of our work is that under some assumptions, we provide pointwise second-order necessary conditions which are new for the stochastic continuous-singular case and are natural extension of their deterministic counterparts. When the first-order necessary conditions of optimality are singular in some sense, the second-order necessary conditions will come naturally. The second-order maximum principle established in this chaptre can be used to choose the candidates from the singularity of our stochastic controls for optimal ones. Usually, in order to derive the second-order maximum principle for optimality, one needs to assume that the first-order condition degenerates in some sense. The results obtained in Chapter §3 are all new and are the subject of a second article entitled :

Nour El Houda Abada , Mokhtar Hafayed : Stochastic pointwise second-order maximum principle for optimal continuous-singular control using variational approach, *International Journal Modelling Identification and Control,* accepté, 2022

Chapitre 1

Introduction

Optimal control theory can be described as the study of strategies to optimally influence a system x with dynamics evolving over time according to a differential equation. The influence on the system is modeled as a vector of parameters, u, called the control. It is allowed to take values in some set U, which is known as the action space. For a control to be optimal, it should minimize a cost functional (or maximize a reward functional), which depends on the whole trajectory of the system x and the control u over some time interval $[0, T]$. The infimum of the cost functional is known as the value function (as a function of the initial time and state). This minimization problem is infinite dimensional, since we are minimizing a functional over the space of functions $u(t), t \in [0, T]$. Optimal control theory essentially consists of different methods of reducing the problem to a less transparent, but more manageable problem.

1.1 Formulation of stochastic optimal control problem

It is well-known that control theory was founded by *N. Wiener in 1948*. After that, this theory was greatly extended to various complicated settings and widely used in sciences and technologies. Clearly, control means a suitable manner for people to change the dynamics of a system under consideration. Let $\left(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\in[0,T]}, \mathbb{P}\right)$ be a given filtered probability space.

1.1.1 Stochastic process

Let $\mathbb{T}$ be a nonempty index set and $(\Omega, \mathcal{F}, \mathbb{P})$ a probability space. A family $\{X(t) : t \in \mathbb{T}\}$ of random variables from $(\Omega, \mathcal{F}, \mathbb{P})$ to $\mathbb{R}^n$ is called a stochastic process. For any $w \in \Omega$ the map $t \mapsto X(t, w)$ is called a sample path.

1.1.2 Natural fitration

Let $X = (X_t, t \geq 0)$ a stochastic process defined on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$. The natural filtration of X , denoted by $\mathcal{F}_t^X$, is defined by $\mathcal{F}_t^X = \sigma(X_s, 0 \leq s \leq t)$. Also, we called the filtaration generated by X.

1.1.3 Brownian motion

The stochastic process $(W(t),\ t \geq 0)$ is a brownian motion (standard) iff :

1. $\mathbb{P}[W(0) = 0] = 1$.
2. $t \rightarrow W(t, w)$ is continuous.$\mathbb{P}-p.s.$
3. $\forall s \leq t$, $W(t) - W(s)$ is normally distributed; center with variation $(t-s)$ i.e $W(t) - W(s) \sim \mathcal{N}(0, t-s)$.
4. $\forall n$, $\forall\ 0 \leq t_0 \leq t_1 \leq ... \leq t_n$, the variables $\left(W_{t_n} - W_{t_{n-1}}, ..., W_{t_1 -} W_{t_0}, W_{t_0}\right)$ are independents.The following result gives special case of the Itô formula for jump diffusions.

1.1.4 Integration by parts formula

Suppose that the processes $x_i(t)$ are given by : for $i = 1, 2$, $t \in [0, T]$:

$$\begin{cases} dx_i(t) = f\left(t, x_i(t)\right) dt + \sigma\left(t, x_i(t)\right) dW(t) \\ x_i(0) = 0. \end{cases}$$

Then we get

$$\begin{aligned} \boldsymbol{E}\left(x_1(T)x_2(T)\right) &= \boldsymbol{E}\left[\int_0^T x_1(t)dx_2(t) + \int_0^T x_2(t)dx_1(t)\right] \\ &\quad + \boldsymbol{E}\int_0^T \sigma^{\intercal}\left(t, x_1(t)\right)\sigma\left(t, x_2(t)\right) dt. \end{aligned}$$

In this section, we present two mathematical formulations (strong and weak formulations) of stochastic optimal control problems in the following two subsections, respectively.

1.1.5 Strong formulation

Let $\left(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\in[0,T]}, \mathbb{P}\right)$ be a given filtered probability space satisfying the usual condition, on which an d-dimensional standard Brownian motion $W(\cdot)$ is defined, consider the following controlled stochastic differential equation :

$$\begin{cases} dx(t) &= f(t, x(t), u(t))dt + \sigma(t, x(t), u(t))dW(t), \\ x(0) &= x_0 \in \mathbb{R}^n, \end{cases} \tag{1.1}$$

where

$$f : [0, T] \times \mathbb{R}^n \times A \longrightarrow \mathbb{R}^n,$$
$$\sigma : [0, T] \times \mathbb{R}^n \times A \longrightarrow \mathbb{R}^{n\times d},$$

and $x(\cdot)$ is the variable of state.

The function $u(\cdot)$ is called the control representing the action of the decision-makers (controller). At any time instant the controller has some information (as specified by the information field $\{\mathcal{F}_t\}_{t\in[0,T]}$) of what has happened up to that moment, but not able to foretell what is going to happen afterwards due to the uncertainty of the system (as a consequence, for any t the controller cannot exercise his/her decision $u(t)$ before the time t really comes), This nonanticipative restriction in mathematical terms can be expressed as "$u(\cdot)$ is $\{\mathcal{F}_t\}_{t\in[0,T]}-$adapted".

The control $u(\cdot)$ is an element of the set

$$\mathcal{U}[0,T] = \{u(\cdot) : [0,T]\times\Omega \longrightarrow \mathbb{A} \text{ such that } u(\cdot) \text{ is } \{\mathcal{F}_t\}_{t\in[0,T]} - \text{adapted}\}.$$

We introduce the cost functional as follows

$$J(u(\cdot)) \doteq \boldsymbol{E}\left[\int_0^T l(t,x(t),u(t))dt + g(x(T))\right], \tag{1.2}$$

where

$$l : [0,T]\times\mathbb{R}^n\times A \longrightarrow \mathbb{R},$$

$$g : \mathbb{R}^n \longrightarrow \mathbb{R}.$$

Definition 1.1. Let $\left(\Omega,\mathcal{F},\{\mathcal{F}_t\}_{t\in[0,T]},\mathbb{P}\right)$ be given satisfying the usual conditions and let $W(t)$ be a given d-dimensional standard $\{\mathcal{F}_t\}_{t\in[0,T]}$-Brownian motion.

A control $u(\cdot)$ is called an admissible control, and $(x(\cdot),u(\cdot))$ an admissible pair, if

i) $u(\cdot)\in\mathcal{U}[0,T]$; $x(\cdot)$ is the unique solution of equation (1.1) ;

ii) $l(\cdot,x(\cdot),u(\cdot))\in\mathbb{L}^1_{\mathcal{F}}([0,T];\mathbb{R})$ and $g(x(T))\in\mathbb{L}^1_{\mathcal{F}_T}(\Omega;\mathbb{R})$.

The set of all admissible controls is denoted by $\mathcal{U}([0,T])$. Our stochastic optimal control problem under strong formulation can be stated as follows :

Problem 1.1 Minimize (1.2) over $\mathcal{U}([0,T])$. The goal is to find $u^*(\cdot)\in\mathcal{U}([0,T])$, such that

$$J(u^*(\cdot)) \quad = \inf_{u(\cdot)\in\mathcal{U}([0,T])} J(u(\cdot)). \tag{1.3}$$

For any $u^*(\cdot) \in \mathcal{U}^s([0,T])$ satisfying (1.3) is called an strong optimal control. The corresponding state process $x^*(\cdot)$ and the state control pair $(x^*(\cdot), u^*(\cdot))$ are called an strong optimal state process and an strong optimal pair, respectively.

1.1.6 Weak formulation

In stochastic control problems, there exists for the optimal control problem another formulation of a more mathematical aspect, it is the weak formulation of the stochastic optimal control problem. Unlike in the strong formulation the filtered probability space $\left(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\in[0,T]}, \mathbb{P}\right)$ on which we define the Brownian motion $W(\cdot)$ are all fixed, but it is not the case in the weak formulation, where we consider them as a parts of the control.

Definition 1.2. A 6-tuple $\left(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\in[0,T]}, \mathbb{P}, W(\cdot), u(\cdot)\right)$ is called weak-admissible control and $(x(\cdot), u(\cdot))$ an weak admissible pair, if

Définition 1.1.1 **1.** $\left(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\in[0,T]}, \mathbb{P}\right)$ *is a filtered probability space satisfying the usual conditions ;*

2. $W(\cdot)$ *is an d-dimensional standard Brownian motion defined on* $\left(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\in[0,T]}, \mathbb{P}\right)$;

3. $u(\cdot)$ *is an* $\{\mathcal{F}_t\}_{t\in[0,T]}-$*adapted process on* $(\Omega, \mathcal{F}, \mathbb{P})$ *taking values in* U *;*

4. $x(\cdot)$ *is the unique solution of equation (1.1),*

5. $l(\cdot, x(\cdot), u(\cdot)) \in \mathbb{L}^1_{\mathcal{F}}([0,T];\mathbb{R})$ *and* $g(x(T)) \in \mathbb{L}^1_{\mathcal{F}}(\Omega;\mathbb{R})$.

The set of all weak admissible controls is denoted by $\mathcal{U}^w([0,T])$. Sometimes, might write $u(\cdot)) \in \mathcal{U}^w([0,T])$ instead of $\left(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\in[0,T]}, \mathbb{P}, W(\cdot), u(\cdot)\right) \in \mathcal{U}^w([0,T])$.

Our stochastic optimal control problem under weak formulation can be formulated as follows :

Problem 1.2. The objective is to minimize the cost functional given by equation (1.2) over the of admissible controls $\mathcal{U}^w([0,T])$. Namely, one seeks $v^*(\cdot) = \left(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\in[0,T]}, \mathbb{P}, W(\cdot), u(\cdot)\right) \in \mathcal{U}^w([0,T])$ such that

$$J(v^*(\cdot)) = \inf_{v(\cdot)\in\mathcal{U}^w([0,T])} J(v(\cdot)).$$

1.2 Methods to solving optimal control problem

In optimal control problems, two major tools for studing optimal control are Pontryagin's maximum principle and Bellman's dynamic programming method.

1.2.1 The Dynamic Programming (*Bellman* Principle)

We present an approach to solving optimal control problems, namely, the method of dynamic programming. Dynamic programming, originated by R. Bellman (*Bellman, R. : Dynamic programming, Princeton Univ. Press., (1957)*) is a mathematical technique for making a sequence of interrelated decisions, which can be applied to many optimization problems (including optimal control problems). The basic idea of this method applied to optimal controls is to consider a family of optimal control problems with different initial times and states, to establish relationships among these problems via the so-called Hamilton-Jacobi-Bellman equation (HJB, for short), which is a nonlinear first-order (in the deterministic case) or second-order (in the stochastic case) partial differential equation. If the HJB equation is solvable (either analytically or numerically), then one can obtain an optimal feedback control by taking the maximize/minimize of the Hamiltonian or generalized Hamiltonian involved in the HJB equation. This is the so-called verification technique. Note that this approach actually gives solutions to the whole family of problems (with different initial times and states).

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space with filtration $\{\mathcal{F}_t\}_{t\in[0,T]}$, satisfying the usual conditions, $T > 0$ a finite time, and W a d-dimensional Brownian motion defined on the filtered probability space $\left(\Omega, \mathcal{F}, \mathbb{P}, \{\mathcal{F}_t\}_{t\in[0,T]}\right)$.

The Bellman dynamic programming principle. We consider the following stochastic differential equation

$$dx(s) = f(s, x(s), u(s))ds + \sigma(s, x(s), u(s))dW(s), \ s \in [0, T]. \tag{1.4}$$

The control $u = u(s)_{0\leq s\leq T}$ is a progressively measurable process valued in the control set U, a subset of $\mathbb{R}^k$, satisfies a square integrability condition. We denote by $\mathcal{U}([t, T])$ the set of control processes u.

Conditions. To ensure the existence of the solution to SDE-(1.4), the Borelian functions

$$f : [0, T] \times \mathbb{R}^n \times U \longrightarrow \mathbb{R}^n$$

$$\sigma : [0, T] \times \mathbb{R}^n \times U \longrightarrow \mathbb{R}^{n\times d}$$

satisfy the following conditions :

$$|f(t, x, u) - f(t, y, u)| + |\sigma(t, x, u) - \sigma(t, y, u)| \leq C\,|x - y|\,,$$

$$|f(t, x, u)| + |\sigma(t, x, u)| \leq C\,[1 + |x|]\,,$$

for some constant $C > 0$. We define the gain function as follows :

$$J(t, x, u) = \boldsymbol{E}\left[\int_t^T l(s, x(s), u(s))ds + g(x\,(T))\right], \tag{1.5}$$

where

$$l : [0, T] \times \mathbb{R}^n \times U \longrightarrow \mathbb{R},$$

$$g : \mathbb{R}^n \longrightarrow \mathbb{R},$$

be given functions. We have to impose integrability conditions on f and g in order for the above expectation to be well-defined, e.g. a lower boundedness or quadratic growth condition. The objective is to maximize this gain function. We introduce the so-called value function :

$$V(t,x) = \sup_{u\in\mathcal{U}([t,T])} J(t,x,u), \tag{1.6}$$

where $x(t)=x$ is the initial state given at time t. For an initial state (t,x), we say that $u^*\in\mathcal{U}([t,T])$ is an optimal control if

$$V(t,x) = J(t,x,u^*).$$

Theorem 1.1. Let $(t,x)\in[0,T]\times\mathbb{R}^n$ be given. Then we have

$$V(t,x) = \sup_{u\in\mathcal{U}([t,T])} \boldsymbol{E}\left[\int_t^{t+h} l(s,x(s),u(s))dt + V(t+h,x(t+h))\right], \text{ for } t\leq t+h\leq T. \tag{1.7}$$

Proof. The proof of the dynamic programming principle is technical and has been studied by different methods, we refer the reader to Yong and Zhou [92].

The Hamilton-Jacobi-Bellman equation. The HJB equation is the infinitesimal version of the dynamic programming principle. It is formally derived by assuming that the value function is $C^{1,2}([0,T]\times\mathbb{R}^n)$, applying Itô's formula to $V(s,x^{t,x}(s))$ between $s=t$ and $s=t+h$, and then sending h to zero into (1.6). The classical HJB equation associated to the stochastic control problem (1.6) is

$$-V_t(t,x) - \sup_{u\in U}\left[\mathcal{L}^u V(t,x) + l(t,x,u)\right] = 0, \text{ on } [0,T]\times\mathbb{R}^n, \tag{1.8}$$

where $\mathcal{L}^u$ is the second-order infinitesimal generator associated to the diffusion x with control u

$$\mathcal{L}^u V = f(x,u).D_x V + \frac{1}{2} tr\left(\sigma(x,u)\,\sigma^{\intercal}(x,u)\,D_x^2 V\right).$$

This partial differential equation (PDE) is often written also as :

$$-V_t(t,x) - H(t,x,D_xV(t,x),D_x^2V(t,x)) \;=\; 0, \quad \forall (t,x) \in [0,T] \times \mathbb{R}^n, \tag{1.9}$$

where for $(t,x,\Psi,Q) \in [0,T] \times \mathbb{R}^n \times \mathbb{R}^n \times \mathcal{S}_n$ ($\mathcal{S}_n$ is the set of symmetric $n \times n$ matrices) :

$$H(t,x,\Psi,Q) \;=\; \sup_{u\in U}\left[f(t,x,u).\Psi + \frac{1}{2}tr\left(\sigma\sigma^{\intercal}\left(t,x,u\right)Q\right) + l(t,x,u)\right]. \tag{1.10}$$

The function H is sometimes called Hamiltonian of the associated control problem, and the PDE (1.8) or (1.9) is the dynamic programming or HJB equation.
There is also an a priori terminal condition :

$$V(T,x) \;=\; g(x), \ \forall x \in \mathbb{R}^n,$$

which results from the very definition of the value function V.

The classical verification approach The classical verification approach consists in finding a smooth solution to the HJB equation, and to check that this candidate, under suitable sufficient conditions, coincides with the value function. This result is usually called a verification theorem and provides as a byproduct an optimal control. It relies mainly on Itô's formula. The assertions of a verification theorem may slightly vary from problem to problem, depending on the required sufficient technical conditions. These conditions should actually be adapted to the context of the considered problem. In the above context, a verification theorem is roughly stated as follows :

Theorem 1.2. Let W be a $C^{1,2}$ function on $[0,T]\times\mathbb{R}^n$ and continuous in T, with suitable growth condition. Suppose that for all $(t,x) \in [0,T] \times \mathbb{R}^n$, there exists $u^*(t,x)$ mesurable, valued in U such that W solves the HJB equation :

$$\begin{aligned} 0 &= -W_t(t,x) - \sup_{u\in U}\left[\mathcal{L}^u W(t,x) + l(t,x,u)\right] \\ &= -W_t(t,x) - \mathcal{L}^{u^*(t,x)} W(t,x) - l(t,x,u^*(t,x)), \text{ on } [0,T]\times\mathbb{R}^n, \end{aligned}$$

together with the terminal condition $W(T,\cdot) = g$ on $\mathbb{R}^n$, and the stochastic differential equation :

$$dx(s) = f(s,x(s),u^*(s,x(s)))ds + \sigma(s,x(s),u^*(s,x(s)))dW(t),$$

admits a unique solution x^*, given an initial condition $x(t) = x$. Then, $W = V$ and $u^*(s,x^*)$ is an optimal control for $V(t,x)$.

A proof of this verification theorem can be found in book, by Yong & Zhou [92].

1.2.2 The pontryagin type stochastic maximum principle

The pioneering works on the stochastic maximum principle were written by Kushner [56, 57]. Since then there have been a lot of works on this subject, among them, in particular, those by Bensoussan [16], Peng [76], and so on. The stochastic maximum principle gives some necessary conditions for optimality for a stochastic optimal control problem. The original version of Pontryagin's maximum principle was first introduced for deterministic control problems in the 1960's by Pontryagin et al. (*Pontryagin,L.S., Boltyanski,V.G., Gamkrelidze, R.V., Mischenko, E.F.*) [1] as in classical calculus of variation. The basic idea is to perturbe an optimal control and to use some sort of Taylor expansion of the state trajectory around the optimal control, by sending the perturbation to zero, one obtains some inequality, and by duality.

The deterministic maximum principle. As an illustration, we present here how the

[1] Pontryagin,L.S., Boltyanski,V.G., Gamkrelidze,R.V., Mischenko, E.F. Mathematical Theory of Optimal Processes, Wiley, New York, 1962.

maximum principle for a deterministic control problem is derived. In this setting, the state of the system is given by the ordinary differential equation (ODE) of the form

$$\begin{cases} dx(t) &= f(t, x(t), u(t))dt, \ t \in [0, T], \\ x(0) &= x_0, \end{cases} \tag{1.11}$$

where

$$f : [0, T] \times \mathbb{R} \times \mathcal{A} \longrightarrow \mathbb{R},$$

and the action space $\mathcal{A}$ is some subset of $\mathbb{R}$. The objective is to minimize some cost function of the form :

$$J(u(\cdot)) = \int_0^T l(t, x(t), u(t)) + g(x(T)), \tag{1.12}$$

where

$$l : [0, T] \times \mathbb{R} \times \mathcal{A} \longrightarrow \mathbb{R},$$
$$g : \mathbb{R} \longrightarrow \mathbb{R}.$$

That is, the function l inflicts a running cost and the function g inflicts a terminal cost. We now assume that there exists a control $u^*(t)$ which is optimal, i.e.

$$J(u^*(\cdot)) = \inf_u J(u(\cdot)).$$

We denote by $x^*(t)$ the solution to (1.11) with the optimal control $u^*(t)$. We are going to derive necessary conditions for optimality, for this we make small perturbation of the optimal control. Therefore we introduce a so-called spike variation, i.e. a control which is equal to u^* except on some small time interval :

$$u^{\varepsilon}(t) = \begin{cases} v & \text{for } \tau - \varepsilon \leq t \leq \tau, \\ u^*(t) & \text{otherwise.} \end{cases} \tag{1.13}$$

We denote by $x^{\varepsilon}(t)$ the solution to (1.11) with the control $u^{\varepsilon}(t)$. We set that $x^*(t)$ and $x^{\varepsilon}(t)$ are equal up to $t = \tau - \varepsilon$ and that

$$\begin{aligned} x^{\varepsilon}(\tau) - x^*(\tau) &= (f(\tau, x^{\varepsilon}(\tau), v) - f(\tau, x^*(\tau), u^*(\tau)))\varepsilon + o(\varepsilon) \\ &= (f(\tau, x^*(\tau), v) - f(\tau, x^*(\tau), u^*(\tau)))\varepsilon + o(\varepsilon), \end{aligned} \tag{1.14}$$

where the second equality holds since $x^{\varepsilon}(\tau) - x^*(\tau)$ is of order ε. We look at the Taylor expansion of the state with respect to ε. Let

$$z(t) = \frac{\partial}{\partial \varepsilon} x^{\varepsilon}(t) \mid_{\varepsilon=0},$$

i.e. the Taylor expansion of $x^{\varepsilon}(t)$ is

$$x^{\varepsilon}(t) = x^*(t) + z(t)\varepsilon + o(\varepsilon). \tag{1.15}$$

Then, by (1.14)

$$z(\tau) = f(\tau, x^*(\tau), v) - f(\tau, x^*(\tau), u^*(\tau)). \tag{1.16}$$

Moreover, we can derive the following differential equation for $z(t)$.

$$\begin{aligned} dz(t) &= \frac{\partial}{\partial \varepsilon} dx^{\varepsilon}(t) \mid_{\varepsilon=0} \\ &= \frac{\partial}{\partial \varepsilon} f(t, x^{\varepsilon}(t), u^{\varepsilon}(t))dt \mid_{\varepsilon=0} \\ &= f_x(t, x^{\varepsilon}(t), u^{\varepsilon}(t)) \frac{\partial}{\partial \varepsilon} x^{\varepsilon}(t)dt \mid_{\varepsilon=0} \\ &= f_x(t, x^*(t), u^*(t))z(t)dt, \end{aligned}$$

where f_x denotes the derivative of f with respect to x. If we for the moment assume that $l=0$, the optimality of $u^*(t)$ leads to the inequality

$$
\begin{aligned}
0 \leq \frac{\partial}{\partial \varepsilon} J(u^\varepsilon)\Big|_{\varepsilon=0} &= \frac{\partial}{\partial \varepsilon} g\left(x^\varepsilon(T)\right)|_{\varepsilon=0} \\
&= g_x\left(x^\varepsilon(T)\right) \frac{\partial}{\partial \varepsilon} x^\varepsilon(T)\,|_{\varepsilon=0} \\
&= g_x\left(x^*(T)\right) z(T).
\end{aligned}
$$

We shall use duality to obtain a more explicit necessary condition from this. To this end we introduce the adjoint equation :

$$
\begin{cases}
d\Psi(t) &= -f_x(t, x^*(t), u^*(t))\Psi(t)dt, t \in [0, T], \\
\Psi(T) &= g_x(x^*(T)).
\end{cases}
$$

Then it follows that

$$
d(\Psi(t)z(t)) = 0,
$$

i.e. $\Psi(t)z(t)) = \text{constant}$. By the terminal condition for the adjoint equation we have

$$
\Psi(t)z(t) = g_x(x^*(T))z(T) \geq 0, \text{ for all } 0 \leq t \leq T.
$$

In particular, by (1.16)

$$
\Psi(\tau)\left(f(\tau, x^*(\tau), v) - f(\tau, x^*(\tau), u^*(\tau))\right) \geq 0.
$$

Since τ was chosen arbitrarily, this is equivalent to

$$
\Psi(t)f(t, x^*(t), u^*(t)) = \inf_{v \in \mathcal{U}} \Psi(t)f(t, x^*(t), v), \text{ for all } 0 \leq t \leq T.
$$

By repeating the calculations above for this two-dimensional system, one can derive the necessary condition

$$H(t,x^*(t),u^*(t),\Psi(t)) = \inf_{v\in\mathcal{U}} H(t,x^*(t),v,\Psi(t)) \text{ for all } 0\leq t\leq T, \tag{1.17}$$

where H is the so-called Hamiltonian (sometimes defined with a minus sign which turns the minimum condition above into a maximum condition) :

$$H(x,u,\Psi) = l(x,u)+\Psi f(x,u),$$

and the adjoint equation is given by

$$\begin{cases} d\Psi(t) = -(l_x(t,x^*(t),u^*(t))+f_x(t,x^*(t),u^*(t))\Psi(t))dt, \\ \Psi(T) = g_x(x^*(T)). \end{cases} \tag{1.18}$$

The minimum condition (1.17) together with the adjoint equation (1.18) specifies the Hamiltonian system for our control problem.

The stochastic maximum principle. Stochastic control is the extension of optimal control to problems where it is of importance to take into account some uncertainty in the system. One possibility is then to replace the differential equation by an SDE :

$$dx(t) = f(t,x(t),u(t))dt+\sigma(t,x(t))dW(t), t\in[0,T], \tag{1.19}$$

where f and σ are deterministic functions and the last term is an Itô integral with respect to a Brownian motion W defined on a probability space $\left(\Omega,\mathcal{F},\{\mathcal{F}_t\}_{t\in[0,T]},\mathbb{P}\right)$.
More generally, the diffusion coefficient σ may has an explicit dependence on the control : $t\in[0,T]$.

$$dx(t) = f(t,x(t),u(t))dt+\sigma(t,x(t),u(t))dW(t), \tag{1.20}$$

The cost function for the stochastic case is the expected value of the cost function (1.12), i.e. we want to minimize

$$J(u(\cdot)) = \boldsymbol{E}\left[\int_0^T l(t, x(t), u(t)) + g(x(T))\right].$$

For the case (1.19) the adjoint equation is given by the following Backward SDE :

$$\begin{cases} -d\Psi(t) & = & \{f_x(t, x^*(t), u^*(t))\Psi(t) + \sigma_x(t, x^*(t))Q(t) \\ & & +(l_x(t, x^*(t), u^*(t))\}dt - Q(t)dW(t), \\ \Psi(T) & = & g_x(x^*(T)). \end{cases} \tag{1.21}$$

A solution to this backward SDE is a pair $(\Psi(t), Q(t))$ which fulfills (1.21). The Hamiltonian is

$$H(x, u, \Psi(t), Q(t)) \quad = \quad l(t, x, u) + \Psi(t)f(t, x, u) + Q(t)\sigma(t, x),$$

and the maximum principle reads for all $0 \leq t \leq T$,

$$H(t, x^*(t), u^*(t), \Psi(t), Q(t)) \quad = \quad \inf_{u \in \mathcal{U}} H(t, x^*(t), u, \Psi(t), Q(t)) \ \ \mathbb{P} - \text{a.s.} \tag{1.22}$$

Noting that there is also third case : if the state is given by (1.20) but the action space $\mathcal{A}$ is assumed to be convex, it is possible to derive the maximum principle in a local form. This is accomplished by using a convex perturbation of the control instead of a spike variation, see Bensoussan 1983 [16]. The necessary condition for optimality is then given by the following : for all $0 \leq t \leq T$

$$\boldsymbol{E}\int_0^T H_u(t, x^*(t), u^*(t), \Psi^*(t), Q^*(t))\,(u - u^*(t))\,dt \geq 0.$$

1.3 Control classes

Let $(\Omega, \mathcal{F}, \mathcal{F}_{t\geq 0}, P)$ be a complete filtred probability space.

1. Admissible control An admissible control is $\mathcal{F}_t$-adapted process $u(t)$ with values in a borelian $A \subset \mathbb{R}^n$

$$\mathcal{U} := \{u(\cdot) : [0,T] \times \Omega \to A : u(t) \text{ is } \mathcal{F}_t\text{-adapted}\}. \tag{1.23}$$

2. Optimal control The optimal control problem consists to minimize a cost functional $J(u)$ over the set of admissible control $\mathcal{U}$. We say that the control $u^*(\cdot)$ is an optimal control if

$$J(u^*(t)) \leq J(u(t)), \text{ for all } u(\cdot) \in \mathcal{U}.$$

3. Near-optimal control Let $\varepsilon > 0$, a control $u^\varepsilon(\cdot)$ is a near-optimal control (or ε-optimal) if for all control $u(\cdot) \in \mathcal{U}$ we have

$$J(u^\varepsilon(t)) \leq J(u(t)) + \varepsilon. \tag{1.24}$$

See for some applications.

4. Singular control. An admissible control is a pair $(u(\cdot), \xi(\cdot))$ of measurable $\mathbb{A}_1 \times \mathbb{A}_2-$valued, $\mathcal{F}_t-$adapted processes, such that $\xi(\cdot)$ is of bounded variation, non-decreasing continuous on the left with right limits and $\xi(0_-) = 0$. Since $d\xi(t)$ may be singular with respect to Lebesgue measure dt, we call $\xi(\cdot)$ the singular part of the control and the process $u(\cdot)$ its absolutely continuous part.

5. Feedback control : We say that $u(\cdot)$ is a feedback control if $u(\cdot)$ depends on the state variable $X(\cdot)$.

If $\mathcal{F}_t^X$ the natural filtration generated by the process X, then $u(\cdot)$ is a feedback control if $u(\cdot)$ is $\mathcal{F}_t^X-$adapted.

6. Robust control. In the problems formulated above, the dynamics of the control system is assumed to be known and fixed. Robust control theory is a method to measure the performance changes of a control system with changing system parameters. This is of course important in engineering systems, and it has recently been used in finance in relation with the theory of risk measure.

Indeed, it is proved that a coherent risk measure for an uncertain payoff $x(T)$ at time T is represented by :

$$\rho(-X(t)) = \sup_{Q\in\mathcal{M}} \boldsymbol{E}^{Q}(X(T)),$$

where $\mathcal{M}$ is a set of absolutly continuous probability measures with respect to the original probability P.

7. Partial observation control problem It is assumed so far that the controller completely observes the state system. In many real applications, he is only able to observe partially the state via other variables (called observed variable) and there is noise in the observation system. For example in financial models, one may observe the asset price but not completely its rate of return and/or its volatility, and the portfolio investment is based only on the asset price information. This may be formulated in a general form as follows : we have a controlled (unobserved) process governed by the following SDE :

$$dx\left(t\right) = f\left(t, x\left(t\right), y(t), u\left(t\right)\right) dt + \sigma\left(t, x\left(t\right), y\left(t\right), u\left(t\right)\right) dW\left(t\right),$$

and $y\left(t\right)$ an observation process defined by

$$dy\left(t\right) = h\left(t, x\left(t\right), u\left(t\right)\right) dW\left(t\right),$$

where $B\left(t\right)$ is another Brownian motion, eventually correlated with $W\left(t\right)$. The control $u(t)$ is adapted with respect to the filtration generated by the observation F_t^Y and the cost functional to optimize is :

$$J\left(u\left(\cdot\right)\right) = \boldsymbol{E}\left[h\left(x\left(T\right), y(T)\right) + \int_0^T g\left(t, x\left(t\right), y(t), u\left(t\right)\right) dt\right].$$

8. Ergodic control Some stochastic systems may exhibit over a long period a stationary behavior characterized by an invariant measure. This measure, if it does exists, is obtained by the average of the states over a long time. An ergodic control problem consists in optimizing over the long term some criterion taking into account this invariant measure. (See Pham [75], Borkar [18]). The cost functional is given by

$$\lim \sup_{T\to+\infty} \frac{1}{T}\boldsymbol{E}\int_0^T f(x(t),u(t))dt.$$

9. Random horizon In classicla problem, the time horizon is fixed until a deterministic terminal time T. In some real applications, the time horizon may be random, the cost functional is given by the following :

$$J\left(u\left(\cdot\right)\right)=\boldsymbol{E}\left[h\left(x\left(\tau\right)\right)+\int_0^{\tau} g\left(t,x\left(t\right),y(t),u\left(t\right)\right)dt\right],$$

where τ s a finite random time.

10. Relaxed control The idea is then to compactify the space of controls $\mathcal{U}$ by extending the definition of controls to include the space of probability measures on U. The set of relaxed controls $\mu_t\left(du\right)dt$, where μ_t is a probability measure, is the closure under weak* topology of the measures $\delta_{u(t)}(du)dt$ corresponding to usual, or strict, controls. This notion of relaxed control is introduced for deterministic optimal control problems in Young (*Young, L.C. Lectures on the calculus of variations and optimal control theory, W.B. Saunders Co., 1969.*) (See Borkar [18]).

11. Impulsive control. Impulse control : Here one is allowed to reset the trajectory at stopping times τ_i from $X_{\tau_{i-}}$ (the value immediately before i) to a new (non-anticipative) value X_{τ_i}, resp., with an associated cost $M\left(X_{\tau_{i-}},X_{\tau_i}\right)$. The aim of the controller is to minimizes the cost functional :

$$\begin{aligned}&\boldsymbol{E}\int_0^T \exp\left[-\int_0^t C(X(s),u(s))ds\right]K(X(t),u(t))\\&+\sum_{\tau_i<T}\exp\left[-\int_0^{\tau_i} C(X(s),u(s))ds\right]M(X_\tau,X_{\tau_{i-}})\\&+\exp\left[-\int_0^{\tau_i} C(X(s),u(s))ds\right]h(X(T)).\end{aligned}$$

In this model, we should assume that $M(X_\tau, X_{\tau_{i-}}) > \delta$ for some $\delta > 0$ to avoid infinitely many jumps in a finite time interval. Some recent examples and applications on control classes can be found in [18], [51], [75] and [92].

Chapitre 2

Partially observed optimal singular control of McKean-Vlasov stochastic systems

2.1 Introduction

2.1.1 Singular optimal control problems

Stochastic singular control problems have received considerable attention in the literature. The first version of maximum principle for stochastic singular control problem was obtained by Cadenillas and Haussmann [19]. Stochastic maximum principle where the singular part has a linear form was proved by Dufour and Miller [24]. Sufficient conditions for existence of optimal singular control and the connection between the singular control and optimal stopping problems have been investigated by Dufour and Miller [25]. Necessary conditions for general optimal singular stochastic control problems have been derived by Dufour and Miller [26]. Maximum principle for optimal stochastic singular stochastic control was investigated by many authors. Under partial-information, optimal singular control problem for mean-field stochastic differential equations driven by Teugels martingales measures has been studied in Hafayed et al. [48]. Necessary and sufficient conditions for near-optimal McKean-Vlasov stochastic singular control have been studied in [33]. The first-order local maximum principle for singular optimal control for mean-field SDEs has been derived in Hafayed [39]. Maximum principle for optimal singular control problem for general controlled nonlinear McKean-Vlasov SDEs has been obtained by Hafayed et al [32]. A class of solvable singular stochastic control problems have been studied in Alvarez [6]. Singular stochastic control problem for linear diffusions and optimal stopping have been derived by Alvarez [4]. An extensive list of references to the stochastic singular control problem, called also *intervention control*, in which the optimal control has both absolutely continuous and singular components, with some applications in finance and economics can be found in [33, 55, 69]. Some recent examples on singular stochastic control have been investigated by Shreve [80].

2.1.2 Partially observed stochastic control problem

With the development of nonlinear filtering theory, partially observed stochastic control problem has been one of the most important and well established topics in control theory. Maximum principle for partially observed control problems have received much attention and became a powerful tool in many recent fields, such as mathematical finance, optimal control, etc. From the viewpoint of reality, many situations, full information is not always available to controllers, but the partial one with noise, see e.g , Fleming [29], Bensoussan [16], Baras, Elliott and Kohlmann, [9] and the references therein for the explanation. The necessary conditions of optimality for forward-backward stochastic control systems with correlated state and observation noise have been obtained by Wang, Wu and Xiong [83]. A class of linear-quadratic optimal control problem of forward-backward stochastic differential equations with partial information has been studied by Wang, Wu and Xiong [84]. General maximum principles for partially observed risk-sensitive optimal control problems with some applications to finance have been studied by Wang and Wu [85]. Recently, maximum principle for mean-field optimal stochastic control with partial-information has been discussed in Wang, Zhang, and Zhang [86]. An optimal control problem for systems governed by mean-field forward-backward stochastic differential equation with noisy observation has been studied by Wang, Xiao and Xing [87]. In a recent paper [88], Wang and Wu established a maximum principle for mean-field stochastic control system, where the state is partially observed via a noisy process. Risk sensitive mean-field type control problem under partial observation has been studied by Djehiche and Tempine [23]. Partially observed optimal control problem for forward-backward stochastic systems with jump has been investugated by Wang, Shi and Meng [89].

McKean-Vlasov dynamics are Itô's stochastic differential equations (SDEs), where the coefficients of the state equation depend on the state of the solution process as well as of its probability law. This kind of equations was studied by Kac [58] as a stochastic model for the Vlasov-Kinetic equation of plasma and the study of which was initiated by McKean [65] to provide a rigorous treatment of special nonlinear partial differential equations. Optimal control problems for McKean-Vlasov SDEs have been investigated by many authors, for example, Buckdahn, Li and Ma, [14] proved the necessary conditions for general mean-field systems by applying second order derivatives with respect to measures. Maximum principle for optimal control of McKean-Vlasov forward-backward stochastic differential equations (FBSDEs) with Lévy process via the differentiability with respect to probability law has been proved by Meherrem and Hafayed [67]. Necessary and sufficient optimality conditions of optimal singular control problem for general Mckean-Vlasov differential equations have been discussed by Hafayed et al., [50]. Maximum principle for stochastic continuous-singular control of McKean-Vlasov type systems, where the control domain is not assumed convex has been proved by Guenane et al., [31]. Necessary conditions for optimal partially observed control problems of general controlled mean-field differential systems have been established by Lakhdari, Miloudi and Hafayed [59]. Necessary conditions for partially observed optimal control of general McKean-Vlasov dynamics with Poisson jumps have been studied in Miloudi et al [68]. A necessary condition for mean-field type stochastic differential equations with correlated state and observation noises have been obtained in Zhang [52].

In this chaptre, we establish a set of necessary conditions in the form of stochastic maximum principle for partially observed optimal singular control problems of McKean-Vlasov type. The stochastic system under consideration is governed by Itô stochastic differential equation of general McKean-Vlasov type, with correlated noises between the system and the observation allowing both classical and singular control. The coefficients of our McKean-Vlasov dynamic depend nonlinearly on both the state process as well as of its probability law. The derivatives with respect to probability measure in P.L Lions' sense and the associate Itô-formula are applied to derive our main results. Since the control domain is assumed to be convex, the proof of our partially observed maximum principle based on convex perturbation for both continuous and singular parts of the control process, and Girsanov's theorem.

Our general McKean-Vlasov partially observed singular control problem occur naturally in the probabilistic analysis of financial optimization problems. Moreover, the changes of probability measure are the cornerstone of the rational pricing of derivatives and are used for converting actual probabilities into those of the risk-neutral probabilities.

Our class of partially observed singular control problem is strongly motivated by the recent study of the McKean-Vlasov games and recently play an important role in different fields of economics and finance with an intervention controls. As an illustration, by applying our partially observed maximum principle, McKean-Vlasov type linear quadratic singular control problem is discussed, where the partially observed optimal singular control is established explicitly in feedback form.

This chapter is organized as follows. Sect. 2 begins with a formulation of the partially observed singular stochastic control problem. We give the notations and definitions of the derivatives with respect to probability measure and assumptions used throughout the paper. In Sect. 3, we prove the necessary conditions of optimality which are our main results. A linear quatratic control problem of this kind of partially observed control problem is also given in Sect. 4. At the end of this chapter, some discussions with concluding remarks and future developments are presented in the last Section.

2.2 Assumptions and statement of the control problem

Let us formulate the optimal mixed control. Let T be a fixed strictly positive real number and $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ be a complete filtered probability space satisfying the usual conditions in which one-dimentional Brownian motion $W(t) = \{W(t) : 0 \leq t \leq T\}$ and $W(0) = 0$ is defined, where $\mathbb{F} = (\mathcal{F}_t^W)_{t\in[0,T]}$ is the natural filtration generated by $W(\cdot)$,

$$\mathcal{F}_t^W = \sigma\left\{W(s) : s \in [0,t]\right\}$$

augmented by all the $\mathbb{P}$-null sets.

Let $\mathbb{A}_1$ be a closed convex and bounded subset of $\mathbb{R}$ and $\mathbb{A}_2 := ([0, +\infty))$. Let $\mathcal{A}_1\left([0,T]\right)$ be the class of $\mathcal{B}([0,T]) \otimes \mathcal{F}$ measurable, $\mathbb{F}-$adapted processes $u(\cdot) : [0,T] \times \Omega \to \mathbb{A}_1$ and $\mathcal{A}_2\left([0,T]\right)$ is the class of $\mathcal{B}([0,T]) \otimes \mathcal{F}$-measurable, $\mathbb{F}-$adapted processes $\xi(\cdot) : [0,T] \times \Omega \to \mathbb{A}_2$.

We give here the precise definition of the complete observed continuous-singular control.

Definition 2.1. An admissible continuous-singular control is a pair $(u(\cdot), \xi(\cdot))$ of measurable $\mathbb{A}_1 \times \mathbb{A}_2-$valued, $\mathcal{F}^W-$adapted processes, such that the process $\xi(\cdot) : [0,T] \times \Omega \to \mathbb{A}_2$ is of bounded variation, non-decreasing continuous on the right with left limits and $\xi(0_-) = 0$.Moreover, $\boldsymbol{E}(|\xi(T)|^2) < \infty$.

Notation. Throughout what follows, $\mathcal{N}$ denotes the totality of $\mathbb{P}$-null sets. We denote by $\langle \cdot, \cdot \rangle$ (resp. $|\cdot|$) the scalar product (resp., norm), $\boldsymbol{E}\left(\cdot\right)$ denotes the expectation on

$(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbb{P})$. Moreover, we denote by

1. $L^2([0,T];\mathbb{R}^n)$ the space of $\mathbb{R}^n$-valued deterministic function $\beta(\cdot)$, such that $\int_0^T |\beta(t)|^2 \mathrm{d}t < +\infty$.

2. $L^2(\mathcal{F}_t;\mathbb{R}^n)$ the space of $\mathbb{R}^n$-valued $\mathcal{F}_t$-measurable random variable X, such that $\boldsymbol{E}\left(|X|^2\right) < +\infty$.

3. $L^2_{\mathcal{F}}([0,T];\mathbb{R}^n)$ the space of $\mathbb{R}^n$-valued $\mathcal{F}_t$-adapted processes X, such that $\boldsymbol{E}\int_0^T |X(t)|^2 \mathrm{d}t < +\infty$.

4. $\mathbb{L}^2\left(\mathcal{F};\mathbb{R}^d\right)$ is the Hilbert space with inner product $(X,Y)_2 = \boldsymbol{E}[X.Y]$, $X,Y \in \mathbb{L}^2\left(\mathcal{F};\mathbb{R}^d\right)$ and the norm $\|X\|_2^2 = (X,X)_2$.

5. $\Gamma_2\left(\mathbb{R}^d\right)$ the space of all probability measures μ on $\left(\mathbb{R}^d, \mathcal{B}\left(\mathbb{R}^d\right)\right)$ with finite second moment, i.e, $\int_{\mathbb{R}^d} |x|^2 \mu(\mathrm{d}x) < +\infty$, endowed with the following 2-Wasserstein metric, for $\mu, \nu \in \Gamma_2\left(\mathbb{R}^d\right)$,

$$\mathbb{W}_2(\mu,\nu) = \inf\left[\int_{\mathbb{R}^d} |x-y|^2 \delta(\mathrm{d}x, \mathrm{d}y)\right]^{\frac{1}{2}}, \tag{2.1}$$

where $\delta \in \Gamma_2\left(\mathbb{R}^{2d}\right)$, $\delta\left(\cdot, \mathbb{R}^d\right) = \mu$, and $\delta\left(\mathbb{R}^d, \cdot\right) = \nu$.

2.2.1 *Differentiability with respect to probability measures*

Now, we recall briefly the main results of the differentiability with respect to probability measures, which have been studied by P.L Lions [61]. The main idea is to identify a distribution $\mu \in \Gamma_2\left(\mathbb{R}^d\right)$ with a random variables $\vartheta \in \mathbb{L}^2\left(\mathcal{F};\mathbb{R}^d\right)$ so that $\mu = \mathbb{P}_\vartheta$. To be more precise, we assume that probability space $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbb{P})$ is rich enough in the sense that for every $\mu \in \Gamma_2\left(\mathbb{R}^d\right)$, there is a random variable $\vartheta \in \mathbb{L}^2\left(\mathcal{F};\mathbb{R}^d\right)$ such that $\mu = \mathbb{P}_\vartheta$, see Buckdahn, Li and Ma [13].

Definition 2.2 (*Lift function*) Let f be a given function such that $f : \Gamma_2\left(\mathbb{R}^d\right) \to \mathbb{R}$. We define the lift function $\widetilde{f} : \mathbb{L}^2\left(\mathcal{F};\mathbb{R}^d\right) \to \mathbb{R}$ such that

$$\widetilde{f}(X) := f \circ \mathbb{P}_X = f(\mathbb{P}_X), \; X \in \mathbb{L}^2(\mathcal{F};\mathbb{R}^d).$$

Clearly, the lift function $\widetilde{f}$ of f, depends only on the law of $Z \in \mathbb{L}^2\left(\mathcal{F};\mathbb{R}^d\right)$ and is independent of the choice of the representative Z.

Definition 2.3 A function $f : \Gamma_2\left(\mathbb{R}^d\right) \to \mathbb{R}$ is said to be differentiable at $\mu_0 \in \Gamma_2\left(\mathbb{R}^d\right)$ if there exists $\vartheta_0 \in \mathbb{L}^2\left(\mathcal{F};\mathbb{R}^d\right)$ with $\mu_0 = \mathbb{P}_{\vartheta_0}$ such that its lift function $\widetilde{f}$ is Fréchet differentiable at ϑ_0. More precisely, there exists a continuous linear functional $D\widetilde{f}\left(\vartheta_0\right)$: $\mathbb{L}^2\left(\mathcal{F};\mathbb{R}^d\right) \to \mathbb{R}$ such that

$$\widetilde{f}\left(\vartheta_0+\beta\right)-\widetilde{f}\left(\vartheta_0\right)=\left\langle D\widetilde{f}\left(\vartheta_0\right),\beta\right\rangle+O\left(\|\beta\|_2\right)=D_\beta f\left(\mu_0\right)+O\left(\|\beta\|_2\right), \tag{2.2}$$

where $\langle\cdot,\cdot\rangle$ is the dual product on $\mathbb{L}^2(\mathcal{F};\mathbb{R}^d)$, and we will refer to $D_\beta f\left(\mu_0\right)$ as the Fréchet derivative of f at μ_0 in the direction β. In this case, for $\mu_0 = \mathbb{P}_{\vartheta_0}$ we have

$$D_\beta f\left(\mu_0\right)=\left\langle D\widetilde{f}\left(\vartheta_0\right),\beta\right\rangle=\left.\frac{\mathrm{d}}{\mathrm{d}t}\widetilde{f}\left(\vartheta_0+t\beta\right)\right|_{t=0}.$$

By applying the Riesz' representation theorem, there is a unique random variable $z_0 \in \mathbb{L}^2\left(\mathcal{F};\mathbb{R}^d\right)$ such that $\left\langle D\widetilde{f}\left(\vartheta_0\right),\beta\right\rangle=\left(z_0,\beta\right)_2=\boldsymbol{E}\left[\left(z_0,\beta\right)_2\right]$, where $\beta \in \mathbb{L}^2\left(\mathcal{F};\mathbb{R}^d\right)$. It was shown, see the works of Buckdahn Li and Ma [13] and Lions [61] that there exists a Boral function $\varphi\left[\mu_0\right] : \mathbb{R}^d \to \mathbb{R}^d$, depending only on the law $\mu_0 = \mathbb{P}_{\vartheta_0}$ but not on the particular choice of the representative ϑ_0 such that $z_0 = \varphi\left[\mu_0\right]\left(\vartheta_0\right)$.

Thus, we can write (2.2) as $\forall\vartheta \in \mathbb{L}^2(\mathcal{F};\mathbb{R}^d)$.

$$f\left(\mathbb{P}_\vartheta\right)-f\left(\mathbb{P}_{\vartheta_0}\right)=\left(\varphi\left[\mu_0\right]\left(\vartheta_0\right),\vartheta-\vartheta_0\right)_2+O\left(\|\vartheta-\vartheta_0\|_2\right).$$

We denote

$$\partial_\mu f\left(\mathbb{P}_{\vartheta_0},x\right)=\varphi\left[\mu_0\right](x),\ x \in \mathbb{R}^d.$$

Moreover, we have the following identities

$$D\widetilde{f}\left(\vartheta_0\right)=z_0=\varphi\left[\mu_0\right]\left(\vartheta_0\right)=\partial_\mu f\left(\mathbb{P}_{\vartheta_0},\vartheta_0\right),$$

and

$$D_{\beta} f\left(\mathbb{P}_{\vartheta_0}\right)=\left\langle\partial_{\mu} f\left(\mathbb{P}_{\vartheta_0}, \vartheta_0\right), \beta\right\rangle,$$

where $\xi=\vartheta-\vartheta_0$. We note that for each $\mu \in \Gamma_2\left(\mathbb{R}^d\right), \partial_{\mu} f\left(\mathbb{P}_{\vartheta}, \cdot\right)=\varphi\left[\mathbb{P}_{\vartheta}\right](\cdot)$ is only defined in a $\mathbb{P}_{\vartheta}(\mathrm{d} x)$-$a.\boldsymbol{E}$ sense, where $\mu=\mathbb{P}_{\vartheta}$.

Definition 2.4 (Space of differentiable functions in $\Gamma_2\left(\mathbb{R}^d\right)$.) We say that the function $f \in \mathbb{C}_b^{1,1}\left(\Gamma_2\left(\mathbb{R}^d\right)\right)$ if for all $\vartheta \in \mathbb{L}^2\left(\mathcal{F} ; \mathbb{R}^d\right)$, there exists a $\mathbb{P}_{\vartheta}$-modification of $\partial_{\mu} f\left(\mathbb{P}_{\vartheta}, \cdot\right)$ such that $\partial_{\mu} f: \Gamma_2\left(\mathbb{R}^d\right) \times \mathbb{R}^d \rightarrow \mathbb{R}^d$ is bounded and Lipchitz continuous. That is for some $C>0$, it holds that

(i) $\left|\partial_{\mu} f(\mu, x)\right| \leq C, \forall \mu \in \Gamma_2\left(\mathbb{R}^d\right), \forall x \in \mathbb{R}^d$;

(ii) $\left|\partial_{\mu} f\left(\mu_1, x_1\right)-\partial_{\mu} f\left(\mu_2, x_2\right)\right| \leq C\left(\mathbb{W}_2(\mu_1, \mu_2)+\left|x_1-x_2\right|\right), \forall \mu_1, \mu_2 \in \Gamma_2\left(\mathbb{R}^d\right), \forall x_1, x_2 \in \mathbb{R}^d$.

We would like to point out that the version of $\partial_{\mu} f\left(\mathbb{P}_{\vartheta}, \cdot\right), \vartheta \in \mathbb{L}^2\left(\mathcal{F} ; \mathbb{R}^d\right)$ indicated in in the above definition is unique (see Remark 2.2 in Buckdahn, Li and Ma [13] for more information).

Let $(\widehat{\Omega}, \widehat{\mathcal{F}}, \widehat{\mathcal{F}}_t, \widehat{\mathbb{P}})$ be a copy of the probability space $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbb{P})$. For any pair of random variable $(\vartheta, \xi) \in \mathbb{L}^2\left(\mathcal{F} ; \mathbb{R}^d\right) \times \mathbb{L}^2\left(\mathcal{F} ; \mathbb{R}^d\right)$, we let $(\widehat{\vartheta}, \widehat{\xi})$ be an independent copy of (ϑ, ξ) defined on $(\widehat{\Omega}, \widehat{\mathcal{F}}, \widehat{\mathcal{F}}_t, \widehat{\mathbb{P}})$. We consider the product probability space $(\Omega \times \widehat{\Omega}, \mathcal{F} \otimes \widehat{\mathcal{F}}, \mathcal{F}_t \otimes \widehat{\mathcal{F}}_t, \mathbb{P} \otimes \widehat{\mathbb{P}})$ and setting $(\widehat{\vartheta}, \widehat{\xi})(W, \widehat{W})=(\vartheta(W), \xi(\widehat{W}))$ for any $(W, \widehat{W}) \in \Omega \times \widehat{\Omega}$. Let $(\widehat{u}(t), \widehat{x}(t))$ be an independent copy of $(u(t), x(t))$ so that $\mathbb{P}_{x(t)}=\widehat{\mathbb{P}}_{\widehat{x}(t)}$.

Throughout this chapter, we denote by $\widehat{\boldsymbol{E}}$ the expectation under probability measure $\widehat{\mathbb{P}}$ and $\mathbb{P}_X=\mathbb{P} \circ X^{-1}$ denotes the law of the random variable X.

2.2.2 Partially observed optimal control Model

In this chaptre, we study partially observed optimal stochastic singular control problem of general Mckean-Vlasov type with correlated noises between the system and the observation. The control variable has two components, the first being absolutely continuous and the second is a bounded variation, non decreasing continuous on the right with left limits. The dynamic system is governed by Itô-type controlled stochastic differential equation. The coefficients of the dynamic depend on the state process as well as of its probability law and the continuous control variable. In terms of a classical convex variational techniques, we establish a set of necessary contiditions of optimal singular control in the form of maximum principle. Our main result is proved by applying Girsanov's theorem and the derivatives with respect to probability law in P.L. Lions' sense. To illustrate our theoretical result, we study partially observed linear quadratic singular control problem of McKean-Vlasov type.

We consider the partially observed optimal stochastic singular control problem for systems governed by nonlinear controlled McKean-Vlasov stochastic differential equations (SDEs) with correlated noisy between the system and the observation, allowing both classical and singular control of the form : $t \in [0,T]$

$$\begin{cases} \mathrm{d}x^{v,\xi}(t) = & f(t, x^{v,\xi}(t), \mathbb{P}_{x^{v,\xi}(t)}, v(t))\mathrm{d}t + \sigma(t, x^{v,\xi}(t), \mathbb{P}_{x^{v,\xi}(t)}, v(t))\mathrm{d}W(t) \\ & +g(t, x^{v,\xi}(t), \mathbb{P}_{x^{v,\xi}(t)}, v(t))\mathrm{d}\widetilde{W}(t) + G(t)\mathrm{d}\xi(t), \\ x^{v,\xi}(0) = & x_0, \quad t \in [0,T], \end{cases} \tag{2.3}$$

where $\mathbb{P}_{x^{v,\xi}} = \mathbb{P}\circ \left(x^{v,\xi}\right)^{-1}$ denotes the law of the random variable $x^{v,\xi}(\cdot)$. The coefficients $f : [0,T] \times \mathbb{R}^n \times \Gamma_2\left(\mathbb{R}^d\right) \times \mathbb{A}_1 \to \mathbb{R}^n$, $\sigma : [0,T] \times \mathbb{R}^n \times \Gamma_2\left(\mathbb{R}^d\right) \times \mathbb{A}_1 \to \mathbb{R}^{n\times d}$, $g : [0,T] \times \mathbb{R}^n \times \Gamma_2\left(\mathbb{R}^d\right) \times \mathbb{A}_1 \to \mathbb{R}^{n\times d}$, and $G(\cdot) : [0,T] \times \Omega \to \mathbb{R}$ are given functions.
We assume that the state process $x^{v,\xi}(\cdot)$ cannot be observed directly, but the controllers can observe a related noisy process $Y(\cdot)$, which is governed by the following equation :

$$\begin{cases} dY(t) & = h(t, x^{v,\xi}(t), v(t))dt + \mathrm{d}\widetilde{W}(t) \\ Y(0) & = 0, \end{cases} \tag{2.4}$$

where $h : [0,T] \times \mathbb{R}^n \times \mathbb{A}_1 \to \mathbb{R}^r$ and $\widetilde{W}(\cdot)$ is a stochastic process depending on the control

$v(\cdot)$, and $Y(\cdot)$ the observation process.

We give here the precise definition of the partially observed continuous-singular control.

2.2.3 Partially observed continuous-singular control

Throughout this chapter $\mathcal{F}_t^Y$ is the natural filtration generated by $Y(\cdot)$,

$$\mathcal{F}_t^Y = \sigma \{Y(s) : s \in [0,t]\} .$$

Definition 2.5. Let $\mathcal{A}_1^Y([0,T])$ be the class of $\mathcal{B}([0,T]) \otimes \mathcal{F}$ measurable, $\mathcal{F}_t^Y-$adapted processes $u(\cdot) : [0,T] \times \Omega \to \mathbb{A}_1$ satisfies

$$\sup_{t\in[0,T]} \boldsymbol{E}(|u(t)|^2) < \infty$$

and $\mathcal{A}_2^Y([0,T])$ is the class of $\mathcal{B}([0,T]) \otimes \mathcal{F}$-measurable, $\mathcal{F}_t^Y-$adapted processes $\xi(\cdot)$: $[0,T] \times \Omega \to \mathbb{A}_2$ such that $\xi(\cdot)$ is of bounded variation, non-decreasing continuous on the right with left limits (càdlàg) and $\xi(0_-) = 0$. Moreover, $\boldsymbol{E}(\,|\xi(T)|^2) < \infty$.

Denote by $\mathcal{A}_1^Y \times \mathcal{A}_2^Y([0,T])$ the set of $\mathcal{B}([0,T]) \otimes \mathcal{F}$-measurable and $\mathcal{F}_t^Y$-adapted stochastic processes valued in $\mathbb{A}_1 \times \mathbb{A}_2$. Any $(u(\cdot), \xi(\cdot)) \in \mathcal{A}_1 \times \mathcal{A}_2([0,T])$ is called partially observed admissible control. Notice that the jumps of a singular control $\xi(\cdot)$ at any jumping time τ_j denote by $\Delta\xi(\tau_j) := \xi(\tau_j) - \xi(\tau_{j-})$.

We should note that since $\mathrm{d}\xi(t)$ may be singular with respect to Lebesgue measure $\mathrm{d}t$, we call $\xi(\cdot)$ the singular part of the control variable and the process $u(\cdot)$ its absolutely continuous part.

Our partially observed optimal singular control problem is to minimize the cost functional

$$J(v(\cdot), \xi(\cdot)) = \boldsymbol{E}^{v}\left[\int_0^T l(t, x^{v,\xi}(t), \mathbb{P}_{x^{v,\xi}(t)}, v(t))\mathrm{d}t + \psi(x^{v,\xi}(T), \mathbb{P}_{x^{u,\xi}(T)}) + \int_{[0,T]} M(t)\mathrm{d}\xi(t)\right]. \tag{2.5}$$

Here, $l : [0,T]\times\mathbb{R}^n\times\Gamma_2(\mathbb{R})\times\mathbb{A}_1 \to \mathbb{R}$, $\psi : \mathbb{R}^n\times\Gamma_2(\mathbb{R}) \to \mathbb{R}$ and $M : [0,T]\times\Omega \to ([0,\infty))$. Moreove, $\boldsymbol{E}^{v}(\cdot)$ stands for the mathematical expectation on $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbb{P}^v)$ given by

$$\boldsymbol{E}^{v}(X) = \boldsymbol{E}_{\mathbb{P}^v}(X) = \int_\Omega X(w)\mathrm{d}\mathbb{P}^v(w).$$

The partially observed stochastic optimal control problem considered in this paper is to find a couple of $\mathcal{F}_t^Y-$adapted processes $(u^*(\cdot), \xi^*(\cdot)) \in \mathcal{A}_1^Y \times \mathcal{A}_2^Y([0,T])$ such that

$$J(u^*(\cdot), \xi^*(\cdot)) = \inf_{(u(\cdot),\xi(\cdot))\in\mathcal{A}_1^Y\times\mathcal{A}_2^Y([0,T])} J(u(\cdot), \xi(\cdot)). \tag{2.6}$$

Any partially observed admissible control $(u^*(\cdot), \xi^*(\cdot)) \in \mathcal{A}_1^Y \times \mathcal{A}_2^Y([0,T])$ satisfying (2.6) is called an optimal control. The corresponding state $x^*(\cdot) = x^{u^*,\xi^*}(\cdot)$ is called an partially observed optimal state, and $(x^*(\cdot), u^*(\cdot), \xi^*(\cdot))$ is called an optimal solution of the partially observed control problem (2.3)-(2.5).

In this chapter, the following hypothesis will be in force throughout this paper.

Hypothesis (C1) The maps $f, \sigma, g, l : [0,T]\times\mathbb{R}\times\Gamma_2(\mathbb{R})\times\mathbb{A}_1 \to \mathbb{R}$ and $\psi : \mathbb{R}\times\Gamma_2(\mathbb{R}) \to \mathbb{R}$ are measurable in all variables. Moreover, $f(t,\cdot,\cdot,v)$, $\sigma(t,\cdot,\cdot,v)$, $g(t,\cdot,\cdot,v)$, $l(t,\cdot,\cdot,v) \in \mathbb{C}_b^{1,1}(\mathbb{R}\times\Gamma_2(\mathbb{R}), \mathbb{R})$ and $\psi(\cdot,\cdot) \in \mathbb{C}_b^{1,1}(\mathbb{R}\times\Gamma_2(\mathbb{R}), \mathbb{R})$ for all $v \in \mathbb{A}_1$.

Hypothesis (C2) The functions $\varphi(x,\mu) = f(t,x,\mu,v), \sigma(t,x,\mu,v), g(t,x,\mu,v), l(t,x,\mu,v)$, $\psi(x,\mu)$ satisfies the following properties.

(1) For fixed $x \in \mathbb{R}$ and $\mu \in \Gamma_2(\mathbb{R})$, the function $\varphi(\cdot,\mu) \in \mathbb{C}_b^1(\mathbb{R})$ and $\varphi(x,\cdot) \in \mathbb{C}_b^{1,1}(\Gamma_2(\mathbb{R}^d), \mathbb{R})$.

(**2**) The functions f, σ, c and l are continuously differentiable with respect to control variable v, and all their derivatives are continuous and bounded. All the derivatives φ_x and $\partial_\mu \varphi$, for $\varphi = f, \sigma, g, l, \psi$ are bounded and Lipschitz continuous, with Lipschitz constants independent of $v \in \mathbb{A}_1$.

(**3**) The function h is continuously differentiable in x and continuous in v, its derivatives and h are all uniformly bounded such that

$$\boldsymbol{E}\left(\exp\left[\frac{1}{2}\int_0^t \left|h(s, x^{v,\xi}(s), v(s))\right|^2 \mathrm{d}s\right]\right) < \infty. \tag{2.7}$$

Hypothesis (C3) The functions $G(\cdot) : [0,T] \times \Omega \to \mathbb{R}$, and $M(\cdot) : [0,T] \times \Omega \to \mathbb{R}^+$ are continuous and bounded.

Clearly, hypothesis (C3) allows us to define integrals of the form

$$\int_{[0,T]} G(t)\mathrm{d}\xi(t) \text{ and } \int_{[0,T]} M(t)\mathrm{d}\xi(t).$$

Moreover, under hypothesis (C1), (C2), (C3) and for any $(v(\cdot), \xi(\cdot)) \in \mathcal{A}_1^Y \times \mathcal{A}_2^Y([0,T])$, the McKean-Vlasov system (2.3) admits a unique strong solution, and the cost functional (2.5) is well defined on $\mathcal{A}_1^Y \times \mathcal{A}_2^Y([0,T])$.

We define the $\mathcal{F}_t^Y-$martingale $\rho^v(t)$ which is the solution of the equation

$$\begin{cases} \mathrm{d}\rho^v(t) = \rho^v(t) h\left(t, x^v(t), v(t)\right) \mathrm{d}Y(t), \\ \rho^v(0) = 1. \end{cases} \tag{2.8}$$

This martingale allowed to define a new probability, denoted by $\mathbb{P}^v$ on the space $(\Omega, \mathcal{F})$, to emphasize the fact that it depend on the control $v(\cdot)$. It is given by the Radon-Nikodym derivative :

$$\left.\frac{\mathrm{d}\mathbb{P}^v}{\mathrm{d}\mathbb{P}}\right|_{\mathcal{F}_t^Y} = \rho^v(t). \tag{2.9}$$

From the linear equatiion (2.8), and by a simple computation, we can get

$$\rho^v(t) = \exp\left\{ \int_0^t h(s, x^{v,\xi}(s), v(s))\mathrm{d}Y(s) - \frac{1}{2}\int_0^t \left| h(s, x^{v,\xi}(s), v(s)) \right|^2 \mathrm{d}s \right\}. \tag{2.10}$$

We note that $\boldsymbol{E}^v(X)$ refers to the expected value of X with respect to the probabilily law $\mathbb{P}^v$. Moreover, since $\mathrm{d}\mathbb{P}^v = \rho^v(t)\mathrm{d}\mathbb{P}$, we have

$$\begin{aligned}
\boldsymbol{E}^v(X) &:= \boldsymbol{E}_{\mathbb{P}^v}(X) \\
&= \int_\Omega X(w)\mathrm{d}\mathbb{P}^v(w) \\
&= \int_\Omega X(w)\rho^v(t)\mathrm{d}\mathbb{P}(w) \\
&= \boldsymbol{E}_{\mathbb{P}}(\rho^v(t)X) \\
&= \boldsymbol{E}(\rho^v(t)X).
\end{aligned}$$

Note that the condition in (2.7) is called "*conditions of Novikov*" and equation (2.10) is called "*exponential of Doléan-Dade*". Such changes of probability measure (2.9) are the cornerstone of the rational pricing of derivatives and are used for converting actual probabilities into those of the risk-neutral probabilities.

By applying Itô's formula, we can prove that

$$\sup_{t\in[0,T]} \boldsymbol{E}\left(|\rho^v(t)|^n\right) < +\infty, \ n > 1.$$

Hence, by Girsanov's theorem and hypothesis (C1) and (C2), $\mathbb{P}^v$ is a new probability measure of density $\rho^v(t)$. The process

$$\widetilde{W}(t) = Y(t) - \int_0^t h(s, x^{v,\xi}(s), v(s))\mathrm{d}s,$$

is a standard Brownian motion independent of $W(\cdot)$ and x_0 on the new probability space $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbb{P}^v)$.

By using *Radon-Nikodym derivative* (2.9), and the martingale property of $\rho^v(t)$, the cost functional (2.5) can be written as

$$J(v(\cdot),\xi(\cdot)) = \boldsymbol{E}\left[\int_0^T \rho^v(t) l(t, x^{v,\xi}(t), \mathbb{P}_{x^{u,\xi}(t)}, v(t))\mathrm{d}t + \rho^v(T)\psi(x^{v,\xi}(T), \mathbb{P}_{x^{v,\xi}(T)}) \right. \tag{2.11}$$
$$\left. + \int_{[0,T]} \rho^v(t) M(t)\mathrm{d}\xi(t)\right].$$

So deduce that the first original optimization problem is equivalent to minimizing (2.11) over $(v(\cdot),\xi(\cdot)) \in \mathcal{A}_1^Y \times \mathcal{A}_2^Y([0,T])$, subject to (2.3)-(2.8).

2.3 Necessary conditions for optimal partially observed singular control

Our aim in this section is to establish the necessary conditions of optimality in the form of stochastic maximum principle for our partially observed singular optimal control problem. Our main result is derived by applying the differentiability with respect to probability measure and Girsanov's theorem.

In our study, since the control domain is assumed to be convex, the proof of our partially observed maximum principle based on convex perturbation for both continuous and singular parts of the control process.

Hamiltonian. We define the Hamiltonian function

$$H : [0,T] \times \mathbb{R} \times \Gamma_2(\mathbb{R}) \times \mathbb{A}_1 \times \mathbb{R} \times \mathbb{R} \times \mathbb{R} \times \mathbb{R} \times \mathbb{R} \to \mathbb{R},$$

associated with our control problem by

$$H(t,x,\mu,v,p,q,\overline{q},k) = l(t,x,\mu,v) + f(t,x,\mu,v)p + \sigma(t,x,\mu,v)q$$
$$+ g(t,x,\mu,v)\overline{q} + h(t,x,v)k. \tag{2.12}$$

Adjoint equation. We introduce the adjoint equations involved in the stochastic maximum principle for our singular McKean-Vlasov control problem. The adjoint equation turns out to be a linear McKean-Vlasov BSDE. So for any $(u(\cdot),\xi(\cdot)) \in \mathcal{A}_1^Y \times \mathcal{A}_2^Y$ and the corresponding state trajectory $x(t) = x^{u,\xi}(t)$, we consider the following adjoint equation :

$$\begin{cases} -dy(t) &= l(t)dt - z\,(t)\,\mathrm{d}W(t) - k\,(t)\,\mathrm{d}\widetilde{W}\,(t)\,, \\ y(T) &= \psi(x(T),\mathbb{P}_{x(T)}), \end{cases} \tag{2.13}$$

and

$$\begin{cases} -\mathrm{d}p\,(t) &= \Big\{ f_x\,(t)\,p\,(t) + \widehat{\boldsymbol{E}}(\partial_\mu \widehat{f}\,(t)\,\widehat{p}\,(t)) + \sigma_x\,(t)\,q\,(t) + \widehat{\boldsymbol{E}}(\partial_\mu \widehat{\sigma}\,(t)\,\widehat{q}\,(t)) \\ & \quad + g_x\,(t)\,\overline{q}\,(t) + \widehat{\boldsymbol{E}}(\partial_\mu \widehat{g}\,(t)\,\widehat{\overline{q}}\,(t)) + l_x\,(t) + \widehat{\boldsymbol{E}}(\partial_\mu \widehat{l}\,(t)) + h_x\,(t)\,k(t) \Big\}\,\mathrm{d}t \\ & \quad - q(t)\mathrm{d}W(t) - \overline{q}(t)\mathrm{d}\widetilde{W}(t), \\ p(T) &= \psi_x(x\,(T)\,,\mathbb{P}_{x(T)}) + \widehat{\boldsymbol{E}}\left[\partial_\mu \psi(\widehat{x}\,(T)\,,\mathbb{P}_{x(T)}; x(T))\right]. \end{cases} \tag{2.14}$$

Clearly, under hypothesis (C1) and (C2), it is easy to prove that Eqs (2.13) and (2.14) admits a unique strong solution. Since the coefficients $G\,(\cdot)$ and $M\,(\cdot)$ are not related to $x(\cdot)$, then the adjoint process $(p\,(\cdot)\,, q\,(\cdot)\,, \overline{q}(\cdot), k\,(\cdot))$ are independent to singular control $\xi(\cdot)$.

The main result of this paper is stated in the following theorem.

Theorem 3.1 (Maximum principle) Let hypothesis (C1), (C2) and (C3) hold. Let $(u^*(\cdot), \xi^*(\cdot), x^*(\cdot))$ be the optimal solution of the control problem (2.3)-(2.5). Then there exists $(p\,(\cdot)\,, q\,(\cdot)\,, \overline{q}(\cdot), k\,(\cdot))$ solution of (2.14), such that for any $(u,\xi) \in \mathbb{A}_1 \times \mathbb{A}_2$, we have

$$\begin{aligned} &\boldsymbol{E}^u\left[H_u(t, x^*(t), \mathbb{P}_{x^*(t)}, u^*\,(t)\,, p\,(t)\,, q\,(t)\,, \overline{q}\,(t)\,, k\,(t))\,(u\,(t) - u^*\,(t)) \mid \mathcal{F}_t^Y\right] \\ &+ \boldsymbol{E}^u\left[\int_{[0,T]} (M(t) + G(t)p(t))\mathrm{d}\,(\xi - \xi^*)\,(t) \mid \mathcal{F}_t^Y\right] \geq 0. \end{aligned} \tag{2.15}$$

$\mathbb{P}-a.s.,\ a.e.t \in [0,T]\,,$

where the Hamiltonian function H is defined by (2.12).

To prove our main result, the approach that we use is based on a double perturbation of the optimal control. This perturbation is described as follows :
Let $(u(\cdot),\xi(\cdot)) \in \mathcal{A}_1^Y \times \mathcal{A}_2^Y([0,T])$, be any given admissible control. Let $\varepsilon \in (0,1)$, and write

$$u^{\varepsilon}(\cdot) = u^*(\cdot) + \varepsilon v(\cdot) \quad \text{where } v(\cdot) = u(\cdot) - u^*(\cdot), \tag{2.16}$$

and

$$\xi^{\varepsilon}(t) = \xi^*(t) + \varepsilon\zeta(t) \text{ where } \zeta(t) = \xi(t) - \xi^*(t), \tag{2.17}$$

where ε a sufficiently small $\varepsilon > 0$.
Here the admissible control $(u^{\varepsilon}(\cdot),\xi^{\varepsilon}(\cdot))$ is the so called convex perturbation of $(u^*(\cdot),\xi^*(\cdot))$ defined as follows : $t \in [0,T]$

$$(u^{\varepsilon}(t),\xi^{\varepsilon}(t)) = (u^*(t),\xi^*(t)) + \varepsilon\left[(u(t),\xi(t)) - (u^*(t),\xi^*(t))\right],$$

In this chapter, we denote by $x^{\varepsilon}(\cdot) = x^{u^{\varepsilon},\xi^{\varepsilon}}(\cdot)$ the solution of (2.3) associated with $(u^{\varepsilon}(\cdot),\xi^{\varepsilon}(\cdot))$ and by $\rho^{\varepsilon}(\cdot)$ the solution of (2.8) corresponding to $u^{\varepsilon}(\cdot)$.
Hereinafter, we use the following short-hand notations :

$$\varphi(t) = \varphi\left(t, x^*(t), \mathbb{P}_{x^*(t)}, u^*(t)\right), \qquad h(t) = h\left(t, x^*(t), u^*(t)\right),$$

$$\varphi^{\varepsilon}(t) = \varphi(t, x^{\varepsilon}(t), \mathbb{P}_{x^{\varepsilon}(t)}, u^{\varepsilon}(t)), \qquad h^{\varepsilon}(t) = h\left(t, x^{\varepsilon}(t), u^{\varepsilon}(t)\right),$$

where $\varphi := f,\sigma,g,l$ as well as their partial derivatives with respect to x and v. Also, we denote for $\varphi = f,\sigma,g,l$:

$$\partial_\mu\varphi(t) = \partial_\mu\varphi\left(t, x(t), \mathbb{P}_{x(t)}, u(t); \widehat{x}(t)\right),$$

$$\partial_\mu\widehat{\varphi}(t) = \partial_\mu\varphi\left(t, \widehat{x}(t), \mathbb{P}_{\widehat{x}(t)}, \widehat{u}(t); x(t)\right).$$

Under hypothesis (C1), (C2) and (C3), Eqs (2.22) and (2.21) which are a linear SDEs with bounded coefficients, admits a unique adapted solutions $p(\cdot)$ and $\rho_1(\cdot)$, respectively.

In order to prove our main result in Theorem 3.1, we need the following results which we have to translate to our partially observed singular problem.

Lemma 3.2 Let hypothesis (C1), (C2) and (C3) hold. Then, we have

$$\lim_{\varepsilon\rightarrow 0}\boldsymbol{E}\left[\sup_{0\leq t\leq T}\left|x^{\varepsilon}(t)-x^{*}(t)\right|^{2}\right]=0.$$

Proof Applying standard estimates and *Burkholder-Davis-Gundy inequality*, we have

$$\begin{aligned}
&\boldsymbol{E}(\sup_{t\in[0,T]}\left|x^{\varepsilon}(s)-x^{*}(s)\right|^{2})\\
&\leq \boldsymbol{E}\int_0^t\left|f\left(s,x^{\varepsilon}(s),P_{x^{\varepsilon}(s)},u^{\varepsilon}(s)\right)-f\left(s,x^{*}(s),P_{x^{*}(s)},u^{*}(s)\right)\right|^{2}\mathrm{d}s\\
&+\boldsymbol{E}\int_0^t\left|\sigma\left(s,x^{\varepsilon}(s),P_{x^{\varepsilon}(s)},u^{\varepsilon}(s)\right)-\sigma\left(s,x^{*}(s),P_{x^{*}(s)},u^{*}(s)\right)\right|^{2}\mathrm{d}s\\
&+\boldsymbol{E}\int_0^t\left|g\left(s,x^{\varepsilon}(s),P_{x^{\varepsilon}(s)},u^{\varepsilon}(s)\right)-g\left(s,x^{*}(s),P_{x^{*}(s)},u^{*}(s)\right)\right|^{2}\mathrm{d}s\\
&+\left|\int_{[0,t]}G(s)d\left(\xi^{\varepsilon}-\xi^{*}\right)(s)\right|^{2},
\end{aligned}$$

Applying hypothesis (C1), (C2), (C3) and from to the Lipschitz conditions on the coefficients f,σ and g with respect to x,μ and u, we get

$$\begin{aligned}
\boldsymbol{E}\left[\sup_{0\leq t\leq T}\left|x^{\varepsilon}(t)-x^{*}(t)\right|^{2}\right]&\leq C_T\boldsymbol{E}\int_0^t\left[\left|x^{\varepsilon}(s)-x^{*}(s)\right|^{2}+\left|\mathbb{B}_2\left(\mathbb{P}_{x^{\varepsilon}(s)},\mathbb{P}_{x^{*}(s)}\right)\right|^{2}\right]\mathrm{d}s\\
&+C_T\varepsilon^2\boldsymbol{E}\int_0^t\left|u^{\varepsilon}(s)-u^{*}(s)\right|^{2}\mathrm{d}s \qquad (2.18)\\
&+C_T\varepsilon^2\boldsymbol{E}\left|\xi^{\varepsilon}(T)-\xi^{*}(T)\right|^{2},
\end{aligned}$$

from the definition of Wasserstein metric $\mathbb{W}_2(\cdot,\cdot)$, we have

$$\begin{aligned}\mathbb{W}_2\left(\mathbb{P}_{x^\varepsilon(t)},\mathbb{P}_{x^*(t)}\right) &= \inf\Big\{\left[\boldsymbol{E}\left|\widetilde{x}^\varepsilon(t)-\widetilde{x}(t)\right|^2\right]^{\frac{1}{2}}, \text{ for all } \widetilde{x}^\varepsilon(\cdot),\widetilde{x}(\cdot)\in\mathbb{L}^2\left(\mathcal{F};\mathbb{R}^d\right), \\ &\qquad \text{with } \mathbb{P}_{x^\varepsilon(t)}=\mathbb{P}_{\widetilde{x}^\varepsilon(t)} \text{ and } \mathbb{P}_{x^*(t)}=\mathbb{P}_{\widetilde{x^*}(t)}\Big\} \\ &\leq \left[\boldsymbol{E}\left|x^\varepsilon(t)-x^*(t)\right|^2\right]^{\frac{1}{2}}. \end{aligned} \tag{2.19}$$

By Definition 2.5 , then from (2.18) and (2.19), we obtain

$$\boldsymbol{E}\left[\sup_{0\leq t\leq T}\left|x^\varepsilon(t)-x^*(t)\right|^2\right]\leq C_T\boldsymbol{E}\int_0^t\sup_{r\in[0,s]}\left|x^\varepsilon(r)-x^*(r)\right|^2\mathrm{d}s+C_T\varepsilon^2.$$

By applying Gronwall's inequality, the desired result follows immediately by letting ε go to zero. □

Let $\mathcal{Z}(t)$ and $\rho_1(t)$ be the solutions of the following linear SDEs

$$\begin{cases}\mathrm{d}\mathcal{Z}(t) &= \left[f_x(t)\,\mathcal{Z}(t)+\widehat{\boldsymbol{E}}\left[\partial_\mu f(t)\,\widehat{\mathcal{Z}}(t)\right]+f_u(t)(u(t)-u^*(t))\right]\mathrm{d}t \\ &\quad+\left[\sigma_x(t)\mathcal{Z}(t)+\widehat{\boldsymbol{E}}\left[\partial_\mu\sigma(t)\,\widehat{\mathcal{Z}}(t)\right]+\sigma_u(t)(u(t)-u^*(t))\right]\mathrm{d}W(t) \\ &\quad+\left[g_x(t)\mathcal{Z}(t)+\widehat{\boldsymbol{E}}\left[\partial_\mu g(t)\,\widehat{\mathcal{Z}}(t)\right]+g_u(t)(u(t)-u^*(t))\right]\mathrm{d}\widetilde{W}(t) \\ &\quad+G(t)\mathrm{d}\left(\xi-\xi^*\right)(t), \\ \mathcal{Z}(0) &= 0,\end{cases} \tag{2.20}$$

and

$$\begin{cases}\mathrm{d}\rho_1(t)= & \left[\rho_1(t)h(t)+\rho(t)h_x(t)\mathcal{Z}(t)+\rho(t)h_v(t)v(t)\right]\mathrm{d}Y(t), \\ \rho_1(0)= & 0.\end{cases} \tag{2.21}$$

If we put $v(\cdot) = u(\cdot) - u^*(t)$, and $\zeta = \xi - \xi^*$ thus we can write Eq-(2.20) in the form

$$\begin{cases} \mathrm{d}\mathcal{Z}(t) &= \left[f_x(t)\,\mathcal{Z}(t) + \widehat{\boldsymbol{E}}\left[\partial_\mu f(t)\,\widehat{\mathcal{Z}}(t)\right] + f_v(t)v(t)\right]\mathrm{d}t \\ &\quad + \left[\sigma_x(t)\mathcal{Z}(t) + \widehat{\boldsymbol{E}}\left[\partial_\mu \sigma(t)\,\widehat{\mathcal{Z}}(t)\right] + \sigma_v(t)v(t)\right]\mathrm{d}W(t) \\ &\quad + \left[g_x(t)\mathcal{Z}(t) + \widehat{\boldsymbol{E}}\left[\partial_\mu g(t)\,\widehat{\mathcal{Z}}(t)\right] + g_v(t)\,v(t)\right]\mathrm{d}\widetilde{W}(t) \\ &\quad + G(t)\mathrm{d}\zeta(t), \\ \mathcal{Z}(0) &= 0, \end{cases} \tag{2.22}$$

Lemma 3.3 Suppose that hypothesis (C1), (C2) and (C3) hold. Then, we have

$$\lim_{\varepsilon\to 0}\boldsymbol{E}\left[\sup_{0\le t\le T}\left|\frac{1}{\varepsilon}\left[x^\varepsilon(t) - x^*(t)\right] - \mathcal{Z}(t)\right|^2\right] = 0. \tag{2.23}$$

Proof Under hypothesis (C1), (C2) and (C3), Eqs (2.22) and (2.21) which are a linear SDEs with bounded coefficients, admits a unique adapted solutions $\mathcal{Z}(\cdot)$ and $\rho_1(\cdot)$, respectively.

We put

$$\eta^\varepsilon(t) = \frac{x^\varepsilon(t) - x^*(t)}{\varepsilon} - \mathcal{Z}(t),\ t \in [0,T].$$

To simplify, we will use the following notations, for $\varphi = f, \sigma, g$ and l :

$$\varphi_x^{\alpha,\varepsilon}(t) = \varphi_x\left(t, x^{\alpha,\varepsilon}(t), \mathbb{P}_{x^\varepsilon(t)}, v^\varepsilon(t)\right),$$

$$\partial_\mu^{\alpha,\varepsilon}\varphi(t) = \partial_\mu\varphi(s, x^\varepsilon(t), \mathbb{P}_{\widehat{x}^{\alpha,\varepsilon}(t)}, v^\varepsilon(t); \widehat{x}(t)),$$

and

$$x^{\alpha,\varepsilon}(t) = x^*(t) + \alpha\varepsilon\left(\eta^\varepsilon(t) + \mathcal{Z}(t)\right),$$

$$\widehat{x}^{\alpha,\varepsilon}(t) = x^*(t) + \alpha\varepsilon(\widehat{\eta}^\varepsilon(t) + \widehat{\mathcal{Z}}(t)),$$

$$v^{\alpha,\varepsilon}(t) = u^*(t) + \alpha\varepsilon v(t).$$

Since $D_{\theta} f(\mu_0) = \left\langle D\widetilde{f}(\vartheta_0), \theta \right\rangle = \frac{d}{dt}\widetilde{f}(\vartheta_0 + t\theta)\Big|_{t=0}$, we have the following form of the Taylor expansion

$$f(\mathbb{P}_{\vartheta_0+\theta}) - f(\mathbb{P}_{\vartheta_0}) = D_{\xi} f(\mathbb{P}_{\vartheta_0}) + \mathcal{R}(\theta),$$

where $\mathcal{R}(\theta)$ is of order $O(\|\theta\|_2)$ with $O(\|\theta\|_2) \to 0$ for $\theta \in \mathbb{L}^2(\mathcal{F}; \mathbb{R}^d)$.

$$\begin{aligned}
\eta^{\varepsilon}(t) &= \frac{1}{\varepsilon}\int_0^t [f^{\varepsilon}(s) - f(s)]\,\mathrm{d}s + \frac{1}{\varepsilon}\int_0^t [\sigma^{\varepsilon}(s) - \sigma(s)]\,\mathrm{d}W(s) \\
&+ \frac{1}{\varepsilon}\int_0^t [c^{\varepsilon}(s) - c(s)]\,\mathrm{d}\widetilde{B}(s) + \frac{1}{\varepsilon}\int_{[0,t]} G(s)\mathrm{d}(\xi^{\varepsilon} - \xi^*)(s), \\
&- \int_0^t \left[f_x(s)\mathcal{Z}(s) + \widehat{\boldsymbol{E}}\left[\partial_{\mu} f(s)\widehat{\mathcal{Z}}(s)\right] + f_v(s)v(s)\right]\mathrm{d}s \\
&- \int_0^t \left[\sigma_x(s)\mathcal{Z}(s) + \widehat{\boldsymbol{E}}\left[\partial_{\mu}\sigma(s)\widehat{\mathcal{Z}}(s)\right] + \sigma_v(s)v(s)\right]\mathrm{d}W(s) \\
&- \int_0^t \left[g_x(s)\mathcal{Z}(s) + \widehat{\boldsymbol{E}}\left[\partial_{\mu} g(s)\widehat{\mathcal{Z}}(s)\right] + g_v(s)v(s)\right]\mathrm{d}\widetilde{W}(s) \\
&- \int_{[0,t]} G(s)d(\xi - \xi^*)(s).
\end{aligned}$$

We decompose $\frac{1}{\varepsilon}\int_0^t \left[f(s, x^{\varepsilon}(s), \mathbb{P}_{x^{\varepsilon}(s)}, u^{\varepsilon}(s)) - f(s, x^*(s), \mathbb{P}_{x^*(s)}, u^*(s))\right]\mathrm{d}s$ into the following parts

$$\begin{aligned}
&\frac{1}{\varepsilon}\int_0^t \left[f(s, x^{\varepsilon}(s), \mathbb{P}_{x^{\varepsilon}(s)}, u^{\varepsilon}(s)) - f(s, x^*(s), \mathbb{P}_{x^*(s)}, u^*(s))\right]\mathrm{d}s \\
&= \frac{1}{\varepsilon}\int_0^t \left[f(s, x^{\varepsilon}(s), \mathbb{P}_{x^{\varepsilon}(s)}, u^{\varepsilon}(s)) - f(s, x^*(s), \mathbb{P}_{x^{\varepsilon}(s)}, u^{\varepsilon}(s))\right]\mathrm{d}s \\
&+ \frac{1}{\varepsilon}\int_0^t \left[f(s, x^*(s), \mathbb{P}_{x^{\varepsilon}(s)}, u^{\varepsilon}(s)) - f(s, x^*(s), \mathbb{P}_{x^*(s)}, u^{\varepsilon}(s))\right]\mathrm{d}s \\
&+ \frac{1}{\varepsilon}\int_0^t \left[f(s, x^*(s), \mathbb{P}_{x^*(s)}, u^{\varepsilon}(s)) - f(s, x^*(s), \mathbb{P}_{x^*(s)}, u^*(s))\right]\mathrm{d}s.
\end{aligned}$$

We notice that

$$\begin{aligned}
\int_0^t \left[f^{\varepsilon}(s) - f(s, x^*(s), \mathbb{P}_{x^{\varepsilon}(s)}, u^{\varepsilon}(s))\right]\mathrm{d}s &= \varepsilon\int_0^t\int_0^1 \left[f_x^{\alpha,\varepsilon}(s)\left(\eta^{\varepsilon}(s) + \mathcal{Z}(s)\right)\right]\mathrm{d}\alpha\mathrm{d}s, \\
\int_0^t \left[f^{\varepsilon}(s) - f(s, x^*(s), \mathbb{P}_{x^{\varepsilon}(s)}, u^*(s)\right]\mathrm{d}s &= \varepsilon\int_0^t\int_0^1 \widehat{\boldsymbol{E}}\left[\partial_{\mu}^{\alpha,\varepsilon} f(s)\left(\widehat{\eta}^{\varepsilon}(s) + \widehat{\mathcal{Z}}(s)\right)\right]\mathrm{d}\alpha\mathrm{d}s,
\end{aligned}$$

and

$$\int_0^t \left[f\left(s, x^*(s), \mathbb{P}_{x^*(s)}, u^\varepsilon(s)\right) - f(s)\right] \mathrm{d}s = \varepsilon \int_0^t \int_0^1 \left[f_v\left(s, x(s), \mathbb{P}_{x(s)}, v^{\alpha,\varepsilon}(s)\right) v(s)\right] \mathrm{d}\alpha \mathrm{d}s.$$

By the similar method, the analogue relations hold for the coefficients σ and g. Moreover, from (2.17), we have

$$\frac{1}{\varepsilon} \int_{[0,t]} G(s) \mathrm{d}\left(\xi^\varepsilon - \xi^*\right)(s) - \int_{[0,t]} G(s) \mathrm{d}\left(\xi - \xi^*\right)(s) = 0.$$

Therefore, we obtain

$$\begin{aligned}
\boldsymbol{E}\left[\sup_{s\in[0,t]} |\eta^\varepsilon(s)|^2\right] = C_t \boldsymbol{E}\Bigg[&\int_0^t \int_0^1 \left|f_x^{\alpha,\varepsilon}(s)\, \eta^\varepsilon(s)\right|^2 \mathrm{d}\alpha \mathrm{d}s \\
&+ \int_0^t \int_0^1 \widehat{\boldsymbol{E}} \left|\partial_\mu^{\alpha,\varepsilon} f(s)\, \widehat{\eta}^\varepsilon(s)\right|^2 \mathrm{d}\alpha \mathrm{d}s \\
&+ \int_0^t \int_0^1 \left|\sigma_x^{\alpha,\varepsilon}(s)\, \eta^\varepsilon(s)\right|^2 \mathrm{d}\alpha \mathrm{d}s \\
&+ \int_0^t \int_0^1 \widehat{\boldsymbol{E}} \left|\partial_\mu^{\alpha,\varepsilon} \sigma(s) \widehat{\eta}^\varepsilon(s)\right|^2 \mathrm{d}\alpha \mathrm{d}s \\
&+ \int_0^t \int_0^1 \left|g_x^{\alpha,\varepsilon}(s)\, \eta^\varepsilon(s)\right|^2 \mathrm{d}\alpha \mathrm{d}s \\
&+ \int_0^t \int_0^1 \widehat{\boldsymbol{E}} \left|\partial_\mu^{\alpha,\varepsilon} g(s) \widehat{\eta}^\varepsilon(s))\right|^2 \mathrm{d}\alpha \mathrm{d}s \\
&+ C_t \boldsymbol{E}\left[\sup_{s\in[0,t]} |\gamma^\varepsilon(s)|^2\right],
\end{aligned}$$

where

$$
\begin{aligned}
\gamma^{\varepsilon}(t) = & \int_0^t \int_0^1 \left[f_x^{\alpha,\varepsilon}(s) - f_x(s) \right] \mathcal{Z}(s) \mathrm{d}\alpha \mathrm{d}s \\
& + \int_0^t \int_0^1 \widehat{\boldsymbol{E}} \left[\left(\partial_\mu^{\alpha,\varepsilon} f(s) - \partial_\mu f(s) \right) \widehat{\mathcal{Z}}(s) \right] \mathrm{d}\alpha \mathrm{d}s \\
& + \int_0^t \int_0^1 \left[f_v\left(s, x(s), \mathbb{P}_{x(s)}, v^{\alpha,\varepsilon}(s)\right) - f_v(s) \right] v(s) \mathrm{d}\alpha \mathrm{d}s \\
& + \int_0^t \int_0^1 \left[\sigma_x^{\alpha,\varepsilon}(s) - \sigma_x(s) \right] \mathcal{Z}(s) \mathrm{d}\alpha \mathrm{d}W(s) \\
& + \int_0^t \int_0^1 \widehat{\boldsymbol{E}} \left[\left(\partial_\mu^{\alpha,\varepsilon} \sigma(s) - \partial_\mu \sigma(s) \right) \widehat{\mathcal{Z}}(s) \right] \mathrm{d}\alpha \mathrm{d}W(s) \\
& + \int_0^t \int_0^1 \left[\sigma_v\left(s, x(s), \mathbb{P}_{x(s)}, v^{\alpha,\varepsilon}(s)\right) - \sigma_v(s) \right] v(s) \mathrm{d}\alpha \mathrm{d}W(s)
\end{aligned}
$$

$$
\begin{aligned}
& + \int_0^t \int_0^1 \left[g_x^{\alpha,\varepsilon}(s) - g_x(s) \right] \mathcal{Z}(s) \mathrm{d}\alpha \mathrm{d}\widetilde{W}(s) \\
& + \int_0^t \int_0^1 \widehat{\boldsymbol{E}} \left[\left(\partial_\mu^{\alpha,\varepsilon} g(s) - \partial_\mu g(s) \right) \widehat{\mathcal{Z}}(s) \right] \mathrm{d}\alpha \mathrm{d}\widetilde{W}(s) \\
& + \int_0^t \int_0^1 \left[g_v\left(s, x(s), \mathbb{P}_{x(s)}, v^{\alpha,\varepsilon}(s)\right) - g_v(s) \right] v(s) \mathrm{d}\alpha \mathrm{d}\widetilde{W}(s).
\end{aligned}
$$

Now, the derivatives of f, σ and g with respect to (x, μ, v) are Lipschitz continuous in (x, μ, v), we get

$$
\lim_{\varepsilon \to 0} \boldsymbol{E} \left[\sup_{s \in [0,T]} \left| \gamma^{\varepsilon}(s) \right|^2 \right] = 0.
$$

Since the derivatives of the coefficients f, σ, g and γ are bounded with respect to (x, μ, v), we have

$$
\boldsymbol{E}\left(\sup_{s \in [0,t]} \left| \eta^{\varepsilon}(s) \right|^2 \right) \leq C_t \left[\boldsymbol{E} \int_0^t \left| \eta^{\varepsilon}(s) \right|^2 \mathrm{d}s + \boldsymbol{E}\left(\sup_{s \in [0,t]} \left| \gamma^{\varepsilon}(s) \right|^2 \right) \right].
$$

Finally, by applying Gronwall's Lemma, then by putting $t = T$ and letting ε go to 0, the proof of Lemma 3.3 is complete. □

Now, we introduce the following lemma which play an important role in computing the variational inequality for the cost functional (2.11) subject to (2.3) and (2.8).

Lemma 3.4. Let hypothesis (C1) and (C2) hold. Then, we have

$$\lim_{\varepsilon\to 0}\sup_{0\leq t\leq T}\boldsymbol{E}\left|\frac{1}{\varepsilon}\left[\rho^{\varepsilon}(t)-\rho(t)\right]-\rho_1(t)\right|^2=0. \tag{2.24}$$

Proof. From (2.8) and (2.21), we have

$$\begin{aligned}
\rho(t)+\varepsilon\rho_1(t) &= \rho(0)+\int_0^t \rho(s)h(s)\mathrm{d}Y(s) \\
&+\varepsilon\int_0^t \left[\rho_1(s)\,h(s)+\rho(s)h_x(s)\,\mathcal{Z}(s)+\rho(s)h_v(s)\,v(s)\right]\mathrm{d}Y(s) \\
&= \rho(0)+\varepsilon\int_0^t \rho_1(s)h(s)\mathrm{d}Y(s) \\
&+\int_0^t \rho(s)\,h(s,x(s)+\varepsilon\mathcal{Z}(s)\,,u(s)+\varepsilon v(s))\mathrm{d}Y(s) \\
&-\varepsilon\int_0^t \rho(s)A^{\varepsilon}(s)\mathrm{d}Y(s),
\end{aligned}$$

where

$$\begin{aligned}
A^{\varepsilon}(s) &= \mathcal{Z}(s)\int_0^1 \left[h_x(s,x(s)+\alpha\varepsilon\mathcal{Z}(s),u(s)+\alpha\varepsilon v(s))-h_x(s)\right]\mathcal{Z}(s)\mathrm{d}\alpha \\
&+v(s)\int_0^1 \left[h_v(s,x(s)+\alpha\varepsilon\mathcal{Z}(s),u(s)+\alpha\varepsilon v(s))-h_v(s)\right]v(s)\mathrm{d}\alpha.
\end{aligned}$$

Then, we have

$$
\begin{aligned}
&\rho^{\varepsilon}(t)-\rho(t)-\varepsilon\rho_1(t)\\
&=\int_0^t \rho^{\varepsilon}(s)\, h^{\varepsilon}(t)\, dY(s)-\varepsilon\int_0^t \rho_1(s)h(s)\mathrm{d}Y(s)\\
&-\int_0^t \rho(s)h\left(s,x(s)+\varepsilon\mathcal{Z}(s),u(s)+\varepsilon v(s)\right)\mathrm{d}Y(s)+\varepsilon\int_0^t \rho(s)A^{\varepsilon}(s)\,\mathrm{d}Y(s)\\
&=\int_0^t \left(\rho^{\varepsilon}(s)-\rho(s)-\varepsilon\rho_1(s)\right)h^{\varepsilon}(s)\,\mathrm{d}Y(s)\\
&+\int_0^t \left(\rho(s)+\varepsilon\rho_1(s)\right)\left[h^{\varepsilon}(s)-h\left(s,x(s)+\varepsilon\mathcal{Z}(s),u(s)+\varepsilon v(s)\right)\right]\mathrm{d}Y(s)\\
&+\varepsilon\int_0^t \rho_1(s)\, h(s,x(s)+\varepsilon\mathcal{Z}(s),u(s)+\varepsilon v(s))\mathrm{d}Y(s)\\
&-\varepsilon\int_0^t \rho_1(s)\, h(s)\mathrm{d}Y(s)+\varepsilon\int_0^t \rho(s)A^{\varepsilon}(s)\,\mathrm{d}Y(s)
\end{aligned}
$$

By simple computations, we obtain

$$
\begin{aligned}
&\rho^{\varepsilon}(t)-\rho(t)-\varepsilon\rho_1(t)\\
&=\int_0^t \left(\rho^{\varepsilon}(s)-\rho(s)-\varepsilon\rho_1(s)\right)h^{\varepsilon}(s)\,\mathrm{d}Y(s)\\
&+\int_0^t (\rho(s)+\varepsilon\rho_1(s))B_1^{\varepsilon}(s)\mathrm{d}Y(s)+\varepsilon\int_0^t \rho_1(s)B_2^{\varepsilon}(s)\mathrm{d}Y(s)\\
&+\varepsilon\int_0^t \rho(s)A^{\varepsilon}(s)\mathrm{d}Y(s),
\end{aligned}
$$

where

$$
\begin{aligned}
B_1^{\varepsilon}(s)&=h^{\varepsilon}(s)-h\left(s,x(s)+\varepsilon\mathcal{Z}(s),u(s)+\varepsilon v(s)\right),\\
B_2^{\varepsilon}(s)&=h(s,x(s)+\varepsilon\mathcal{Z}(s),u(s)+\varepsilon v(s))-h(s).
\end{aligned}
$$

Note that

$$B_1^{\varepsilon}(s) = \int_0^1 [h_x(s, x\,(s) + \varepsilon \mathcal{Z}\,(s) + \alpha(x^{\varepsilon}\,(s) - x\,(s) - \varepsilon \mathcal{Z}\,(s)), v^{\varepsilon}\,(s))] \\ \times (x^{\varepsilon}\,(s) - x\,(s) - \varepsilon \mathcal{Z}\,(s)) \mathrm{d}\alpha.$$

By Lemma 3.3 , we have

$$\boldsymbol{E} \int_0^t |(\rho(s) + \varepsilon \rho_1(s)) B_1^{\varepsilon}(s)|^2 \,\mathrm{d}s \leq C_{\varepsilon} \varepsilon^2, \tag{2.25}$$

where C_{ε} nonnegative constant such that $C_{\varepsilon} \to 0$ as $\varepsilon \to 0$.

Moreover, it is easy to see that

$$\sup_{0 \leq t \leq T} \boldsymbol{E} \left[\varepsilon \int_0^t \rho(s) A^{\varepsilon}(s) \mathrm{d}Y(s) \right]^2 \leq C_{\varepsilon} \varepsilon^2, \tag{2.26}$$

and

$$\sup_{0 \leq t \leq T} \boldsymbol{E} \left[\varepsilon \int_0^t \rho_1(s) B_2^{\varepsilon}(s) \mathrm{d}Y(s) \right]^2 \leq C_{\varepsilon} \varepsilon^2. \tag{2.27}$$

From (2.25), (2.26) and (2.27), we get

$$\begin{aligned}
&\boldsymbol{E} \left|(\rho^{\varepsilon}(t) - \rho(t)) - \varepsilon \rho_1(t)\right|^2 \\
&\leq C \left[\int_0^t \boldsymbol{E} \left|(\rho^{\varepsilon}\,(s) - \rho\,(s)) - \varepsilon \rho_1(s)\right|^2 + \boldsymbol{E} \int_0^t |(\rho\,(s) + \varepsilon \rho_1\,(s)) B_1^{\varepsilon}(s)|^2 \,\mathrm{d}s \right. \\
&\quad \left. + \sup_{0 \leq s \leq t} \boldsymbol{E} \left(\varepsilon \int_0^t \rho(s) A^{\varepsilon}(s) \mathrm{d}Y(s) \right)^2 + \sup_{0 \leq s \leq t} \boldsymbol{E} \left(\varepsilon \int_0^t \rho_1(s) B_2^{\varepsilon}(s) \mathrm{d}Y(s) \right)^2 \right] \\
&\leq C \int_0^t \boldsymbol{E} \left|\rho^{\varepsilon}(s) - \rho(s) - \varepsilon \rho_1(s)\right|^2 \mathrm{d}s + C_{\varepsilon} \varepsilon^2.
\end{aligned}$$

Finally, by using Gronwall's inequality, the proof of Lemma 3.4 is complete. □

Lemma 3.5. Let hypothesis (C1), (C2) and (C3) hold. Then, we have

$$\begin{aligned}
0 \leq\ & \boldsymbol{E} \int_0^T \Big[\rho_1(t)\, l(t) + \rho(t)\, l_x(t)\mathcal{Z}(t) + \rho(t)\widehat{\boldsymbol{E}}\left[\partial_\mu l(t)\right] \mathcal{Z}(t) + \rho(t) l_v(t) v(t)\Big] \mathrm{d}t \\
& + \boldsymbol{E}\left[\rho_1(T)\, \psi(x^*(T), \mathbb{P}_{x^*(T)})\right] + \boldsymbol{E}\left[\rho(T)\, \psi_x(x^*(T), \mathbb{P}_{x^*(T)})\mathcal{Z}(T)\right] \\
& + \boldsymbol{E}\Big[\rho(T)\, \widehat{\boldsymbol{E}}\left[\partial_\mu \psi(x^*(T), \mathbb{P}_{x^*(T)}; \widehat{x}(T))\right] \mathcal{Z}(T)\Big] \\
& + \boldsymbol{E}\int_{[0,T]} \rho(t)\, M(t) \mathrm{d}(\xi - \xi^*)(t).
\end{aligned} \tag{2.28}$$

Proof. From (2.6), we have

$$\begin{aligned}
0 &\leq \frac{1}{\varepsilon}\left[J(v^\varepsilon(t), \xi^\varepsilon(t)) - J(u^*(t), \xi^*(t))\right] \\
&= \frac{1}{\varepsilon}\left[J(v^\varepsilon(t), \xi^\varepsilon(t)) - J(u^*(t), \xi^\varepsilon(t))\right] \\
&+ \frac{1}{\varepsilon}\left[J(u^*, \xi^\varepsilon(t)) - J(u^*(t), \xi^*(t))\right],
\end{aligned} \tag{2.29}$$

From (2.5), we get

$$\begin{aligned}
& \frac{1}{\varepsilon}\left[J(v^\varepsilon(t), \xi^\varepsilon(t)) - J(u^*(t), \xi^\varepsilon(t))\right] \\
&= \frac{1}{\varepsilon}\boldsymbol{E}\int_0^T \left[\rho^\varepsilon(t) l^\varepsilon(t) - \rho(t) l(t)\right] \mathrm{d}t \\
&+ \frac{1}{\varepsilon}\boldsymbol{E}\left[\rho^\varepsilon(T)\, \psi(x^\varepsilon(T), \mathbb{P}_{x^\varepsilon(T)}) - \rho(T)\, \psi(x(T), \mathbb{P}_{x(T)})\right],
\end{aligned} \tag{2.30}$$

and

$$\begin{aligned}
& \frac{1}{\varepsilon}\left[J(u^*, \xi^\varepsilon(t)) - J(u^*(t), \xi^*(t))\right] \\
&= \frac{1}{\varepsilon}\left[\boldsymbol{E}\int_{[0,T]} \rho(t) M(t) \mathrm{d}\xi^\varepsilon(t) - \int_{[0,T]} \rho(t) M(t) \mathrm{d}\xi^*(t)\right].
\end{aligned} \tag{2.31}$$

By applying Taylor expansion, Lemma 3.3 and Lemma 3.4 , we obtain

$$\begin{aligned}
&\lim_{\varepsilon\to 0}\frac{1}{\varepsilon}\boldsymbol{E}\left[\rho^{\varepsilon}\left(T\right)\psi(x^{\varepsilon}\left(T\right),\mathbb{P}_{x^{\varepsilon}(T)})-\rho\left(T\right)\psi(x\left(T\right),\mathbb{P}_{x(T)})\right]\\
&=\boldsymbol{E}\left[\rho_1(T)\psi(x(T),\mathbb{P}_{x(T)})+\rho(T)\psi_x(x(T),\mathbb{P}_{x(T)})\mathcal{Z}\left(T\right)\right]\\
&+\boldsymbol{E}\left[\rho\left(T\right)\widehat{\boldsymbol{E}}\left[\partial_\mu\psi(x(T),\mathbb{P}_{x(T)};\widehat{x}(T))\right]\mathcal{Z}\left(T\right)\right],
\end{aligned}\tag{2.32}$$

and

$$\begin{aligned}
&\lim_{\varepsilon\to 0}\frac{1}{\varepsilon}\boldsymbol{E}\int_0^T\left[\rho^{\varepsilon}(t)l^{\varepsilon}(t)-\rho(t)l(t)\right]\mathrm{d}t\\
&=\boldsymbol{E}\int_0^T\left[\rho_1(t)l(t)+\rho(t)l_x(t)\mathcal{Z}(t)+\rho(t)\widehat{\boldsymbol{E}}\left[\partial_\mu l\left(t\right)\right]\widehat{\mathcal{Z}}(t)+\rho(t)l_v(t)v(t)\right]\mathrm{d}t.
\end{aligned}\tag{2.33}$$

From (2.17), and since $\xi^{\varepsilon}(t)-\xi^{*}(t)=\varepsilon(\xi(t)-\xi^{*}(t))$, we get

$$\begin{aligned}
&\lim_{\varepsilon\to 0}\frac{1}{\varepsilon}\left[\boldsymbol{E}\int_{[0,T]}\rho(t)M(t)\mathrm{d}\xi^{\varepsilon}(t)-\int_{[0,T]}\rho(t)M(t)\mathrm{d}\xi^{*}(t)\right]\\
&=\lim_{\varepsilon\to 0}\frac{1}{\varepsilon}\left[\boldsymbol{E}\int_{[0,T]}\rho(t)M(t)\mathrm{d}(\xi^{\varepsilon}-\xi^{*})(t)\right]\\
&=\lim_{\varepsilon\to 0}\frac{1}{\varepsilon}\left[\boldsymbol{E}\int_{[0,T]}\varepsilon\rho(t)M(t)\mathrm{d}(\xi-\xi^{*})(t)\right]\\
&=\boldsymbol{E}\int_{[0,T]}\rho(t)M(t)\mathrm{d}(\xi-\xi^{*})(t).
\end{aligned}\tag{2.34}$$

Finally, by substituting (2.30), (2.31), (2.32), (2.33) and (2.34) into (2.29), the desired result (2.28) fulfilled immediately. This achieve the proof of Lemma 3.5. □

Note that

$$\begin{cases}
d\widetilde{\rho}(t)=\left(h_x(t)\mathcal{Z}(t)+h_v(t)v(t)\right)\mathrm{d}\widetilde{W}(t),\\
\widetilde{\rho}(0)=0,
\end{cases}\tag{2.35}$$

where $\widetilde{\rho}(t)=\frac{\rho_1(t)}{\rho(t)}$.

Lemma 3.6 Let $p(\cdot)$ and $\mathcal{Z}\left(\cdot\right)$ be the solutions of (2.14) and (2.22) respectively. Then we have

$$
\begin{aligned}
\boldsymbol{E}^{u}\left[p\left(T\right)\mathcal{Z}\left(T\right)\right] &= \boldsymbol{E}^{u}\int_{0}^{T}p\left(t\right)f_{v}(t)v(t)\mathrm{d}t+\boldsymbol{E}^{u}\int_{0}^{T}q(t)\sigma_{v}(t)v(t)\mathrm{d}t \\
&+\boldsymbol{E}^{u}\int_{0}^{T}\overline{q}(t)g_{v}(t)v(t)\mathrm{d}t-\boldsymbol{E}^{u}\int_{0}^{T}\mathcal{Z}\left(t\right)\left(l_{x}\left(t\right)+\widehat{\mathbf{E}}(\partial_{\mu}\widehat{l}\left(t\right))\right)\mathrm{d}t \\
&+\boldsymbol{E}^{u}\int_{0}^{T}p(t)G(t)\mathrm{d}(\xi-\xi^{*})(t),
\end{aligned}
\tag{2.36}
$$

and

$$
\begin{aligned}
\boldsymbol{E}^{u}\left[y\left(T\right)\widetilde{\rho}\left(T\right)\right] &= \boldsymbol{E}^{u}\int_{0}^{T}k\left(t\right)\left[h_{x}(t)\mathcal{Z}(t)+h_{v}(t)v(t)\right]\mathrm{d}t. \\
&-\boldsymbol{E}^{u}\int_{0}^{T}\widetilde{\rho}\left(t\right)l(t)\mathrm{d}t.
\end{aligned}
\tag{2.37}
$$

Proof. By applying Itô's formula to $p\left(t\right)\mathcal{Z}\left(t\right)$ and taking expectation, with $\mathcal{Z}(0)=0$, we obtain

$$
\begin{aligned}
\boldsymbol{E}^{u}\left[p\left(T\right)\mathcal{Z}\left(T\right)\right] &= \boldsymbol{E}^{u}\int_{0}^{T}p\left(t\right)\mathrm{d}\mathcal{Z}\left(t\right)+\boldsymbol{E}^{u}\int_{0}^{T}\mathcal{Z}\left(t\right)\mathrm{d}p\left(t\right) \\
&+\boldsymbol{E}^{u}\int_{0}^{T}q(t)\left[\sigma_{x}(t)\mathcal{Z}(t)+\widehat{\mathbf{E}}\left[\partial_{\mu}\sigma(t)\widehat{\mathcal{Z}}(t)\right]+\sigma_{v}(t)v(t\right]\mathrm{d}t \\
&+\boldsymbol{E}^{u}\int_{0}^{T}\overline{q}(t)\left[g_{x}(t)\mathcal{Z}(t)+\widehat{\mathbf{E}}\left[\partial_{\mu}g(t)\widehat{\mathcal{Z}}(t)\right]+g_{v}(t)v(t)\right]\mathrm{d}t \\
&= \mathcal{I}_{1}+\mathcal{I}_{2}+\mathcal{I}_{3}+\mathcal{I}_{4}.
\end{aligned}
\tag{2.38}
$$

First, from equation (2.22), we obtain

$$
\begin{aligned}
\mathcal{I}_1 &= \boldsymbol{E}^u \int_0^T p(t)\, \mathrm{d}\mathcal{Z}(t) \\
&= \boldsymbol{E}^u \int_0^T p(t) \left[f_x(t)\mathcal{Z}(t) + \widehat{\boldsymbol{E}} \left[\partial_\mu f(t) \widehat{\mathcal{Z}}(t) \right] + f_v(t)v(t) \right] \mathrm{d}t \\
&+ \boldsymbol{E}^u \int_0^T p(t)G(t)\mathrm{d}(\xi - \xi^*)(t) \\
&= \boldsymbol{E}^u \int_0^T p(t)\, f_x(t)\mathcal{Z}(t)\mathrm{d}t + \boldsymbol{E}^u \int_0^T p(t)\, \widehat{\boldsymbol{E}} \left[\partial_\mu f(t) \widehat{\mathcal{Z}}(t) \right] \mathrm{d}t \\
&+ \boldsymbol{E}^u \int_0^T p(t)\, f_v(t)v(t)\mathrm{d}t + \boldsymbol{E}^u \int_0^T p(t)G(t)\mathrm{d}(\xi - \xi^*)(t).
\end{aligned} \tag{2.39}
$$

We proceed to estimate $\mathcal{I}_2$, From equation (2.14), we have

$$
\begin{aligned}
\mathcal{I}_2 &= \boldsymbol{E}^u \int_0^T \mathcal{Z}(t)\, \mathrm{d}p(t) \\
&= -\boldsymbol{E}^u \int_0^T \mathcal{Z}(t) \left[f_x(t)\, p(t) + \widehat{\boldsymbol{E}} \left(\partial_\mu \widehat{f}(t)\, \widehat{p}(t) \right) + \sigma_x(t)\, q(t) + \widehat{\boldsymbol{E}} \left(\partial_\mu \widehat{\sigma}(t)\, \widehat{q}(t) \right) \right. \\
&\left. + g_x(t)\, \overline{q}(t) + \widehat{\boldsymbol{E}} \left(\partial_\mu \widehat{g}(t)\, \widehat{\overline{q}}(t) \right) + l_x(t) + \widehat{\boldsymbol{E}} \left(\partial_\mu \widehat{l}(t) \right) + h_x(t)\, k(t) \right] \mathrm{d}t.
\end{aligned}
$$

By simple computation, we get

$$
\begin{aligned}
\mathcal{I}_2 &= -\boldsymbol{E}^u \int_0^T \mathcal{Z}(t)\, f_x(t)\, p(t)\, \mathrm{d}t - \boldsymbol{E}^u \int_0^T \mathcal{Z}(t)\, \widehat{\boldsymbol{E}}(\partial_\mu \widehat{f}(t)\, \widehat{p}(t))\mathrm{d}t \\
&- \boldsymbol{E}^u \int_0^T \mathcal{Z}(t)\, \sigma_x(t)\, q(t)\mathrm{d}t - \boldsymbol{E}^u \int_0^T \mathcal{Z}(t)\, \widehat{\boldsymbol{E}}(\partial_\mu \widehat{\sigma}(t)\, \widehat{q}(t))\mathrm{d}t \\
&- \boldsymbol{E}^u \int_0^T \mathcal{Z}(t)\, g_x(t)\, \overline{q}(t)\mathrm{d}t - \boldsymbol{E}^u \int_0^T \mathcal{Z}(t)\, \widehat{\boldsymbol{E}}(\partial_\mu \widehat{g}(t)\, \widehat{\overline{q}}(t))\mathrm{d}t \\
&- \boldsymbol{E}^u \int_0^T \mathcal{Z}(t)\, l_x(t)\, \mathrm{d}t - \boldsymbol{E}^u \int_0^T \mathcal{Z}(t)\, \widehat{\boldsymbol{E}}(\partial_\mu \widehat{l}(t))\mathrm{d}t \\
&- \boldsymbol{E}^u \int_0^T \mathcal{Z}(t)\, h_x(t)\, k(t)\mathrm{d}t.
\end{aligned} \tag{2.40}
$$

Similarly, we obtain

$$\begin{aligned}\mathcal{I}_3 = \boldsymbol{E}^u \int_0^T q(t)\sigma_x(t)\mathcal{Z}(t)dt + \boldsymbol{E}^u \int_0^T q(t)\widehat{\boldsymbol{E}}(\partial_\mu \sigma(t)\widehat{\mathcal{Z}}(t))\mathrm{d}t \\ + \boldsymbol{E}^u \int_0^T q(t)\sigma_v(t)v(t)\mathrm{d}t,\end{aligned} \tag{2.41}$$

and

$$\begin{aligned}\mathcal{I}_4 = \boldsymbol{E}^u \int_0^T \overline{q}(t)g_x(t)\mathcal{Z}(t)\mathrm{d}t + \boldsymbol{E}^u \int_0^T \overline{q}(t)\widehat{\boldsymbol{E}}(\partial_\mu g(t)\widehat{\mathcal{Z}}(t))\mathrm{d}t \\ + \boldsymbol{E}^u \int_0^T \overline{q}(t)g_v(t)v(t)\mathrm{d}t.\end{aligned} \tag{2.42}$$

Thus desired result (2.36) follows immediately by substituting (2.39), (2.40), (2.41) and (2.42) into (2.38) with the helps of Fubini's theorem.

Now, by applying Itô's formula to $y(t)\widetilde{\rho}(t)$ and taking expectation, we get

$$\begin{aligned}\boldsymbol{E}^u\left[y(T)\widetilde{\rho}(T)\right] &= \boldsymbol{E}^u \int_0^T y(t)\,\mathrm{d}\widetilde{\rho}(t) + \boldsymbol{E}^u \int_0^T \widetilde{\rho}(t)\,\mathrm{d}y(t) \\ &+ \boldsymbol{E}^u \int_0^T k(t)\left(h_x(t)\mathcal{Z}(t) + h_v(t)v(t)\right)\mathrm{d}t \\ &= \mathcal{J}_1 + \mathcal{J}_2 + \mathcal{J}_3.\end{aligned} \tag{2.43}$$

From (2.35), we have

$$\begin{aligned}\mathcal{J}_1 &= \boldsymbol{E}^u \int_0^T y(t)\,\mathrm{d}\widetilde{\rho}(t) \\ &= \boldsymbol{E}^u \int_0^T y(t)\left(h_x(t)\mathcal{Z}(t) + h_v(t)v(t)\right)\mathrm{d}\widetilde{W}(t),\end{aligned} \tag{2.44}$$

which is a martingale with zero expectation. Moreover, by a simple computations, we get

$$\mathcal{J}_2 = \boldsymbol{E}^u \int_0^T \widetilde{\rho}(t)\,\mathrm{d}y(t) = -\boldsymbol{E}^u \int_0^T \widetilde{\rho}(t)\,l(t)\mathrm{d}t, \tag{2.45}$$

and

$$\mathcal{J}_3 = \boldsymbol{E}^u \int_0^T k(t)\left[h_x(t)\mathcal{Z}(t) + h_v(t)v(t)\right] \mathrm{d}t. \tag{2.46}$$

Finally, substituting (2.44), (2.45), (2.46), into (2.43), the desired result (2.37) fulfilled. This completes the proof of Lemma 3.6. □

Proof of Theorem 3.1. Since $p(T) = \psi_x(x(T), \mathbb{P}_{x(T)}) + \widehat{\boldsymbol{E}}\left[\partial_\mu \psi(\widehat{x}(T), \mathbb{P}_{x(T)}; x(T))\right]$ and $y(T) = \psi_x(x(T), \mathbb{P}_{x(T)})$, then from Lemma 3.5, we have

$$\begin{aligned}
0 &\leq \boldsymbol{E}\int_0^T \Big[\rho_1(t)\, l(t) + \rho(t)\, l_x(t)\mathcal{Z}(t) + \rho(t)\widehat{\boldsymbol{E}}\left[\partial_\mu l(t)\right]\mathcal{Z}(t) + \rho(t) l_v(t) v(t)\Big]\, \mathrm{d}t \\
&+ \boldsymbol{E}\left[\rho_1(T)\, y(T)\right] + \boldsymbol{E}\left[\rho(T)\, p(T)\, \mathcal{Z}(T)\right] \\
&+ \boldsymbol{E}\int_{[0,T]} \rho(t)\, M(t) \mathrm{d}(\xi - \xi^*)(t).
\end{aligned} \tag{2.47}$$

Substituting (2.36) and (2.37) of Lemma 3.6 into (2.47), and since

$$\begin{aligned}
\boldsymbol{E}\left[\rho_1(T)\, y(T)\right] &= \boldsymbol{E}\left[\rho(T)\, \widetilde{\rho}(T)\, y(T)\right] = \boldsymbol{E}^u\left[y(T)\, \widetilde{\rho}(T)\right], \\
\boldsymbol{E}\left[\rho(T)\, p(T)\, \mathcal{Z}(T)\right] &= \boldsymbol{E}^u\left[p(T)\, \mathcal{Z}(T)\right], \\
\boldsymbol{E}\int_{[0,T]} \rho(t)\, M(t)\mathrm{d}(\xi - \xi^*)(t) &= \boldsymbol{E}^u \int_{[0,T]} M(t) \mathrm{d}(\xi - \xi^*)(t),
\end{aligned}$$

we get,

$$\begin{aligned}
0 &\leq \boldsymbol{E}\int_0^T \rho(t)\left[p(t)\, f_v(t) + q(t)\, \sigma_v(t) + \overline{q}(t)\, g_v(t) + K(t) h_v(t) + l_v(t)\right] v(t) \mathrm{d}t \\
&+ \boldsymbol{E}\int_{[0,T]} \rho(t)\left(M(t) + p(t) G(t)\right) \mathrm{d}(\xi - \xi^*)(t).
\end{aligned}$$

This completes the proof of Theorem 3.1. □

2.4 Partially observed McKean-Vlasov singular linear quadratic control problem

In this section, to illustrate our theoretical results we study partially observed optimal singular control problem for Mckean-Vlasov linear quadratic control problem, where the stochastic system is described by linear McKean-Vlasov stochastic differential equations with correlated noisy between the system and the observation.and the cost is described by a quadratic function. By applying our stochastic maximum principle established in Sect. 3 and classical filtering theory, we obtain an explicit expression of the optimal control represented in feedback form involving both controlled state process as well as its law represented by its expectation, via the solutions of ordinary differential equations (ODEs). Consider the following partially observed control system

$$\begin{cases} \mathrm{d}x^{v,\xi}(t) = & f\left(t, x^{v,\xi}(t), \mathbb{P}_{x^{v,\xi}(t)}, v(t)\right) dt + \sigma\left(t, x^{v,\xi}(t), \mathbb{P}_{x^{v,\xi}(t)}, v(t)\right) \mathrm{d}W(t) \\ & +g\left(t, x^{v,\xi}(t), \mathbb{P}_{x^{v,\xi}(t)}, v(t)\right) \mathrm{d}\widetilde{W}(t) + G(t)\mathrm{d}\xi(t), \\ x^{v,\xi}(0) = & x_0, \quad t \in [0, T], \end{cases} \tag{2.48}$$

where the coefficients are given by

$$\begin{aligned} f\left(t, x^{v}(t), \mathbb{P}_{x^{v}(t)}, v(t)\right) &= A_1(t)\, x(t) + A_2(t)\, \boldsymbol{E}\left(x(t)\right) + A_3(t) v(t), \\ \sigma\left(t, x^{v}(t), \mathbb{P}_{x^{v}(t)}, v(t)\right) &= A_4(t), \\ h(t, x^{v}(t), v(t)) &= A_5(t), \\ \psi(x(t), \mathbb{P}_{x(t)}) &= N(t) x^2(t) \\ g &\equiv 0, \end{aligned}$$

with an observation

$$\begin{cases} \mathrm{d}Y(t) & = A_5(t)\, dt + \mathrm{d}\widetilde{W}(t), \\ Y(0) & = 0, \end{cases} \tag{2.49}$$

and the quadratic cost functional $J(\cdot,\cdot)$ has the form

$$J\left(v\left(\cdot\right),\xi\left(\cdot\right)\right)=\boldsymbol{E}^{u}\left[\int_{0}^{T}\tau\left(t\right)v^{2}\left(t\right)\mathrm{d}t+N(T)x^{2}\left(T\right)\right]. \tag{2.50}$$

Here, the coefficients $A_1(\cdot)$, $A_2(\cdot)$, $A_3(\cdot)$, $A_4(\cdot)$, $A_5(\cdot)$, $\tau(\cdot)$ and $N(\cdot)$ are bounded continuous deterministic functions and $N(T)\geq 0$. For any $(v(\cdot),\xi(\cdot))\in\mathcal{A}_1^Y\times\mathcal{A}_2^Y([0,T])$, equations (2.48) and (2.49) have a unique solutions respectively. Our goal is to find an explicit optimal observed control to minimize the cost functional $J(v(\cdot),\xi(\cdot))$ over $\mathcal{A}_1^Y\times\mathcal{A}_2^Y([0,T])$, subject to (2.48) and (2.49). From (2.12) the Hamiltonian function H :

$$\begin{aligned} H(t,x,v,p,q,\overline{q}) &= \left[A_1(t)\,x(t)+A_2(t)\,\boldsymbol{E}\left[x(t)\right]+A_3(t)\,v(t)\right]p(t)+A_4(t)\,q(t) \\ &\quad + A_5(t)\,k(t)+\tau(t)\,v^2(t), \end{aligned} \tag{2.51}$$

From (2.51), then by a simple computational, we have

$$H_v(t,x,u,p,q,\overline{q})=A_3(t)\,p(t)+2\tau(t)u(t).$$

By applying Theorem 3.1 , and from the linearity of the conditional expectation, the optimal observed control satisfies the following expression

$$\widehat{u}(t)=-\frac{A_3(t)}{2\tau(t)}\boldsymbol{E}\left[p(t)\mid\mathcal{F}_t^Y\right], \tag{2.52}$$

$\boldsymbol{E}\left[p(t)\mid\mathcal{F}_t^Y\right]$ is conditional expectation of $p(t)$ with respect to $\mathcal{F}_t^Y$, and $(p(\cdot),q(\cdot),\overline{q}(\cdot))$ is the solution of the following BSDE

$$\begin{cases} -\mathrm{d}p(t)=\left[A_1(t)\,p(t)+A_2(t)\,\boldsymbol{E}\left[p(t)\right]\right]\mathrm{d}t \\ \qquad\qquad -q(t)\mathrm{d}W(t)-\overline{q}(t)\mathrm{d}\widetilde{W}(t), \\ p(T)=2N(T)x(T). \end{cases} \tag{2.53}$$

We note that the conditional expectation $\boldsymbol{E}\left[p(t)\mid\mathcal{F}_t^Y\right]$ is a random process, $\mathcal{F}_t^Y-$measurable for any $t\in[0,T]$.

The filtering estimates for optimal trajectories. We obtain the explicit expression of the optimal observed control in (2.52) via the filtering method with some proprieties of conditional expectation $\boldsymbol{E}\left[\cdot\mid\mathcal{F}_t^Y\right]$. From Liptser and Shiryayev [62, Theorems 8.1], Wang et al., [86, Theorem 3.1], and since $\widetilde{W}(\cdot)$ is $\mathcal{F}_t^Y-$measurables, $\widetilde{W}(\cdot)$ is independant to $W(\cdot)$, we obtain the following filtering equations :

$$\begin{cases} \mathrm{d}\widehat{x}(t)=\left[A_1(t)\,\widehat{x}(t)+A_2(t)\,\boldsymbol{E}\left[\widehat{x}(t)\right]-\frac{A_3^2(t)}{2\tau(t)}\widehat{p}(t)\right]\mathrm{d}t \\ -\mathrm{d}\widehat{p}(t)=\left[A_1(t)\,\widehat{p}(t)+A_2(t)\,\boldsymbol{E}\left[\widehat{p}(t)\right]\right]\mathrm{d}t-\widehat{\overline{q}}(t)\,\mathrm{d}\widetilde{W}(t), \\ \widehat{x}(0)=x_0,\ \widehat{p}(T)=2N(T)\widehat{x}(T),\ \widehat{\overline{q}}(t)=0, \end{cases} \tag{2.54}$$

where $\widehat{z}(t)=\boldsymbol{E}^u\left[z(t)\mid\mathcal{F}_t^Y\right]$ is the filtering estimate of the random state process $z(t)$ depending on the observable filtration $\mathcal{F}_t^Y$, for $z=x,p,\overline{q}$. Moreover, the random process $\boldsymbol{E}^u\left[z(t)\mid\mathcal{F}_t^Y\right]$ is $\mathcal{F}_t^Y-$measurable for any $t\in[0,T]$, such that

$$\int_A\boldsymbol{E}^u\left[z(t)\mid\mathcal{F}_t^Y\right](w)\,\mathrm{d}\mathbb{P}^u(w)=\int_A z(t)\,\mathrm{d}\mathbb{P}^u(w),\ \forall A\in\mathcal{F}_t^Y.$$

Now, for this purpose and to solve the above equation (2.54), noting the terminal condition of (2.54), we conjecture the observed adjoint process $\widehat{p}(\cdot)$ of the form

$$\widehat{p}(t)=\varphi_1(t)\,\widehat{x}(t)+\varphi_2(t)\,\boldsymbol{E}\left[\widehat{x}(t)\right], \tag{2.55}$$

where $\varphi_1(\cdot)$ and $\varphi_2(\cdot)$ are deterministic differential functions. Now, we derive equation

(2.55) by comparing it with equation (2.54), we obtain

$$
\begin{aligned}
&-\left[A_1(t)\left(\varphi_1(t)\widehat{x}(t)+\varphi_2(t)\boldsymbol{E}[\widehat{x}(t)]\right)+A_2(t)\boldsymbol{E}\left[\varphi_1(t)\widehat{x}(t)+\varphi_2(t)\boldsymbol{E}[\widehat{x}(t)]\right]\right] \\
&=\dot{\varphi}_1(t)\widehat{x}(t)+\dot{\varphi}_2(t)\boldsymbol{E}[\widehat{x}(t)] \\
&+\varphi_1(t)\left[A_1(t)\widehat{x}(t)+A_2(t)\boldsymbol{E}[\widehat{x}(t)]-\frac{A_3^2(t)}{2\tau(t)}\left(\varphi_1(t)\widehat{x}(t)+\varphi_2(t)\boldsymbol{E}[\widehat{x}(t)]\right)\right] \\
&+\varphi_2(t)\left[\left(A_1(t)+A_2(t)\right)\boldsymbol{E}[\widehat{x}(t)]-\frac{A_3^2(t)}{2\tau(t)}\boldsymbol{E}\left[\varphi_1(t)\widehat{x}(t)+\varphi_2(t)\boldsymbol{E}[\widehat{x}(t)]\right]\right]. \qquad (2.56)
\end{aligned}
$$

By comparing the coefficients of $\widehat{x}(t)$ and $\boldsymbol{E}[\widehat{x}(t)]$ in equation (2.56), we have the following ordinary differential equations (**ODEs**) :

$$
\begin{cases}
\dot{\varphi}_1(t)+2A_1(t)\varphi_1(t)-\frac{A_3^2(t)}{2\tau(t)}\varphi_1^2(t)=0, \\
\varphi_1(T)=2N(T),
\end{cases} \qquad (2.57)
$$

and

$$
\begin{cases}
\dot{\varphi}_2(t)+2\left(A_1(t)+A_2(t)\right)\varphi_2(t)+2A_2(t)\varphi_1(t) \\
\quad -\frac{A_3^2(t)}{\tau(t)}\varphi_1(t)\varphi_2(t)-\frac{A_3^2(t)}{2\tau(t)}\varphi_2^2(t)=0, \\
\varphi_2(T)=0.
\end{cases} \qquad (2.58)
$$

Note that equations (2.57) and (2.58) are Bernoulli type equation and Riccati type equation respectively.

To solve (2.57) and (2.58), we can use the similar method in [59, Sect. 4]. Then, the optimal continuous control for the problem (2.50) is given in the feedback form

$$
\widehat{u}(t)=\widehat{u}(t,\widehat{x}(t))=-\frac{A_3(t)}{2\tau(t)}\left[\varphi_1(t)\widehat{x}(t)+\varphi_2(t)\boldsymbol{E}[\widehat{x}(t)]\right], \qquad (2.59)
$$

where $\varphi_1(\cdot)$, and $\varphi_2(\cdot)$ determined by (2.57) and (2.58) respectively.

Let $\xi^*(\cdot)$ satisfies the maximum condition (2.52), we get : for any $\xi(\cdot) \in \mathcal{A}_2^Y([0,T])$:

$$\begin{aligned} &\boldsymbol{E}^u\left[\int_{[0,T]}(M(t)+G(t)p(t))\mathrm{d}\xi^*(t) \mid \mathcal{F}_t^Y\right] \\ &\leq \boldsymbol{E}^u\left[\int_{[0,T]}(M(t)+G(t)p(t))\mathrm{d}\xi(t) \mid \mathcal{F}_t^Y\right]. \end{aligned} \tag{2.60}$$

Now, we define a set $\mathbb{U} \subset [0,T] \times \Omega$ such that

$$\mathbb{U} = \left\{(t,w) \in [0,T] \times \Omega : M(t) + G(t)\widehat{p}(t) > 0\right\}, \tag{2.61}$$

where $\widehat{p}(t)$ is the adjoint process corresponding to optimal observed control $\widehat{u}(\cdot)$. Let $\xi(\cdot) \in \mathcal{A}_2^Y([0,T])$ such that

$$\mathrm{d}\xi(t) = \begin{cases} 0 : \text{if } \ (t,w) \in \mathbb{U}, \\ \mathrm{d}\widehat{\xi}(t) : \text{if } \ (t,w) \in \overline{\mathbb{U}}, \end{cases} \tag{2.62}$$

where $\overline{\mathbb{U}}$ is the complement of the set $\mathbb{U}$. We denote by $\mathbf{1}_{\mathbb{U}}(\cdot,\cdot)$ the indicator function of $\mathbb{U}$. Then from (2.15), we obtain

$$\begin{aligned} 0 &\leq \boldsymbol{E}^u\int_{[0,T]}(M(t)+G(t)\widehat{p}(t))\mathrm{d}(\xi(t)-\widehat{\xi}(t)) \\ &= \boldsymbol{E}^u\int_{[0,T]}(M(t)+G(t)\widehat{p}(t))\mathbf{1}_{\mathbb{U}}(t,w)\mathrm{d}(\xi-\widehat{\xi})(t) \\ &+ \boldsymbol{E}^u\int_{[0,T]}(M(t)+G(t)\widehat{p}(t))\mathbf{1}_{\overline{\mathbb{U}}}(t,w)\mathrm{d}(\xi-\widehat{\xi})(t). \end{aligned}$$

From (2.62), and since

$$\boldsymbol{E}^u\int_{[0,T]}(M(t)+G(t)\widehat{p}(t))\mathbf{1}_{\overline{\mathbb{U}}}(t,w)\mathrm{d}(\xi-\widehat{\xi})(t) = 0,$$

we have

$$
\begin{aligned}
0 &\leq \boldsymbol{E}^{u} \int_{[0,T]} (M(t) + G(t)\widehat{p}(t)) \mathrm{d}(\xi(t) - \widehat{\xi}(t)) \\
&= \boldsymbol{E}^{u} \int_{[0,T]} (M(t) + G(t)\widehat{p}(t)) \mathbf{1}_{\mathbb{U}}(t,w) \mathrm{d}(-\widehat{\xi})(t) \\
&= -\boldsymbol{E}^{u} \int_{[0,T]} (M(t) + G(t)\widehat{p}(t)) \mathbf{1}_{\mathbb{U}}(t,w) \mathrm{d}\widehat{\xi}(t).
\end{aligned}
$$

This shows that $\widehat{\xi}(\cdot)$ satisfies for any $t \in [0,T]$:

$$
\boldsymbol{E}^{u} \int_{[0,T]} (M(t) + G(t)\widehat{p}(t)) \mathbf{1}_{\mathbb{U}}(t,w) \mathrm{d}\widehat{\xi}(t) = 0.
$$

From (2.61) and (2.62), we can easy shows that the optimal observed singular control $\widehat{\xi}(\cdot)$ has the form :

$$
\widehat{\xi}(t) = \xi(t) + \int_0^t \mathbf{1}_{\overline{\mathbb{U}}}(s,w) \mathrm{d}s, \ \ t \in [0,T].
$$

Finally, we give the explicit optimal observed continuous-singular in feedback form by :

$$
\begin{aligned}
\widehat{u}\,(t, \widehat{x}(t)) &= -\frac{A_3\,(t)}{2\tau\,(t)}[\varphi_1\,(t)\,\widehat{x}\,(t) + \varphi_2\,(t)\,\boldsymbol{E}\,[\widehat{x}\,(t)]], \\
\widehat{\xi}(t) &= \xi(t) + \int_0^t \mathbf{1}_{\overline{\mathbb{U}}}(s,w) \mathrm{d}s, \ \ t \in [0,T]
\end{aligned}
$$

where $\varphi_1(\cdot)$ and $\varphi_2(\cdot)$ are given by (2.57), (2.58) respectively. This completes the proof.$\square$

Chapitre 3

Pointwise second-order stochastic maximum principle for optimal continuous-singular control

3.1 Introduction

In this chapter, we study a stochastic optimization problem. We establish a second-order necessary conditions for optimal continuous-singular stochastic control, where the systems is governed by nonlinear controlled Itô stochastic differential systems. The control process has two components, the first being absolutely continuous and the second is a bounded variation, non decreasing continuous on the right with left limits. Pointwise second order maximum principle in terms of the martingale with respect to the time variable is proved. The control domain is assumed to be convex.

In this chapter, the continuous control variable enters into both the drift and the diffusion coefficients of the control systems. Our result is proved by using a classical variational techniques, stochastic calculs under some convexity conditions.

Throughout this thesis, $(\Omega, \mathcal{F}, \mathbb{F}, P)$ be a complete filtered probability space, on which a d-dimensional Brownian motion $W(\cdot)$ is defined such that $\mathbb{F} = \{\mathcal{F}_t\}_{0\leq t\leq T}$ is the natural filtration generated by $W(\cdot)$ augmented by all the P-null sets. In this work, we study pointwise optimal stochastic singular control problem for systems governed by nonlinear controlled stochastic differential equations (SDEs) allowing both classical and singular control of the form : $t \in [0, T]$

$$\begin{cases} \mathrm{d}x^{u,\xi}(t) = f\left(t, x^{u,\xi}(t), u(t)\right)\mathrm{d}t + \sigma\left(t, x^{u,\xi}(t), u(t)\right)\mathrm{d}W(t) + G(t)\mathrm{d}\xi(t), \\ x^{u,\xi}(0) = x_0. \end{cases} \tag{3.1}$$

The expected cost to be minimized over the class of admissible controls has the form

$$J\left(u(\cdot), \xi(\cdot)\right) = \boldsymbol{E}\left[h(x^{u,\xi}(T)) + \int_0^T \ell(t, x^{u,\xi}(t), u(t))\mathrm{d}t + \int_{[0,T]} M(t)\mathrm{d}\xi(t)\right]. \tag{3.2}$$

Here the control variable is a pair $(u(\cdot), \xi(\cdot))$ of measurable $\mathbb{A}_1 \times \mathbb{A}_2$-valued, $\mathbb{F}-$adapted processes, where $\mathbb{A}_1$ is a closed convex subset of $\mathbb{R}^m$ and $\mathbb{A}_2 := [0, \infty)^m$ such that $\xi(\cdot)$ is of bounded variation, nondecreasing continuous on the right with left limits. The process $x^{u,\xi}(\cdot)$ is the state variable valued in $\mathbb{R}^n$ associated to $(u(\cdot), \xi(\cdot))$. This construction allows us to define integrals of the form $\int_{[0,T]} G(t)\mathrm{d}\xi(t)$ and $\int_{[0,T]} M(t)\mathrm{d}\xi(t)$.

The maps

$$\begin{aligned} &f : [0, T] \times \mathbb{R}^n \times \mathbb{A}_1 \rightarrow \mathbb{R}^n, \\ &\sigma : [0, T] \times \mathbb{R}^n \times \mathbb{A}_1 \rightarrow \mathcal{M}_{n\times d}(\mathbb{R}), \\ &\ell : [0, T] \times \mathbb{R}^n \times \mathbb{A}_1 \rightarrow \mathbb{R}, \\ &h : \mathbb{R}^n \rightarrow \mathbb{R}, \\ &G : [0, T] \times \Omega \rightarrow \mathbb{R} \\ &M : [0, T] \times \Omega \rightarrow [0, \infty)^m \end{aligned}$$

are given functions.

Denote by $\mathcal{A}_1 \times \mathcal{A}_2\left([0,T]\right)$ the set of $\mathcal{B}\left([0,T]\right) \otimes \mathcal{F}$-measurable and $\mathbb{F}$-adapted stochastic processes valued in $\mathbb{A}_1 \times \mathbb{A}_2$. Any $(u(\cdot), \xi(\cdot)) \in \mathcal{A}_1 \times \mathcal{A}_2\left([0,T]\right)$ is called an admissible control.

The stochastic optimal control problem considered in this paper is to find a couple of adapted processes $(u^*(\cdot), \xi^*(\cdot)) \in \mathcal{A}_1 \times \mathcal{A}_2\left([0,T]\right)$ such that

$$J\left(u^*(\cdot), \xi^*(\cdot)\right) = \inf_{(u(\cdot),\xi(\cdot)) \in \mathcal{A}_1 \times \mathcal{A}_2([0,T])} J\left(u(\cdot), \xi(\cdot)\right). \tag{3.3}$$

Any admissible control $(u^*(\cdot), \xi^*(\cdot)) \in \mathcal{A}_1 \times \mathcal{A}_2\left([0,T]\right)$ satisfying (3.3) is called an optimal control. The corresponding state $x^*(\cdot)$ is called an optimal state, and $(x^*(\cdot), u^*(\cdot), \xi^*(\cdot))$ is called an optimal solution of the control problem (3.1)-(3.3).

Stochastic control problems in which the systems are governed by a nonlinear controlled Itô stochastic differential equation have been studied extensively in the last two decades, both by the dynamic programming method and by the Pontryagin maximum principle. Maximum principle is a powerful tool to investigate optimal stochastic control problems. Stochastic singular control problems have received considerable attention in the literature. The first version of maximum principle for singular stochastic control problem has been derived by Cadenillas and Haussmann [19]. Sufficient conditions for existence of optimal singular control and the connection between the singular control and optimal stopping problems have been investigated by Dufour and Miller [25]. Necessary conditions for general optimal singular stochastic control problems have been derived by Dufour and Miller [26]. Maximum principle for optimal stochastic singular stochastic control was investigated by many authors. Under partial-information, optimal singular control problem for mean-field stochastic differential equations driven by Teugels martingales measures has been studied in Hafayed et al. [48]. Necessary and sufficient conditions for near-optimal mean-field stochastic singular control have been established in [33]. The first-order local maximum principle for singular optimal control for mean-field SDEs has been derived in Hafayed [39]. Maximum principle for optimal singular control problem for general controlled nonlinear McKean-Vlasov SDEs has been obtained by Hafayed et al [32]. A class of solvable singular stochastic control problems have been studied in Alvarez [6]. Singular stochastic control problem for linear diffusions and optimal stopping have been derived by Alvarez [4]. A various maximum principles for optimal regular control with applications to finance can be found in Wang and Wu [85], and the book by Zhou and Yong [92]. An extensive list of references to the stochastic singular control problem, called also *intervention control*, in which the optimal control has both absolutely continuous and singular components, with some applications to finance and economics can be found in [33, 55, 69, 73]. Some examples on singular stochastic control have been obtained in Shreve [80].

A pointwise second-order maximum principle for stochastic optimal controls was established by Zhang and Zhang [95] where both drift and diffusion terms may contain the control variable, and the control domain is assumed to be convex. The method was further developed in Zhang and Zhang [96] to derive a general pointwise second-order maximum principle, where the control domain is not assumed to be convex. First and second-order necessary conditions for stochastic optimal controls have been studied by Frankowska et al. [30] and Bonnans, Silva [17]. A second-order maximum principle for singular optimal control for SDEs with uncontrolled diffusion coefficient has been obtained by Tang [81]. Second-order maximum principle for optimal control with recursive utilities has been obtained by Dong and Meng [27]. A second-order necessary conditions for singular optimal controls with recursive utilities of stochastic delay systems have been proved by Huo and Meng [53]. Second-order necessary conditions for singular optimal controls with recursive utilities of mean-field control systems have been investigated in Huo and Meng [54].

In this work, we establish second-order necessary conditions for optimal continuous-singular control problem. The control region is necessary convex. A pointwise second-order maximum principle in terms of the martingale with respect to the time variable is proved. Variational techniques, some Lebesgue theorems in differentiations, measure and integrations, with some appropriate estimates are applied to derive our results. Our continuous-singular control problem under studied provides also an interesting models in many applications such as economics and mathematical finance.

This work extends the results obtained in Zhang and Zhang [95] to a class of continuous-singular stochastic control problems. The main novelty of our work is that under some assumptions, we provide pointwise second-order necessary conditions which are new for the stochastic continuous-singular case and are natural extension of their deterministic counterparts. When the first-order necessary conditions of optimality are singular in some sense, the second-order necessary conditions will come naturally. The second-order maximum principle established in this paper can be used to choose the candidates from the singularity of our stochastic controls for optimal ones.

Usually, in order to derive the second-order maximum principle for optimality, one needs to assume that the first-order condition degenerates in some sense. In our class of second-order stochastic control problem, there are two types of singularity :

1. A singularity in the control variable ; where the control variable has two components $(u(\cdot), \xi(\cdot))$, the first $u(\cdot)$ being absolutely continuous and the second $\xi(\cdot)$ is singular. This singularity come since $\mathrm{d}\xi(t)$ may be singular with respect to Lebesgue measure $\mathrm{d}t$. More precisely $\xi(\cdot)$ is of bounded variation, non-decreasing continuous on the right with left limits (see Definition 2.1).

2. Following the ideas considered in [27, 53, 54, 95, 96], and in order to derive a second-order necessary conditions, one needs to assume that the first order condition degenerates in some sense. So we define a new type of singularity ; in the classical sense for the continuous control part and in maximum principle sense for the singular part of the control, (see Definition 2.2).

Organization : The rest of the chapter is organized as follows. The formulation of the mixed control problem, and basic notations are given in Section 2. In Sections 3 and 4, we prove our main results. The final section concludes the chapter and outlines some of the possible future developments.

3.2 Assumptions and Problem Statement

Let us formulate the optimal mixed control. Let T be a fixed strictly positive real number and $(\Omega, \mathcal{F}, \mathbb{F}, P)$ be a complete filtered probability space satisfying the usual conditions in which one-dimentional Brownian motion $W(t) = \{W(t) : 0 \leq t \leq T\}$ and $W(0) = 0$ is defined, where $\mathbb{F} = \{\mathcal{F}_t\}_{t\in[0,T]}$ is the natural filtration generated by $W(\cdot)$, augmented by all the P-null sets.

Let $\mathbb{A}_1$ be a closed convex and bounded subset of $\mathbb{R}$ and $\mathbb{A}_2 := [0, \infty)^m$. Let $\mathcal{A}_1([0,T])$ be the class of $\mathcal{B}([0,T]) \otimes \mathcal{F}$ measurable, $\mathbb{F}-$adapted processes $u(\cdot) : [0,T] \times \Omega \to \mathbb{A}_1$ and $\mathcal{A}_2([0,T])$ is the class of $\mathcal{B}([0,T]) \otimes \mathcal{F}$-measurable, $\mathbb{F}-$adapted processes $\xi(\cdot) : [0,T] \times \Omega \to \mathbb{A}_2$.

We give here the precise definition of the continuous-singular control.

Definition 2.1. An admissible continuous-singular control is a pair $(u(\cdot), \xi(\cdot))$ of measurable $\mathbb{A}_1 \times \mathbb{A}_2-$valued, $\mathbb{F}-$adapted processes, such that $\xi(\cdot)$ is of bounded variation, non-decreasing continuous on the right with left limits and $\xi(0_-) = 0$. Moreover,

$$\boldsymbol{E}\left[\sup_{t\in[0,T]} |u(t)|^2 + |\xi(T)|^2\right] < \infty.$$

Notice that the jumps of a singular control $\xi(\cdot)$ at any jumping time τ_j denote by

$$\Delta\xi(\tau_j) := \xi(\tau_j) - \xi(\tau_{j-}).$$

We should note that since $\mathrm{d}\xi(t)$ may be singular with respect to Lebesgue measure $\mathrm{d}t$, we call $\xi(\cdot)$ the singular part of the control and the process $u(\cdot)$ its absolutely continuous part.

Notations. In this subsection, we introduce some notation which will be used in what follows. We denote by $\mathcal{B}(F)$: the Borel $\sigma-$field of a metric space F. Let $\varphi : [0,T] \times \mathbb{R}^n \times \mathbb{A}_1 \rightarrow \mathbb{R}^d$ be a given function. We denote by $\varphi_x(t,x,u)$ and $\varphi_u(t,x,u)$ respectively, the first-order partial derivatives of φ with respect to x and u at (t,x,u), by $\varphi_{(x,u)^2}(t,x,u)$ the Hessian of φ with respect to (x,u) at (t,x,u) and by $\varphi_{xx}(t,x,u), \varphi_{xu}(t,x,u)$, and $\varphi_{uu}(t,x,u)$ the second-order partial derivatives of φ at (t,x,u). We denote $\mathcal{A}_1 \times \mathcal{A}_2([0,T])$ the set of all admissible controls. We denote by $\mathbb{L}^2_{\mathbb{F}}([0,T];\mathbb{R})$: the space of $\mathbb{R}$-valued, $\mathcal{B}([0,T]) \otimes \mathcal{F}$-measurable, $\mathbb{F}$-adapted processes ψ such that

$$\|\psi\|_{\mathbb{L}^2_{\mathbb{F}}([0,T];\mathbb{R})} := \left[\boldsymbol{E}\left(\int_0^T |\psi(t)|^2 dt\right)\right]^{\frac{1}{2}} < \infty.$$

Assumptions. Usually, one has to impose more regularity on the data for the second-order necessary conditions than that for the first-order ones. The following assumptions will be in force throughout this paper.

Assumption (H1) The functions f, σ, g and h satisfy the following conditions : for any $(x,u) \in \mathbb{R}^n \times \mathbb{A}_1$, the function $f(\cdot,x,u) : [0,T] \times \Omega \rightarrow \mathbb{R}^n$ and $\sigma(\cdot,x,u) : [0,T] \times \Omega \rightarrow \mathbb{R}^n$ are $\mathcal{B}([0,T]) \otimes \mathcal{F}$ measurable and $\mathbb{F}-$adapted. The functions $f(t,\cdot,\cdot) : \mathbb{R}^n \times \mathbb{A}_1 \rightarrow \mathbb{R}^n$ and $\sigma(t,\cdot,\cdot) : \mathbb{R}^n \times \mathbb{A}_1 \rightarrow \mathbb{R}^n$ are continuously differentiable up to the second order, and all their partial derivatives are uniformly bounded. There exists a constant $C > 0$ such that for a.e. $(t,w) \in [0,T] \times \Omega$ and for any $x, y \in \mathbb{R}^n$ and $u, v \in \mathbb{A}_1$,

$$\begin{aligned}
&|f(t,0,u)| + |\sigma(t,0,u)| \leq C, \\
&|f(t,x,u) - f(t,y,u)| + |\sigma(t,x,u) - \sigma(t,y,u)| \leq C|x-y|, \\
&\left|f_{(x,u)^2}(t,x,u) - f_{(x,u)^2}(t,y,v)\right| \leq C(|x-y| + |u-v|), \\
&\left|\sigma_{(x,u)^2}(t,x,u) - \sigma_{(x,u)^2}(t,y,v)\right| \leq C(|x-y| + |u-v|).
\end{aligned}$$

Assumption (H2) For any $(x,u) \in \mathbb{R}^n \times \mathbb{A}_1$, the function $\ell(\cdot,x,u) : [0,T] \times \Omega \rightarrow \mathbb{R}$ is $\mathcal{B}([0,T]) \otimes \mathcal{F}$ measurable and $\mathbb{F}-$adapted, and the random variable $h(x)$ is $\mathcal{F}_T$ -measurable.

For a.e. $(t,\omega) \in [0,T] \times \Omega$, the functions $\ell(t,\cdot,\cdot) : \mathbb{R}^n \times \mathbb{A}_1 \to \mathbb{R}$ and $h(\cdot) : \mathbb{R}^n \to \mathbb{R}$ are continuously differentiable up to the second order, and for any $x, y \in \mathbb{R}^n$ and $u, v \in \mathbb{A}_1$,

$$\begin{aligned}
&|\ell(t,x,u)| \leq C(1+|x|^2+|u|^2),\\
&|\ell_x(t,x,u)| + |\ell_u(t,x,u)| \leq C(1+|x|+|u|),\\
&\left|\ell_{(x,u)^2}(t,x,u) - \ell_{(x,u)^2}(t,y,v)\right| \leq C\left(|x-y|+|u-v|\right),\\
&|\ell_{xx}(t,x,u)| + |\ell_{xu}(t,x,u)| + |\ell_{uu}(t,x,u)| \leq C,
\end{aligned}$$

and

$$\begin{aligned}
&|h(x)| \leq C(1+|x|^2), \quad |h_x(x)| \leq C(1+|x|),\\
&|h_{xx}(x)| \leq C, \quad |h_{xx}(x) - h_{xx}(y)| \leq C\,|x-y|.
\end{aligned}$$

Assumption (H3) The functions $G(\cdot) : [0,T] \times \Omega \to \mathbb{R}$, and $M(\cdot) : [0,T] \times \Omega \to \mathbb{R}^+$ are continuous and bounded.

We note that the nonlinear controlled stochastic differential equation (**??**) occur naturally in the probabilistic analysis of financial optimization problems, see [6, 55, 69, 73] and the references cited therein. Under assumptions (H1), (H2) and (H3), Eq-(3.1) has a unique strong solution. By standard arguments it is easy to show that for any $k > 0$, it holds that

$$\boldsymbol{E}\big(\sup_{t\in[0,T]} \left|x^{u,\xi}(t)\right|^k\big) < C_k,$$

where C_k is a constant depending only on k. Moreover, the cost functional (3.2) is well defined on $\mathcal{A}_1 \times \mathcal{A}_2\left([0,T]\right)$.

Remark 2.1 In order not to over complicate the already notational heavy presentation of this paper, in what follows we shall assume all processes are one-dimensional. We should note that the higher dimensional cases can be argued along the same lines without substantial difficulties, except for even heavier notations.

We define the Hamiltonian

$$H\left(t,x,p,q\right):=f\left(t,x,u\right)p+\sigma\left(t,x,u\right)q-\ell(t,x,u), \tag{3.4}$$

where $(t,x,u,p,q)\in[0,T]\times\mathbb{R}\times\mathbb{A}_1\times\mathbb{R}\times\mathbb{R}$. We introduce respectively the following two adjoint equations :

$$\begin{cases} \mathrm{d}p(t)=-\left[f_x(t)p(t)+\sigma_x(t)q(t)-\ell_x(t)\right]\mathrm{dt}+q(t)\mathrm{d}W(t), \\ p(T)=-h_x\left(x^*\left(T\right)\right), \end{cases} \tag{3.5}$$

and

$$\begin{cases} \mathrm{d}P(t)=-\left[2f_x(t)P(t)+2\sigma_x(t)Q(t)+\sigma_x(t)^2P(t)+H_{xx}(t)\right]\mathrm{d}t \\ \qquad\qquad +Q(t)\mathrm{d}W(t), \\ P(T)=-h_{xx}\left(x^*\left(T\right)\right), \end{cases} \tag{3.6}$$

where

$$H_{xx}(t)=H_{xx}(t,x(t),u(t),p(t),q(t))=f_{xx}\left(t,x,u\right)p(t)+\sigma_{xx}\left(t,x,u\right)q(t)-\ell_{xx}(t,x,u).$$

It is easy to prove that under assumptions (H1)-(H2), the BSDEs (3.5) and (3.6) admits a unique strong $\mathbb{F}-$adapted solution $(p(t),q(t))$ and $(P(t),Q(t))$ respectively,

$$p(t)=-h_{xx}\left(x^*\left(T\right)\right)-\int_t^T\left[f_x(s)p(s)+\sigma_x(s)q(s)-\ell_x(s)\right]\mathrm{ds}+\int_t^T q(s)\mathrm{d}W(s),$$

and

$$P(t)=-h_{xx}\left(x^*\left(T\right)\right)-\int_t^T\left[2f_x(s)P(s)+2\sigma_x(s)Q(s)+\sigma_x(s)^2P(s)+H_{xx}(s)\right]\mathrm{d}s \\ +\int_t^T Q(s)\mathrm{d}W(s),$$

which satisfies

$$\boldsymbol{E}\left[\sup_{t\in[0,T]}|p(t)|^2+\int_0^T|q(t)|^2\,\mathrm{d}t\right]<\infty,$$
$$\boldsymbol{E}\left[\sup_{t\in[0,T]}|P(t)|^2+\int_0^T|Q(t)|^2\,\mathrm{d}t\right]<\infty.$$

Remark 2.2 Since the coefficients $G(\cdot)$ and $M(\cdot)$ are not related to $x(\cdot)$, then the adjoint process $(p(\cdot), q(\cdot))$ and $(P(\cdot), Q(\cdot))$ are independent to singular control $\xi(\cdot)$, and it is readily seen that the adjoint equations (3.5) and (3.6) coincides with [95, Eqs-(3.5), (3.6)]. Also, we define

$$\begin{aligned}\mathbb{K}(t,x,u,p,q,P,Q) &:= H_{xu}(t,x,u,p,q)+f_u(t,x,u)\,P(t) \\ &\quad +\sigma_u(t,x,u)\,Q(t)+\sigma_u(t,x,u)\,P(t)\sigma_x(t,x,u),\end{aligned}\tag{3.7}$$

where $(t,x,u,p,q,P,Q)\in[0,T]\times\mathbb{R}^n\times\mathbb{A}_1\times\mathbb{R}^n\times\mathbb{R}^n\times\mathbb{R}^{n\times n}\times\mathbb{R}^{n\times n}$, and

$$H_{xu}(t,x,u,p,q)=f_{xu}(t,x,u)\,p(t)+\sigma_{xu}(t,x,u)\,q(t)-\ell_{xu}(t,x,u).$$

In this chapter, we denote

$$\mathbb{K}(t)=\mathbb{K}(t,x^*(t),u^*(t),p(t),q(t),P(t),Q(t)),\quad t\in[0,T]. \tag{3.8}$$

Note that the new function $\mathbb{K}(\cdot)$ is not affected by the singular control $\xi(\cdot)$. This is because the adjoint process $(p(\cdot),q(\cdot))$ and $(P(\cdot),Q(\cdot))$ are not related to $\xi(\cdot)$. The main reason is that the coefficients $G(\cdot)$ and $M(\cdot)$ not related to $x(\cdot)$. It is worth mentioning that if $G(\cdot)$ and $M(\cdot)$ depend to $x(\cdot)$ everything changes and this is an open problem left unsolved.

We define a new type of singularity in the classical sense for the continuous control part and in Pontryging-type maximum principle sense for the singular part of the control.

Definition 2.2 We call $(u_*(\cdot),\xi_*(\cdot))\in\mathcal{A}_1\times\mathcal{A}_2([0,T])$ a singular if the pair $(u_*(\cdot),\xi_*(\cdot))$

satisfies

$$\begin{cases} H_u(t, x_*(t), u_*(t), p_*(t), q_*(t)) = 0, \quad a.s.\ a.e. t \in [0, T], \\ H_{uu}(t, x_*(t), u_*(t), p_*(t), q_*(t)) + P_*(t)\sigma_u\left(t, x_*(t), u_*(t)\right)^2 = 0, \\ \qquad\qquad a.s.\ a.e.\ t \in [0, T], \end{cases} \tag{3.9}$$

$$\boldsymbol{E}\int_{[0,T]} (M(t) - p_*(t)G(t))\, \mathrm{d}\xi(t) = \boldsymbol{E}\int_{[0,T]} (M(t) - p_*(t)G(t))\, \mathrm{d}\xi_*(t), \tag{3.10}$$

for any $\xi(\cdot) \in \mathcal{A}_2([0, T])$. Here $x_*(t) = x^{u_*,\xi_*}(t)$ is the state with respect to $(u_*(\cdot), \xi_*(\cdot))$ and $(p_*(\cdot), q_*(\cdot))$, $(p_*(\cdot), q_*(\cdot))$ are the adjoint processes given respectively by (3.5) and (3.6) associated to $(u_*(\cdot), \xi_*(\cdot))$. If $(u_*(\cdot), \xi_*(\cdot))$ is also optimal, satisfies (3.5), then we call it a *singular optimal control.*

Other type of singularity have been studied by some authors. Singularity in classical sense has been considered in [54, Definition 2.4] and [95, Definition 3.3], singularity in Pontyagin-type maximum principle sense has been investigated in [96, Definition 3.2] and partially singular control in classical sense in [30, Definition 4.1].

3.3 Second-order maximum principle in integral form

In this section, our aim is to establish second-order necessary conditions in integral form for optimality satisfied by an optimal mixed control, where the system evolves according to nonlinear controlled SDEs. We are now ready to state the main theorem of the paper.

Theorem 3.1 (Second-order necessary condition). Let assumptions (H1), (H2) and (H3) hold. If $(u^*(\cdot), \xi^*(\cdot))$ is an optimal stochastic control that satisfy (3.9), then we have

$$\boldsymbol{E}\int_0^T \mathbb{K}(t)Y(t)(u(t) - u^*(t))\mathrm{d}t + \boldsymbol{E}\int_{[0,T]} Y(t)P(t)G(t)\mathrm{d}(\xi - \xi^*)(t) \leq 0, \tag{3.11}$$

$$\boldsymbol{E}\int_{[0,T]} (M(t) - p(t)G(t))\, I_{\{(w,t)\in\Omega\times[0,T]:(M(t)-p(t)G(t))\geq 0\}}\mathrm{d}\xi^*(t) = 0. \tag{3.12}$$

for any $(u(\cdot), \xi(\cdot)) \in \mathcal{A}_1 \times \mathcal{A}_2([0,T])$, where $\mathbb{K}(t)$ is given by equation (3.7).

To prove our main result, the approach that we use is based on a double perturbation of the optimal continuous-singular control. This perturbation is described as follows : Let $(x^*(\cdot), u^*(\cdot), \xi^*(\cdot))$ be an optimal solution and $(u(\cdot), \xi(\cdot)) \in \mathcal{A}_1 \times \mathcal{A}_2([0,T])$ be any given admissible control. Let $\varepsilon \in (0,1)$, and write

$$u^{\varepsilon}(\cdot) = u^*(\cdot) + \varepsilon v(\cdot) \quad \text{where } v(\cdot) = u(\cdot) - u^*(\cdot), \tag{3.13}$$

and

$$\xi^{\varepsilon}(t) = \xi^*(t) + \varepsilon(\xi(t) - \xi^*(t)) \text{ where } \zeta(t) = \xi(t) - \xi^*(t). \tag{3.14}$$

where ε a sufficiently small $\varepsilon > 0$. Denote by $x^{\varepsilon} = x^{u^{\varepsilon}, \xi^{\varepsilon}}$ the state of (3.1) with respect to $(u^{\varepsilon}(\cdot), \xi^{\varepsilon}(\cdot))$, and put $\triangle x(\cdot) = x^{\varepsilon}(\cdot) - x^*(\cdot)$.

For simplicity, we let for $\varphi = f, \sigma, g$:

$$\begin{aligned}
\varphi_x(t) &= \varphi_x(t, x^*(t), u^*(t)), \\
\varphi_u(t) &= \varphi_u(t, x^*(t), u^*(t)) \\
\varphi_{xu}(t) &= \varphi_{xu}(t, x^*(t), u^*(t)), \\
\varphi_{xx}(t) &= \varphi_{xx}(t, x^*(t), u^*(t)), \\
\varphi_{uu}(t) &= \varphi_{uu}(t, x^*(t), u^*(t))
\end{aligned}$$

We introduce the following two variational equations : $t \in [0,T]$

$$\begin{cases}
\mathrm{d}Y(t) = [f_x(t)Y(t) + f_u(t)v(t)]\,\mathrm{d}t \\
\qquad\qquad + [\sigma_x(t)Y(t) + \sigma_u(t)v(t)]\,\mathrm{d}W(t) \\
\qquad\qquad + G(t)\mathrm{d}\zeta(t), \\
Y(0) = 0.
\end{cases} \tag{3.15}$$

Here $Y(\cdot)$ is called the *first-order variational process*, which is depend explicitly to singular control. Since the coefficients f_x, σ_x, f_u, σ_u and G are bounded, then the linear stochastic differential equation (3.15) admits a unique $\mathbb{F}$-adapted strong solution such that

$$\begin{aligned} Y(t) &= \int_0^t \left[f_x(s)Y(s) + f_u(s)v(s)\right] \mathrm{d}s + \int_0^t \left[\sigma_x(s)Y(s) + \sigma_u(s)v(s)\right] \mathrm{d}W(s) \\ &\quad + \int_{[0,t]} G(s)\mathrm{d}\zeta(s), \end{aligned}$$

which satisfies the following estimate

$$\boldsymbol{E}\Big(\sup_{t\in[0,T]} |Y(t)|^k\Big) \leq C_k. \tag{3.16}$$

Second-order variational equation :

$$\begin{cases} \mathrm{d}Z(t) = \left[f_x(t)Z(t) + f_{xx}(t)Y(t)^2 + 2f_{xu}(t)v(t)Y(t) + f_{uu}(t)v(t)^2\right] \mathrm{d}t \\ \qquad + \left[\sigma_x(t)Z(t) + \sigma_{xx}(t)Y(t)^2 + 2\sigma_{xu}(t)v(t)Y(t) + \sigma_{uu}(t)v(t)^2\right] \mathrm{d}W(t), \\ Z(0) = 0. \end{cases} \tag{3.17}$$

Here the process $Z(\cdot)$ is called the *second-order variational process.* Moreover, similar to [95, Proposition 3.1], equation (3.17) admits a unique $\mathbb{F}$-adapted strong solution

$$\begin{aligned} Z(t) &= \int_0^t \left[f_x(s)Z(s) + f_{xx}(s)Y(s)^2 + 2f_{xu}(s)v(s)Y(s) + f_{uu}(s)v(s)^2\right] \mathrm{d}s \\ &\quad + \int_0^t \left[\sigma_x(t)Z(s) + \sigma_{xx}(s)Y(s)^2 + 2\sigma_{xu}(s)v(s)Y(s) + \sigma_{uu}(s)v(s)^2\right] \mathrm{d}W(s). \end{aligned}$$

such that : for any $k \geq 1$ we have

$$\boldsymbol{E}\Big(\sup_{t\in[0,T]} |Z(t)|^k\Big) \leq C_k. \tag{3.18}$$

We note that unless specified, for each $k \in \mathbb{R}_+$, we denote by $C_k > 0$ a generic positive constant depending only on k, which may vary from line to line.

We shall establish some fundamental estimates that will play the crucial roles in our discussion.

Proposition 3.1 Let assumptions (H1), (H2) and (H3) hold. Then, for any $k \geq 1$ the following estimates hold :

$$\boldsymbol{E}\big(\sup_{t\in[0,T]} |x^\varepsilon(t) - x^*(t)|^{2k}\big) \leq C_k \varepsilon^{2k}, \tag{3.19}$$

$$\boldsymbol{E}\left[\sup_{t\in[0,T]} |x^\varepsilon(t) - x^*(t) - \varepsilon Y(t)|^{2k}\right] \leq C_k \varepsilon^{4k}, \tag{3.20}$$

$$\boldsymbol{E}\left[\sup_{t\in[0,T]} \left|x^\varepsilon(t) - x^*(t) - \varepsilon Y(t) - \frac{\varepsilon^2}{2} Z(t)\right|^{2k}\right] \leq C_k \varepsilon^{6k}. \tag{3.21}$$

Proof. Let $x^*(\cdot)$ and $x^\varepsilon(\cdot)$ be the trajectory of (3.1) corresponding to $u^*(\cdot)$ and $u^\varepsilon(\cdot)$ resp. Let $Y(\cdot)$ and $Z(\cdot)$ be the solution of first and second order adjoint equations (3.15)-(3.17) corresponding to $u^*(\cdot)$.

Proof of (3.19) : Let $k = 1$. By a simple computation, we have

$$\begin{aligned} E\left(\sup_{0\leq t\leq T} |x^\varepsilon(t) - x^*(t)|^2\right) &\leq C\mathbf{E}\left[\sup_{0\leq t\leq T}\left|\int_0^t \left[f\left(s, x^\varepsilon(s), u^\varepsilon(s)\right) - f\left(s, x^*(s), u^*(s)\right)\right] \mathrm{d}s\right|^2\right] \\ &+ C\mathbf{E}\left[\sup_{0\leq t\leq T}\left|\int_0^t \left[\sigma\left(s, x^\varepsilon(s), u^\varepsilon(s)\right) - \sigma\left(s, x^*(s), u^*(s)\right)\right] \mathrm{d}W(s)\right|^2\right] \\ &+ C\mathbf{E}\left|\int_{[0,t]} G(s)\mathrm{d}\left(\xi^\varepsilon - \xi^*\right)(s)\right|^2, \end{aligned}$$

by *Burkholder-Davis-Gundy inequality*, we get

$$\begin{aligned}E\left(\sup_{0\leq t\leq T}\left|x^{\varepsilon}(t)-x^{*}(t)\right|^{2}\right) &\leq C\mathbf{E}\int_{0}^{t}\left|f\left(s,x^{\varepsilon}(s),u^{\varepsilon}(s)\right)-f\left(s,x^{*}(s),u^{*}(s)\right)\right|^{2}\mathrm{d}s\\&+C\mathbf{E}\int_{0}^{t}\left|\sigma\left(s,x^{\varepsilon}(s),u^{\varepsilon}(s)\right)-\sigma\left(s,x^{*}(s),u^{*}(s)\right)\right|^{2}\mathrm{d}s\\&+C\mathbf{E}\left|\int_{[0,t]}G(s)\mathrm{d}\left(\xi^{\varepsilon}-\xi^{*}\right)(s)\right|^{2},\end{aligned}$$

by assumption (H1) and the Lipschitz conditions on the coefficients f, σ with respect to x,μ, we get

$$\begin{aligned}E\left(\sup_{0\leq t\leq T}\left|x^{\varepsilon}(t)-x^{*}(t)\right|^{2}\right) &\leq C_{T}E\int_{0}^{t}\sup_{\tau\in[0,s]}\left|x^{\varepsilon}(\tau)-x^{*}(\tau)\right|^{2}\mathrm{d}s\\&+CE\left|\int_{[0,t]}G(s)\mathrm{d}\left(\xi^{\varepsilon}-\xi^{*}\right)(s)\right|^{2},\end{aligned}$$

by assumption (H3), and since $\xi^{\varepsilon}(t)-\xi^{*}(t)=\varepsilon\left(\xi(t)-\xi^{*}(t)\right)$, we deduce

$$E(\sup_{0\leq t\leq T}\left|x^{\varepsilon}(t)-x^{*}(t)\right|^{2})\leq C_{T}E\int_{0}^{t}\sup_{\tau\in[0,s]}\left|x^{\varepsilon}(\tau)-x^{*}(\tau)\right|^{2}\mathrm{d}s+C_{T}\varepsilon^{2},$$

by applying *Gronwall's Lemma*, the desired result follows. Similar for $k>1$.

Proof of (3.20) : From (3.1) and (3.15) we have

$$\begin{aligned}&\left|x^{\varepsilon}(t)-x^{*}(t)-\varepsilon Y(t)\right|^{2k}\\&=\left|\int_{0}^{t}\left[f\left(s,x^{\varepsilon}(s),u^{\varepsilon}(s)\right)-f\left(s,x^{*}(s),u^{*}(s)\right)-\varepsilon\left[f_{x}(s)Y(s)+f_{u}(s)v(s)\right]\right]ds\right.\\&+\int_{0}^{t}\left[\sigma\left(s,x^{\varepsilon}(s),u^{\varepsilon}(s)\right)-\sigma\left(s,x^{*}(s),u^{*}(s)\right)-\varepsilon\left[\sigma_{x}(s)Y(s)+\sigma_{u}(s)v(s)\right]\right]dW(s)\\&\left.+\int_{[0,t]}G(s)\mathrm{d}\left(\xi^{\varepsilon}-\xi^{*}\right)(s)-\varepsilon\int_{[0,t]}G(s)\mathrm{d}\zeta(s)\right|^{2k}.\end{aligned}$$

Since $\zeta(t)=\xi(t)-\xi^*(t)$, then a straightforward calculation shows that

$$\int_{[0,t]} G(s)\mathrm{d}\left(\xi^\varepsilon-\xi^*\right)(s)-\varepsilon\int_{[0,t]} G(s)\mathrm{d}\zeta(s)=0,$$

the rest of the proof is very close to Bensoussan [15, Lemma 4.1, page 26], we omit the details.

Proof of (3.21) : From (3.1), (3.15) and (3.17), then by a straightforward calculation, we obtain

$$\begin{aligned}
&\left|x^\varepsilon(t)-x^*(t)-\varepsilon Y(t)-\frac{\varepsilon^2}{2}Z(t)\right|^{2k}\\
&=\left|\int_0^t\left[f\left(s,x^\varepsilon(s),u^\varepsilon(s)\right)-f\left(s,x^*(s),u^*(s)\right)-\varepsilon\left[f_x(s)Y(s)+f_u(s)v(s)\right]\right.\right.\\
&-\frac{\varepsilon^2}{2}\left[f_x(s)Z(s)+f_{xx}(s)Y(s)^2+2f_{xu}(s)Y(s)v(s)+f_{uu}(s)v(s)^2\right]\Big]\mathrm{d}s\\
&+\int_0^t\left[\sigma\left(s,x^\varepsilon(s),u^\varepsilon(s)\right)-\sigma\left(s,x^*(s),u^*(s)\right)-\varepsilon\left[\sigma_x(s)Y(s)+\sigma_u(s)v(s)\right]\right.\\
&-\frac{\varepsilon^2}{2}\left[\sigma_x(s)Z(s)+\sigma_{xx}(s)Y(s)^2+2\sigma_{xu}(s)Y(s)v(s)\right.\\
&\left.\left.+\;\sigma_{uu}(s)v(s)^2\right]\Big]\mathrm{d}W(s)\right|^{2k}.
\end{aligned}\tag{3.22}$$

Since the right hand side of (3.22) is independent to singular control $\xi(\cdot)$, the rest of the proof is similar to Zhang and Zhang [95, Proposition 3.1], then the desired result (3.21) is fulfilled. This completes the proof of Proposition 3.1. □

To prove the main theorem we need the following technical Lemmas.

Lemma 3.1 Let (p,q) and (P,Q) be the solution to the adjoint equation (3.5) and (3.6) respectively. Let Y and Z be the solutions to the first and second order variational equations (3.15) and (3.17), respectively associated to $(u^*(\cdot),\xi^*(\cdot))$. Then the following duality relations hold :

$$
\begin{aligned}
&\boldsymbol{E}\left[h_x(x^*(T))Y(T)\right] \\
&= -\boldsymbol{E}\int_0^T \left[p(t)\, f_u(t)\, v(t) + q(t)\sigma_u(t)\, v(t) + \ell_x(t) Y(t)\right] \mathrm{d}t \\
&- \boldsymbol{E}\int_{[0,T]} p(t)G(t)\mathrm{d}\zeta(t),
\end{aligned}
\tag{3.23}
$$

$$
\begin{aligned}
&\boldsymbol{E}\left[h_x(x^*(T))Z(T)\right] \\
&= -\boldsymbol{E}\left[p(T)\, Z(T)\right] \\
&= -\boldsymbol{E}\int_0^T \Big[p(t)\, f_{xx}(t)\, Y^2(t) + 2p(t)\, f_{xx}(t)\, Y^2(t) \\
&+ p(t)\, f_{uu}(t)\, v^2(t) + q(t)\sigma_{xx}(t)\, Y^2(t) \\
&+ 2q(t)\sigma_{xu}(t)\, v(t)\, Y(t) + q(t)\sigma_{uu}(t)\, v^2(t) \\
&+ \ell_x(t) Z(t)\Big] \mathrm{d}t,
\end{aligned}
\tag{3.24}
$$

and

$$
\begin{aligned}
&\boldsymbol{E}\left[h_{xx}(x^*(T))Y^2(T)\right] \\
&= -\boldsymbol{E}\int_0^T \big[2P(t)\, Y(t)\, f_u(t)\, v(t) + 2P(t)\, \sigma_x(t)\, Y(t)\, \sigma_u(t)\, v(t) \\
&+ P(t)\, \sigma_u^2(t)\, v^2(t) + 2Q(t)\, \sigma_u(t)\, v(t)\, Y(t) - H_{xx}(t)\, Y^2(t)\big] \mathrm{d}t \\
&- \boldsymbol{E}\int_{[0,T]} 2Y(t)P(t)G(t)\mathrm{d}\zeta(t).
\end{aligned}
\tag{3.25}
$$

Proof.

Proof of (3.23). By Itô's formula to $p(T)Y(T)$, we have

$$\begin{aligned}
&\boldsymbol{E}\left[p(T)Y(T)\right]-\boldsymbol{E}\left[p(0)Y(0)\right] \\
&=\boldsymbol{E}\int_0^T p(t)\mathrm{d}Y(t)+\boldsymbol{E}\int_0^T Y(t)\mathrm{d}p(t)+\boldsymbol{E}\int_0^T q(t)(\sigma_x(t)Y(t)+\sigma_u(t)v(t))\mathrm{d}t \\
&=I_1+I_2+I_3.
\end{aligned} \tag{3.26}$$

From (3.15), we get

$$\begin{aligned}
I_1&=\boldsymbol{E}\int_0^T p(t)\mathrm{d}Y(t) \\
&=\boldsymbol{E}\int_0^T p(t)\left[f_x(t)Y(t)+f_u(t)v(t)\right]\mathrm{d}t+\boldsymbol{E}\int_{[0,T]} p(t)G(t)\mathrm{d}\zeta(t),
\end{aligned} \tag{3.27}$$

and from (3.5), we get

$$\begin{aligned}
I_2&=\boldsymbol{E}\int_0^T Y(t)\mathrm{d}p(t) \\
&=-\boldsymbol{E}\int_0^T Y(t)\left[p(t)f_x(t)+q(t)\sigma_x(t)-\ell_x(t)\right]\mathrm{d}t.
\end{aligned} \tag{3.28}$$

Similarly, we have

$$I_3=\boldsymbol{E}\int_0^T q(t)\left[\sigma_x(t)Y(t)+\sigma_u(t)v(t)\right]\mathrm{d}t. \tag{3.29}$$

Substituting (3.27), (3.28), and (3.29) into (3.26), with the fact that $Y(0)=0$, we get

$$\begin{aligned}
&\boldsymbol{E}\left[p(T)Y(T)\right] \\
&=\boldsymbol{E}\int_0^T \left[p(t)f_u(t)v(t)+q(t)\sigma_u(t)v(t)+\ell_x(t)Y(t)\right]\mathrm{d}t \\
&+\boldsymbol{E}\int_{[0,T]} p(t)G(t)\mathrm{d}\zeta(t).
\end{aligned}$$

Since $p(T) = -h_x(x^*(T))$, we get

$$\begin{aligned}
&\boldsymbol{E}\left[h_x(x^*(T))Y(T)\right] \\
&= -\boldsymbol{E}\left[p(T)\,Y(T)\right] \\
&= -\boldsymbol{E}\int_0^T \left[p(t)\,f_u(t)\,v(t) + q(t)\sigma_u(t)\,v(t) + \ell_x(t)Y(t)\right]dt \\
&- \boldsymbol{E}\int_{[0,T]} p(t)G(t)\mathrm{d}\zeta(t),
\end{aligned}$$

then the desired result (3.23) is fulfilled

Proof of (3.24). By applying Itô's formula to $p(T)\,Z(T)$, we have

$$\begin{aligned}
&\boldsymbol{E}\left[p(T)\,Z(T)\right] - \boldsymbol{E}\left[p(0)\,Z(0)\right] \\
&= \boldsymbol{E}\int_0^T p(t)\mathrm{d}Z(t) + \boldsymbol{E}\int_0^T Z(t)\mathrm{d}p(t) \\
&+ \boldsymbol{E}\int_0^T q(t)\left[\sigma_x(t)Z(t) + \sigma_{xx}(t)Y^2(t) + 2\sigma_{xu}(t)Y(t)v(t)\right. \\
&\left.+\ \sigma_{uu}(t)v^2(t)\right]\mathrm{d}t \\
&= J_1 + J_2 + J_3.
\end{aligned} \tag{3.30}$$

From (3.17), we have

$$\begin{aligned}
J_1 &= \boldsymbol{E}\int_0^T p(t)\mathrm{d}Z(t) \\
&= \boldsymbol{E}\int_0^T p(t)\left[f_x(t)Z(t) + f_{xx}(t)Y^2(t) + 2f_{xu}(t)Y(t)v(t)\right. \\
&\left.+\ f_{uu}(t)v^2(t)\right]\mathrm{d}t.
\end{aligned} \tag{3.31}$$

From (3.5), it is easy to show that

$$\begin{aligned} J_2 &= \boldsymbol{E}\int_0^T Z(t)\mathrm{d}p(t) \\ &= -\boldsymbol{E}\int_0^T Z(t)\left[p(t)f_x(t)+q(t)\sigma_x(t)-\ell_x(t)\right]\mathrm{d}t, \end{aligned} \tag{3.32}$$

and similarly, we get

$$\begin{aligned} J_3 = \boldsymbol{E}\int_0^T q(t)\Big[&\sigma_x(t)Z(t)+\sigma_{xx}(t)Y^2(t)+2\sigma_{xu}(t)Y(t)v(t) \\ &+\sigma_{uu}(t)v^2(t)\Big]\,dt. \end{aligned} \tag{3.33}$$

Combining (3.31), (3.32), and (3.33) into (3.30), with the fact that $Z(0)=0$, we get

$$\begin{aligned} &\boldsymbol{E}\left[p\left(T\right)Z\left(T\right)\right] \\ &= \boldsymbol{E}\int_0^T \Big[p\left(t\right)f_{xx}\left(t\right)Y^2\left(t\right)+2p\left(t\right)f_{xu}\left(t\right)Y\left(t\right)v(t) \\ &+p\left(t\right)f_{uu}\left(t\right)v^2\left(t\right)+q(t)\sigma_{xx}\left(t\right)Y^2\left(t\right) \\ &+2q(t)\sigma_{xu}\left(t\right)Y\left(t\right)v\left(t\right)+q(t)\sigma_{uu}\left(t\right)v^2\left(t\right) \\ &+\ell_x(t)Z\left(t\right)\Big]\mathrm{d}t, \end{aligned}$$

this completes the proof of (3.24).

Proof of (3.25). By Itô's formula to $P\left(t\right)Y^2\left(t\right)$, we have

$$\begin{aligned} &\boldsymbol{E}\left[P\left(T\right)Y^2\left(T\right)\right]-\boldsymbol{E}\left[P\left(0\right)Y^2\left(0\right)\right] \\ &= \boldsymbol{E}\int_0^T P(t)Y(t)\mathrm{d}Y(t)+\boldsymbol{E}\int_0^T Y(t)\mathrm{d}(P(t)Y(t)) \\ &+\boldsymbol{E}\int_0^T \left[P(t)(\sigma_x(t)Y(t)+\sigma_u(t)v(t))+Y(t)Q(t)\right]\left[\sigma_x(t)Y(t)+\sigma_u(t)v(t)\right]\mathrm{d}t \\ &= A_1+A_2+A_3. \end{aligned} \tag{3.34}$$

$$\begin{aligned} A_1 &= \boldsymbol{E}\int_0^T P(t)Y(t)\mathrm{d}Y(t) \\ &= \boldsymbol{E}\int_0^T P(t)Y(t)\left[f_x(t)Y(t)+f_u(t)v(t)\right]\mathrm{d}t \\ &+ \boldsymbol{E}\int_{[0,T]} P(t)Y(t)G(t)\mathrm{d}\zeta(t). \end{aligned} \tag{3.35}$$

Analogously, we can have a similar estimate for A_2.We again use *Itô's formula*, we have

$$\begin{aligned} A_2 &= \boldsymbol{E}\int_0^T Y(t)\mathrm{d}(P(t)Y(t)) \\ &= \boldsymbol{E}\int_0^T Y(t)P(t)\mathrm{d}Y(t) + \boldsymbol{E}\int_0^T Y^2(t)\mathrm{d}P(t) \\ &+ \boldsymbol{E}\int_0^T Y(t)\left[Q(t)\left(\sigma_x(t)Y(t)+\sigma_u(t)v(t)\right)\right]\mathrm{d}t \\ &= \boldsymbol{E}\int_0^T Y(t)P(t)\left[f_x(t)Y(t)+f_u(t)v(t)\right]\mathrm{d}t + \boldsymbol{E}\int_0^T Y(t)P(t)G(t)\mathrm{d}\zeta(t) \\ &- \boldsymbol{E}\int_0^T Y^2(t)\left[f_x(t)P(t)+P(t)f_x(t)+\sigma_x(t)P(t)\sigma_x(t)+\sigma_x(t)Q(t)\right. \\ &+ \left. Q(t)\sigma_x(t)+H_{xx}(t)\right]\mathrm{d}t \\ &+ \boldsymbol{E}\int_0^T Y(t)\left[Q(t)\left(\sigma_x(t)Y(t)+\sigma_u(t)v(t)\right)\right]\mathrm{d}t. \end{aligned}$$

By simple computations, we obtain

$$\begin{aligned} A_2 &= \boldsymbol{E}\int_0^T Y(t)P(t)\left[f_u(t)v(t)\right]\mathrm{d}t + \boldsymbol{E}\int_0^T Y(t)P(t)G(t)\mathrm{d}\zeta(t) \\ &- \boldsymbol{E}\int_0^T Y^2(t)\left[f_x(t)P(t)+\sigma_x^2(t)P(t)+\sigma_x(t)Q(t)+H_{xx}(t)\right]\mathrm{d}t \\ &+ \boldsymbol{E}\int_0^T Y(t)Q(t)\sigma_u(t)v(t)\mathrm{d}t. \end{aligned} \tag{3.36}$$

and it is easy to show that

$$A_3 = \boldsymbol{E}\int_0^T \left[P(t)\sigma_x(t)Y(t)+P(t)\sigma_u(t)v(t)+Y(t)Q(t)\right]\left[\sigma_x(t)Y(t)+\sigma_u(t)v(t)\right]\mathrm{d}t. \tag{3.37}$$

Now, by substituting (3.35), (3.36), and (3.37) into (3.34), with the fact that $Y(0)=0$, we get

$$\begin{aligned}
&\boldsymbol{E}\left[P(T)Y^{2}(T)\right] \\
&=\boldsymbol{E}\int_{0}^{T}[P(t)f_{u}(t)Y(t)v(t)+P(t)f_{u}(t)Y(t)v(t) \\
&+P(t)\sigma_{x}(t)\sigma_{u}(t)Y(t)v(t)+P(t)\sigma_{u}(t)\sigma_{x}(t)Y(t)v(t) \\
&+P(t)\sigma_{u}^{2}(t)v^{2}(t)+Q(t)\sigma_{u}(t)Y(t)v(t) \\
&+Q(t)Y(t)\sigma_{u}(t)v(t)-H_{xx}(t)Y^{2}(t)\big]\,\mathrm{d}t \\
&+\boldsymbol{E}\int_{[0,T]}2Y(t)P(t)G(t)\mathrm{d}\zeta(t).
\end{aligned}\tag{3.38}$$

Finally, since $P(T)=-h_{xx}(x^{*}(T))$, then the desired result (3.25) is fulfilled, which completes the proof of Lemma 3.1 □

To prove the main theorem we need the following technical result.

Proposition 3.2 Let assumptions (H1), (H2) and (H3) hold. Then, the following variational equality holds : for any $(u(\cdot),\xi(\cdot))\in\mathcal{A}_{1}\times\mathcal{A}_{2}([0,T])$,

$$\begin{aligned}
&J(u^{\varepsilon}(\cdot),\xi^{\varepsilon}(\cdot))-J(u^{*}(\cdot),\xi^{*}(\cdot)) \\
&=-\boldsymbol{E}\int_{0}^{T}\Big[\varepsilon H_{u}(t)v(t)+\frac{\varepsilon^{2}}{2}H_{uu}(t)v^{2}(t) \\
&+\frac{\varepsilon^{2}}{2}P(t)\sigma_{u}^{2}(t)v^{2}(t)+\varepsilon^{2}\mathbb{K}(t)Y(t)v(t)\Big]dt \\
&+\varepsilon\boldsymbol{E}\int_{[0,T]}(M(t)-p(t)G(t))\,\mathrm{d}\zeta(t) \\
&-\varepsilon^{2}\boldsymbol{E}\int_{[0,T]}Y(t)P(t)G(t)\mathrm{d}\zeta(t)+o\left(\varepsilon^{2}\right),\quad\left(\varepsilon\to0^{+}\right).
\end{aligned}\tag{3.39}$$

where $H_{u}(t)=H_{u}(t,x^{*},u^{*},p,q)$ and $H_{uu}(t)=H_{uu}(t,x^{*},u^{*},p,q)$, with $v(\cdot)=u(\cdot)-u^{*}(\cdot)$ and $\zeta(\cdot)=\xi(\cdot)-\xi^{*}(\cdot)$.

Proof. From (3.2), we have

$$\begin{aligned}
&J\left(u^{\varepsilon}\left(\cdot\right),\xi^{\varepsilon}(\cdot)\right)-J\left(u^{*}\left(\cdot\right),\xi^{*}(\cdot)\right)\\
&=\boldsymbol{E}\left[h(x^{\varepsilon}(T)-h(x^{*}(T)\right]\\
&+\boldsymbol{E}\int_{0}^{T}\left[\ell(t,x^{\varepsilon}(t),u^{\varepsilon}(t))-\ell(t,x^{*}(t),u^{*}(t))\right]\mathrm{d}t\\
&+\boldsymbol{E}\int_{[0,T]}M(t)\mathrm{d}(\xi^{\varepsilon}-\xi^{*})(t).
\end{aligned}\tag{3.40}$$

Applying Taylor-Young's formula for the function $\ell\left(t,\cdot,\cdot\right)$, we get

$$\begin{aligned}
&\ell(t,x^{\varepsilon}(t),u^{\varepsilon}(t))-\ell(t,x^{*}(t),u^{*}(t))\\
&=\ell_{x}(t,x^{*}(t),u^{*}(t))\left(x^{\varepsilon}(t)-x^{*}(t)\right)\\
&+\ell_{u}(t,x^{*}(t),u^{*}(t))\left(u^{\varepsilon}(t)-u^{*}(t)\right)\\
&+\frac{1}{2}\left[\ell_{xx}(t,x^{*}(t),u^{*}(t))\left(x^{\varepsilon}(t)-x^{*}(t)\right)^{2}+\ell_{uu}(t,x^{*}(t),u^{*}(t))\left(u^{\varepsilon}(t)-u^{*}(t)\right)^{2}\right.\\
&\left.+\,2\ell_{xu}(t,x^{*}(t),u^{*}(t))\left(x^{\varepsilon}(t)-x^{*}(t)\right)\left(u^{\varepsilon}(t)-u^{*}(t)\right)\right].
\end{aligned}\tag{3.41}$$

Substituting (3.41) into (3.40), we obtain

$$\begin{aligned}
&J\left(u^{\varepsilon}\left(\cdot\right),\xi^{\varepsilon}(\cdot)\right)-J\left(u^{*}\left(\cdot\right),\xi^{*}(\cdot)\right)\\
&=\boldsymbol{E}\int_{0}^{T}\left[\ell_{x}(t)\delta x\left(t\right)+\varepsilon\ell_{u}(t)v(t)+\frac{1}{2}\ell_{xx}(t)\delta x\left(t\right)^{2}\right.\\
&\left.+\,\varepsilon\ell_{xu}(t)\delta x\left(t\right)v(t)+\frac{\varepsilon^{2}}{2}\ell_{uu}(t)v^{2}(t)\right]\mathrm{d}t\\
&+\boldsymbol{E}\left[h_{x}(x^{*}\left(T\right))\delta x\left(T\right)+\frac{1}{2}h_{xx}\left(x^{*}\left(T\right)\right)\delta x\left(T\right)^{2}\right]\\
&+\boldsymbol{E}\int_{[0,T]}M(t)\mathrm{d}\left(\xi^{\varepsilon}-\xi^{*}\right)(t)+o\left(\varepsilon^{2}\right),\quad\left(\varepsilon\rightarrow0^{+}\right).
\end{aligned}\tag{3.42}$$

From Proposition 3.1 and since $\xi^{\varepsilon}(t)-\xi^{*}(t)=\varepsilon\left(\xi(t)-\xi^{*}(t)\right)=\varepsilon\zeta(t)$, we deduce

$$\begin{aligned}
&J\left(u^{\varepsilon}\left(\cdot\right),\xi^{\varepsilon}(\cdot)\right)-J\left(u^{*}\left(\cdot\right),\xi^{*}(\cdot)\right)\\
&=\boldsymbol{E}\int_{0}^{T}\left[\varepsilon\ell_{x}(t)Y\left(t\right)+\frac{\varepsilon^{2}}{2}\ell_{x}(t)Z\left(t\right)+\varepsilon\ell_{u}(t)v\left(t\right)\right.\\
&\left.+\frac{\varepsilon^{2}}{2}\left(\ell_{xx}(t)Y^{2}\left(t\right)+2\ell_{xu}(t)Y\left(t\right)v\left(t\right)+\ell_{uu}(t)v^{2}\left(t\right)\right)\right]\mathrm{d}t\\
&+\boldsymbol{E}\left[\varepsilon h_{x}(x^{*}\left(T\right))Y\left(T\right)+\frac{\varepsilon^{2}}{2}h_{x}\left(x^{*}\left(T\right)\right)Z\left(T\right)+\frac{\varepsilon^{2}}{2}h_{xx}\left(x^{*}\left(T\right)\right)Y^{2}\left(T\right)\right]\\
&+\boldsymbol{E}\int_{[0,T]}\varepsilon M(t)\mathrm{d}\zeta(t)+o\left(\varepsilon^{2}\right).\ \ \left(\varepsilon\rightarrow0^{+}\right).
\end{aligned}\tag{3.43}$$

Substituting (3.23), (3.24), and (3.25) into (3.43), we obtain

$$\begin{aligned}
&J\left(u^{\varepsilon}\left(\cdot\right),\xi^{\varepsilon}(\cdot)\right)-J\left(u^{*}\left(\cdot\right),\xi^{*}(\cdot)\right)\\
&=-\boldsymbol{E}\int_{0}^{T}\left[\varepsilon\left(p\left(t\right)f_{u}\left(t\right)v\left(t\right)+q(t)\sigma_{u}\left(t\right)v\left(t\right)+\ell_{u}(t)v\left(t\right)\right)\right.\\
&+\frac{\varepsilon^{2}}{2}\left(p\left(t\right)f_{uu}\left(t\right)v^{2}\left(t\right)+q(t)\sigma_{uu}\left(t\right)v^{2}\left(t\right)\right.\\
&\left.-\ell_{uu}(t)v^{2}\left(t\right)\right)+\frac{\varepsilon^{2}}{2}P(t)\sigma_{u}^{2}\left(t\right)v^{2}\left(t\right)\\
&+\varepsilon^{2}\left(p\left(t\right)f_{xu}\left(t\right)Y\left(t\right)v\left(t\right)+q(t)\sigma_{xu}\left(t\right)Y\left(t\right)v\left(t\right)\right.\\
&-\ell_{xu}(t)Y\left(t\right)v\left(t\right)+f_{u}(t)P(t)Y\left(t\right)v\left(t\right)\\
&\left.\left.+\sigma_{u}\left(t\right)\sigma_{x}\left(t\right)P(t)Y\left(t\right)v\left(t\right)+\ \sigma_{u}\left(t\right)Q(t)Y\left(t\right)v\left(t\right)\right)\right]\mathrm{d}t\\
&+\varepsilon\boldsymbol{E}\int_{[0,T]}\left(M(t)-p(t)G(t)\right)\mathrm{d}\zeta(t)-\varepsilon^{2}\boldsymbol{E}\int_{[0,T]}Y(t)P(t)G(t)\mathrm{d}\zeta(t)\\
&+o\left(\varepsilon^{2}\right),\quad\left(\varepsilon\rightarrow0^{+}\right)
\end{aligned}$$

From (3.4), and (3.7), we get

$$
\begin{aligned}
& J\left(u^{\varepsilon}(\cdot), \xi^{\varepsilon}(\cdot)\right)-J\left(u^{*}(\cdot), \xi^{*}(\cdot)\right) \\
& =-\boldsymbol{E} \int_{0}^{T}\left[\varepsilon H_{u}(t) v(t)+\frac{\varepsilon^{2}}{2} H_{u u}(t) v^{2}(t)\right. \\
& \left.+\frac{\varepsilon^{2}}{2} P(t) \sigma_{u}^{2}(t) v^{2}(t)+\varepsilon^{2} \mathbb{K}(t) Y(t) v(t)\right] \mathrm{d} t \\
& +\varepsilon \boldsymbol{E} \int_{[0, T]}(M(t)-p(t) G(t)) \mathrm{d} \zeta(t)-\varepsilon^{2} \boldsymbol{E} \int_{[0, T]} Y(t) P(t) G(t) \mathrm{d} \zeta(t)+o\left(\varepsilon^{2}\right), \\
& \left(\varepsilon \rightarrow 0^{+}\right).
\end{aligned}
$$

Thus, we finish the proof of Proposition 3.2 $\square$

Proof of Theorem 3.1 From Proposition 3.2, we have

$$
\begin{aligned}
& \frac{1}{\varepsilon^{2}}\left[J\left(u^{\varepsilon}(\cdot), \xi^{\varepsilon}(\cdot)\right)-J\left(u^{*}(\cdot), \xi^{*}(\cdot)\right)\right] \\
& =-\boldsymbol{E} \int_{0}^{T}\left[\frac{1}{\varepsilon} H_{u}(t) v(t)+\frac{1}{2}\left[H_{u u}(t)+P(t) \sigma_{u}^{2}(t)\right] v^{2}(t)+\mathbb{K}(t) Y(t) v(t)\right] \mathrm{d} t \\
& +\frac{1}{\varepsilon} \boldsymbol{E} \int_{[0, T]}(M(t)-p(t) G(t)) \mathrm{d} \zeta(t) \\
& -\boldsymbol{E} \int_{[0, T]} Y(t) P(t) G(t) \mathrm{d} \zeta(t)+o\left(\varepsilon^{2}\right), \quad\left(\varepsilon \rightarrow 0^{+}\right).
\end{aligned} \tag{3.44}
$$

Applying (3.3) and Definition 2.2, we shows that

$$
\begin{aligned}
0 & \leq \lim_{\varepsilon \rightarrow 0} \frac{1}{\varepsilon^{2}}\left[J\left(u^{\varepsilon}(\cdot), \xi^{\varepsilon}(\cdot)\right)-J\left(u^{*}(\cdot), \xi^{*}(\cdot)\right)\right] \\
& =-\boldsymbol{E} \int_{0}^{T} \mathbb{K}(t) Y(t) v(t) d t-\boldsymbol{E} \int_{[0, T]} Y(t) P(t) G(t) \mathrm{d} \zeta(t),
\end{aligned} \tag{3.45}
$$

then the desired result (3.11) is fulfilled.

Now let us turn to prove (3.12). From the singularity in (3.9) holds for any $\xi(\cdot) \in \mathcal{A}_{2}([0, T])$.

$$
\boldsymbol{E} \int_{[0, T]}(M(t)-p(t) G(t)) \mathrm{d}\left(\xi-\xi^{*}\right)(t)=0.
$$

Let $\xi(\cdot) \in \mathcal{A}_2([0,T])$ be defined by

$$\mathrm{d}\xi(t) = \begin{cases} 0 \text{ if } (M(t) - p(t)G(t)) \geq 0, \\ \mathrm{d}\xi^*(t) \text{ if } (M(t) - p(t)G(t)) < 0, \end{cases} \tag{3.46}$$

Let $\mathcal{N}$ a set be defined by

$$\mathcal{N} = \{(t,w) \in [0,T] \times \Omega : (M(t) - p(t)G(t)) \geq 0\}.$$

This means that

$$\begin{aligned} \mathrm{d}\xi(t) &= I_{\mathcal{N}}\mathrm{d}\xi(t) + I_{\mathcal{N}^c}\mathrm{d}\xi(t) \\ &= I_{\{(t,w)\in[0,T]\times\Omega:(M(t)-p(t)G(t))<0\}}(t)\mathrm{d}\xi^*(t). \end{aligned} \tag{3.47}$$

By a simple computations, it is easy to see that $\xi(\cdot)$ is in $\mathcal{A}_2([0,T])$. Moreover, we have

$$\begin{aligned} 0 &= \boldsymbol{E}\int_{[0,T]} (M(t) - p(t)G(t))\,\mathrm{d}(\xi - \xi^*)(t) \\ &= \boldsymbol{E}\int_{[0,T]} (M(t) - p(t)G(t))\, I_{\{(t,w)\in[0,T]\times\Omega:(M(t)-p(t)G(t))<0\}}\mathrm{d}(\xi^* - \xi^*)(t), \\ &+ \boldsymbol{E}\int_{[0,T]} (M(t) - p(t)G(t))\, I_{\{(t,w)\in[0,T]\times\Omega:(M(t)-p(t)G(t))\geq 0\}}\mathrm{d}(-\xi^*)(t), \end{aligned}$$

then we conclude that

$$\boldsymbol{E}\int_{[0,T]} (M(t) - p(t)G(t))\, I_{\mathcal{N}}(t)\mathrm{d}\xi^*(t) = 0. \tag{3.48}$$

This completes the proof of Theorem 3.1. □

Theorem 3.2. Let assumptions (H1), (H2) and (H3) hold. If $(u^*(\cdot), \xi^*(\cdot))$ is an optimal continuous-singular stochastic control that satisfy (3.9), then for any $u(\cdot) \in \mathcal{A}_1([0,T])$ and for any $\xi(\cdot) \in \mathcal{A}_2([0,T])$ we have

$$\boldsymbol{E}\int_0^T \mathbb{K}(t)Y(t)(u(t)-u^*(t))\mathrm{d}t \leq 0, \tag{3.49}$$

$$\boldsymbol{E}\int_{[0,T]} Y(t)P(t)G(t)\mathrm{d}\xi^*(t) \geq \boldsymbol{E}\int_{[0,T]} Y(t)P(t)G(t)\mathrm{d}\xi(t), \tag{3.50}$$

where $\mathbb{K}(\cdot)$ is given by equation (3.7) and $Y(\cdot)$ is the solution of the first variational equation (3.15)

Proof The inequality (3.11) is valid for every $(u(\cdot),\xi(\cdot)) \in \mathcal{A}_1 \times \mathcal{A}_2([0,T])$. If we choose $\xi = \xi^*$ in inequality (3.11), we see that for every measurable $\mathbb{F}-$adapted process $u(\cdot):[0,T]\times\Omega \to \mathbb{A}_1$, the inequality

$$\boldsymbol{E}\int_0^T \mathbb{K}(t)Y(t)(u(t)-u^*(t))\mathrm{d}t \leq 0.$$

holds. Then the desired result (3.49) is fulfilled.

Further, if instead we choose $u = u^*(t)$ in inequality (3.11), we see that for every measurable $\mathbb{F}-$adapted process $\xi(\cdot):[0,T]\times\Omega \to \mathbb{A}_2$.satisfies *Definition 2.1*, the inequality

$$\boldsymbol{E}\int_{[0,T]} Y(t)P(t)G(t)\mathrm{d}(\xi-\xi^*)(t) \leq 0. \tag{3.51}$$

holds. The desired result (3.50) follows immediately from (3.51), which completes the proof of Theorem 4.2. □

3.4 Pointwise second-order necessary condition in terms of martingale

Our purpose in this section is to establish a pointwise second-order necessary conditions for optimal controls. Note that the solution $Y(\cdot)$ to the first variational equation (3.15) appears in the integral-type second-order condition (3.11).

Lemma 4.1 The first variational equation (3.15) admits a unique strong solution $Y(\cdot)$, which is represented by the following stochastic differential equation :

$$\begin{aligned} Y(t) = \Phi(t) \bigg[\int_0^t \Psi(s) \left[f_u(s) - \sigma_x(s)\sigma_u(s) \right] v(s) \mathrm{d}s \\ + \int_0^t \Psi(s)\sigma_u(s)v(s)\mathrm{d}W(s) + \int_{[0,t]} \Psi(s)G(s)\mathrm{d}\zeta(s) \bigg], \end{aligned} \tag{3.52}$$

where $\Phi(t)$ is a defined by the following linear stochastic differential equation :

$$\begin{cases} \mathrm{d}\Phi(t) = f_x(t)\Phi(t)\,\mathrm{d}t + \sigma_x(t)\Phi(t)\,\mathrm{d}W(s) \\ \Phi(0) = 1. \end{cases} \tag{3.53}$$

and $\Psi(t)$ its inverse.

Proof. Equation (3.15) is linear with bounded coefficients, then it admits a unique strong solution. Moreover, this solution is invertible and its inverse $\Psi(t) = \Phi^{-1}(t)$ given by the following equation :

$$\begin{cases} \mathrm{d}\Psi(t) = \left[\sigma_x^2(t)\Psi(t) - f_x(t)\Psi(t)\right]\mathrm{d}t - \left[\sigma_x(t)\Psi(t)\right]\mathrm{d}W(t) \\ \Psi(0) = 1. \end{cases} \tag{3.54}$$

By applying Itô's formula to $\Psi(t)Y(t)$ we get

$$\begin{aligned} \mathrm{d}\left[\Psi(t)Y(t)\right] &= Y(t)\,\mathrm{d}\Psi(t) + \Psi(t)\mathrm{d}Y(t) + \mathrm{d}\left\langle \Psi(t), Y(t) \right\rangle, \\ &= I_1 + I_2 + I_3, \end{aligned} \tag{3.55}$$

where

$$\begin{aligned} I_1 &= Y(t)\,\mathrm{d}\Psi(t) \\ &= \left[Y(t)\,\sigma_x^2(t)\Psi(t) - Y(t)\,f_x(t)\Psi(t)\right]\mathrm{d}t \\ &- Y(t)\,\sigma_x(t)\Psi(t)\mathrm{d}W(t), \end{aligned} \tag{3.56}$$

By simple computations, we obtain

$$\begin{aligned} I_2 &= \Psi(t)\mathrm{d}Y(t) \\ &= \left[\Psi(t)f_x(t)Y(t) + \Psi(t)f_u(t)v(t)\right]\mathrm{d}t \\ &+ \left[\Psi(t)\sigma_x(t)Y(t) + \Psi(t)\sigma_u(t)v(t)\right]\mathrm{d}W(t) \\ &+ \Psi(t)G(t)\mathrm{d}\zeta(t), \end{aligned} \tag{3.57}$$

and

$$\begin{aligned} I_3 &= \mathrm{d}\left\langle \Psi(t), Y(t)\right\rangle \\ &= -\left[\sigma_x(t)\Psi(t)\right]\left[\sigma_x(t)Y(t) + \sigma_u(t)v(t)\right]\mathrm{d}t \end{aligned} \tag{3.58}$$

Substituting (3.56), (3.57), and (3.58) into (3.55), we get

$$\begin{aligned} &\Psi(t)Y(t) - \Psi(0)Y(0) \\ &= \int_0^t \Psi(s)\left[f_u(s) - \sigma_x(s)\sigma_u(s)\right]v(s)\mathrm{d}s \\ &+ \int_0^t \Psi(s)\sigma_u(s)v(s)\mathrm{d}W(s) + \int_{[0,t]} \Psi(s)G(s)\mathrm{d}\zeta(s), \end{aligned}$$

Since $Y(0) = 0$, and $\Psi^{-1}(t) = \Phi(t)$ then the desired result (3.52) is fulfilled. Thus, we finish the proof of Lemma 4.1 □

We need the following simple lemma, proved in [95, Lemma 3.8] by applying martingale representation theorem.

Lemma 4.2 Let (H1)-(H2) hold. Then $\mathbb{K}(\cdot) \in \mathbb{L}^2_{\mathcal{F}}([0,T],\mathbb{R})$ and for any $v \in \mathbb{A}_1$, there exists a $\phi_v(\cdot,t) \in \mathbb{L}^2_{\mathcal{F}}([0,t],\mathbb{R})$ such that

$$\mathbb{K}(t)(v-u^*(t)) = \boldsymbol{E}\left[\mathbb{K}(t)(v-u^*(t))\right] + \int_0^t \phi_v(s,t)\,\mathrm{d}W(s) \tag{3.59}$$
$$a.\boldsymbol{e}.\ t\in[0,T],\ \ P-a.s.$$

We note that, for every $k \geq 1$ it follows from the *Burkholder-Davis-Gundy* and *Hölder's inequalities* that there exists a constant C_k independent of t such that

$$\sup_{t\in[0,T]} \boldsymbol{E}\left[\left(\int_0^t \left|\phi_v(s,t)\right|^2 ds\right)^{\frac{k}{2}}\right] \leq C_k.$$

In this section, our aim is to prove pointwise second-order maximum principle form for optimality. The following theorem constitutes the second main contribution of the paper

Theorem 4.1. Let assumptions (H1), (H2) and (H3) hold. If $(u^*(\cdot),\xi^*(\cdot))$ is an optimal continuous-singular stochastic control that satisfy (3.9), then for any $(v,\zeta) \in \mathbb{A}_1 \times \mathbb{A}_2$ it holds that

$$\begin{aligned}0 \geq\ & \boldsymbol{E}\big[\mathbb{K}(\tau)\, f_u(\tau)(v-u^*(\tau))^2 \\ & + \left[(\mathbb{K}(\tau)+P(\tau)(f_u(\tau)))\,G(\tau)(v-u^*(\tau))\zeta(\tau)\right] \\ & + \left[P(\tau)G^2(\tau)\zeta^2(\tau)\right]\big] + \mathcal{D}^+_\tau(\mathbb{K}(\tau)\,\sigma_u(\tau)\,(v-u^*(\tau))^2)\end{aligned} \tag{3.60}$$
$$a.e.\ \tau\in[0,T].$$

$$\boldsymbol{E}\int_{[0,T]} (M(t)-p(t)G(t))\, I_{\{(w,t)\in\Omega\times[0,T]:(M(t)-p(t)G(t))\geq 0\}}\mathrm{d}\xi^*(t) = 0, \tag{3.61}$$

where $\mathbb{K}(t)$ is given by equation (3.7) and $\mathcal{D}_{\tau}^{+}(\mathbb{K}(\tau)\,\sigma_u(\tau)\,(v-u^*(\tau))^2)$ is given by

$$\begin{aligned}
&\mathcal{D}_{\tau}^{+}(\mathbb{K}(\tau)\,\sigma_u(\tau)\,(v-u^*(\tau))^2) \\
&= 2\lim_{\theta\to 0^+}\sup\frac{1}{\theta^2}\boldsymbol{E}\int_{\tau}^{\tau+\theta}\int_{\tau}^{t}\left[\Phi(\tau)\,\Psi(s)\,\sigma_u(s)\phi_v(s,t)(v-u^*(s))\right]\mathrm{d}s\mathrm{d}t.
\end{aligned}$$

Here $\Phi(\cdot)$ is given by (3.53), $\Psi(\cdot)$ is given by (3.54) and $\phi_v(\cdot,\cdot)$ is defined by (3.59).

Proof. In order to establish a pointwise second-order necessary condition from the integral one (3.11), we need to choose the following *needle variation* for the optimal control $(u^*(\cdot),\xi^*(\cdot))$ by the form :

$$(u(t),\xi(t)) = \begin{cases} (v,\xi^*(t)+\theta\zeta(t)), & t\in\boldsymbol{E}_\theta \\ (u^*(t),\xi^*(t)+\theta\zeta(t)), & t\in[0,T]\mid\boldsymbol{E}_\theta. \end{cases} \tag{3.62}$$

For any $(v,\zeta)\in\mathbb{A}_1\times\mathbb{A}_2$, $\tau\in[0,T)$, and $\theta\in(0,T-\tau)$, let $\boldsymbol{E}_\theta=[\tau,\tau+\theta)$, and define $u(\cdot)$ as that in (3.62). Then $v(\cdot)=u(\cdot)-u^*(\cdot)=(v-u^*(\cdot))\,I_{\boldsymbol{E}_\epsilon}(\cdot)$.

We note that from Lemma 4.1, we can rewrite equation (3.52) in the form

$$\begin{aligned}
Y(t) &= \Phi(t)\int_0^t\Psi(s)\left[f_u(s)-\sigma_x(s)\sigma_u(s)\right]v(s)\mathrm{d}s+\Phi(t)\int_0^t\Psi(s)\sigma_u(s)v(s)\mathrm{d}W(s) \\
&\quad+\Phi(t)\int_{[0,t]}\Psi(s)G(s)\mathrm{d}\zeta(s), \\
&= y_1(t)+\Phi(t)\int_{[0,t]}\Psi(s)G(s)\mathrm{d}\zeta(s).
\end{aligned}\tag{3.63}$$

where $y_1(t)$ is defined as in [95, Eq-(3.21)] by the following equation :

$$\begin{aligned}
y_1(t) = \Phi(t)\int_0^t\Psi(s)\left[f_u(s)-\sigma_x(s)\sigma_u(s)\right]v(s)\mathrm{d}s \\
+\Phi(t)\int_0^t\Psi(s)\sigma_u(s)v(s)\mathrm{d}W(s).
\end{aligned}$$

Substituting $v(\cdot)=(v-\bar{u}(\cdot))I_{\boldsymbol{E}_\theta}(\cdot)$ and (3.63) into (3.11), we have

$$\begin{aligned}
0 &\geq \frac{1}{\theta^2}\boldsymbol{E}\int_\tau^{\tau+\theta}\mathbb{K}(t)\,y_1(t)(v-u^*(t))dt \\
&+\frac{1}{\theta^2}\boldsymbol{E}\int_\tau^{\tau+\theta}\mathbb{K}(t)\,\Phi(t)\,(v-u^*(t))\int_{[\tau,t]}\Psi(s)G(s)\mathrm{d}\zeta(s)dt \\
&+\frac{1}{\theta^2}\boldsymbol{E}\int_{[\tau,\tau+\theta]}Y(t)P(t)G(t)\mathrm{d}(\xi-\xi^*)(t) \\
&= I_1(\theta)+I_2(\theta)+I_3(\theta),
\end{aligned} \tag{3.64}$$

where

$$I_1(\theta)=\frac{1}{\theta^2}\boldsymbol{E}\int_\tau^{\tau+\theta}\mathbb{K}(t)\,y_1(t)(v-u^*(t))dt, \tag{3.65}$$

$$I_2(\theta)=\frac{1}{\theta^2}\boldsymbol{E}\int_\tau^{\tau+\theta}\mathbb{K}(t)\,\Phi(t)\,(v-u^*(t))\int_{[\tau,t]}\Psi(s)G(s)\mathrm{d}\zeta(s)dt, \tag{3.66}$$

$$I_3(\theta)=\frac{1}{\theta^2}\boldsymbol{E}\int_{[\tau,\tau+\theta]}Y(t)P(t)G(t)\mathrm{d}(\xi-\xi^*)(t). \tag{3.67}$$

Estimate of (3.65). Using the similar arguments developed in [95, Theorem 3.10], we obtain

$$\begin{aligned}
\limsup_{\theta\to0^+} I_1(\theta) &= \frac{1}{2}\boldsymbol{E}\left[\mathbb{K}(\tau)\,f_u(\tau)(v-u^*(\tau))^2\right] \\
&+\frac{1}{2}\mathcal{D}_\tau^+(\mathbb{K}(\tau)\,\sigma_u(\tau)\,(v-u^*(\tau))^2).
\end{aligned} \tag{3.68}$$

where

$$\begin{aligned}
&\mathcal{D}_\tau^+(\mathbb{K}(\tau)\,\sigma_u(\tau)\,(v-u^*(\tau))^2) \\
&=2\lim_{\theta\to0^+}\sup\frac{1}{\theta^2}\boldsymbol{E}\int_\tau^{\tau+\theta}\int_\tau^t\left[\Phi(\tau)\,\Psi(s)\,\sigma_u(s)\phi_v(s,t)(v-u^*(s))\right]dsdt.
\end{aligned} \tag{3.69}$$

We note that by the Martingale Representation Theorem in Lemma 4.2, we only know that $\phi_v(\cdot,t)\in\mathbb{L}^2_{\mathcal{F}}([0,t];\mathbb{R})$ for any $v\in\mathbb{A}_1$, and hence, for each $\tau\in[0,T]$, the function

$\varphi_t(\cdot)$

$$\varphi_t(s) = \boldsymbol{E}\left[\Phi(\tau)\Psi(s)\sigma_u(s)(v-u^*(s))\phi_v(s,t)\right], \quad s\in[0,t], \; t\in[0,T],$$

is in $\mathbb{L}^1_{\mathcal{F}}([0,t],\mathbb{R})$. See [95] for more details of integrals $\int_\tau^{\tau+\theta}\int_\tau^t \varphi_t(s)\,dsdt$, and its superior limit $\lim_{\theta\to 0^+}\frac{1}{\theta^2}\int_\tau^{\tau+\theta}\int_\tau^t \varphi_t(s)\,dsdt$.

Estimate of (3.66). From [95, Lemma 4.1] and Dominate Convergence Theorem, we have

$$\begin{aligned}\lim_{\theta\to 0^+} I_2(\theta) &= \lim_{\theta\to 0^+}\frac{1}{\theta^2}\boldsymbol{E}\int_\tau^{\tau+\theta}\mathbb{K}(t)\Phi(t)(v-u^*(t))\int_{[\tau,t]}\Psi(s)G(s)\mathrm{d}\zeta(s)dt \\ &= \frac{1}{2}\boldsymbol{E}\left[\mathbb{K}(\tau)G(\tau)(v-u^*(\tau))\zeta(\tau)\right].\end{aligned} \tag{3.70}$$

Estimate of (3.67). Now, let us turn to estimate $I_3(\theta)$. From (3.63), we have

$$\begin{aligned}I_3(\theta) &= \frac{1}{\theta^2}\boldsymbol{E}\int_{[\tau,\tau+\theta]}Y(t)P(t)G(t)\mathrm{d}(\xi-\xi^*)(t) \\ &= I_3^1(\theta)+I_3^2(\theta)+I_3^3(\theta),\end{aligned} \tag{3.71}$$

where

$$I_3^1(\theta) = \frac{1}{\theta^2}\boldsymbol{E}\int_{[\tau,\tau+\theta]}\Phi(t)\left[\int_\tau^t\Psi(s)\sigma_u(s)(v-u^*(t))\mathrm{d}W(s)\right]P(t)G(t)\mathrm{d}\zeta(t) \tag{3.72}$$

$$I_3^2(\theta) = \frac{1}{\theta^2}\boldsymbol{E}\int_{[\tau,\tau+\theta]}\Phi(t)\left[\int_\tau^t\Psi(s)\left[f_u(s)-\sigma_x(s)\sigma_u(s)\right](v-u^*(t))\mathrm{d}s\right] \tag{3.73}$$

$$\times P(t)G(t)\mathrm{d}\zeta(t)$$

$$I_3^3(\theta) = \frac{1}{\theta^2}\boldsymbol{E}\int_{[\tau,\tau+\theta]}\Phi(t)\left[\int_\tau^t\Psi(s)G(s)\mathrm{d}\zeta(s)\right]P(t)G(t)\mathrm{d}\zeta(t). \tag{3.74}$$

Estimate of (3.72). From [95, Eq-(3.21)] and (3.53), we have

$$\Phi(t) = \Phi(\tau)+\int_\tau^t\Phi(s)f_x(s)\mathrm{d}s+\int_\tau^t\Phi(s)f_x(s)\mathrm{d}W(s). \tag{3.75}$$

Substituting (3.75) into (3.72), we obtain

$$I_3^1(\theta) = \frac{1}{\theta^2}\boldsymbol{E}\int_{[\tau,\tau+\theta]}\Phi(t)\left[\int_\tau^t \Psi(s)\sigma_u(s)(v-u^*(s))\mathrm{d}W(s)\right]P(t)G(t)\mathrm{d}\zeta(t) \tag{3.76}$$
$$= I_3^{1,1}(\theta) + I_3^{1,2}(\theta) + I_3^{1,3}(\theta),$$

where

$$I_3^{1,1}(\theta) = \frac{1}{\theta^2}\boldsymbol{E}\int_{[\tau,\tau+\theta]}\Phi(\tau)\left[\int_\tau^t \Psi(s)\sigma_u(s)(v-u^*(s))\mathrm{d}W(s)\right]P(t)G(t)\mathrm{d}\zeta(t)$$
$$I_3^{1,2}(\theta) = \frac{1}{\theta^2}\boldsymbol{E}\int_{[\tau,\tau+\theta]}\left[\int_\tau^t \Phi(s)f_x(s)\mathrm{d}s\right] \times\left[\int_\tau^t \Psi(s)\sigma_u(s)(v-u^*(s))\mathrm{d}W(s)\right]P(t)G(t)\mathrm{d}\zeta(t)$$
$$I_3^{1,3}(\theta) = \frac{1}{\theta^2}\boldsymbol{E}\int_{[\tau,\tau+\theta]}\left[\int_\tau^t \Phi(s)\sigma_x(s)\mathrm{d}W(s)\right] \times\left[\int_\tau^t \Psi(s)\sigma_u(s)(v-u^*(s))\mathrm{d}W(s)\right]P(t)G(t)\mathrm{d}\zeta(t)$$

By [95, Eq-(3.23)], we have

$$\limsup_{\theta\to0^+} I_3^{1,1}(\theta) = \limsup_{\theta\to0^+}\frac{1}{\theta^2}\boldsymbol{E}\int_{[\tau,\tau+\theta]}\left[\int_\tau^t \Phi(\tau)\Psi(s)\sigma_u(s)(v-u^*(s))\mathrm{d}W(s)\right] \times P(t)G(t)\mathrm{d}\zeta(t) = 0. \tag{3.77}$$

Applying as in [95, p 2288], with the helps of Cuachy Schwartz inequality, we can prove that

$$\limsup_{\theta\to0^+} I_3^{1,2}(\theta) = \limsup_{\theta\to0^+}\frac{1}{\theta^2}\boldsymbol{E}\int_{[\tau,\tau+\theta]}\left[\int_\tau^t \Phi(s)f_x(s)\mathrm{d}s\right] \times\left[\int_\tau^t \Psi(s)\sigma_u(s)(v-u^*(s))\mathrm{d}W(s)\right]P(t)G(t)\mathrm{d}\zeta(t) = 0. \tag{3.78}$$

By [95, Lemma 4.1 and Eq-(4.10)], and Dominate convergence theorem, we have

$$\begin{aligned}\lim_{\theta\to0^+} I_3^{1,3}(\theta) &= \lim_{\theta\to0^+}\frac{1}{\theta^2}\boldsymbol{E}\int_{[\tau,\tau+\theta]}\left[\int_\tau^t \Phi(s)\sigma_x(s)\mathrm{d}W(s)\right] \\ &\quad\times\left[\int_\tau^t \Psi(s)\sigma_u(s)(v-u^*(s))\mathrm{d}W(s)\right]P(t)G(t)\mathrm{d}\zeta(t) \\ &= \lim_{\theta\to0^+}\frac{1}{\theta^2}\int_{[\tau,\tau+\theta]}\boldsymbol{E}\left[\int_\tau^t \sigma_x(s)\sigma_u(s)(v-u^*(s))\mathrm{d}s\right]P(t)G(t)\mathrm{d}\zeta(t) \\ &= \frac{1}{2}\boldsymbol{E}\left[P(\tau)G(\tau)\sigma_x(\tau)\sigma_u(\tau)(v-u^*(\tau))\zeta(\tau)\right].\end{aligned} \tag{3.79}$$

Substituting (3.77), (3.78), (3.79) into (3.76), we obtain

$$\lim_{\theta\to0^+} I_3^{1}(\theta) = \frac{1}{2}\boldsymbol{E}\left[P(\tau)G(\tau)\sigma_u(\tau)\sigma_x(\tau)(v-u^*(\tau))\zeta(\tau)\right]. \tag{3.80}$$

Estimate of (3.73). We proceed to estimate the second term $I_3^2(\theta)$. By Lemma 4.1 in [95], we have

$$\begin{aligned}\lim_{\theta\to0^+} I_3^{2}(\theta) &= \lim_{\theta\to0^+}\frac{1}{\theta^2}\boldsymbol{E}\int_{[\tau,\tau+\theta]}\Phi(t)P(t)G(t) \\ &\quad\times\left[\int_\tau^t \Psi(s)\left[f_u(s)-\sigma_x(s)\sigma_u(s)\right](v-u^*(t))\mathrm{d}s\right]\mathrm{d}\zeta(t) \\ &= \frac{1}{2}\boldsymbol{E}\left[P(\tau)G(\tau)\left[f_u(\tau)-\sigma_x(\tau)\sigma_u(\tau)\right](v-u^*(\tau))\zeta(\tau)\right].\end{aligned} \tag{3.81}$$

Estimate of (3.74). Applying [95, Lemma 4.1, Eq (3.21)], we obtain

$$\begin{aligned}\lim_{\theta\to0^+} I_3^{3}(\theta) &= \lim_{\theta\to0^+}\frac{1}{\theta^2}\boldsymbol{E}\int_{[\tau,\tau+\theta]}\Phi(t)P(t)G(t)\left[\int_\tau^t \Psi(s)G(s)\mathrm{d}\zeta(s)\right]\mathrm{d}\zeta(t) \\ &= \frac{1}{2}\boldsymbol{E}\left[P(\tau)G^2(\tau)\zeta^2(\tau)\right].\end{aligned} \tag{3.82}$$

By substituting (3.80), (3.81), (3.82) into (3.71), we have

$$\lim_{\theta \to 0^+} I_3(\theta) = \frac{1}{2}\boldsymbol{E}\left[P(\tau)G(\tau)(f_u(\tau) + \sigma_u(\tau))(v - u^*(\tau))\zeta(\tau)\right] + \frac{1}{2}\boldsymbol{E}\left[P(\tau)G^2(\tau)\zeta^2(\tau)\right]. \tag{3.83}$$

Now, by substituting (3.68), (3.70), (3.83) into (3.64), we have

$$\begin{aligned} 0 \geq\; & \frac{1}{2}\boldsymbol{E}\left[\mathbb{K}(\tau) f_u(\tau)(v - u^*(\tau))^2\right] \\ & + \frac{1}{2}\boldsymbol{E}\left[(\mathbb{K}(\tau) + P(\tau)(f_u(\tau)) G(\tau)(v - u^*(\tau))\zeta(\tau)\right] \\ & + \frac{1}{2}\boldsymbol{E}\left[P(\tau)G^2(\tau)\zeta^2(\tau)\right] + \frac{1}{2}\mathcal{D}^+_\tau(\mathbb{K}(\tau)\sigma_u(\tau)(v - u^*(\tau))^2), \end{aligned}$$

where $\frac{1}{2}\mathcal{D}^+_\tau(\mathbb{K}(\tau)\sigma_u(\tau)(v - u^*(\tau))^2)$ is given by (3.69). This completes the proof of Theorem 4.1 $\square$

Example. We show that a singular control via *Definition 2.2* does not necessary optimal. Let $\mathbb{A}_1 = [-1, 1]$ and $\mathbb{A}_2 = [0, +\infty)$. Consider the following SDEs :

$$\begin{cases} \mathrm{d}x^{u,\xi}(t) = u(t)\mathrm{d}t + u(t)\mathrm{d}W(t) + \mathrm{d}\xi(t), \\ x^{u,\xi}(0) = 0. \end{cases} \tag{3.84}$$

The expected cost to be minimized has the form :

$$J(u(\cdot), \xi(\cdot)) = \frac{1}{2}\boldsymbol{E}\int_0^1 |u(t)|^2\,\mathrm{d}t - \frac{1}{2}\boldsymbol{E}\left|x^{u,\xi}(1)\right|^2 + \boldsymbol{E}(\xi(1)). \tag{3.85}$$

Note that Eq-(3.84) has a unique strong solution $x^{u,\xi}(\cdot)$ given by $t \in [0, 1]$, $\xi(0) = 0$.

$$x^{u,\xi}(t) = \int_0^t u(s)\mathrm{d}s + \int_0^t u(s)\mathrm{d}W(s) + \xi(t).$$

The Hamiltonian function (3.4) gets the form

$$H(t,x,u,p,q) = p(t)u(t) + q(t)u(t) - \frac{1}{2}u^2(t).$$

Let $(u_*(\cdot), \xi_*(\cdot)) = (0,0)$, then the corresponding trajectory is $x_*(t) = x^{u_*,\xi_*}(t) = 0$. The corresponding adjoint processes are defined by the following adjoint equations :

$$\begin{cases} \mathrm{d}p_*(t) = q_*(t)\mathrm{d}W(t),\ t \in [0,1]. \\ p_*(1) = 0, \end{cases} \tag{3.86}$$

and

$$\begin{cases} \mathrm{d}P_*(t) = Q_*(t)\mathrm{d}W(t),\ t \in [0,1]. \\ P_*(1) = 1. \end{cases} \tag{3.87}$$

By a simple computations, the BSDEs (3.86) and (3.87) admits a unique strong $\mathbb{F}-$adapted solution $(p_*(t), q_*(t)) = (0,0)$ and $(p_*(t), q_*(t)) = (1,0)$.

By a simple computations, we have

$$\begin{aligned} H_u(t,x,u,p,q) &= p(t) + q(t) - u(t), \\ H_{\mathrm{u}u}(t,x,u,p,q) &= -1 \\ \sigma_u(t, x(t), u(t)) &= 1 \\ M(t) &= 1, \\ G(t) &= 1, \end{aligned}$$

then the admissible control $(u_*(\cdot), \xi_*(\cdot)) = (0,0)$ satisfies (3.9) and (3.10). Now, applying *Definition 2.2*, the control $(u_*(\cdot), \xi_*(\cdot)) \in \mathcal{A}_1 \times \mathcal{A}_2([0,1])$ is a singular control. Noting that $J(u_*(\cdot), \xi_*(\cdot)) = J(0,0) = 0$.

However, if we choose $u(t) = 1 \in [-1,1]$ and $\xi(t) = 0 \in [0,+\infty)$, and by a simple computation, we have

$$J(1,0) = -\frac{1}{2} < J(0,0) = 0.$$

This implied that the control $(u_*(\cdot), \xi_*(\cdot)) = (0,0)$ is not optimal for the control problem (3.84)-(3.85).

Conclusion, perspectives and future Developments

In this thesis, we establish a set of necessary conditions of optimal stochastic for different stochastic models. More precisely, in the second chapter, we have developed a necessary conditions for partially observed singular stochastic optimal control problem, where the controlled state dynamics is influenced by unobserved uncertainties. The system is governed by general McKean-Vlasov differential equations. By transforming the partial observation problem to a related problem with full information, a stochastic maximum principle for optimal singular control has been established via the derivative with respect to probability measure in P.Lions' sense. The main feature of these results is to explicitly solve some new mathematical finance problems such as general conditional mean-variance portfolio selection problem in incomplete market.

Apparently, there are many problems left unsolved :

1. One possible problem is to establish some optimality conditions (or near-optimality) for partially observed singular stochastic optimal control for systems governed forward-backward stochastic differential equations of general McKean-Vlasov type with some recent applications.

2. The partially observed singular control in the case when the control domain is not necessarily convex.

3. It would be quite interesting to derive a general maximum principle for partially observed optimal control for fully coupled forward-backward stochastic differential equations FBSEDs following Yong's maximum principle.

In the third chapter, pointwise second-order necessary conditions, in the form of Pontryagin maximum principal for optimal stochastic singular control have been established. The control dynamic system was governed by nonlinear controlled stochastic differential equation. In our class of control problem, we have studied two types of singularity, the predictable ones which come from the singular control part and the second ones which come from the irregularity in some senses.

We note that if the coefficients $G(t) = M(t) = 0$ our results coincides with second-order maximum principle developed in [95, Theorem 3.5]. Apparently, there are many problems left unsolved such as :

1. The case when the control domain is not assumed to be convex (general action space).

2. One possible problem is to study the second-order maximum principle for optimal singular control for McKean-Vlasov stochastic differential equations.

3. Another challenging problem left unsolved is to derive a various second-order maximum principles in the case where the coefficients G and M depend on the state of the solution process $x^{u,\xi}(\cdot)$.

4. It would be quite interesting to establish second order maximum principle for systems governed by forward-backward stochastic differential equations with some applications.

   We plane to study these interesting problems in forthcoming papers.

Bibliographie

[1] Abada N.E.H, Hafayed M, Meherrem S. : On Partially observed optimal singular control of McKean-Vlasov stochastic systems : maximum principle approach, *Mathematical Methods in the Applied Sciences,* Math Meth Appl Sci. 2022 ;1-21.DOI : 10.1002/mma.8373.

[2] Abada N.E.H , Hafayed M., Stochastic pointwise second-order maximum principle for optimal continuous-singular control using variational approach. IJMIC-345743 *International Journal of Modelling, Identification and Control.* Accept 2022.

[3] Ahmed NU. : Nonlinear diffusion governed by McKean-Vlasov equation on Hilbert space and optimal control. SIAM J. Control Optim. 46, (2007) 356–378.

[4] Alvarez L. : Singular stochastic control linear diffusion and optimal stopping : A class of solvable problems, SIAM J. Control Optim., 39 (2001) 1697-1710.

[5] Alvarez L. and T.A. Rakkolainen : On singular stochastic control and optimal stopping of spectrally negative jump diffusions, Stochastics An International Journal of Probability and Stochastics Processes 81(1) (2009) 55-78.

[6] Alvarez, L.H.R. : A class of solvable singular stochastic control problems. *Stoch. Stoch. Rep.* **67**, 83-122 (1999)

[7] An T.T K. : Combined optimal stopping and singular stochastic control, Stochastic Analysis and Applications, 28 (2010) 401-414.

[8] Andersson D, Djehiche B. : A maximum principle for SDEs of mean-field type, Appl Math Optim, **63**, 341-356 (2011)

[9] Baras, J.S., Elliott, R.J., Kohlmann, M. The partially observed stochastic minimum principle. *SIAM J. Control Optim.* **27**, 1279–1292. (1989)

[10] Bellman, R. : Dynamic programming, Princeton Univ. Press., (1957)

[11] Buckdahn, R., Li, J, Peng, S. : Mean-field backward stochastic differential equations and related partial differential equations. Stochastic Processes and their Applications, **119**, 3133-3154 (2009).

[12] Buckdahn, R., Djehiche, B., Li, J. : A general stochastic maximum principle for SDEs of mean-field type. Appl. Math. Optim. **64**, 197-216 (2011).

[13] Buckdahn R., Li, J., Ma J. : A stochastic maximum principle for general mean-field system, Appl Math Optim.(**74**) (2016) 507-534.

[14] Buckdahn R., Li J., Ma J. A stochastic maximum principle for general mean-field systems. *Appl. Math. Optim.* **74**(3), 507–534. (2016)

[15] Bensoussan, A. : *Lectures on stochastic control.* In : Lecture Notes in Mathematics, **972**,1-62, Springer, Berlin (1981).

[16] Bensoussan, A. Maximum principle and dynamic programming approaches of the optimal control of partially observed diffusions. Stochastic 9, 169-222. (1983)

[17] Bonnans J.F., Silva, : F.J. : First and second order necessary conditions for stochastic optimal control problems, *Appl. Math. Optim.*, **65**, 403-439 (2012)

[18] Borkar V. : Controlled diffusion processes, Probability Surveys, **2** (2005) 213-244.

[19] Cadenillas A., Haussman U. : The stochastic maximum principle for singular control problem, *Stochastics, Stochastics Rep.,* **49**, N 3-4, 211-237 (1994)

[20] Carmona, R., Delarue, F : Forward-backward stochastic differential equations and controlled McKean-Vlasov dynamics. The annals of Probability, **43**(5), (2015), 2647-2700.

[21] Cardaliaguet, P. : Notes on mean field games (from P.-L. Lions' lectures at Collège de France). https ://www.ceremade.dauphine.fr/cardalia/ (2013)

[22] Carrnona, R., Delarue, F., Lachapelle, A. : Control of McKean–Vlasov dynamics versus mean field games. Math. Financ. Econ. 7(2), (2013) 131–166.

[23] Djehiche B., Tembine H. Risk sensitive mean-field type control under partial observation. *Stochastics of Environmental and Financial Economics, Springer, Cham.,* 243-263,.(2016)

[24] Dufour, F., Miller B. : Maximum principle for singular stochastic control problem. *SIAM J. Control Optim.* **45**(2), 668-698 (2006)

[25] Dufour F., B. Miller B. : Singular stochastic control problem. *SIAM J. Control Optim.,* **43**(2), 705-730 (2004)

[26] Dufour F., Miller B. : Necessary conditions for optimal singular stochastic control problems, *Stochastics* **79**(5), 469-504 (2007)

[27] Dong Y, Meng Q.X. : Second-order necessary conditions for optimal control with recursive utilities, *J. Optim Theory Appl,* **182**(2), 494-524 (2019)

[28] Elliott R.J., Li X. and Ni. Y.H. : (2013) Discrete time mean-field stochastic linear-quadratic optimal control problems. Automatica, Vol 49, No 11, pp. 3222-3233.

[29] Fleming W.H. Optimal control of partially observable diffusions. *SIAM J. Control* **6**(2), 194-214. (1968)

[30] Frankowska, H., Zhang, H., Zhang, X. : First and second order necessary conditions for stochastic optimal controls. *Journal of Differential Equations* 262(6), 3689-3736 (2017)

[31] Guenane L., Hafayed M., Meherrem S., Abbas S. On optimal solutions of general continuous-singular stochastic control problem of McKean-Vlasov type. *Mathematical Methods in the Applied Sciences.* **43**(10), 6498-6516. (2020)

[32] Hafayed M., Meherrem S., Eren S. Guoclu D.H. : *On optimal singular control problem for general McKean-Vlasov differential equations : Necessary and sufficient optimality conditions,* Optim Control Appl Meth ; (39)1202–1219.(2018)

[33] Hafayed M, Abbas S. : On near-optimal mean-field stochastic singular controls : necessary and sufficient conditions for near-optimality, J. Optim Theory Appl, **160**(3), 778-808 (2014).

[34] Hafayed M., Abba A and Abbas S. : On mean-field stochastic maximum principle for near-optimal controls for Poisson jump diffusion with applications, Int. J. Dynam. Control, **2**, 262–284 (2014).

[35] Hafayed M, and Abbas S. : Stochastic near-optimal singular controls for jump diffusions : necessary and sufficient conditions, Journal of Dynamical and Control Systems, **19**(4), 503-517 (2013).

[36] Hafayed M., Abbas S., and Veverka P. : On necessary and sufficient conditions for near-optimal singular stochastic controls. Optim. Lett., (**7**)5, 949-966, (2013).

[37] Hafayed M., Veverka P., and Abbas S. : On maximum principle of near-optimality for diffusions with jumps, with application to Consumption-investment problem, Differ. Equ. Dyn. Syst. , **20**(2), 111-125 (2012).

[38] M. Hafayed, S. Abbas : Stochastic near-optimal singular controls for jump diffusions : Necessary and sufficient conditions, Journal of Dynamical and Control Systems, 19(4) (2013) 503-517.

[39] Hafayed M. : (2013) A mean-field necessary and sufficient conditions for optimal singular stochastic control, Commun. Math. Stat. Vol 1, No 4, pp. 417–435.

[40] Hafayed M. : (2013). A mean-field maximum principle for optimal control of forward-backward stochastic differential equations with Poisson jump processes, Int. J. Dynam. Control, Vol 1 No 4, pp. 300-315.

[41] Hafayed M. : (2014) Singular mean-field optimal control for forward-backward stochastic systems and applications to finance, Int. J. Dynam. Control, Vol 2 No 4, pp. 542–554

[42] Hafayed M, Abbas S. : (2013) A general maximum principle for stochastic differential equations of mean-field type with jump processes. Technical report, arXiv : 1301.7327v4.

[43] Hafayed M, Abbas, S., Abba A. : On mean-field partial information maximum principle of optimal control for stochastic systems with Lévy processes, J. Optim Theory Appl, 10.1007/s10957-015-0762-4, (2015).

[44] Hafayed M, Veverka P and Abbas A. : (2014) On Near-optimal Necessary and Sufficient Conditions for Forward-backward Stochastic Systems with Jumps, with Applications to Finance. Applications of Mathematics, Vol 59 No.4, pp. 407-440.

[45] Hafayed, M., Tabet. M., Boukaf S. : Mean-field maximum principle for optimal control of forward-backward stochastic systems with jumps and its application to mean-variance portfolio problem, Commun. Math. Stat (3) (2015) 163–186.

[46] Hafayed, M. Ghebouli M., Boukaf S., Shi Y. : Partial information optimal control of mean-field forward-backward stochastic system driven by Teugels martingales with applications (2016) DOI 10.1016/j.neucom. 2016.03.002. Neurocomputing Vol **200** pages 11–21 (2016).

[47] Hafayed M., Boukaf M., Shi Y. Meherrem S. : A McKean-Vlasov optimal mixed regular-singular control problem, for nonlinear stochastic systems with Poisson jump processes (2016) Neurocomputing. Doi 10.1016/j.neucom.2015.11.082, Volume **182**, 19, pages 133-144 (2016).

[48] Hafayed, M,.Abba A, Abbas S : On partial-information optimal singular control problem for mean-field stochastic differential equations driven by Teugels martingales measures, Internat. J. Control **89** (2016), no. 2, 397–410.

[49] Hafayed, M, Abba A Boukaf S : On Zhou's maximum principle for near-optimal control of mean-field forward-backward stochastic systems with jumps and its applications,"International Journal of Modelling, Identification and Control.25 (1), 1-16, (2016).

[50] Hafayed M., Meherrem. S., Eren, Ş., Guçoglu, D.H. On optimal singular control problem for general Mckean–Vlasov differential equations : necessary and sufficient optimality conditions. *Optim. Control Appl. Methods* **39**(3), 1202-1219. (2019)

[51] Hafayed M. Gradient generalisé et contrôle stochastique, Thèse de doctorat, Université Mohamed Khider, Biskra.(2009).

[52] Haiyan Zhang A necessary condition for mean-field type stochastic differential equations with correlated state and observation noises, *Journal of Industrial and Management Optimization,* **12**(4), 1287-1301. (2016)

[53] Hao T., Meng Q.X. : A second-order maximum principle for singular optimal controls with recursive utilities of stochastic delay systems, *European Journal of Control*, **50** 96–106 (2019)

[54] Hao T., Meng Q.X : Singular optimal control problems with recursive utilities of mean-field type, *Asian Journal of Control*, **23**(3) 1524-1535 (2021)

[55] Haussmann, U.G., Suo W. : Singular optimal control I, II, *SIAM J. Control Optim.,* **33**(3), 916-936, 937-959 (1995)

[56] Kushner, H.J. : Optimal stochastic control, IRE Trans. Auto. Control, AC-7 (1962), 120-122.

[57] Kushner, H.J. :. On the stochastic maximum principle : Fixed time of control, J. Math. Anal. Appl., 11 (1965), 78-92.

[58] Kac M. : Foundations of kinetic theory, Proc. 3-rd Berkeley Sympos. Math. Statist. Prob. **3** : 171-197 (1956)

[59] Lakhdari I.E., Miloudi H., Hafayed M. Stochastic maximum principle for partially observed optimal control problems of general McKean-Vlasov differential equations. *Bull. Iran. Math. Soc.*, **47**, 1021-1043. (2020)

[60] Lasry, J.M., Lions, P.L. : Mean field games. Japan Jour. Math. **2**, 229-260 (2007).

[61] Lions P.L. Cours au Collège de France : Théorie des jeu à champs moyens. *http ://www. college-de-france.fr/default/EN/all/equ[1]der/audiovideo.jsp.* (2013)

[62] Liptser, R.S., Shiryayev, A.N. Statistics of Random Process. Springer-Verlag, New York. (1977)

[63] Li, J. : Stochastic maximum principle in the mean-field controls. Automatica, **48**, 366-373 (2012)

[64] Li T., and Zhang J.F. : (2013) Adaptive mean field games for large population coupled ARX Systems with unknown coupling strength, Dyn Games Appl, Vol 3 : pp. 489–507.

[65] McKean, H.P. : A class of Markov processes associated with nonlinear parabolic equations. Proc. Natl. Acad. Sci. **56**, 1907-1911 (1966).

[66] Markowitz, H. : Portfolio selection, J. of Finance, **7** : 77-91 (1952).

[67] Meherrem S., Hafayed M. Maximum principle for optimal control of McKean–Vlasov FBSDEs with Lévy process via the differentiability with respect to probability law. *Optim. Control Appl. Methods,* **40**(3), 499-516. (2019)

[68] Miloudi H, Meherrem S, Lakhdari I.E, Hafayed M. Necessary conditions for partially observed optimal control of general McKean-Vlasov stochastic differential equations with jumps. *International Journal of Control,* Doi10.1080/00207179.2021.1961020. (2021)

[69] Mundaca G., ∅ksendal B. : Optimal stochastic intervention control with application to the exchange rate, *Journal of Mathematical Economics*, **29**, 225-243 (1988)

[70] Menaldi J. and M. Robin : On singular stochastic control problems for diffusions with jumps, IEEE Transactions on Automatic Control 29(11), (1984) 991-1004.

[71] Ni. Y.H, Zhang J.F., and Li X., (2014) Indefinite mean-field stochastic linear-quadratic optimal control, IEEE Transactions on automatic control, Doi : 10.1109/TAC.2014.2385253

[72] Ni. Y.H., Zhang J.F. and Li X. : Indefinite mean-field stochastic linear-quadratic optimal Control, IEEE Trans. Autom. Control, DOI : 10.1109/TAC.2014.2385253 (2014)

[73] ∅ksendal B., Sulem A. : Applied Stochastic Control of Jump Diffusions, Springer-Verlag Berlin Heidelberg (2005).

[74] Pham H. : Linear quadratic optimal control of conditional McKean-Vlasov equation with random coefficients and applications, Probability, Uncertainty and Quantitative Risk **1**(7), 1-26 (2016)

[75] Pham H. On some recent aspects of stochastic control and their applications, Probability Surveys Vol. 2 (2005) 506-549.

[76] Peng. S, A general stochastic maximum principle for optimal control problems, SIAM J. Control Optim., **28**, 966-979 (1990)

[77] Shen Y, Siu T K. The maximum principle for a jump-diffusion mean-field model and its application to the mean-variance problem, Nonlinear Analysis, **86**, 58-73 (2013)

[78] Shen Y, Meng Q, Shi P. : Maximum principle for mean-field jump-diffusions to stochastic delay differential equations and its applicationt to finance, Automatica **50**, 1565-1579 (2014)

[79] Shi J. : Sufficient conditions of optimality for mean-field stochastic control problems. 12^{-th} International Conference on Control, Automation, Robotics & Vision Guangzhou, China, 5-7^{th} December, ICARCV 2012, 747-752 (2012)

[80] Shreve S.E. An Introduction to singular stochastic control. stochastic differential systems, *Stochastic Control Theory and Applications*, **30** 513–528, (1988)

[81] Tang S. : A second-order maximum principle for singular optimal stochastic controls. *Discrete and continuous dynamical systems, Series B.,* **14**(4), 1581-1599 (2010)

[82] Tang M. : (2014) Stochastic maximum principle of near-optimal control of fully coupled forward-backward stochastic differential equation, Abstract and Applied Analysis Volume 2014, Article ID 361259, 12 pages

[83] Wang G., Wu Z., Xiong J. Maximum principles for forward-backward stochastic control systems with correlated state and observation noise. *SIAM Journal on Control and Optimization,* 51, 491–524. (2013)

[84] Wang G., Wu Z., Xiong J. A linear-quadratic optimal control problem of forward-backward stochastic differential equations with partial information. *IEEE Transactions on Automatic Control,* **60**, 2904-2916. (2015)

[85] Wang G., Wu Z. General maximum principles for partially observed risk-sensitive optimal control problems and applications to finance. *J. Optim. Theory Appl.* **141**(3), 677-700. (2009)

[86] Wang G., Zhang C., Zhang, W. Stochastic maximum principle for mean-field type optimal control with partial information. *IEEE Transactions on Automatic Control,* **59**, 522-528. (2014)

[87] Wang G., Xiao,H., Xing G. An optimal control problem for mean-field forward-backward stochastic differential equation with noisy observation, *Automatica* **86** (2017) 104-109. (2017)

[88] Wang G., Wu Z. A maximum principle for mean-field stochastic control syste with noisy observation. *Automatica* **137**(3), 110135. (2022)

[89] Wang M., Shi Q., Meng, Q. Optimal control of forward-backward stochastic jump-diffusion differential systems with observation noises : Stochastic Maximum Principle. *Asian Journal of Control.* **23** (1), 241-254 .(2021)

[90] Wang B.C. and Zhang J.F. : (2012).Mean-field games for larg-population multiagent systems with Markov jump parameters, SIAM J. Control. Vol 50 No. 4, pp. 2308–2334.

[91] Wu, Z and F. Zhang : Stochastic maximum principle for optimal control problems of forward-backward systems involving impulse controls, IEEE Transactions on Automatic Control, 56(6) (2011) 1401-1406.

[92] Yong J, Zhou X.Y. : Stochastic Controls. Hamiltonian Systems and HJB Equations. Springer-Verlag. New York, (1999)

[93] Yong J. : Optimality variational principle for controlled forward-backward stochastic differential equations with mixed intial-terminal conditions. SIAM J. Control Optim. **48**(6), 4119-4156 (2010)

[94] Yong J. : A linear-quadratic optimal control problem for mean-field stochastic differential equations. SIAM J. Control Optim. Vol 51 No 4, 2809-2838. (2013)

[95] Zhang H., Zhang X. : Pointwise second-order necessary conditions for stochastic optimal controls, Part I : The case of convex control constraint, *SIAM J. Control Optim.* **53**(4), 2267-2296 (2015)

[96] Zhang, H., Zhang, X. : Pointwise second-order necessary conditions for stochastic optimal controls, part II : The general case. *SIAM J. Control Optim.,* **55**(5), 2841-2875 (2017)